

**UNIVERSIDAD COMPLUTENSE DE MADRID**

FACULTAD DE FARMACIA

Departamento de Bioquímica y Biología Molecular II



**TESIS DOCTORAL**

**SÍNDROMES HEREDITARIOS DE CÁNCER DE MAMA  
FAMILIAR:  
VARIANTES DE SIGNIFICADO CLÍNICO INCIERTO Y  
CONSEJO GENÉTICO**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Gorka Ruiz de Garibay Ponce**

Director

Miguel de la Hoya Mantecón

**Madrid, 2014**

**UNIVERSIDAD COMPLUTENSE DE MADRID**

FACULTAD DE FARMACIA

Departamento de Bioquímica y Biología Molecular II



**SÍNDROMES HEREDITARIOS DE CÁNCER DE MAMA FAMILIAR:  
VARIANTES DE SIGNIFICADO CLÍNICO INCIERTO Y CONSEJO  
GENÉTICO**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

**GORKA RUIZ DE GARIBAY PONCE**

Director:

Dr. Miguel de la Hoya Mantecón

**Madrid, 2014**



Don Miguel de la Hoya Mantecón, Doctor en Ciencias e Investigador I3-SNS del Instituto de Investigaciones Sanitarias San Carlos (IdISSC)

**CERTIFICA:**

Que el trabajo titulado: “Síndromes hereditarios de cáncer de mama familiar: variantes de significado clínico incierto y consejo genético” que presenta el Licenciado en Bioquímica D. Gorka Ruiz de Garibay Ponce, 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 Complutense de Madrid.

Para que así conste, y a los efectos oportunos, firma la presente en Madrid a 12 de Diciembre de 2013.

Dr. Miguel de la Hoya Mantecón



A quién bautizó a esta tesis como *“la tesis más larga del mundo”*,

Gracias al esfuerzo, sacrificio, inteligencia, paciencia y trabajo, este proyecto ha llegado a  
buen puerto.



A menudo, en ciencia los mejores descubrimientos suceden al azar y de manera totalmente inesperada. Esto es lo que me sucedió a mí cuando decidí presentarme a los exámenes del BIR. Descubrí un hospital, el Clínico San Carlos, totalmente desconocido para mí. Gracias al azar me formé en su Servicio de Inmunología y tuve la magnífica oportunidad de realizar investigación en el Laboratorio de Oncología Molecular.

En primer lugar, me gustaría agradecer a la Institución del Hospital Clínico por promover y favorecer que los residentes realicemos investigación como parte de nuestra formación. Gracias a ello, una gran cantidad de los compañeros de residencia hemos iniciado una carrera científica. Espero que en un futuro, como especialistas y doctores, muchos de nosotros seamos capaces de dirigir grandes equipos investigadores desde los mejores hospitales del mundo.

Desearía agradecer a la Dra. Trinidad Caldés y al Dr. Eduardo Diaz-Rubio por haberme acogido en el Laboratorio de Oncología Molecular durante estos años y ayudado en este proyecto.

Me gustaría agradecer al Servicio de Inmunología por facilitar unas condiciones que favorecen que los residentes decidamos realizar investigación. En mi caso en particular, quisiera agradecer a mi tutor de residentes, por presentarme como accesibles una diversidad de líneas de investigación del hospital.

Quiero agradecer a la Dra. Conchi Nuñez por haberme puesto en contacto con mi director y por su apoyo y consejos durante todos estos años.

Mi mayor agradecimiento al Dr. Miguel de la Hoya por la confianza depositada en mí desde el inicio y por su apoyo. Gracias por haberme enseñado, orientado e implicado en este proyecto a pesar de toda la carga de trabajo que mantienes de manera rutinaria. Gracias por haber logrado que esta colaboración haya dado unos grandes frutos. Por tu generosidad. Por haber comprendido que las obligaciones como residente quitan tiempo a la investigación. He de reconocer que durante mucho tiempo he dudado de que fuera capaz de realizar una tesis completa y estoy seguro que sin tu dirección no habría sido posible.

Me gustaría agradecer también al resto de miembros del Laboratorio de Oncología Molecular, en particular a aquellos que han tenido una implicación más directa en el trabajo de esta tesis. Gracias a Trinidad Caldés por los consejos recibidos y la confianza depositada en mí. Gracias a Alicia, por tu gran trabajo con las muestras, con las familias y por tu ayuda con la extracción de muestras de ARN. Gracias a Paula, por haberme permitido trabajar con tus materiales, sabiendo lo meticulosa que eres en tu trabajo. Gracias a Beatriz por tus consejos y por acogerme en tu espacio durante la redacción de mi tesis. También quería agradecer a Pedro y a Helena por permitirnos el acceso a las historias familiares y por mediar con los pacientes para la recolección de muestras de RNA.

Quisiera agradecer a todos los pacientes que han decidido colaborar con nuestra investigación, sin vuestra cooperación la mayoría de nuestros estudios no habrían sido posibles.

No me gustaría olvidarme de agradecer a todo el personal del Servicio de Inmunología. A todos los adjuntos y técnicos por haber comprendido que los residentes que hacemos investigación necesitamos disponer de algo más de tiempo para esta tarea. A todos los residentes del Servicio de Inmunología con los que he coincidido durante mi formación y a los residentes consortes, ya que a pesar de que todos somos diversos, habéis logrado ser como una pequeña familia para mí. También a todos los becarios del servicio, que a pesar de no formar parte del mismo equipo de investigación me habéis acogido como uno más. Ana, Manmen, Carlos, Laura, Arturo, Luzma, Romina, Miguel Angel, Felix, Lidia, Virginia, Maria. Barbara, Nieves, María y Belén.

A todos los profesionales que han colaborado en mi formación como residente y como investigador. Al grupo de Linfomas del CNIO y en especial a Miguel Ángel Piris. A la Sección de Inmunología Celular del Hospital Ramón y Cajal y en especial a Ernesto Roldán. A la Dra. Olopade, al Dr. Fackenthal y a Thosio Yoshimatsu de la Universidad de Chicago por acogerme durante mi estancia de investigación.

A Francesca Consolaro, por prestarme su acceso a la biblioteca de la Universidad de Chicago, ya que la mayoría de la documentación de la Tesis la he obtenido a través de ella.

A todos los residentes de International House y a Evelyn, por hacer mucho más acogedora mi estancia en Chicago.

A Spotify, Groovespark y Rdio por proporcionar la banda sonora que me ha permitido aguantar tantas horas de trabajo. A todas las demos que siguieron funcionando a pesar de estar caducadas y a mi viejito ordenador, que ha aguantado a pesar de recalentarse. A Dropbox, sin este servicio habría perdido toda la tesis a mitad de trabajo.

A todas las personas del IDIBELL que me han dado consejos durante el proceso de redacción de esta tesis, en especial a Maria y a Dori. A Miquel Àngel Pujana, por permitirme durante todo este tiempo dedicar parte del tiempo a la escritura de mi tesis.

A la ciudad de Madrid, por ser una ciudad que acoge e integra a cualquiera de las personas que vive en ella. Por estar llena de posibilidades a nivel científico y también a nivel personal. A todas las personas que han contribuido que Madrid sea como mi segunda casa, en especial a Roberto.

A mi familia por todos los esfuerzos invertidos en mi formación, por haber sufrido el tener al hijo lejos (incluso en Estados Unidos), por inculcarme el esfuerzo, por respetar mis decisiones y por acompañarme durante este proceso. Mari Carmen, Santiago, Dani, Dina y Ayannita (eres la nueva ilusión de la familia).

A mi abuelo Jesús por ser un gran referente.

# Índice

<b>ÍNDICE</b>	<b>1</b>
<b>ÍNDICE DE FIGURAS</b>	<b>4</b>
<b>ÍNDICE DE TABLAS</b>	<b>5</b>
<b>GLOSARIO DE ABREVIATURAS</b>	<b>7</b>
<b>RESUMEN</b>	<b>17</b>
<b>SUMMARY</b>	<b>19</b>
<b>INTRODUCCIÓN</b>	<b>21</b>
1. EPIDEMIOLOGÍA DEL CÁNCER DE MAMA	21
2. CÁNCER DE MAMA FAMILIAR	23
<i>Síndrome de cáncer de mama-ovario hereditario (HBOC)</i>	24
<i>Otros síndromes</i>	25
<i>Epidemiología del cáncer de mama familiar.</i>	27
3. GENÉTICA DE LA SUSCEPTIBILIDAD AL CÁNCER DE MAMA	29
4. BRCA1 Y BRCA2 EN LA SUSCEPTIBILIDAD AL CÁNCER DE MAMA	33
<i>Estructura de las proteínas BRCA1 y BRCA2.</i>	33
<i>Funciones de las proteínas BRCA1 y BRCA2</i>	36
<i>Variantes genéticas en BRCA1 y BRCA2.</i>	39
<i>Mutaciones recurrentes y efecto fundador</i>	40
<i>Penetrancia de las mutaciones en BRCA1 y BRCA2.</i>	41
<i>Implicación de BRCA1 y BRCA2 en la carcinogénesis</i>	43
5. BÚSQUEDA DE GENES ADICIONALES EN LA SUSCEPTIBILIDAD AL CÁNCER DE MAMA	45
<i>Primeros estudios de resecuenciación de genes candidatos</i>	45
<i>Segunda generación de estudios de resecuenciación de genes candidatos.</i>	46
<i>Tercera generación: secuenciación de exomas</i>	47
<i>Estudios de asociación: Genes de baja penetrancia</i>	48
6. CONSEJO GENÉTICO EN CÁNCER DE MAMA FAMILIAR	50
<i>Consideraciones previas a la realización de un test genético</i>	50
<i>Realización de un test genético</i>	51
7. CLASIFICACIÓN DE VARIANTES GENÉTICAS EN BRCA1/2	53
<i>Evidencias directas: análisis de la cosegregación</i>	53
<i>Evidencias directas: estudios caso-control</i>	54
<i>Evidencias directas: historia personal y familiar</i>	54
<i>Evidencias directas: Co-ocurrencia</i>	55

<i>Evidencias indirectas: análisis in silico del efecto de los cambios de aminoácido</i>	55
<i>Evidencias indirectas: análisis in silico y experimental del splicing</i>	56
<i>Evidencias indirectas: Ensayos funcionales in vitro</i>	58
<i>Modelo probabilístico multifactorial</i>	58
<b>OBJETIVOS</b>	<b>63</b>
<b>1. ANÁLISIS DE LA RELEVANCIA DEL ESTUDIO DEL <i>SPLICING</i> EN LA CLASIFICACIÓN DE VARIANTES GENÉTICAS EN <i>BRCA2</i></b>	<b>65</b>
<i>Análisis mediante electroforesis capilar de ensayos convencionales de splicing: Clasificación analítica y clínica de 31 variantes en el gen BRCA2 mediante las normas IARC.</i>	65
<b>2. BÚSQUEDA DE VARIANTES RELACIONADAS CON EL SÍNDROME HBOC EN FAMILIAS BRCA1/2</b>	<b>103</b>
2.A BÚSQUEDA DE VARIANTES PATOGENICAS EN EL GEN BRCA2 QUE NO SE DETECTAN MEDIANTE LAS TÉCNICAS BASADAS EN PCR	105
<i>Caracterización de cuatro nuevos grandes reordenamientos genómicos en el locus BRCA2 en familias españolas con cáncer de mama/ovario. Revisión de la literatura y reevaluación de los mecanismos genéticos involucrados en su origen.</i>	105
<i>Análisis de la mutación c.6937+594T&gt;G en el gen BRCA2 en una cohorte de 302 familias HBOC españolas negativas para mutación en BRCA1/2</i>	124
2.B BÚSQUEDA DE VARIANTES PATOGENICAS IMPLICADAS EN HBOC EN GENES CANDIDATOS ADICIONALES A BRCA1/2	133
<i>Baja prevalencia de mutaciones inactivantes de SLX4 en familias con cáncer de mama y/o ovario no portadoras de mutación en BRCA1/2.</i>	133
<b>DISCUSIÓN</b>	<b>143</b>
RELEVANCIA DEL ESTUDIO DEL <i>SPLICING</i> EN LA CLASIFICACIÓN DE VARIANTES GENÉTICAS EN <i>BRCA2</i>	143
Aspectos Metodológicos	143
Clasificación de variantes mediante el análisis de <i>splicing</i> .	145
BÚSQUEDA DE VARIANTES RELACIONADAS CON EL SÍNDROME HBOC EN FAMILIAS BRCA1/2	147
<i>Búsqueda de variantes patogénicas en el gen BRCA2 que no se detectan mediante las técnicas basadas en PCR</i>	147
Grandes reordenamientos genómicos en el gen BRCA2	147
Variantes intrónicas alejadas del exón	149
<i>Sobre la búsqueda de variantes patogénicas implicadas en HBOC en genes candidatos adicionales a BRCA1/2</i>	151
SLX4	151
RAD51D	159
PERSPECTIVAS FUTURAS	163
<b>CONCLUSIONES</b>	<b>167</b>
<b>ANEXO I</b>	<b>169</b>

<i>Comparación de protocolos de análisis del splicing del ARNm entre múltiples laboratorios: Guía de buenas prácticas para el análisis clínico estandarizado.</i>	169
<i>Aproximadamente el 1% de las familias españolas con cáncer de mama y ovario negativas para BRCA1 y BRCA2 son portadoras de variantes patogénicas en RAD51D</i>	187
<b>ANEXOII</b>	<b>205</b>
<b>REFERENCIAS</b>	<b>219</b>

## Índice de figuras

<b>Figura 1:</b> Descripción del primer árbol genealógico con agregación de cáncer de mama.....	23
<b>Figura 2:</b> Modelo poligénico de la susceptibilidad al cáncer de mama .....	30
<b>Figura 3:</b> Secuencias repetitivas en los genes BRCA1 y BRCA2 .....	33
<b>Figura 4:</b> Dominios de la proteína BRCA1 .....	34
<b>Figura 5:</b> Dominios de la proteína BRCA2 .....	35
<b>Figura 6:</b> BRCA1 forma 4 complejos macromoleculares .....	37
<b>Figura 7:</b> Esquema de la ruta de anemia de Fanconi .....	38
<b>Figura 8:</b> Penetrancia de cáncer de mama y ovario en portadores de mutación en BRCA1/2 .....	43
<b>Figura 9:</b> Funciones del complejo MRN.....	45
<b>Figura 10:</b> Señales que definen los límites intrón-exón.....	57
<b>Figura 11:</b> Señales potenciadoras y silenciadoras del splicing .....	57
<b>Figura 12:</b> Localización de las variantes intrónicas en el gen BRCA2 presentes en las base de datos BIC y dbSNP. .....	150

## Índice de tablas

<b>Tabla 1:</b> Probabilidad a priori según el modelo multifactorial.....	59
<b>Tabla 2:</b> Clasificación de variantes propuesta por IARC.....	61
<b>Tabla 3:</b> Reordenamientos genómicos en el gen BRCA1 descritos en población española. ....	148
<b>Tabla 4:</b> Predicciones y estudios funcionales en las variantes de SLX4 descritas por primera vez en casos HBOC. .....	153
<b>Tabla 5:</b> Descripción de las variantes raras en SLX4 presentes en las series de HBOC y en ESP6500. ....	155
<b>Tabla 6 :</b> Predicciones y estudios funcionales en las variantes de RAD51D descritas por primera vez en casos HBOC y esporádicos .....	160
<b>Tabla 7:</b> Descripción de las variantes raras en RAD51D presentes en las series de HBOC/casos esporádicos y en ESP6500.....	161
<b>Tabla 8:</b> Frecuencia de las mutaciones encontradas en el gen RAD51D.....	162
<b>Tabla 9:</b> Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios.....	205
<b>Tabla 10:</b> Variantes encontradas en el gen RAD51D en casos HBOC y esporádicos en siete estudios.....	215



## Glosario de abreviaturas

BRCA1/2: Familias con síndrome HBOC no portadoras de variante patogénica en los genes BRCA1/2

DSB: Roturas de doble cadena (Double Strand Break)

eQTL: Expression Quantitative Trait Locus (Locus regulador de la expresión)

DHPLC: Denaturing High Pressure Liquid Chromatography

ENIGMA: Evidence-based Network for the Interpretation of Germline Mutant Alleles

ENT: Enfermedades no transmisibles

ESE: Exonic Splicing Enhancer (Potenciador del splicing exónico)

ESS: Exonic Splicing Silencer (Silenciador del splicing exónico)

HBOC: Hereditary breast and ovarian cancer (cáncer de mama-ovario hereditario)

HR: Recombinación Homóloga (Homologous Recombination)

HRM: High Resolution Melting

ISE: Intronic Splicing Enhancer (Potenciador del splicing intrónico)

ISS: Intronic Splicing Silencer (Silenciador del splicing intrónico)

OCCR: Ovarian Cancer Cluster Region

NCCN: National Comprehensive Cancer Network

NHEJ: Unión de extremos no homólogos (Non-Homologous End Joining)

NICE: National Institute for Health and Clinical Excellence

SEOM: Sociedad Española de Oncología Médica

UKCFSG: UK Cancer Family Study Group

SNP: Single Nucleotide Polymorphism

sQTL: splicing quantitative trait locus (locus que regula de manera cuantitativa los niveles de splicing alternativo)

UTR: Untranslated Region (region del RNA localizada en 5' o 3' y que no es traducida a proteína)

UV: Unclassified Variant



El síndrome de cáncer de mama y ovario hereditario es el síndrome de predisposición hereditaria al cáncer más frecuente, con una incidencia estimada de 1/500 – 1/2500. Está caracterizado por la agregación familiar de múltiples casos de cáncer de mama y/o de ovario con un patrón de herencia autosómica dominante. El síndrome es genéticamente complejo y actualmente se explica mediante un modelo poligénico, según el cual variantes genéticas con distinto grado de penetrancia y de prevalencia en la población darían lugar a un fenotipo común, el riesgo a padecer cáncer de mama y/o de ovario.

Debido a la alta prevalencia y penetrancia de sus alteraciones, los genes *BRCA1* y *BRCA2* se consideran los principales factores genéticos del síndrome y se han incorporado mediante el consejo genético a la práctica clínica. Variantes deletéreas en estos genes explican aproximadamente el 20% de los casos que acuden a las unidades clínicas de cáncer familiar. Sin embargo, un porcentaje variable comprendido entre el 5 al 20 % de los casos dependiendo de la población, es portador de variantes de significado clínico desconocido en estos mismos genes, con frecuencia denominadas UVs (del inglés Unclassified Variants). Este hecho representa un reto para la aplicación del consejo genético.

Se han descrito variantes deletéreas en genes adicionales que interaccionan con *BRCA1/2* o están implicados en las mismas rutas moleculares que *BRCA1/2*, como por ejemplo *RAD51C*, *PALB2*, o *BRIP1*. Aunque, individualmente, variantes en estos genes explican un porcentaje muy bajo de casos, en total explicarían un porcentaje destacable pero difícil de estimar.

Con el objetivo de mejorar el proceso de consejo genético, en esta tesis se han analizado diversos factores genéticos que contribuyen al síndrome de cáncer de mama y ovario hereditario en población española. Por una parte, se ha analizado la relevancia clínica de UVs en el gen *BRCA2* encontradas durante el proceso de consejo genético en el Hospital Clínico San Carlos. El estudio se ha llevado a cabo mediante el análisis de alteraciones en el proceso de *splicing*.

Por otra parte, en aquellas familias sin defectos detectables en los genes *BRCA1* y *BRCA2* tras el estudio de mutaciones puntuales, se han llevado a cabo estudios adicionales. En primer lugar se ha analizado el papel de los reordenamientos genómicos en el locus *BRCA2* en población española. Además, se ha analizado la presencia, en población española, de la variante patogénica intrónica c.6937+594T>G. En segundo lugar, se ha analizado la posible contribución de genes adicionales al síndrome en población española. En concreto se han resecuenciado los genes *SLX4* y *RAD51D* en aquellos casos no explicados por variantes en *BRCA1/2*.

Los estudios llevados a cabo en esta tesis han permitido:

1. Clasificar como variantes claramente patogénicas en el gen *BRCA2* a c.7617+1G>A, c.682-2A>G y c.8954-5A>G.
2. Identificar reordenamientos genómicos en el locus *BRCA2* en 7 de 813 familias española con síndrome HBOC previamente negativas para pequeñas alteraciones genómicas (substituciones e inserciones/deleciones).
3. Describir una variante claramente inactivante de la proteína *SLX4* (p.Glu1517\*) y un cambio de aminoácido (p.Arg372Trp) que presentó cosegregación con los casos de cáncer.
4. Describir dos variantes claramente inactivantes de la proteína *RAD51D*, p.Arg231\* y p.Val193Alafs\*4 y una probablemente inactivante, c.1A>T p.Met1?.

Los estudios realizados en esta tesis demuestran la existencia de variantes genéticas causales previamente no detectadas tanto en los genes de susceptibilidad principales, *BRCA1/2*, como en genes adicionales. Este resultado sugiere que la introducción de nuevas técnicas en el cribado de variantes genéticas, la inclusión de nuevos genes y una mejora en las técnicas utilizadas para la clasificación de variantes, podrían producir una mejora en el consejo genético que se realiza actualmente en familias con el síndrome de cáncer de mama y ovario hereditario.

## Summary

Hereditary breast and ovary cancer syndrome (HBOC) is the most frequent hereditary cancer predisposition syndrome, with a stimated incidence of 1/500 – 1/2500. It is characterized by a familial aggregation of multiple breast/ovary cancer cases with a pattern of autosomic dominant inheritance. HBOC is genetically complex and nowadays is defined by a polygenic model. In this model genetic variants with different degree of prevalence and penetrance in a population explain a common phenotype, the susceptibility to breast and ovarian cancer.

Due to the high prevalence and penetrance of their variants, *BRCA1* and *BRCA2* are considered the main genetic factors of the syndrome and have been incorporated to the clinical practice of genetic counseling. Deleterious variants in these genes account approximately for 20% of the cases attending familial cancer clinical units. However, a variable proportion of cases, ranging between 5 to 10% according to the population, are carriers of unclassified variants (UV), which poses a challenge to clinical counselors.

Deleterious variants have been described in additional genes that either interact with *BRCA1/2* or play a role in the same pathways, such as *RAD51C*, *PALB2*, or *BRIP1*. Individually these genes explain a very low proportion of cases, but as a whole could explain a remarkable but difficult to stimate percentage of the HBOC cases.

With the aim of improving the process of genetic counseling, this thesis has analysed some genetic factors that contribute to HBOC in Spanish population.

On the one hand, the clinical relevance of the *BRCA2* UVs found during the process of genetic counselling was analyzed. The study was carried out through the analysis of defects on the *splicing* process. On the other hand, in the families without a genetic explanation for the syndrome (BRCAX), additional studies have been performed with the objective of the identification of the causal variant. First, novel genetic techniques have been applied to the process of *BRCA2* screening. Second, the putative contribution of additional genes to HBOC syndrome has been analyzed in the Spanish population. *SLX4* and *RAD51D* genes have been resequenced in BRCAX cases.

The studies performed in this thesis have contributed to:

1. Classification of the *BRCA2* c.7617+1G>A, c.682-2A>G and c.8954-5A>G variants as clearly pathogenic.

2. Identification of genomic rearrangements in *BRCA2* locus in 7 of 813 spanish HBOC families, previously negative for small genomic alterations (substitutions and insertions/deletions)
3. Description of a clearly deleterious *SLX4* variant (p.Glu1517\*) and a missense (p.Arg372Trp) that cosegregated with the cancer cases.
4. Description of two clearly deleterious *RAD51D* variants, p.Arg231\* and p.Val193Alafs\*4 and a probably deleterious variant, c.1A>T p.Met1?.

The studies performed in this thesis prove the existence of previously non described deleterious genetic variants either in the mayor susceptibility genes, *BRCA1/2*, and in additional genes. Theses results suggest that the introduction of novel screening techniques and the inclusion of additional genes could contribute to an improvement of the genetic counseling process currently offerd to HBOC families.

### 1. Epidemiología del cáncer de mama

De los 57 millones de defunciones acaecidas en el año 2008, se calcula que 36 millones (63%) se deben a enfermedades no transmisibles (ENT)(WHO, 2011). De acuerdo a las estadísticas mundiales sanitarias de la OMS del año 2012, se prevé que el número de defunciones anuales por ENT alcance los 55 millones en el año 2030. Las enfermedades cardiovasculares son la principal causa de defunción por ENT (48%), seguidas del cáncer (21%) y de las enfermedades respiratorias crónicas (12%) (WHO, 2012). Actualmente, el cáncer es la principal causa de defunción en los países desarrollados y la segunda en países en vías de desarrollo. La incidencia de cáncer en los países en vías de desarrollo está experimentando un incremento como resultado del aumento poblacional y del envejecimiento de la población, así como de la adopción de estilos de vida que predisponen al riesgo; como son: el consumo de tabaco, la inactividad física y la adopción de una dieta occidentalizada(Jemal et al., 2011). En el año 2008, el cáncer de mama fue el cáncer más frecuentemente diagnosticado en mujeres, tanto en países desarrollados como en países en desarrollo. Así mismo, el cáncer de mama fue la principal causa de defunción por cáncer entre las mujeres (Jemal et al., 2011).

Se estima que en el año 2008 en Europa, se diagnosticaron 3.2 millones de casos nuevos de cáncer, de los cuales el 47% afectó a mujeres (1.5 millones.) La incidencia del cáncer de mama fue de 421,000 nuevos casos. La incidencia del cáncer de mama en Europa presenta una tendencia al alza, pasándose de una tasa de incidencia ajustada por edad de 76 en el año 1995 a 88.4 en el año 2008 (Ferlay et al., 2010). La mortalidad por cáncer de mama se estima para el año 2012 en 14.9 de cada 100.000 mujeres, registrándose un descenso del 9% respecto a la observada en el año 2007 ((Malvezzi et al. (2012)). Este descenso en la mortalidad podría estar explicado por la introducción de estrategias de cribado, que permiten un diagnóstico temprano y por mejoras en el tratamiento (Bray et al., 2004; Hery et al., 2008). En España se diagnostican unos 22.000 casos al año, lo que representa casi el 30% de todos los tumores diagnosticados en mujeres. La mayoría de los casos se diagnostican entre los 35 y los 80 años, con un máximo entre los 45 y los 65. La tasa de mortalidad por cáncer de mama estimada para el año 2012 en España es 11,74, la más baja de todos los países Europeos (Malvezzi et al., 2012).

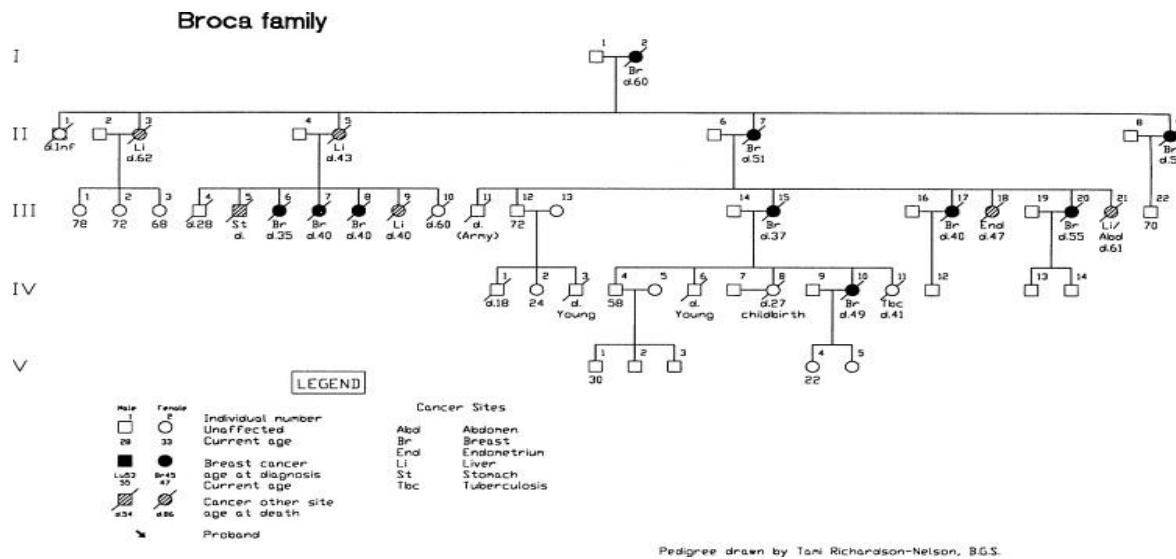
Los estudios epidemiológicos han sido útiles en la identificación de factores que incrementan el riesgo de desarrollar cáncer de mama. La probabilidad de desarrollar cáncer de

mama aumenta con la edad, y el riesgo acumulado a la edad de 70 años se cuantifica en 12% (Sasieni et al., 2011). El riesgo también se encuentra influenciado por factores hormonales. Un incremento en la exposición a estrógenos, bien sea debido a una edad temprana de menarquia (García-Closas et al., 2006), una menopausia tardía (Cancer, 1997) o una edad tardía del primer embarazo (Collaborative Group on Hormonal Factors in Breast, 2002), incrementa el riesgo. Como ejemplo, las mujeres nulíparas presentan un incremento del riesgo de 30% (Brewster and Helzlsouer, 2001). El uso de terapia hormonal sustitutiva también es un factor de riesgo, aunque únicamente se ha encontrado asociación en aquellas terapias que incluyen la combinación de estrógenos y progesterona (Chlebowski et al., 2003). Entre los factores de riesgo modificables el consumo de grasa (Boyd et al., 2003), alcohol (Stuckey, 2011) y la obesidad (Reeves et al., 2007) están relacionados con un incremento del riesgo. A modo de ejemplo, un estudio publicado en el año 2011 estimó que aproximadamente 9% de los cánceres de mama diagnosticados en el Reino Unido estuvieron relacionados con el sobrepeso. En la misma línea, el ejercicio moderado y regular presenta un efecto protector (Stuckey, 2011).

La historia personal también ejerce una influencia sobre el riesgo. En mujeres jóvenes diagnosticadas de cáncer de mama, se ha demostrado un incremento en el riesgo de desarrollar un segundo cáncer de mama primario de 4,5 veces y de 3 veces de desarrollar cualquier otro tipo de cáncer (Stuckey, 2011). En cuanto a la historia familiar, se estima que una mujer con un familiar de primer grado diagnosticado de cáncer de mama, tiene el doble de riesgo y este riesgo se incrementa a medida que aumentan el número de familiares afectados o si algunos de los casos fue diagnosticado a edad inferior a 50 años (Pharoah et al., 1997). Sin embargo, el 85% de las mujeres con familiares diagnosticados con cáncer de mama nunca desarrollará la enfermedad y el 85% de las mujeres que padecen cáncer de mama no poseen ningún tipo de historia familiar (Collaborative Group on Hormonal Factors in Breast, 2001). En países desarrollados, se estima que los factores hereditarios contribuyen a una cuarta parte de las diferencias interindividuales en la susceptibilidad al cáncer de mama, los factores medioambientales y el tipo de vida, contribuyen al resto de la susceptibilidad (Key et al., 2001).

## 2. Cáncer de mama familiar

La primera descripción de un árbol genealógico familiar con presencia de cáncer de mama fue realizada en el año 1866 por el médico Paul Broca. En este estudio, realizó un seguimiento de la causa de fallecimiento en la familia de su mujer a lo largo de 5 generaciones (Figura 1). La evidencia del fallecimiento de 10 de las 24 mujeres por cáncer de mama le hizo sugerir la presencia de algún factor hereditario de predisposición al cáncer en la familia.



**Figura 1:** Descripción del primer árbol genealógico con agregación de cáncer de mama

Descripción de la familia de Paul Broca en la cual se evidencia por primera vez la transmisión hereditaria del cáncer de mama a lo largo de 5 generaciones. Imagen adaptada de (Lynch and Lynch, 2002)

A mediados del siglo pasado numerosas publicaciones recogieron la existencia de agrupación familiar de casos de cáncer de mama. En algunas familias la agregación fue únicamente de casos de cáncer de mama (Jacobsen and Fraser, 1946; Lynch, 1967; Macklin, 1959; Stephens et al., 1958), mientras que en otras se acompañaba de otros tipos de cáncer como el de colon, ovario, endometrio, estomago, sarcomas y leucemias (Lynch et al., 1973; Lynch et al., 1972). En 1971 Lynch y Krush (Lynch and Krush, 1971) describieron tres familias con agregación de cáncer de mama y ovario y sugirieron una herencia compatible con un factor autosómico dominante. Posteriores estudios epidemiológicos (Go et al., 1983; Newman et al., 1988; Williams and Anderson, 1984) ofrecieron suficientes evidencias de la existencia de genes de susceptibilidad al cáncer de mama con herencia autosómica dominante. El mapeo (Hall et al., 1990; Wooster et al., 1994) y descubrimiento de los genes *BRCA1* (Miki et al., 1994) y *BRCA2* (Wooster et al., 1995) confirmó la existencia de una base genética en la susceptibilidad al cáncer de mama. Por otro lado,

estudios en gemelos determinaron que la mayor parte de la agregación familiar está más influida por la susceptibilidad genética hereditaria que por factores ambientales (Mack et al., 2002; Peto and Mack, 2000). Sin embargo, el análisis del archivo histórico de la Universidad de Creighton demuestra heterogeneidad tanto genotípica como fenotípica. De las 100 familias estudiadas por Henry Lynch desde los años 60 del siglo pasado, 63 incluyen casos de cáncer de mama y ovario, 22 presentaron casos sólo de cáncer de mama y 15 presentaron casos de cáncer de mama en varón. Genéticamente, 75 de estas familias fueron explicadas por mutaciones en los genes *BRCA1* y *BRCA2* (Sinilnikova et al., 2006). Basándose en la presencia de otros tipos tumorales y otras manifestaciones clínicas segregando en familias con cáncer de mama se ha podido subclasificar el cáncer de mama hereditario en distintos síndromes.

#### *Síndrome de cáncer de mama-ovario hereditario (HBOC)*

Se trata del síndrome de predisposición hereditaria al cáncer más frecuente, con una incidencia estimada de 1/500 – 1/2500. Está caracterizado por la agregación familiar de múltiples casos de cáncer de mama y/o de ovario con un patrón de herencia autosómica dominante. Los tumores aparecen a una edad más temprana que aquellos esporádicos, observándose a menudo casos de cáncer de mama premenopausicos (Claus et al., 1990). Para la identificación de familias con el síndrome HBOC, grupos como el National Comprehensive Cancer Network (NCCN)<sup>39</sup>, UK Cancer Family Study Group (UKCFSG) (Eccles et al., 2000), National Institute for Health and Clinical Excellence (NICE) o la Sección de Cáncer Hereditario de la Sociedad Española de Oncología Médica (SEOM) (de la Hoya, 2004) han establecido una serie de criterios clínicos.

#### Criterios diagnósticos HBOC según SEOM

(Debe cumplirse al menos uno de los criterios)

- Un caso de cáncer de mama menor o igual a 40 años.
- Diagnóstico de cáncer de mama y ovario en el mismo individuo.
- Dos o más casos de cáncer de mama, uno de los cuales en menor de 50 años o bilateral.
- Un caso de cáncer de mama a edad menor o igual a 50 años o bilateral y un caso de cáncer de ovario en familiares de primer o segundo grado.
- Tres casos de cáncer de mama y ovario (al menos un caso de ovario) en familiares de primer o segundo grado.
- Dos casos de cáncer de ovario en familiares de primer o segundo grado.
- Un caso de cáncer de mama en un varón y familiar de primer o segundo grado con cáncer de mama u ovario.

No existe una definición de este síndrome que haya sido aceptada de forma generalizada y distintas sociedades científicas utilizan criterios ligeramente diferentes. La edad temprana y el grado de agregación familiar se definen de manera diferente según cada guía y algunas guías incluyen otros tipos de cáncer como el pancreático o criterios anatómo-patológicos. En una proporción de familias con HBOC la enfermedad sigue un patrón de herencia claramente mendeliano, suele ser el caso de familias grandes y multigeneracionales. Sin embargo en la mayoría de las familias el tamaño pequeño, la elevada proporción de varones y una estructura limitada distorsionan la percepción de un patrón de herencia mendeliana.

### *Otros síndromes*

Hasta la fecha se han descrito entre 100 y 200 síndromes de predisposición hereditaria al cáncer. A parte del HBOC, existen otros síndromes hereditarios de susceptibilidad al cáncer que incluyen entre sus características un incremento en la susceptibilidad al cáncer de mama. Sin embargo, en todos los casos estos síndromes incluyen una manifestación fenotípica totalmente característica y fácilmente diferenciable de síndrome HBOC. Se estima en todo caso que la contribución de esos síndromes a la prevalencia de cáncer de mama hereditario es individualmente inferior al 1%. Dentro de estos síndromes se encuentran:

#### Síndrome de Li-Fraumeni

La primera descripción de este síndrome fue realizada en 1969 (Li and Fraumeni, 1969). Tras la revisión de 600 casos pediátricos diagnosticados de sarcoma se describieron en cuatro familias casos de cáncer de mama asociado a sarcoma de tejido blando, leucemias y linfoma. Se trata de un síndrome autosómico dominante asociado con diagnóstico a edades inferiores a los 45 años de sarcomas de tejidos blandos, cáncer de mama, leucemia infantil, osteosarcoma, melanoma, cáncer de colon, cáncer pancreático, carcinoma adrenocortical y tumores cerebrales. En el 50% de las familias con el síndrome se encuentran mutaciones en el gen *TP53* (Malkin et al., 1990). La incidencia de cáncer de mama en mujeres portadoras que alcanzan la edad adulta es aproximadamente del 22-50 % a la edad de 50 años (Easton et al., 1993).

#### Síndrome de Cowden

El síndrome de Cowden es de herencia autosómica dominante y se caracteriza por la presencia de hamartomas y predisposición al cáncer de mama, tiroides y endometrio entre otros tumores. Su incidencia se estima en 1/200000. Está causado por mutaciones en el gen *PTEN* (Nelen

et al., 1996). Se estima que el riesgo de mama acumulado a lo largo de la vida en mujeres con el síndrome es de 50%

#### Síndrome de Peutz-Jeghers

Se trata de un síndrome caracterizado por la aparición de máculas melanocíticas en los labios y mucosa bucal y perianal junto con múltiples pólipos hamartomatosos gastrointestinales. Su incidencia se estima en 1/120000. Está causado por mutaciones en el gen *STK11* (Hemminki et al., 1998). Se ha demostrado un incremento en el riesgo de cáncer de mama entre 30%–50% en individuos con el síndrome (Giardiello et al., 2000; Hearle et al., 2006; Lim et al., 2004).

#### Síndrome de Cáncer Gástrico Difuso Hereditario

Este síndrome está causado por mutaciones en el gen *CDH1* (Brooks-Wilson et al., 2004; Gayther et al., 1998; Guilford et al., 1998). Los individuos con este síndrome, tienen un riesgo a desarrollar cáncer gástrico difuso, a lo largo de su vida, superior al 70% (Pharoah et al., 2001). Las mujeres tienen, además, un riesgo de padecer cáncer de mama, especialmente del subtipo lobular, que se cifra en aproximadamente un 40% (Pharoah et al., 2001).

#### Ataxia telangiectasia

Se trata de un síndrome autosómico recesivo caracterizado por ataxia cerebelar progresiva, telangiectasias óculo-cutáneas, defectos inmunitarios, inestabilidad cromosómica, mayor sensibilidad a las radiaciones ionizantes, y aumento del riesgo de varios tipos de cáncer, particularmente linfomas y leucemias (1985). Está asociado con mutaciones homocigotas en el gen *ATM*. Estudios en familiares de primer grado de individuos con Ataxia telangiectasia demostraron un incremento en el riesgo a padecer cáncer de mama en portadores heterocigotos (Olsen et al., 2005; Swift et al., 1987). Estudios en familias con cáncer de mama familiar identificaron la presencia de portadores heterocigotos de mutaciones en el gen *ATM* con un riesgo relativo estimado de 2.37 (Broeks et al., 2000; Renwick et al., 2006).

#### Anemia de Fanconi

Se trata de una forma congénita de anemia aplásica que clínicamente se caracteriza por fallo de medula ósea junto con presencia de anomalías congénitas en los sistemas nervioso central, esquelético y gastrointestinal y susceptibilidad al desarrollo de tumores (Green and Kupfer, 2009). Su prevalencia es de 1 en 1 millón (D'Andrea, 2010). Su causa es heterogénea, con patrones de herencia tanto autosómicos como ligados al cromosoma X. Hasta la fecha se han

descrito mutaciones bialelicas en 15 genes (Kee and D'Andrea, 2012) y se estima una frecuencia de portadores heterocigotos de 1 en 300 (D'Andrea, 2010). La esperanza de vida de estos pacientes no sobrepasa la etapa infantil, sin embargo esto está cambiando debido al éxito de la terapia basada en el trasplante de médula osea. En los pacientes con anemia de Fanconi que sobreviven a la edad adulta, se observa un posterior incremento en el desarrollo de tumores sólidos (Tamary et al., 2010), que se cifra entre un 20-30% a la edad de 50 años (Alter et al., 2010; Rosenberg et al., 2008; Rosenberg et al., 2003), sin embargo no se ha demostrado un riesgo específico de cáncer de mama. Por el contrario, en individuos heterocigotos para mutaciones en los genes de Anemia de Fanconi se ha demostrado un riesgo incrementado de cáncer de mama, con una tasa de incidencia estandarizada de 1.7 (Berwick et al., 2007). De hecho, en el año 2002 el gen *BRCA2* fue identificado como uno de los genes implicados en la anemia de Fanconi (Howlett et al., 2002). Los últimos estudios sugieren contribución al riesgo de cáncer de mama en individuos heterocigotos en 5 de los 15 genes implicados en la anemia de Fanconi (Kee and D'Andrea, 2012).

#### *Epidemiología del cáncer de mama familiar.*

Se estima que alrededor del 7% de los casos de cáncer de mama y un 25% de los casos de cáncer ovario en la población general presentan alteraciones en algún gen de susceptibilidad al cáncer de mama de herencia mendeliana (Claus et al., 1996; Walsh et al., 2011; Wickramanyake et al., 2012). Estudios sobre la prevalencia del síndrome HBOC realizados en población no seleccionada en Noruega identificaron un 2,88 % de individuos cumpliendo los criterios HBOC (Stormorken et al., 2006). Cifras similares se obtuvieron en estudios poblacionales en California (2,5%) (Scheuner et al. (2010)), Letonia (2,38%) (Vanags et al., 2010) y Polonia (3,27 %) (Gronwald et al., 2006). Se estima que el resto de los síndromes de cáncer de mama hereditario son mucho menos frecuentes y causan, cada uno de ellos, menos del 1% de los casos de cáncer de mama hereditario.

Debido a la laboriosidad del cribado de mutaciones y a la baja prevalencia esperada, no existen estudios sobre la incidencia de mutaciones en los genes de susceptibilidad al cáncer de mama a nivel poblacional; una excepción son las poblaciones con efecto fundador, de las cuales se hablará más adelante (ver página 40). Modelos genéticos estiman una frecuencia de portadores de mutación de entre 1 en 860 (Antoniou et al., 2002) y 1 en 974 (Antoniou et al., 2000) para el gen *BRCA1* y de 1 en 740 para el gen *BRCA2*. Estudios en series de casos de cáncer de mama esporádicos en distintas poblaciones, ofrecen resultados variables con respecto a la incidencia de mutaciones en los genes *BRCA1* y *BRCA2*. A modo de ejemplo, en población japonesa se han detectado un 0,8 % de mutaciones (Emi et al., 1998), 2 % en población británica(2000), 3,6 % en

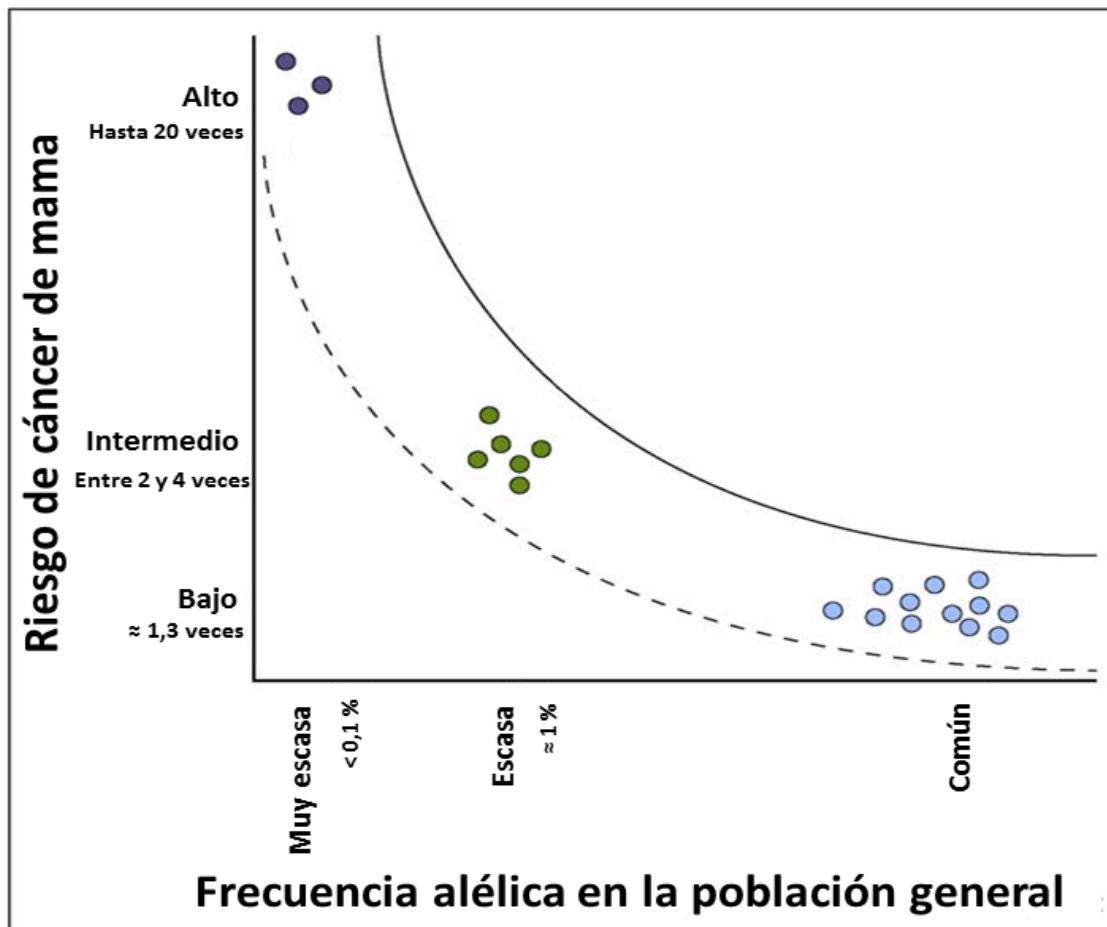
población húngara (Van Der Looij et al., 2000) y 11 % en población nigeriana (Fackenthal et al., 2012b).

En España no se han realizado estudios poblacionales sobre la incidencia de cáncer de mama familiar, esto se debe principalmente a la inexistencia de un registro nacional de tumores. Existen tres estudios sobre la incidencia de mutaciones en los genes *BRCA1/2* en series de casos esporádicos. García-Patiño et al (García-Patino et al., 1998) estudiaron 105 casos de cáncer de mama esporádico, encontrando 5,7% de mutaciones. Por su parte, Martínez-Ferrandis et al (Martínez-Ferrandis et al., 2003) encontraron un 5,6 % de mutaciones en una serie de 124 casos y de Juan Jiménez et al (de Juan Jimenez et al., 2012) encontraron un 1,05 % de mutaciones. Aunque estos datos pueden ofrecer una idea de la prevalencia de mutaciones en la población general española, a falta de datos epidemiológicos, la prevalencia de cáncer de mama hereditario en población española permanece desconocida.

### 3. Genética de la susceptibilidad al cáncer de mama

Como ya se ha mencionado, tras el descubrimiento inicial de los genes *BRCA1* y *BRCA2* se esperaba que éstos explicasen toda la carga genética del cáncer de mama hereditario. Sin embargo, a medida que se han realizado más estudios, quedó demostrado que aunque mutaciones en estos dos genes son los factores genéticos principales en el cáncer de mama hereditario, sólo explican aproximadamente un 16% de la heredabilidad familiar (2000; Peto et al., 1999). Pronto se empezó la búsqueda del posible gen *BRCA3* utilizando estudios de ligamiento (Arason et al., 2010; Gonzalez-Neira et al., 2007; Oldenburg et al., 2008; Rosa-Rosa et al., 2009; Smith et al., 2006). Aunque algunos estudios de ligamiento indicaron asociaciones en ciertas regiones cromosómicas como 8p12-p22 (Seitz et al., 1997) y 13q21 (Kainu et al., 2000), estas asociaciones no fueron corroboradas en estudios posteriores (Rahman et al., 2000; Thompson et al., 2002). Considerando los grandes esfuerzos invertidos para la identificación de *BRCA3* se asume que, si existiese un gen de características (frecuencia y penetrancia) similares a *BRCA1/2*, ya se habría logrado encontrar. Ante esta perspectiva, y considerando que los estudios en gemelos determinaron que la mayor parte de la agregación familiar tiene su origen en la susceptibilidad genética hereditaria, se sugieren dos posibles explicaciones. Por una parte se propone la existencia de mutaciones en genes que confieren una susceptibilidad similar a la de *BRCA1/2* pero que aparecen en una frecuencia mucho menor y que individualmente explicarían solo una proporción muy baja de las familias. Otros investigadores sugieren la existencia de mutaciones que influyen la susceptibilidad al cáncer de mama de una manera más sutil y complicada que las de *BRCA1/2* incluyendo interacciones entre genes o interacciones gen-ambiente.

Actualmente se asume que la susceptibilidad genética al cáncer de mama se ajusta a un modelo poligénico. Según este modelo, variantes genéticas con distinto grado de penetrancia y de prevalencia en la población darían lugar a un fenotipo común, que es el riesgo a cáncer de mama (Figura 2). Un riesgo elevado se puede adquirir por herencia de una única variante con una alta penetrancia, de unas pocas variantes con penetrancia moderada o de un conjunto amplio de variantes con penetrancia baja. De esta manera, a medida que disminuye la penetrancia de cada variante se requerirá un mayor número de variantes de riesgo presentes en un mismo individuo para lograr el mismo nivel de riesgo. Por lo tanto, variantes de alta penetrancia estarán presentes en familias que se ajustan de una manera óptima a un patrón de herencia mendeliana de un único factor, mientras que, según disminuye la penetrancia de las variantes, los patrones de herencia familiar se alejarán de la herencia mendeliana.



**Figura 2:** Modelo poligénico de la susceptibilidad al cáncer de mama

Representación esquemática del modelo poligénico de susceptibilidad al cáncer de mama. Las mutaciones germinales que confieren un alto riesgo al cáncer de mama (hasta 20 veces) son muy raras en la población, aquellas que confieren un riesgo intermedio (entre 2 y 4 veces) aparecen en una frecuencia de 1% aprox. En la población. Las variantes que confieren un riesgo bajo pueden llegar a ser tan frecuentes en la población como un 40%. La línea punteada representa el límite mínimo de frecuencia y riesgo bajo los cuales si existiese una variante se estima que no sería posible detectarla. La línea continua representa los límites de frecuencia y riesgo máximo permitidos compatibles con la teoría evolutiva. Imagen adaptada de (Hollestelle et al., 2010)

El descubrimiento de las variantes genéticas implicadas en la susceptibilidad a cáncer de mama requerirá distintas técnicas según se trate de una variante de penetrancia alta, baja o intermedia. En cualquiera de los casos, la aproximación puede estar basada en dos enfoques. El primero de los enfoques es el estudio de genes candidatos. Se basa en la selección de genes para los cuales exista una alta sospecha de estar relacionados con la susceptibilidad al cáncer de mama, como por ejemplo genes que codifiquen proteínas que interaccionen con BRCA1 Y BRCA2. El segundo enfoque es el estudio con técnicas de análisis genómico. Este enfoque presenta como ventaja el no introducir ninguna hipótesis *a priori* y por lo tanto permite identificar variantes en

genes sobre los cuales no se intuía que estuviesen relacionados con la susceptibilidad al cáncer de mama. Las técnicas a utilizar son las siguientes:

Búsqueda de variantes con penetrancia alta: El método de elección es el análisis de ligamiento. Se analiza la cosegregación en familias de marcadores genéticos, bien sean microsatelites o SNP's junto con el fenotipo de cáncer de mama. Si los marcadores están ligados a la enfermedad, se heredarán de forma conjunta, lo cual permite mapear el gen de interés en una región cromosómica. Posteriores estudios de clonación, secuenciación y estudios funcionales permitirán identificar el gen implicado. También resulta útil para la identificación como agente genético causal de nuevas variantes descritas en genes conocidos de susceptibilidad. Se obtienen buenos resultados en aquellas familias donde la herencia se ajuste bien a un patrón mendeliano, ya que se basa en la premisa de la existencia de un único factor genético causal. Para obtener datos estadísticamente significativos se debe analizar un gran número de individuos dentro de la misma familia, tanto afectados como no afectados. La aplicación del análisis de ligamiento en familias grandes y muy seleccionadas permitió la identificación de los genes *BRCA1* y *BRCA2*. Sin embargo la mayoría de las familias no resultan informativas para la aplicación de estudios de ligamiento.

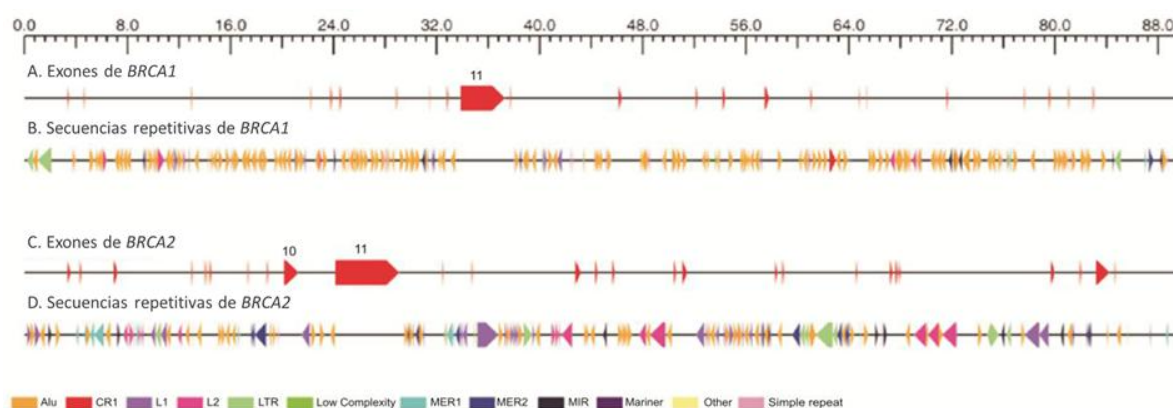
Búsqueda de variantes con penetrancia baja: En este caso el método de elección es el estudio de asociación, en el cual se compara la frecuencia de una variante genética en un grupo de individuos afectados de cáncer de mama con respecto a un grupo de individuos sanos. Cuando la distribución de frecuencias difiere entre los dos grupos de forma estadísticamente significativa, se puede asumir que existe una asociación entre la variante genética y la enfermedad. Esta aproximación se puede realizar mediante el análisis de variantes genéticas que representen todo el genoma, en los llamados GWAS (genome wide association studies) o mediante el estudio de genes candidatos. El estudio de genes candidatos presenta como desventaja el hecho de que se parte *a priori* de la hipótesis de la participación de un gen en la susceptibilidad a cáncer de mama con lo cual el descubrimiento de variantes dependerá de una buena selección de los genes candidatos. Sin embargo los GWAS son un análisis sin hipótesis previa y el descubrimiento de variantes no depende de la selección realizada por el investigador. Esto permite descubrir variantes en rutas sobre las cuales, previamente, no se había sospechado su participación en la susceptibilidad a cáncer de mama. Como desventaja, el gran número de comparaciones que se realiza en un GWAS requiere el estudio de una muestra amplia para lograr significación estadística. En este sentido el establecimiento de grandes consorcios colaborativos como "Breast Cancer Association Consortium" (BCAC) y "Collaborative Oncological Gene-Environment Study" (COGS) ha permitido recopilar el número de muestras necesarias para poder realizar estos estudios. A modo

de ejemplo, COGS incorpora más de 200.000 individuos para el estudio de 1.536 variantes genéticas. Para que una variante se pueda incluir en un estudio de asociación debe aparecer en una frecuencia cercana al 1% de la población general, ya que, de otra manera, el número de individuos portadores de la variante detectados sería muy bajo como para realizar comparaciones estadísticas o se requeriría incluir un número inasumible de individuos en el estudio. A medida que aumenta la frecuencia en la población general, se hace más sencillo detectar la posible asociación de variantes con el riesgo a cáncer de mama.

Búsqueda de variantes con penetrancia intermedia: El análisis de ligamiento no resulta útil para este tipo de variantes genéticas, ya que no todos los individuos que hereden la variante mostrarán el fenotipo de cáncer de mama. Por otro lado, las variantes genéticas de penetrancia intermedia presentan una frecuencia poblacional baja, lo cual desaconseja su detección mediante estudios de asociación. En este caso la metodología recomendada es la resecuenciación. Este método consiste en la secuenciación en series de cáncer de mama familiar de genes candidatos y comparar las variantes halladas con respecto a las encontradas en una población control.

#### 4. BRCA1 y BRCA2 en la susceptibilidad al cáncer de mama

*BRCA1* y *BRCA2* son los principales genes de alta penetrancia implicados en la susceptibilidad al cáncer de mama. *BRCA1* se localiza en la región cromosómica 17q21, a lo largo de 100 kb. Contiene 22 exones codificantes y dos exones que corresponden a dos sitios alternativos de origen de la transcripción. La región 5' de *BRCA1* contiene el pseudogen  $\Psi$ *BRCA1* que consiste en una duplicación parcial de *BRCA1* que contiene los exones 1a, 1b y 2 con sus correspondientes intrones (Brown et al., 1996). Por su parte *BRCA2* se localiza en la región cromosómica 13q12, a lo largo de 70 kb. Está compuesto por 27 exones, de los cuales 26 codifican una proteína de 3418 aminoácidos. Ambos genes poseen exones centrales muy largos: en *BRCA1* el exón 11 y en *BRCA2* los exones 10 y 11 codifican aproximadamente el 60% de la proteína (Figura 3. A y C). Las regiones genómicas de *BRCA1* y *BRCA2* contienen una alta densidad de secuencias repetitivas de ADN (figura 3 B y D). Las regiones del genoma densas en secuencias *Alu* se han visto involucradas en reordenamiento cromosómicos. (Korenberg and Rykowski, 1988)



**Figura 3:** Secuencias repetitivas en los genes *BRCA1* y *BRCA2*

La estructura exónica de los genes *BRCA1* y *BRCA2* se muestra en los paneles A y C. B) La región cromosómica de *BRCA1* contiene un 42 % de secuencias *Alu* y un 5% de secuencias repetitivas no *Alu*. D) La región cromosómica de *BRCA2* contiene un 20% de secuencias *Alu* y un 27 % de secuencias repetitivas no *Alu*. Imagen adaptada de (Welch and King, 2001)

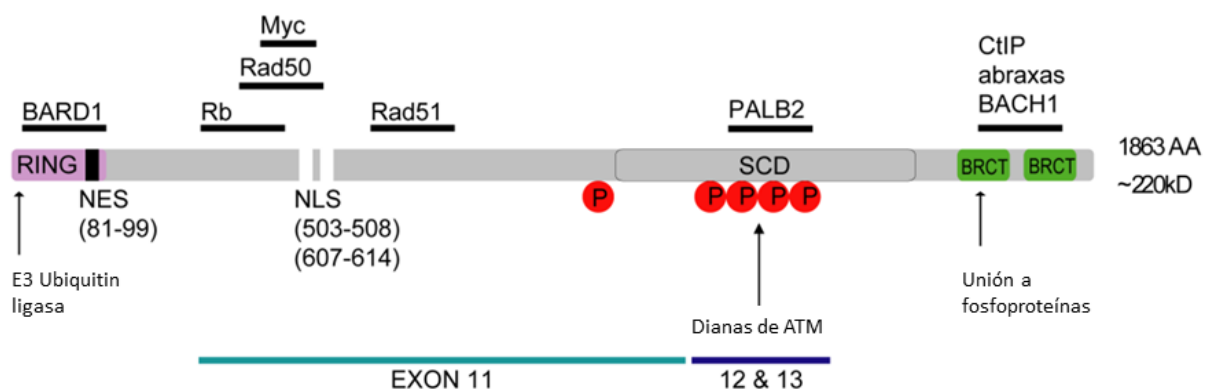
#### *Estructura de las proteínas BRCA1 y BRCA2.*

Las proteínas *BRCA1* y *BRCA2* muestran escasa homología entre sí, así como con otras proteínas conocidas. Su estructura es compleja y contiene múltiples dominios estructurales. El análisis de dichos dominios ofrece una idea sobre las posibles funciones bioquímicas de estas proteínas.

## BRCA1

Se han realizado estudios cristalográficos en dos dominios relevantes de BRCA1, RING y BRCT. Estas regiones se corresponden con un 17% de la secuencia primaria de BRCA1. El dominio RING está localizado en el extremo aminoterminal de BRCA1 (entre los aminoácidos 1 y 109 /exones 2-7) y es responsable de la actividad E3 ubiquitin ligasa de BRCA1 (Lorick et al., 1999) y de la unión a la proteína BARD1. El dominio RING incluye una secuencia de 4 aminoácidos hidrofóbicos que actúan como señal de exportación nuclear (NES) (Rodriguez and Henderson, 2000). La formación del heterodímero BRCA1/BARD1 incrementa de manera sustancial la actividad E3 ubiquitin ligasa (Hashizume et al., 2001) y produce un cambio conformacional que esconde la señal de exportación nuclear provocando que el dímero BRCA1/BARD1 se mantenga en el núcleo (Henderson, 2005). Dentro de las proteínas dianas de ubiquitinación se encuentran entre otras el receptor de estrógenos  $\alpha$ , el receptor de progesterona, la histona H2A y la proteína CtIP (Calvo and Beato, 2011; Heine and Parvin, 2007; Ma et al., 2010; You and Bailis, 2010; Zhu et al., 2011).

El dominio BRCT fue descrito por primera vez en la proteína BRCA1, localizado entre los aminoácidos 1650 y 1863. Posteriormente se ha encontrado en otras proteínas, la mayoría de ellas involucradas en reparación del daño al ADN (Bork et al., 1997; Callebaut and Mornon, 1997). En BRCA1, el dominio BRCT aparece en forma de repetición tándem. Este dominio media la unión con proteínas fosforiladas. BACH1/BRIP1 (Yu et al., 2003), CtIP (Yu and Chen, 2004) y abraxas (Kim et al., 2007) se han detectado unidas al dominio BRCT de BRCA1. También se ha detectado la capacidad de unión del dominio BRCT a roturas de doble cadena de ADN aunque se desconoce la relevancia *in vivo* de esta interacción (Yamane et al., 2000).



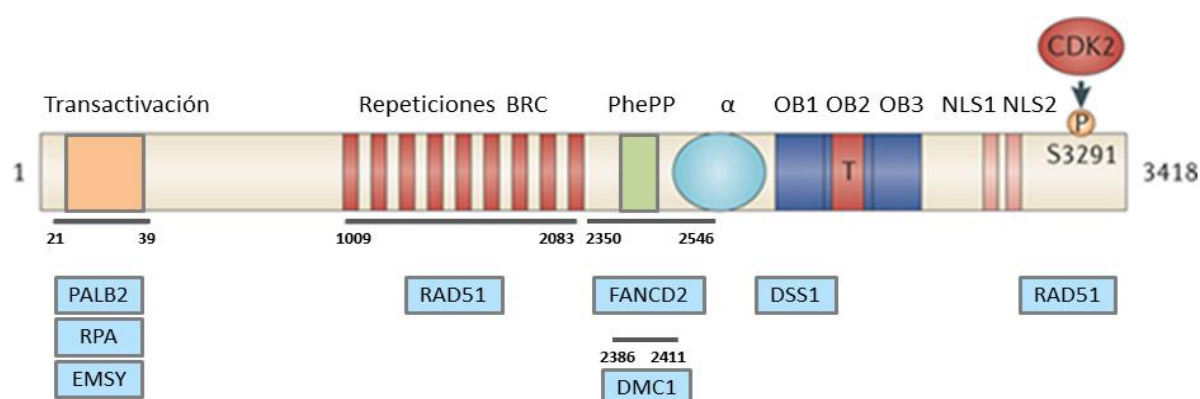
**Figura 4:** Dominios de la proteína BRCA1

Resumen de los dominios de la proteína BRCA1 y de las interacciones que establece con proteínas relevantes. Imagen adaptada de (Clark et al., 2012)

La secuencia localizada entre los aminoácidos 170 y 1649 pertenece a una región intrínsecamente desordenada (no adquiere una configuración tridimensional determinada). Esta región contiene dos señales de localización celular, NLS (aa 501-507 y 607-614) y una región rica en residuos de serina (aa 1280-1524), que son diana de la fosforilación mediada por la proteína ATM. Se han mapeado en esta región interacciones con proteínas como Rb (aa 304-394), cMyc (dos sitios de unión aa 175-303 y 433-511), Rad50 (aa 341-748), Rad51 (aa 758-1064) y PALB2 (aa 1364-1437). La Figura 4 muestra un diagrama de los dominios de BRCA1 y mapea las interacciones proteicas de BRCA1.

### BRCA2

Se han realizado estudios cristalográficos en dos dominios relevantes de BRCA2, el dominio de unión al ADN (DBD) y el dominio BRC. El dominio de unión al ADN está formado por 800 aa localizados en el extremo C-terminal de BRCA2. Su estructura consta de 5 motivos: un primer motivo rico en alfa-hélices, tres motivos de unión a oligosacárido y oligonucleótido (OB) y un motivo helicoidal que sobresale de la estructura, denominado Torre (T) (Yang et al., 2002). Este dominio está involucrado en la unión de ADN de cadena sencilla (OB2) y en la interacción con la proteína DSS1 (OB1). La interacción con DSS1 incrementa la estabilidad de la proteína BRCA2. El dominio BRC está formado por una secuencia central de 30 aa. En BRCA2 el dominio está codificado en el exón 11 y se encuentran 8 repeticiones en tándem a lo largo de una región de 1127 aa. Las repeticiones BRC están implicadas en la unión con la proteína RAD51 (Pellegrini et al., 2002; Rajendra and Venkitaraman, 2010).



**Figura 5:** Dominios de la proteína BRCA2

Resumen de los dominios de la proteína BRCA2 y de las interacciones que establece con proteínas relevantes. Imagen adaptada de (Roy et al., 2012)

En el extremo N-terminal, la región codificada por el exón 3 presenta gran interés. Por una parte se ha visto implicada en la regulación de la transcripción y por ello se denomina dominio de

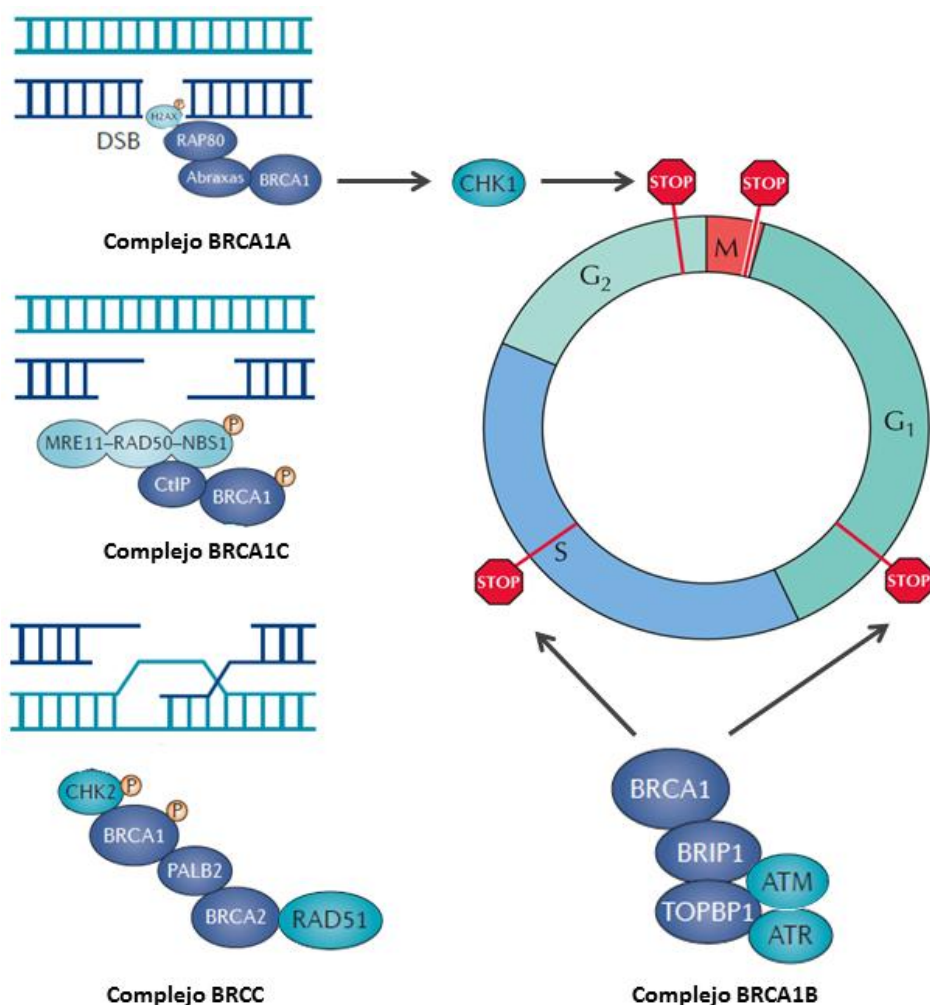
transactivación (Milner et al., 1997) (Shin and Verma, 2003). Adicionalmente esta región está implicada en la interacción con las proteínas PALB2 (Xia et al., 2006), EMSY (Hughes-Davies et al., 2003) y RPA (Wong et al., 2003). En la región central de BRCA2 mapea la zona de interacción con la proteína FANCD2 (Hussain et al., 2004) y el dominio PhePP, involucrado en la interacción con la proteína DMC1 (Thorslund et al., 2007). Por último el extremo C-terminal de BRCA2 contiene dos señales de localización nuclear, un sitio adicional de interacción con RAD51 (Esashi et al., 2005) y un sitio diana de fosforilación (Esashi et al., 2005) regulado por la ciclinkinasa CDK2.

### *Funciones de las proteínas BRCA1 y BRCA2*

Estas proteínas están involucradas en la respuesta al daño celular. La respuesta al daño al ADN implica la activación de mecanismos de control del ciclo celular que evitan la progresión del ciclo mientras se activan las rutas de reparación del daño. Dentro de los distintos daños que puede sufrir el ADN, en concreto, BRCA1 y BRCA2 se encargan, principalmente, de la reparación de las roturas de doble cadena (DSB). Los mecanismos de reparación de estas roturas se pueden llevar a cabo mediante dos vías, la de la recombinación homóloga (HR) y la de la unión de extremos no homólogos (NHEJ). Cualquiera de estas vías consta de tres tipos de moléculas: las moléculas sensoras son aquellas que reconocen el daño en el ADN; las moléculas efectoras son aquellas que se encargan de reparar el daño y las moléculas mediadoras son aquellas que facilitan la interacción entre las dos anteriores.

BRCA1 ejerce sus funciones mediante su capacidad de interactuar con otras proteínas. En este contexto, las interacciones establecidas a través del dominio RING con la proteína BARD1, parecen participar por un lado en la estabilidad de la misma y por otro en potenciar la capacidad ubiquitin ligasa del complejo. Muchas de las señalizaciones ejercidas por BRCA1 están mediadas por residuos de ubiquitina. Sin embargo, el dominio BRCT es el que define en cuál de sus múltiples funciones participará BRCA1, según cuál sea la proteína que esté unida a él. Mecanísticamente se conoce con más profundidad la función de BRCA1 en HR que en NHEJ.

BRCA1 forma parte de 4 macrocomplejos diferentes (Huen et al., 2010; Roy et al., 2012): BRCA1A, BRCA1B, BRCA1C y BRCC (Figura6). El complejo BRCA1A actúa como sensor del daño al ADN y adicionalmente regula el punto de control de ciclo celular G2/M. El complejo BRCA1B participa en los puntos de control del ciclo celular G<sub>1</sub>-S e intra-S, que aseguran la integridad genómica previo a la síntesis del ADN. El complejo BRCA1C media en la reparación del ADN mediante HR, facilitando la formación de ADN de cadena sencilla como paso previo a la HR. El complejo BRCC es capaz de reconocer el ADN de cadena sencilla formado previamente y mediar la invasión de la hebra homóloga, se considera el efector en la recombinación homóloga.



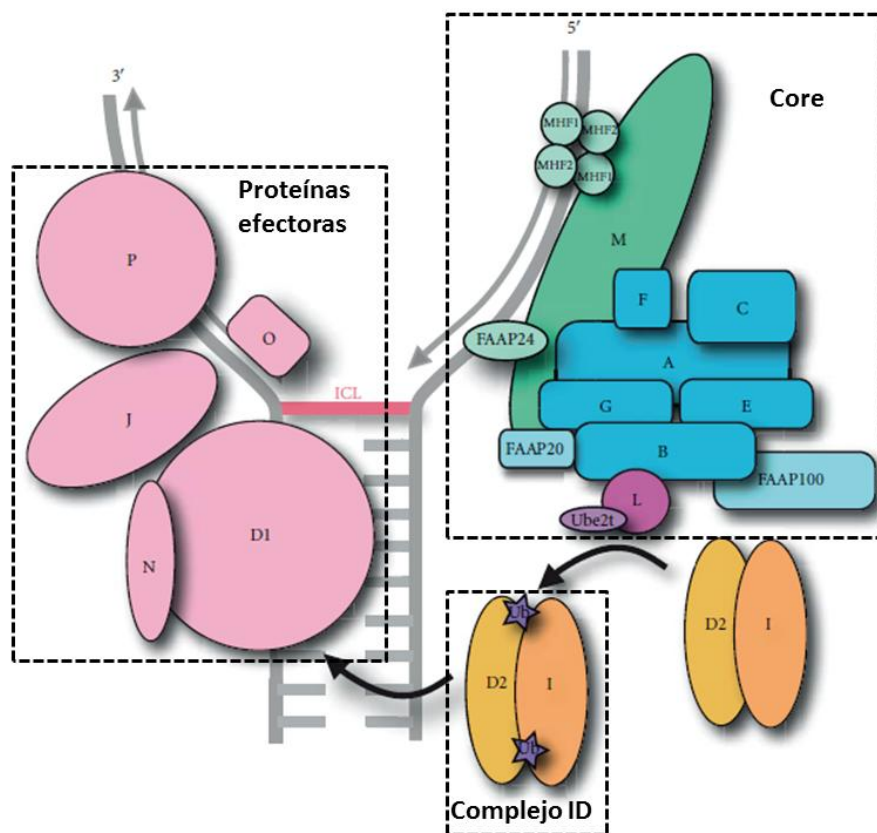
**Figura 6:** BRCA1 forma 4 complejos macromoleculares

En respuesta a la generación de DSB BRCA1 forma 4 complejos macromoleculares diferentes que participan en la regulación del ciclo celular y en la reparación del daño mediante recombinación homóloga

Se han descrito funciones de BRCA1 adicionales a la reparación del ADN, como la capacidad de regular la expresión mediante la interacción con la ARN polimerasa II (Kleiman et al., 2005) y la capacidad de remodelación de la cromatina mediante la interacción con el complejo SWI/SNF (Bochar et al., 2000) y con desacetilasas de histonas (Yarden and Brody, 1999).

En comparación con BRCA1, que posee un papel central en la regulación de la recombinación homóloga, del ciclo celular y de otros procesos celulares, BRCA2 parece estar implicado principalmente y de una manera muy directa, en la recombinación homóloga. Especialmente interesante es la relación de BRCA2 con la ruta de anemia de Fanconi. La ruta de anemia de Fanconi se activa en respuesta a daño al ADN que induce la formación de enlaces covalentes intercatenarios (ICL). La ruta se compone de un "core" multiproteico que ejerce las funciones de sensor; el complejo ID, que ejerce de mediador y una serie de efectores que incluyen a la proteína BRCA2, denominada en este contexto FANCD1. Uno de los eventos más relevantes en

la transmisión de señales es la ubiquitinación de la proteína FANCD2, llevada a cabo por la única proteína con actividad ubiquitin ligasa de esta ruta, FANL. (Figura 7)



**Figura 7:** Esquema de la ruta de anemia de Fanconi

Estructura de la ruta de anemia de Fanconi, aquellas proteínas que se nombran con el código de una letra se ha encontrado mutadas en algún tipo de anemia de Fanconi. (Imagen adaptada de (Hodson and Walden, 2012))

Los patrones de expresión de BRCA2 son dependientes del ciclo celular, observándose un pico de expresión en la fase G<sub>1</sub>/S (Wang et al., 1997). Esta regulación depende de la expresión de la proteína ZAR2. Ambos genes aparecen adyacentes en el genoma y están regulados por un promotor bidireccional. Durante la fase G<sub>1</sub>, la expresión de ZAR2 es máxima. ZAR2 es transportada al núcleo, uniéndose posteriormente al promotor de BRCA2 silenciando su expresión (Misra et al., 2010). Aunque existe también un promotor bidireccional que regula la expresión de los genes BRCA1 y NBR2 (Suen et al., 2005), no se ha descrito para esta proteína funciones similares a las de ZAR2. De hecho, algunos autores han detectado regulación de la expresión de BRCA1 durante el ciclo (Gudas et al., 1996) y otros no (Saunus et al., 2007). La localización de estas proteínas también está regulada a lo largo del ciclo, a modo de ejemplo BRCA1 permanece en el citosol y en el inicio de la fase S es trasladada al interior del núcleo (Glover-Collins and Thompson, 2008).

### *Variantes genéticas en BRCA1 y BRCA2.*

Desde que se inició la aplicación clínica de la búsqueda de mutaciones en ambos genes los resultados se recopilan en la base de datos Breast Cancer Information Core (BIC)(Szabo et al., 2000). A pesar de que la base BIC puede no recoger la totalidad de variantes detectadas en estos genes, la gran cantidad de registros disponibles convierte a BIC en una buena fuente de análisis de la variabilidad genética. Según datos obtenidos en Enero de 2013, existen 14866 y 14778 registros en la base BIC relativos al gen *BRCA1* y *BRCA2* respectivamente. En total se corresponden con 1726 variantes genéticas distintas en *BRCA1* y 1929 en *BRCA2*. De ellas 979 en *BRCA1* y 1075 en *BRCA2* han sido volcadas a la base de datos una única vez, sugiriendo que son variantes que suceden de manera muy poco frecuente. En ambos genes las variantes se distribuyen a lo largo del gen en las regiones codificantes y en las secuencias intrónicas flanqueantes. Dentro de estas variantes se incluyen aquellas que están relacionadas con el síndrome HBOC (variantes patogénicas, en ocasiones también llamadas mutaciones), aquellas que no están relacionadas con el síndrome HBOC (polimorfismos) y aquellas cuya relación con el síndrome HBOC no está ni confirmada ni descartada (variantes de significado incierto, UVs). Acorde con la función de estos genes como supresores de tumores, las variantes patogénicas han demostrado ser aquellas que conducen a una pérdida de función de las proteínas. El 30% de todas las variantes individuales descritas en la base BIC en ambos genes son pequeñas inserciones y deleciones que alteran el marco de lectura de la proteína. Un 10% son sustituciones de una única base que originan un codón de parada prematuro. Estos dos tipos de variantes producen una obvia pérdida de función en la capacidad codificante y por ello son fácilmente clasificables como variantes patogénicas. El 60% de las variantes no producen alteraciones tan obvias de la función y su clasificación en alguna de las tres categorías es más compleja. (ver página 53)

Un tipo especial de variantes genéticas son los llamados Grandes Reordenamientos Genómicos (LGR), que consisten en deleciones o inserciones de gran tamaño que pueden involucrar a múltiples exones o incluso la pérdida del gen completo. El uso inicial en la búsqueda de variantes de técnicas basadas en PCR no permitió detectar este tipo de variantes genéticas. La introducción de técnicas más actuales como el MLPA (Multiplex Ligation and Probe Amplification) ha permitido determinar una contribución de este tipo de variantes al total de variantes genéticas que explican HBOC de entre un 0 a un 30% en el gen *BRCA1* en función de la población analizada (Sluiter and van Rensburg, 2011). En el gen *BRCA2* la implicación de reordenamientos genómicos es mucho menor. En un reciente estudio (Judkins et al., 2012) se analizó la prevalencia de reordenamientos en aproximadamente 50000 individuos. El 10% de los portadores de variante patogénica fue portador de reordenamientos. Mientras que los reordenamientos en *BRCA1* explicaron un 12% de las variantes patogénicas en este gen, los reordenamientos en *BRCA2*

explicaron un 1,2%. Estas diferencias parecen estar relacionadas con la mayor presencia de secuencias repetitivas *Alu* en el gen *BRCA1*, las cuales a menudo están involucradas en los procesos de generación de reordenamientos. Adicionalmente, la existencia del pseudogen  $\Psi$ *BRCA1* también se sugiere como relevante en una mayor incidencia de reordenamientos en el locus de *BRCA1*.

Por supuesto, existen muchos otros tipos de variantes genéticas de potencial relevancia clínica y que, sin embargo, no se analizan de manera rutinaria en los genes *BRCA1/2*. En este tipo de variantes se incluyen aquellas localizadas en regiones promotoras, regiones 5' y 3' UTR y en zonas intrónicas alejadas de los exones (Fackenthal et al., 2012a). En un estudio reciente, el análisis de alteraciones en el *splicing* del gen *BRCA2* en individuos de familias HBOC permitió identificar 9 familias HBOC de una serie de 2000 portadoras de la variante patogénica c.6937+594T>G. El efecto patogénico de esta variante se debe a la inclusión de un exón adicional en el gen *BRCA2* (Anczukow et al., 2012). Este resultado sugiere la existencia de variantes patogénicas aun no detectadas en estos dos genes.

#### *Mutaciones recurrentes y efecto fundador*

Aunque la mayoría de las variantes patogénicas en *BRCA1* y *BRCA2* son individualmente raras, en ciertas poblaciones con alto grado de endogamia se ha observado la aparición de algunas variantes patogénicas de manera recurrente en familias no emparentadas. Análisis genéticos demuestran que las variantes recurrentes provienen de un único antecesor portador, sugiriendo lo que se conoce como efecto fundador. Se entiende por efecto fundador a las consecuencias derivadas de la formación de una nueva población de individuos a partir de un número muy reducido de éstos. En los miembros de esta nueva población aparece una baja variabilidad genética, ya que la población se ha originado a partir de una pequeña muestra de individuos que no era representativa de la diversidad genética original. Adicionalmente, las poblaciones con efecto fundador suelen ser poblaciones con un exceso de alelos raros.

Un ejemplo clásico de efecto fundador es el de la población de judíos del centro y este de Europa conocida como Ashkenazi. En esta población, el 99% de las familias con mutación patogénica se explican por dos variantes patogénicas en *BRCA1* (c.68\_69delAG y c.5263\_5264insC) y una en *BRCA2* (c.5946\_5946delT) (Phelan et al., 2002). Adicionalmente, la frecuencia de estas variantes patogénicas a nivel poblacional es mucho más elevada en este grupo étnico que en otras poblaciones, 1 de cada 40 individuos frente a 1 de cada 500 estimado en otras poblaciones (Roa et al., 1996). Aunque la historia de la población española no parece muy propicia de posibles efectos fundadores, se ha demostrado la presencia de recurrencia en ciertas variantes genéticas. En el gen

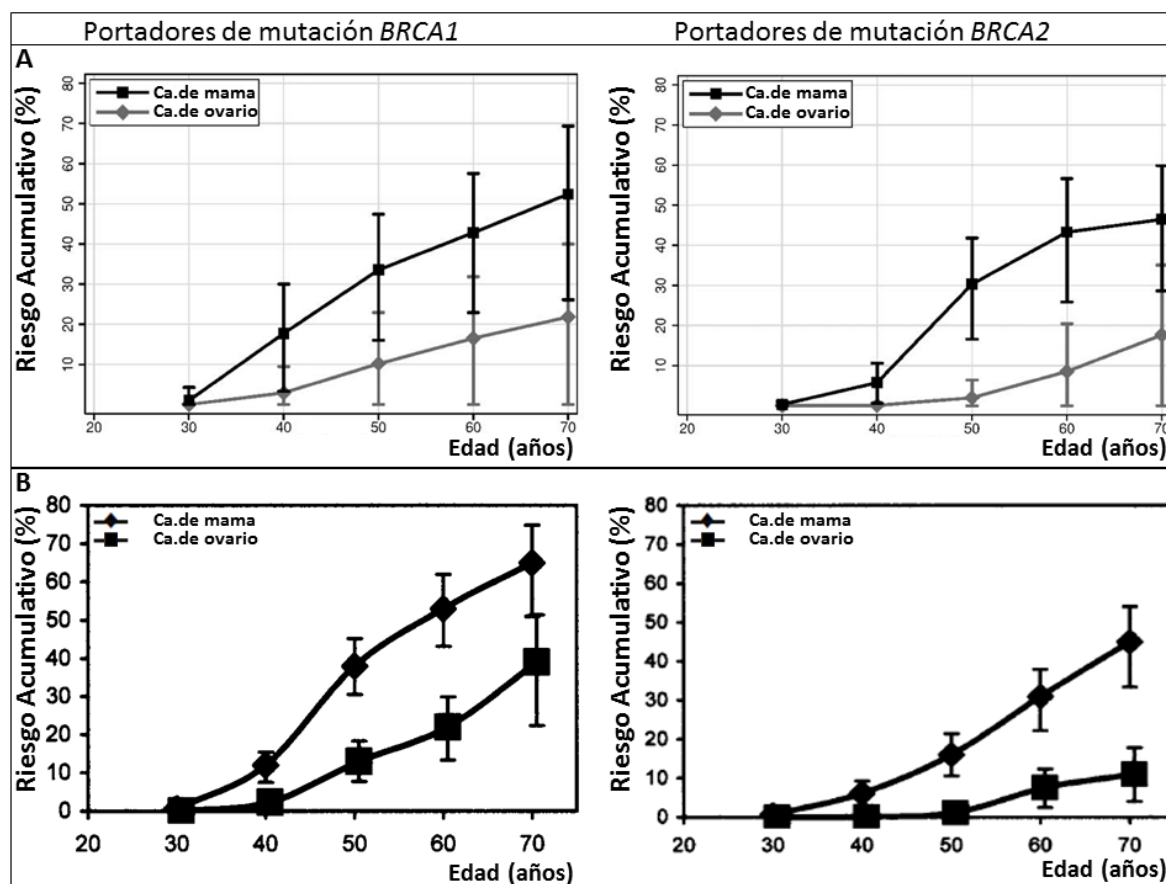
*BRCA1*, 5 variantes patogénicas (c.68\_69delAG, c.211A>G, c.470\_471delCT, c.5117G>A y c.5123C>A) son responsables del 46% de las familias portadoras de mutación en *BRCA1*. En el gen *BRCA2*, 4 variantes patogénicas (c.2808\_2811del4, c.6629\_6630delAA, c.9026\_9030del5 y c.9310\_9311delAA) son responsables del 56 % de las familias portadores de mutación en *BRCA2* (Diez et al., 2003). El análisis de la estructura haplotípica en los individuos portadores de estas variantes ha demostrado un origen común de las variantes c.68\_69delAG, c.211A>G, c.6629\_6630delAA, c.9026\_9030del5 y c.9310\_9311delAA. En el caso de c.68\_69delAG, el haplotipo conservado detectado en población española es similar al encontrado en la población Ashkenazi (Diez et al., 1999). Esta variante estaría relacionada con la presencia histórica de judíos en la península Ibérica. Con respecto al resto de variantes, la incidencia se limita a regiones concretas; como es el caso de c.211A>G encontrada en población gallega (Vega et al., 2001), c.6629\_6630delAA y c.9026\_9030del5, encontradas en población catalana y valenciana (Campos et al., 2004; Campos et al., 2003) y c.9310\_9311delAA, encontrada en Castilla y León (Infante et al., 2010).

#### *Penetrancia de las mutaciones en BRCA1 y BRCA2.*

Los primeros estudios establecieron una estimación de riesgo de desarrollar cáncer de mama de 75-80 % a la edad de 70 años para individuos portadores de mutaciones en *BRCA1/2* (Easton et al., 1997; Ford et al., 1994). Desde entonces, distintas series han obtenido estimaciones de riesgo de cáncer de mama a la edad de 70 años en el rango de 35-69 % en portadores de mutación en *BRCA1* y de 26-56 % en portadores de mutación en *BRCA2* (2000; Hopper et al., 1999; Satagopan et al., 2001; Struewing et al., 1997; Thorlacius et al., 1998; Warner et al., 1999). Un meta-análisis publicado en el año 2003, que incluye más de 8000 casos (Antoniou et al., 2003) combinando datos de 22 estudios, estableció las siguientes estimaciones de riesgo a la edad de 70 años. Para los portadores de mutación en *BRCA1* el riesgo acumulativo de desarrollar cáncer de mama a la edad de 70 años fue 65% (intervalo de confianza 95 %: 44-78 %) y el riesgo de desarrollar cáncer de ovario fue 39% (intervalo de confianza 95 %: 18-54%). Para los portadores de mutación en el gen *BRCA2* los riesgos por su parte fueron 45% (intervalo de confianza 95%: 31-56%) para el cáncer de mama y 11% (intervalo de confianza 95%: 2,4-19%) para el cáncer de ovario. En población española un estudio publicado el año 2008 (Milne et al., 2008) estimó la penetrancia utilizando datos de casos de 155 y 164 familias portadoras respectivamente de mutación en *BRCA1* y *BRCA2*. Los valores de penetrancia varían ligeramente con respecto a los del meta-análisis del año 2003 aunque las diferencias no son significativas. (Figura 8). Esta variabilidad en la penetrancia observada en distintas series podría ser debida a diferencias metodológicas y poblacionales (Fackenthal and Olopade, 2007).

Se han demostrado diferencias de penetrancia en función de características relativas a la mutación, como por ejemplo su posición relativa en el gen. Ciertos estudios han identificado una región de 3.3 kb localizada en el exón 11 del gen *BRCA2* llamada OCCR (**O**varian **C**ancer **C**luster **R**egion). La región OCCR se ha descrito relacionada con una mayor penetrancia para el cáncer de ovario en comparación con el de mama (Gayther et al., 1997; Lubinski et al., 2004; Thompson and Easton, 2001). También se han observado diferencias relacionadas con la identidad específica de algunas mutaciones. Por ejemplo, para la mutación de origen gallego *BRCA1* c.211A>G se sugirió una menor penetrancia para el cáncer de mama que el resto de mutaciones encontradas en población española, aunque el dato no es estadísticamente significativo (Milne et al., 2008). Dentro de los portadores de una misma mutación también se ha encontrado variabilidad en la penetrancia (Levy-Lahad and Friedman, 2007). Esto sugiere la existencia de factores adicionales, genéticos o ambientales que modifican el riesgo en portadores de mutación en *BRCA1/2*.

En la actualidad, existen dos consorcios internacionales, el IBCCS (International BRCA1/2 Carrier Cohort Study) y CIMBA (Consortium of Investigators of Modifiers of BRCA1 and BRCA2) creados para la identificación de factores ambientales y genéticos modificadores del riesgo en portadores de mutación en *BRCA1/2*. Los últimos estudios llevados a cabo por CIMBA han identificado 6 locus que modifican el riesgo de cáncer de mama en portadores de mutación en *BRCA1* y 13 locus en portadores de mutación en *BRCA2* (Antoniou et al., 2012).



**Figura 8:** Penetrancia de cáncer de mama y ovario en portadores de mutación en BRCA1/2

A) Penetrancia del cáncer de mama y ovario en portadores de mutación en los genes BRCA1 (izquierda) y BRCA2 (derecha) según datos obtenidos en población española (Milne et al., 2008). B) Penetrancia del cáncer de mama y ovario según datos del meta-análisis del año 2003 (Antoniou et al., 2003)

### Implicación de BRCA1 y BRCA2 en la carcinogénesis

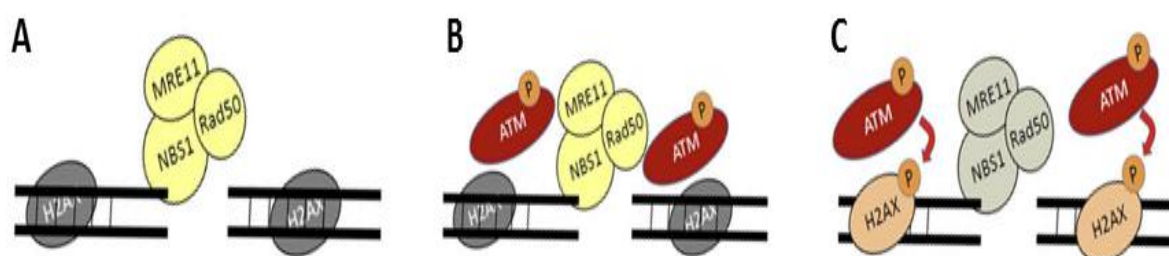
Los tumores de mama que aparecen en portadores de mutaciones en BRCA1 y BRCA2 presentan características patológicas que los diferencian entre sí y de los esporádicos. Los tumores asociados a BRCA1 suelen ser de alto grado y con una tasa de proliferación elevada (Karp et al., 1997; Phillips, 2000), frecuentemente asociados a negatividad para receptor de estrógenos (Foulkes et al., 2004; Johannsson et al., 1997; Palacios et al., 2005). También tienden a asociarse con negatividad para receptor de progesterona y ausencia de sobreexpresión de HER2, clasificándose como fenotipo triple negativo. Molecularmente los tumores asociados a BRCA1 se clasifican en el subtipo basal (Sorlie et al., 2003). En cuanto a los tumores asociados a mutaciones en BRCA2, tienden a ser más heterogéneos y la prevalencia de tumores positivos a receptor de estrógenos parece ser similar a aquella observada en los tumores esporádicos (Armes et al., 1999; Lakhani et al., 2005; Lakhani et al., 2002).

El papel de BRCA1 y BRCA2 en la conservación de la integridad del ADN y de la estructura cromosómica sugiere que su inactivación puede estar relacionada con la carcinogénesis. No obstante, una de las cuestiones clave respecto al papel en la carcinogénesis es la especificidad por el tejido mamario y ovárico. Una característica común de ambos tejidos es la capacidad de sus células epiteliales de responder a estímulos promitóticos a lo largo de los ciclos menstruales. Se ha sugerido la implicación de *BRCA1* en la regulación transcripcional (Fan et al., 1999) y post-traducciona (Ma et al., 2010) del receptor de estrógenos, de forma que la especificidad de tejido podría ser independiente de las funciones en reparación del ADN. Sin embargo esto no explicaría la especificidad de tejido de las mutaciones en *BRCA2*, proteína implicada mayoritariamente en el mantenimiento del ADN. Últimos estudios han implicado a *BRCA1* en los procesos de diferenciación y proliferación de las células madre en el tejido mamario, mostrando que mutaciones en *BRCA1* producen una alteración en el desarrollo del epitelio mamario en modelos *in vitro* y una alteración en los mecanismos de polaridad celular (Maxwell et al., 2011; Santarosa and Maestro, 2012). Se desconoce la implicación de *BRCA2* en estos procesos.

## 5. Búsqueda de genes adicionales en la susceptibilidad al cáncer de mama

### Primeros estudios de resecuenciación de genes candidatos

Tras el descubrimiento de mutaciones en los genes *BRCA1* y *BRCA2*, los primeros estudios de genes candidatos se diseñaron basándose en la hipótesis de la existencia de mutaciones con una frecuencia y penetrancia similar a la de éstos. Por ello, la mayoría de los estudios incluyen tamaños poblacionales pequeños, del orden de la centena. Como posibles candidatos se estudiaron genes que codifican proteínas que interaccionan física o funcionalmente con *BRCA1/2*: *BRIP1* (Cantor et al., 2001), *BARD1* (Ghimenti et al., 2002; Ishitobi et al., 2003; Sauer and Andrulis, 2005) *RAD51* (Kato et al., 2000; Lose et al., 2006; Rapakko et al., 2006) y cuatro de sus cinco genes parálogos: *XRCC2* (Rodríguez-Lopez et al., 2003), *RAD51B* (Johnson et al., 2011), *RAD51C* y *RAD51D* (Rodríguez-Lopez et al., 2004) y los genes del complejo MRN: *MRE11*, *RAD50* y *NBS1* (Bartkova et al., 2008; Bogdanova et al., 2008; Buslov et al., 2005; Gorski et al., 2003; Heikkinen et al., 2006; Steffen et al., 2006; Tommiska et al., 2006). La mayoría de los estudios no fue concluyente en la identificación de variantes de susceptibilidad al cáncer de mama. Excepcionalmente, los genes *RAD50* y *NBS1* del complejo MRN se encontraron asociados con incremento moderado del riesgo de cáncer de mama pero únicamente en poblaciones eslavas y nórdicas y en relación con un efecto fundador de las variantes *NBS1* 657del5 en población eslava (Bogdanova et al., 2008; Buslov et al., 2005; Gorski et al., 2003; Steffen et al., 2006) y *RAD50* 687delT en población nórdica (Heikkinen et al., 2006; Tommiska et al., 2006). Este complejo es el primer sensor en la detección de las roturas de doble cadena, por lo tanto tiene un papel clave en la señalización que media la recombinación homóloga. En la Figura 9 se muestra el papel del complejo MRN en la detección de las roturas de doble cadena.



**Figura 9:** Funciones del complejo MRN

A) El complejo MRN detecta las roturas de doble cadena en el ADN. B) La activación de este complejo fosforila a la quinasa ATM. C) ATM fosforila a la histona H2AX, que se localiza en los sitios de rotura de doble cadena del ADN.

Los resultados de los primeros estudios de genes candidatos no confirman la existencia de mutaciones en genes relacionados con *BRCA1/2* con una frecuencia y penetrancia similar a las de

las mutaciones en *BRCA1/2*. Sin embargo, la asociación de mutaciones en *NBS1* y *RAD50* con una penetrancia moderada en el contexto de poblaciones con efecto fundador sugiere la existencia de mutaciones con una frecuencia y penetrancia menor, que no han podido detectarse debido al limitado tamaño poblacional.

*Segunda generación de estudios de resecuenciación de genes candidatos.*

La identificación del gen *BRCA2* como el responsable del grupo de complementación D1 de la anemia Fanconi (Howlett et al., 2002), estableció un vínculo entre ambas enfermedades y abrió la puerta al estudio de la implicación de otros genes de anemia de Fanconi en la susceptibilidad al cáncer de mama. Además, con el objetivo de la detección de mutaciones con una frecuencia menor que las de *BRCA1/2* la mayoría de los estudios de resecuenciación realizados a partir de este momento involucran un tamaño muestral mayor, con poblaciones que se aproximan al millar de individuos.

En 2006 un estudio detectó 9 familias portadoras de mutaciones inactivantes en *BRIP1* (*FANCF*) de una serie de 1212 familias HBOC, estimándose un riesgo para cáncer de mama intermedio (O.R.= 2) (Seal et al., 2006). En el año 2007, se identificó *PALB2* (*FANCD1*) como uno de los genes causales de la anemia de Fanconi, tras la detección de mutaciones bialélicas en 13 individuos de una serie de 82 familias (Reid et al., 2007). Adicionalmente, 3 de estas familias demostraron una historia familiar de cáncer de mama. De manera simultánea, el mismo grupo identificó mutaciones inactivantes de *PALB2* en 10 casos de una serie de 923 familias con HBOC. Se estimó el riesgo a cáncer de mama en heterocigotos como intermedio (O.R.=2.3)(Rahman et al., 2007). El estudio más grande realizado hasta la fecha detectó 33 mutaciones inactivantes de *PALB2* en una serie de 972 familias (3,4%) (Casadei et al., 2011). En población española, dos estudios con series de 100 individuos cada uno encontraron un portador de mutación en *PALB2* en cada serie (Blanco et al., 2012; Garcia et al., 2009b) .

En 2010 el análisis de 1100 casos índice de HBOC detectó 6 familias portadoras de mutaciones inactivantes de *RAD51C* (cambios de fase del marco de lectura, *splicing*, y cambios de aminoácido). De manera simultánea se identificaron mutaciones bialélicas en el gen *RAD51C* en pacientes con anemia de Fanconi (Vaz et al., 2010), renombrado en ese contexto como *FANCO*. Desde entonces se han realizado múltiples estudios en distintas poblaciones con resultados dispares (Akbari et al., 2010; Clague et al., 2011; Coulet et al., 2012; De Leeneer et al., 2012; Kushnir et al., 2012; Lu et al., 2012; Pang et al., 2011; Pelttari et al., 2011; Romero et al., 2011; Silvestri et al., 2011; Thompson et al., 2012a; Vuorela et al., 2011; Zheng et al., 2010). En algunas de ellas encuentran mutaciones en *RAD51C* y en otras no, probablemente estas diferencias estén

relacionadas con variaciones poblacionales, con distintos criterios de selección o con el pequeño tamaño muestral. Actualmente el estudio más amplio realizado incluye 1132 casos de HBOC y 272 casos de cáncer de ovario no seleccionados por historia familiar. Como resultado se encontraron 12 mutaciones patogénicas y se cuantificó el riesgo de cáncer de ovario en una O.R.= 5,88, lo que se traduce en un 9 % de riesgo acumulativo de cáncer de ovario a la edad de 80 años. No se detectó un incremento significativo en el riesgo de cáncer de mama (Loveday et al., 2012). En población española se estima que la frecuencia de mutaciones en *RAD51C* dentro de los casos HBOC es del 1,3 % (Osorio et al., 2012). Este resultado abre la puerta a la posible implicación del resto de los parálogos de *RAD51* en la susceptibilidad familiar a cáncer de mama/ovario.

Posteriores estudios reanalizaron la implicación de otros parálogos de *RAD51* en el síndrome HBOC. En 2011 se encontraron 8 familias portadoras de mutaciones inactivantes de *RAD51D* (cambios de fase del marco de lectura, codones de parada prematuros y alteraciones en el *splicing*) de una serie de 911 familias HBOC. Se observó además una mayor asociación en familias portadoras de casos de ovario, y se pudo establecer para los portadores de estas mutaciones un incremento del riesgo de desarrollar cáncer de ovario elevado (O.R.=6.30). Sin embargo, no se ha podido establecer una asociación con el riesgo de desarrollar cáncer de mama (Loveday et al., 2011). Posteriores estudios confirmaron la presencia de mutaciones patogénicas en *RAD51D* en distintas poblaciones en asociación únicamente con cáncer de ovario familiar (Osher et al., 2012; Pelttari et al., 2012; Wickramanyake et al., 2012).

### *Tercera generación: secuenciación de exomas*

El desarrollo de técnicas de secuenciación masiva ha supuesto una nueva etapa en la búsqueda de nuevos genes implicados en HBOC. Un estudio reciente involucrando el análisis de 35 exomas de 15 familias HBOC identificó, además de individuos portadores de mutaciones en *BRCA2* y *PTEN*, individuos portadores de mutaciones inactivantes en los genes *FANCC*, *FANCL* y *BLM* (Thompson et al., 2012b). La posterior resecuenciación de *FANCC* y *BLM* en 438 individuos y la búsqueda de mutaciones frecuentes en *FANCC* en otros 957 individuos, encontró 4 individuos portadores de mutaciones inactivantes de *FANCC* y uno de *BLM* (Thompson et al., 2012b). Estudios posteriores confirman la presencia de mutaciones en el gen *BLM* en pacientes con HBOC (Prokofyeva et al., 2013; Sokolenko et al., 2012). La proteína BLM está implicada en la ruta de anemia de Fanconi, interaccionando directamente con SLX4. Sin embargo mutaciones bialélicas en *BLM* no originan anemia de Fanconi, sino un síndrome de inestabilidad cromosómica conocido como Bloom. Este resultado abre la posibilidad a la existencia de factores genéticos comunes a la susceptibilidad del cáncer de mama y otros síndromes de inestabilidad cromosómica adicionales a la anemia de Fanconi.

Otro estudio reciente de secuenciación de exomas que involucra un número mayor de pacientes (13 familias estudiadas mediante secuenciación de exoma y 3490 analizados mediante secuenciación clásica) encontró individuos portadores de 2 variantes patogénicas; c.651\_652del y c.49C>T en el gen *XRCC2* (Park et al., 2012). Existen 5 genes parálogos a *RAD51* de los cuales *RAD51C*, *RAD51D* también se encuentran mutados en familias HBOC. Funcionalmente se agrupan en dos complejos moleculares BCDX2 (*RAD51B*, *RAD51C*, *RAD51D* y *XRCC2*) y CX3 (*RAD51C* y *XRCC3*). El complejo BCDX2 actúa en pasos previos a *RAD51* y CX3 actúa en pasos posteriores (Chun et al., 2013), aunque sus funciones no están claramente establecidas.

#### *Estudios de asociación: Genes de baja penetrancia*

Se han realizado múltiples estudios de asociación tratando de identificar genes que confieran un riesgo bajo al cáncer de mama. Hasta el año 2012 un total de 27 alelos de riesgo han sido identificados y validados en estudios de alta calidad que cumplen las condiciones astringentes de significancia exigidas a los ensayos GWAS ( $p=1 \times 10^{-7}$ ). Estos resultados confirman la existencia de variantes comunes con un papel en la etiología del cáncer de mama, sin embargo la aplicación clínica de estos resultados en el contexto del cáncer de mama hereditario no es clara. En el caso de los genes de alta y moderada susceptibilidad al cáncer de mama, ensayos de genes candidatos y ensayos genómicos han ofrecido como resultados la identificación únicamente de genes implicados en la respuesta al daño al ADN. Por comparación, de los 27 alelos de bajo riesgo identificados en los ensayos GWAS, únicamente dos se localizan en genes relacionados con la respuesta al daño al ADN, *RAD51B* y *MERIT40* y uno relacionado con el mantenimiento de los telómeros, *TERT*. El resto de alelos se localizan en genes con funciones variadas como; la regulación hormonal de estrógenos, andrógenos y parathormona (*ESR1*, *NRIP1*, *ZMIZ1* y *PTHRP*), la apoptosis (*CASP8* y *TOX3*), la regulación del ciclo celular (*CDKN2A*), la respuesta a mitógenos (*FGFR2* y *MAP3K1*), la secreción (*SLC4A7* y *STXBP4*), la migración leucocitaria (*LSP1*), la cadena respiratoria mitocondrial (*COX11*) y la maduración de ARNm (*RALY*). Por su parte, otros 10 alelos se localizan en regiones no asociadas a ningún gen o en genes sin función conocidas (*ZNF365* y *ANKRD16*).

Los resultados recientemente publicados del consorcio COGS (Michailidou et al., 2013) han permitido la confirmación de los 27 alelos previamente identificados y la identificación de 41 alelos adicionales de riesgo. Estos alelos se localizan en la proximidad ( $\approx 50$  kb) de 96 genes. El análisis de enriquecimiento utilizando el método Ingenuity Systems Pathway Analysis demostró un exceso de representación de genes para los cuales se ha sugerido una relación con tumorigénesis (34 genes), cáncer de mama (15 genes) y aparición de tumores en organismos modelo (10 genes).

El mismo análisis demostró una sobrerrepresentación de genes con funciones relacionadas con muerte celular, diferenciación y expresión. Por primera vez para alguno de estos alelos se detectó su función como cis eQTL (regulador de la expresión de genes localizados en las proximidades del SNP  $\approx$ 50 kb) de los genes *PEX14*, *PTPN22*, *DCLRE1B* y *CFL1*, tanto en tejido normal como en tejido tumoral utilizando datos de expresión provenientes de The Cancer Genome Atlas. Asumiendo que el riesgo conferido por cada uno de los alelos se combina de manera únicamente multiplicativa, se estima que un 14 % del riesgo familiar de cáncer de mama estaría explicado por la presencia de alelos de bajo riesgo. Estimaciones adicionales apuntan a la posibilidad de la existencia de 1000 loci adicionales relacionados con la susceptibilidad al cáncer de mama.

## 6. Consejo genético en cáncer de mama familiar

### *Consideraciones previas a la realización de un test genético*

La identificación de los genes de susceptibilidad *BRCA1* y *BRCA2* transformó profundamente el manejo y el tipo de asesoramiento que podían recibir las familias con HBOC. Por primera vez se hizo posible distinguir los individuos portadores de los no portadores y ofrecer así un asesoramiento individualizado a cada miembro de la familia. Sin embargo, el consejo genético presenta una serie de implicaciones clínicas, éticas, psicológicas y también económicas que exigen una cuidadosa valoración de cada caso en particular. Tras una evaluación de la historia médica y familiar se, recomendará la realización de un test genético de *BRCA1/2* a aquellas familias con una probabilidad alta de que sufran un síndrome HBOC debido a estos genes; es lo que se denomina probabilidad pre-test. Existen varios métodos para establecer la probabilidad pre-test:

**Modelos genéticos:** Se parte de un modelo teórico en el cual se define la existencia de uno o varios genes de susceptibilidad al cáncer de mama con una prevalencia, penetrancia y patrón de herencia conocidos. A partir de estos datos se estima la probabilidad de que un individuo con una determinada historia personal y familiar sea portador de mutaciones en esos genes.

**Tablas de prevalencia:** Se basan en clasificar las familias en categorías según su historia de cáncer y asumir que la probabilidad de encontrar una mutación es igual a la frecuencia con la que se han encontrado mutaciones en familias con características similares.

**Modelos empíricos:** Partiendo de una serie de familias a las que se les ha realizado el test genético se identifican una serie de variables que se asocian con la presencia de mutación y se construye un modelo estadístico que predice la probabilidad de que una familia sea portadora de mutación. Existen distintos modelos que, por lo general utilizan como variables la presencia de cáncer de mama unilateral o bilateral, la presencia de cáncer de ovario, la presencia de cáncer de mama en varón y sus edades de diagnóstico. En estos modelos se pueden incluir también características anatomo-patológicas de los tumores

La principal ventaja de los modelos empíricos es que permiten calcular la probabilidad pre-test de un modo sencillo, lo cual es de gran utilidad en el poco tiempo del que normalmente se dispone en una consulta de consejo genético. Una de las características que dificulta el establecimiento de una probabilidad pre-test es la alta frecuencia de cáncer de mama esporádico. A diferencia de otros síndromes de susceptibilidad genética al cáncer, que predisponen a cánceres raros en la población general (como por ejemplo sarcomas o feocromocitomas), en el caso del

HBOC no es inusual la existencia de fenocopias. Se define una fenocopia como la presencia de un cáncer de mama no relacionado con el síndrome, en el seno de una familia HBOC. Un estudio en población israelí ha establecido que la frecuencia de fenocopias en esa población es del 2,26% (Bernholtz et al., 2012).

Además de seleccionar a las familias a las que se les ofrece la posibilidad de la realización de un test genético, también es importante la selección del individuo a testar, dentro de todos los miembros de una familia. El primer individuo de una familia dónde se realiza un test se llama individuo probando. Su elección es clave para la interpretación del resultado del test genético, ya que únicamente se realiza un test genético por familia (en el individuo probando) y el resultado se aplica a todos sus familiares. Debe escogerse al individuo con mayor probabilidad de ser portador dentro de una familia; preferiblemente afecto y, a ser posible, el de mayor afectación y menor edad al diagnóstico.

#### *Realización de un test genético*

Según el último consenso publicado de la sociedad americana de oncología clínica (ASCO), únicamente se recomienda ofrecer un test genético cuando los resultados del test genético ayudan en el diagnóstico o influyen en el manejo médico o quirúrgico del individuo o de sus familiares a riesgo (Robson et al., 2010). A pesar de las evidencias que sugieren la existencia de múltiples genes de susceptibilidad al cáncer de mama, en el momento actual y en el contexto del consejo genético solo se considera de utilidad clínica la realización del test genético en *BRCA1/2*. (National Cancer Institute: PDQ® Genetics of Breast and Ovarian Cancer, Fecha de acceso Enero de 2014, <http://www.cancer.gov/cancertopics/pdq/genetics/breast-and-ovarian/healthprofessional>). El análisis molecular de los genes *BRCA1* y *BRCA2* debe incluir la búsqueda de mutaciones puntuales y pequeñas inserciones/deleciones en toda la región codificante y secuencias intrónicas adyacentes así como el estudio de grandes reordenamientos. A la hora de la realización del test genético, se deben diferenciar dos situaciones:

#### A) Estudio inicial del individuo probando.

En el individuo probando se debe realizar el estudio genético completo. El resultado del test puede ofrecer las siguientes interpretaciones:

-Resultado informativo: Un resultado se clasificará como informativo cuando se haya identificado una variante genética en *BRCA1* o *BRCA2* capaz de explicar el síndrome HBOC. Este resultado proporciona una información precisa sobre el riesgo, permite explicar la agregación

familiar del cáncer y es una oportunidad para expandir la oferta del consejo genético al resto de los miembros de la familia.

-Resultado no informativo: Un resultado se clasificará como no informativo cuando no se haya identificado ninguna variante genética en *BRCA1* ni *BRCA2* capaz de explicar el síndrome HBOC. En ese caso, no se puede descartar la existencia de variantes genéticas en *BRCA1/2* que no se hayan incluido en el test o la existencia de variables genéticas en otros genes adicionales. Dentro de esta categoría se incluyen los resultados de significado incierto. Un resultado se clasificará como de significado incierto cuando se haya identificado alguna variante genética en *BRCA1* o *BRCA2* pero se desconozca su asociación con el síndrome HBOC. Estos resultados además de no resultar informativos, añaden un alto grado de incertidumbre al resultado del test genético. En este caso se sugiere que la variante identificada sea objeto de investigación con el objetivo de dilucidar la relevancia de dicha variante (ver capítulo 7, página 53)

#### B) Estudio de individuos adicionales al probando

Tras un resultado positivo en un individuo probando se ofrece al resto de los miembros de la familia la posibilidad de la realización de un test genético. En este caso el test se centra en la identificación de la variante genética concreta. Se pueden presentar dos resultados;

-Resultado positivo: El individuo es portador de la variante genética que en su familia es causa del síndrome HBOC. Esto lo clasifica como individuo con un riesgo mucho mayor que el poblacional y será objeto de futuras estrategias reductoras del riesgo.

-Resultado verdadero negativo: El individuo no es portador de la variante genética que en su familia es causa del síndrome HBOC y por lo tanto, está expuesto a un riesgo de cáncer de mama (y ovario) similar al de la población general (Korde et al., 2011).

## 7. Clasificación de variantes genéticas en BRCA1/2

Para ilustrar la relevancia de la clasificación de variantes, una revisión de la base de datos BIC (último acceso enero de 2013) reflejó que un 35 y 45 % de las variantes individuales depositadas corresponden respectivamente con cambios de aminoácido en BRCA1 y BRCA2. Con respecto a su clasificación, en *BRCA1* el 5% de estos cambios se clasifican como patogénicos, el 3% como polimorfismos y el 92% como significado incierto. En *BRCA2*, el 1,6% se clasifican como patogénicos, el 3,6% como polimorfismos y el 94,8% como significado incierto. Se estima que en los Estados Unidos de América, un 15% de los individuos testados para *BRCA1/2* es portador de una variante patogénica, mientras que un 5% es portador de una variante de significado incierto (Vallee et al., 2012). El porcentaje de individuos portadores de variante de significado incierto detectados en distintas poblaciones es variable y en ocasiones alcanza hasta el 20% de los individuos analizados. A la hora de la realización del test genético hay que tener en cuenta la posibilidad de encontrar este tipo de variantes y hay que considerar tres puntos:

1. Una vez encontrado un cambio en la secuencia de BRCA1/2, ¿se trata de una variante patogénica, de un polimorfismo o de una variante de significado incierto?
2. Según se trate de una variante de un tipo u otro, ¿cómo debe informarse a los pacientes?
3. En el caso de una variante de significado incierto, ¿qué estrategias se pueden utilizar para tratar de categorizar esta variante como polimorfismo o patogénica?

A continuación se revisarán las evidencias científicas que se pueden utilizar con el objetivo de clasificar variantes en los genes *BRCA1/2*. De especial relevancia es la calidad de la información que cada de las estrategias puede aportar y la cuantificación estadística de esa información. De manera general, las evidencias se pueden clasificar en directas e indirectas. Las directas son aquellas que miden una asociación entre la variante y la enfermedad. Las indirectas, por su parte, se basan en el efecto observado o predicho de la variante sobre aspectos del gen en cuestión, que se asumen como determinantes de la patogénesis (Goldgar et al., 2008).

### *Evidencias directas: análisis de la cosegregación*

Es la evidencia más directa y menos ambigua. De hecho, el método es muy similar al análisis de ligamiento que permitió identificar a los genes *BRCA1* y *BRCA2*. En el análisis de ligamiento se analiza la segregación a lo largo de una familia de marcadores genéticos asociados con una posible variante causal desconocida. Por su parte, en el análisis de la cosegregación se analiza la segregación de la propia variante. Únicamente requiere asumir unos valores de penetrancia y frecuencia, que se

inferen de los valores de aquellas variantes patogénicas conocidas. Como resultado se obtiene una medida estadística de *LOD score*, cuando el valor supera 3, demuestra una evidencia fuerte de cosegregación. Aunque la cosegregación es un método robusto, su poder estadístico depende del número de individuos de una misma familia analizados. Desafortunadamente, la mayoría de las familias en el contexto del consejo genético no poseen el tamaño suficiente para este tipo de análisis y además puede ser complejo el obtener muestras de múltiples individuos afectados. Por otra parte, aunque se confirme la cosegregación, no se puede descartar que la variante sea únicamente un marcador ligado fuertemente a una alteración patogénica no detectada. Como ventaja, el análisis de la cosegregación puede integrar datos de distintas familias portadoras de la misma variante de significado incierto. A pesar de ello, son pocas las veces que las evidencias aportadas por el análisis de cosegregación permiten categorizar una variante. Adicionalmente, el método exige asumir que todas las variantes patogénicas presentan un valor de penetrancia similar y que las UVs se ajustarán también a ese valor de penetrancia. Es posible que cierta variante de significado incierto no se ajuste al patrón de segregación esperado para una variante de penetrancia elevada y en cambio sí que se ajuste a la segregación de una variante con una penetrancia más baja (Lindor et al., 2012; Mohammadi et al., 2009).

#### *Evidencias directas: estudios caso-control*

Se basa en la comparación de la frecuencia de la variante de interés en una serie de casos y de controles. Metodológicamente es un método robusto que ofrece como resultado la medida estadística de riesgo relativo (O.R.). Aunque es un método muy útil en el análisis de variantes frecuentes (>1%), la mayoría de las UVs son individualmente infrecuentes ( $\approx 0,1\%$ ) y el tamaño muestral requerido es tan grande que generalmente se hace inasumible. En la práctica, el estudio caso-control es un método rápido para descartar posibles variantes neutrales. Tras genotipar la variante de interés en un grupo de controles ( $n= 100-200$ ), si su frecuencia es 1% o superior es muy improbable que se trate de una variante patogénica de penetrancia elevada. Se espera que la inclusión de UVs infrecuentes en los grandes estudios caso control, como por ejemplo los de los consorcios BCAC y COGS, permita obtener evidencias que contribuyan a su clasificación (Goldgar et al., 2008).

#### *Evidencias directas: historia personal y familiar*

Se basa en el hecho de que las variantes patogénicas tienden a aparecer de manera más frecuente en familias con una mayor incidencia de los casos de cáncer relacionados con el síndrome. El análisis de una variante de significado incierto en familias con distintos grados de historia familiar

permite analizar su distribución y compararla con la distribución de las variantes patogénicas de la misma población. La ventaja de esta aproximación es que únicamente requiere el análisis del genotipo de un individuo por familia. La desventaja es que requiere información detallada de todas las familias analizadas en una población y no solo de aquellas portadoras de UVs (Gomez Garcia et al., 2009). Dentro de la información de utilidad que se puede obtener de las familias se puede incluir tanto datos que se recogen habitualmente en el consejo genético (edades de diagnóstico y tipos de cáncer), así como datos anatomopatológicos (ER, PR, HER2, etc) que no se recogen habitualmente (Hofstra et al., 2008).

#### *Evidencias directas: Co-ocurrencia*

Experimentos en ratones han demostrado que la delección bialélica de *Brca1* o *Brca2* resulta en letalidad embriónica (Evers and Jonkers, 2006). En base a esta letalidad embriónica, clásicamente se ha asumido que no existen individuos portadores de mutaciones bialélicas en estos genes. Recientemente se han descubierto mutaciones bialélicas en el gen *BRCA2* (Howlett et al., 2002) implicadas en la anemia de Fanconi y el primer caso de individuo portador de mutaciones bialélicas en el gen *BRCA1* (Domchek et al., 2012). Ambos datos sugieren la posibilidad de existencia de mutaciones bialélicas en ambos genes, aunque produciendo fenotipos mucho más extremos y fácilmente diferenciables.

De manera práctica, estas evidencias se traducen en que, en ausencia de fenotipos extremos la demostración de la presencia de una variante patogénica en una copia del gen y de una UV en la otra copia del gen (fase en *trans*) proporciona una evidencia muy fuerte en contra de la patogenicidad de la UV. Como desventaja, para demostrar la fase de dos variantes a menudo se requiere genotipar a múltiples individuos de una familia (Goldgar et al., 2008).

#### *Evidencias indirectas: análisis in silico del efecto de los cambios de aminoácido*

Basándose en la bioquímica básica y la estructura de proteínas, parece lógico pensar en evaluar el efecto de los cambios de aminoácido en función de las diferencias físico-químicas. No obstante, ninguno de los métodos basados únicamente en estas diferencias ha demostrado ser suficientemente robusto. Otra manera de evaluar la relevancia de los cambios de aminoácido se basa en la evidencia de que las posiciones que se mantienen conservadas en una proteína a lo largo de la evolución son relevantes para su función y en que la mayoría de cambios de aminoácido patogénicos conocidos se localizan en estas posiciones conservadas. La herramienta disponible para analizar la conservación de residuos a lo largo de la evolución es el alineamiento múltiple de secuencias proteicas. Una vez identificados los residuos conservados, se debe evaluar el efecto del cambio de aminoácido concreto en esa posición. Esta tarea no es sencilla y se han desarrollado una serie de

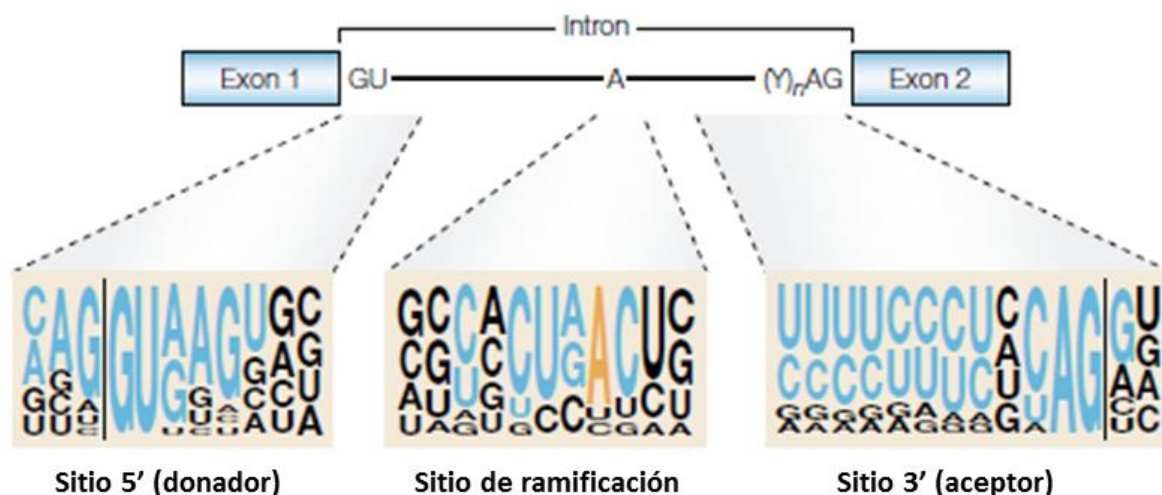
algoritmos que realizan el alineamiento múltiple de secuencias (obteniendo las secuencias del gen de interés de las bases de datos disponibles o partiendo de un alineamiento múltiple aportado por el usuario) y analizan el efecto de los cambios de aminoácido. Algunos ejemplos dentro de la variedad de algoritmos disponibles son SIFT, Panther y align-GVGD. Otra manera de analizar el efecto de los cambios de aminoácido es el análisis del efecto de un cambio de aminoácido sobre la estructura de la proteína. El algoritmo PolyPhen utiliza métodos basados en alineamientos múltiples, a los cuales se les añade información de la estructura cristalina de la proteína (Tavtigian et al., 2008b).

La calidad del alineamiento múltiple utilizado como referencia es clave en la calidad de los resultados de estos algoritmos (dependiendo de la variabilidad de la proteína en cuestión la profundidad –número de especies incluidas en el análisis- requerida será diferente). Por ello align-GVGD incluye una serie de alineamientos múltiples revisados y de alta calidad para los genes de susceptibilidad al cáncer más relevantes. Por otra parte, align-GVGD ha sido calibrado con un conjunto de 70000 individuos testados para mutaciones en *BRCA1/2*, lo cual ha permitido establecer una serie de grados en la predicción del efecto de los cambios de aminoácido. De esta manera cada grado correlaciona con una medida estadística de una probabilidad de patogenicidad (Tavtigian et al., 2008a).

#### *Evidencias indirectas: análisis in silico y experimental del splicing*

El procesamiento correcto de los intrones es un requisito imprescindible para la síntesis de una proteína. Variables genéticas que producen defectos en el *splicing* se han demostrado como causa de múltiples enfermedades hereditarias (Wang et al., 2012). El *splicing* es un proceso complejo en el que intervienen por una parte un complejo riboproteico conocido como *spliceosoma* y por otra parte multitud de señales presentes en la secuencia de ARNm.

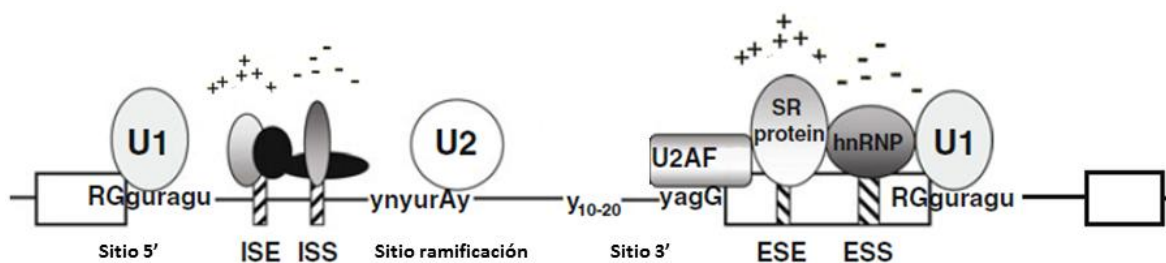
Dentro de las señales presentes en el ARNm, el sitio 5' (o donador), el sitio 3' (o aceptor) y el sitio de ramificación definen los límites entre el exón y el intrón y están implicados directamente en la interacción con el *spliceosoma* y en las reacciones de procesamiento del ARNm. Estas señales, a pesar de su papel clave, están caracterizadas por su bajo grado de conservación (Figura 10).



**Figura 10:** Señales que definen los límites intrón-exón

En la figura se muestra la estructura de los límites intrón-exón. Tres señales son relevantes para definir estos límites y en la parte inferior se muestra las secuencias consenso de estas señales. Como se puede observar en general el grado de conservación de estas secuencias no es muy elevado. Las únicas regiones conservadas completamente son las primeras/últimas 2 pb del intrón. Estas secuencias se conocen como sitios canónicos (adaptado de (Cartegni et al., 2002))

Este bajo grado de conservación permite que a menudo distintas señales estén accesibles al *spliceosoma* de una manera simultánea. Otro tipo de señales conocidas como potenciadores e inhibidores aparecen distribuidos a lo largo del ARNm y actúan como reguladores del *splicing*, permitiendo la existencia de fenómenos de *splicing* alternativo (Cartegni et al., 2002) (Figura 11).



**Figura 11:** Señales potenciadoras y silenciadoras del *splicing*

U1, U2 y U2AF forman parte de los complejos ribonucleoproteicos que conforman el *spliceosoma*. Secuencias del ARNm conocidas como potenciadores presentes en el exón (ESE) o en el intrón (ISE) interactúan con las proteínas SR produciendo para señales positivas que favorecen la inclusión del exón en el ARNm maduro. Por el contrario, otras secuencias presentes en el ARN actúan como silenciadoras (ESS y ISE) mediante el reclutamiento de las ribonucleoproteínas hnRNP. Las señales silenciadoras promueven la exclusión de los exones del ARNm maduro.

El análisis del efecto sobre el *splicing* de una variante puede aportar evidencias en la clasificación de una UV (Spurdle et al., 2008a). Sus ventajas son:

1. La existencia de programas que predicen la probabilidad con la que un cambio de secuencia alterará los patrones normales de *splicing*.
2. La facilidad del diseño experimental en el análisis del *splicing*

3. Permite obtener evidencias para cambios sinónimos o cambios intrónico además de cambios de aminoácido.

No obstante, el análisis del *splicing* no está exento de ciertas desventajas, como son:

1. La necesidad de disponer de una muestra de RNA que, de manera general, no se recoge en el proceso de consejo genético.
2. La posibilidad de detectar patrones de alteración en el *splicing* cuya relevancia clínica no es clara
3. La necesidad de unos criterios claros en la clasificación de los resultados y de estandarización de los métodos de análisis.

La detección de una aberración clara en el *splicing* permite clasificar con mucha confianza a una variante como patogénica. Por el contrario, la detección de un patrón de *splicing* normal o no concluyente no aporta información para clasificar a la variante como polimorfismo.

#### *Evidencias indirectas: Ensayos funcionales in vitro*

El uso de ensayos funcionales se basa en la evidencia de que *BRCA1* y *BRCA2* son genes supresores de tumores. Se asume que una pérdida en la función de un gen supresor de tumores se relaciona con un incremento a la predisposición de padecer cáncer. Debido al carácter multifuncional de *BRCA1* y *BRCA2* se han desarrollado múltiples ensayos *in vitro* diferentes, que a menudo analizan de manera separada la función de cada uno de los dominios. Aún no está claro cuáles son las funciones que específicamente están implicadas en la susceptibilidad al cáncer, por ello la capacidad de discriminar variantes claramente patogénicas de polimorfismos debe ser validada para cada ensayo. Los ensayos que han sido validados para *BRCA1* son el test de activación de la transcripción y el test de actividad ubiquitin ligasa. En *BRCA2* los test validados son el de reparación mediada por recombinación homóloga y el de amplificación del centrosoma (Couch et al., 2008).

#### *Modelo probabilístico multifactorial*

Teniendo en cuenta las ventajas y desventajas de cada una de las evidencias, en la actualidad se propone el uso de un modelo probabilístico bayesiano que integre las múltiples evidencias para crear un valor de la probabilidad de la patogenicidad de una UV. Este modelo parte de una probabilidad *a priori* sobre la cual se integran evidencias para dar lugar a una probabilidad *a posteriori*. La clasificación de variantes se considera dinámica, en el sentido de que a medida que avanzan los conocimientos científicos la relevancia que adquieren ciertas evidencias puede variar. El modelo más frecuentemente utilizado considera la probabilidad derivada de Align-GVGD como

probabilidad *a priori*, integrando también para las secuencias intrónicas la posibilidad de que alteren el *splicing* (Tabla 1).

**Tabla 1:** Probabilidad *a priori* según el modelo multifactorial

Grado Align-GVGD	Probabilidad	Intervalo de Confianza
C65*	0,81	(0,61-0,91)
C35-C55*	0,66	(0,34-0,93)
C15-C25*	0,29	(0,09-0,56)
C0*	0,03	(0,00-0,06)
Fuera de dominios	0,02	(0,00-0,04)
Sitios canónicos de <i>splicing</i>	0,96	(0,91-1,00)
Intrónicas fuera sitio	0,26	(0,15-0,39)

\*El modelo asigna una probabilidad *a priori* basada en el grado align-GVGD para aquellos cambios de aminoácido localizados en los dominios funcionales, reconociendo como tal en BRCA1 al dominio de unión a BARD1 (aa 1-102) y el dominio BRCT (aa 1396-1863) y en BRCA2 al dominio de unión a ADN (aa 2400-3190)

Las evidencias que se utilizan para construir la probabilidad *a posteriori* son: el análisis de la cosegregación, la co-ocurrencia, la historia familiar y personal y los datos anatomopatológicos del tumor (Lindor et al., 2012). Previamente se incluyó en el modelo el análisis de pérdida de heterocigosidad (LOH), según el cual, en el tumor se pierde el alelo normal del gen supresor de tumores de acuerdo con la hipótesis de Knudson. Sin embargo, resultados recientes han demostrado que no todas las variantes patogénicas sufren pérdida del alelo normal y por ello esta evidencia se ha excluido del modelo multifactorial (Beristain et al., 2010; Hofstra et al., 2008; Spurdle et al., 2008b). Por su parte, aunque el análisis del *splicing* se considera una evidencia muy poderosa, hasta el momento no ha podido ser integrada en el modelo multifactorial.

En teoría, con el objetivo de establecer una clasificación objetiva y estandarizada, se debería aplicar el modelo probabilístico multifactorial a cada una de las nuevas variantes encontradas en BRCA1/2, independientemente del tipo que se trate. Sin embargo, el modelo es laborioso y aplicarlo a cada una de las variantes nuevas restaría eficacia y retrasaría la aplicación del consejo genético en muchos individuos. Por ello, únicamente se aplica el modelo multifactorial en aquellas variantes nuevas que presentan dificultad en su clasificación, como es el caso de los cambios de aminoácido, los cambios silenciosos, los cambios en intrones y las inserciones/deleciones de un pequeño número de aminoácidos. Por su parte, aquellas variantes que producen un cambio en la fase del marco de lectura del ARNm, aquellas que introducen un codón de parada prematuro o las localizadas en las posiciones canónicas de los sitios de *splicing* se consideran de fácil clasificación, permitiéndose clasificarlas de manera directa, como patogénicas sin requerir la aplicación del modelo multifactorial.

Sin embargo, hay que tener ciertas precauciones al asumir esta clasificación directa. Con respecto a los codones de parada prematuros, en *BRCA2* se ha descrito la presencia de un codón de parada polimórfico en la Lisina 3326 que, a pesar de eliminar los últimos 93 aminoácidos de la proteína, no segrega con la enfermedad en familias HBOC y se encuentra presente en la población control y en familias HBOC con una frecuencia similar (2%) (Mazoyer et al., 1996). Estudios posteriores sugieren que esta variante es hipomórfica, habiéndose detectado un incremento moderado del riesgo (O.R. =4) en individuos con cáncer pancreático familiar (Martin et al., 2005), así como la presencia de esta variante en individuos con anemia de Fanconi (Howlett et al., 2002). Teniendo en cuenta esta consideración se debe tener especial cautela a la hora de clasificar aquellos codones de parada prematuros o variantes que produzcan un cambio de fase de lectura en posiciones 3' con respecto a Lisina 3326 en *BRCA2*. Con respecto a *BRCA1* no se ha descrito este efecto.

Otra consideración especial se refiere a las variantes en posiciones canónicas de sitios de *splicing*. El hecho de que estas posiciones son las únicas conservadas en los sitios consenso de *splicing* sugiere que cualquier cambio en estas posiciones producirá aberraciones en el procesado de los intrones. No obstante, el resultado concreto de la aberración del *splicing* influirá en la clasificación de la variante. Mientras que aquellas aberraciones del *splicing* que producen un cambio en la fase de lectura del ARNm se clasificarán como patogénicas con un alto grado de evidencia, aquellas que produzcan la delección/inserción de un pequeño número de aminoácidos pasan a considerarse variantes de difícil clasificación. En la base de datos DBASS (<http://www.dbass.org.uk/>, Fecha de acceso: Enero 2013) aparecen recopilados ejemplos de variantes que alteran el *splicing* junto con el efecto concreto que producen. Se puede observar la incidencia tanto de aberraciones que producen cambios en la fase de lectura como de aquellas que no la alteran. Por ello, hay que tener especial precaución a la hora de clasificar de manera directa variantes situadas en sitios canónicos de *splicing*.

El grupo de trabajo de UVs de la IARC sugiere que, tras la aplicación del modelo multifactorial, las variantes se clasifiquen en base a su probabilidad posterior en cinco categorías (Plon et al., 2008). Para cada clase se hace una recomendación con respecto a los puntos más importantes relacionados con el consejo genético, como son:

1. La utilidad de dicha variante como test predictivo para la individualización del riesgo de cáncer de mama y la posibilidad de extender el test al resto de los miembros de la familia.
2. La recomendación de realizar estudios de investigación posteriores, incluyendo el análisis de miembros adicionales de la familia.
3. Las recomendaciones a tener en cuenta en referencia a la susceptibilidad al cáncer en los individuos portadores de dicha variante.

**Tabla 2:** Clasificación de variantes propuesta por IARC

Clase	Definición	Prob. a posteriori	Test predictivo	Investigación	Recomendaciones individuales
5	Patogénica	>0,99	Si	No	Seguimiento como portadores
4	Probablemente patogénica	0,95-0,99	Si	Si	Seguimiento como portadores
3	Incierta	0,005-0,949	No	Si	Seguimiento según historia familiar
2	Probablemente no patogénica	0,001-0,049	No	Si	Similar a la población general
1	No patogénica	<0,001	No	No	Similar a la población general

El proceso de clasificación de variantes es complejo y requiere la recopilación de múltiples evidencias y la interacción de múltiples investigadores, así como la colaboración con otros profesionales. En el año 2008 se creó la base de datos Literature Unclassified Variants (LOVD) (<http://chromium.liacs.nl/LOVD2/cancer/home.php>), con el objetivo de recopilar aquellas evidencias disponibles en la literatura científica sobre UVs en los genes *BRCA1/2*. En la actualidad (junio 2013) esta base de datos contiene evidencias de 528 y 507 UVs de los genes *BRCA1* y *BRCA2* respectivamente. En el año 2012, se creó la base de datos Ex-Uv (<http://brca.iarc.fr/LOVD/home.php>), que recopila todas aquellas variantes que han sido clasificadas mediante la aplicación del modelo probabilístico multifactorial. Con el objetivo de establecer colaboraciones científicas para la clasificación de UVs, en el año 2009 se creó el consorcio ENIGMA (Spurdle et al., 2012). Se espera que este consorcio contribuya de manera notable a la clasificación de UVs gracias a la colaboración de distintos equipos científicos.



## Objetivos

Con el objetivo de mejorar el proceso de consejo genético, en esta tesis se han analizado los factores genéticos que contribuyen al síndrome de cáncer de mama y ovario hereditario en población española.

Esta tesis se presenta como una recopilación de seis artículos. El doctorando es autor principal de cuatro de estos trabajos, que se presentan en el cuerpo de la tesis. Los otros dos artículos son fruto de colaboraciones y se presentan de manera separada en el anexo I (página 169.). Se recomienda la lectura de los artículos en el orden en que se presentan en este apartado de objetivos, con independencia de su localización en el cuerpo de la tesis o en el anexo I.

Los trabajos realizados se centran en dos objetivos concretos:

### **Objetivo 1**

Análisis de la relevancia del estudio del *splicing* en la clasificación de variantes genéticas en *BRCA2*. Este objetivo ha sido tratado en dos artículos. En primer lugar “*Comparison of mRNA splicing assay protocols across multiple laboratories: Guidelines for best practice in standardized clinical*”, publicado en *Clinical Chemistry*, establece una serie de recomendaciones metodológicas para el análisis de *splicing*. En segundo lugar, “*Capillary Electrophoresis analysis of Conventional Splicing Assays: IARC Clinical Classification of 31 BRCA2 genetic variants*”, publicado en *Human Mutation*, realiza en análisis de *splicing* de una serie de variantes del gen *BRCA2*.

### **Objetivo 2**

Búsqueda de variantes relacionadas con el síndrome HBOC en familias BRCA2 tras cribado genético. Este objetivo se divide en dos sub-objetivos:

Objetivo 2 A: Búsqueda de variantes patogénicas en el gen *BRCA2* que no se detectan mediante la estrategia de cribado basada en secuenciación. Este objetivo se ha tratado en dos artículos. Por una parte, “*Characterization of four novel BRCA2 large genomic rearrangements in Spanish breast/ovarian cancer families. Review of the literature, and reevaluation of the genetic mechanisms involved in their origin*” se encuentra publicado en la revista *Breast Cancer Research and Treatment* y se centra en la caracterización de grandes reordenamientos genómicos en el gen *BRCA2*. Por otra parte, “*BRCA2 c.6937+594T>G screening in a cohort of*

*302 BRCA1/2 mutation negative Spanish HBOC families*” se encuentra en proceso de publicación y se centra en el análisis de variantes en regiones intrónicas alejadas de los exones.

Objetivo 2 B: Búsqueda de variantes patogénicas implicadas en HBOC en genes candidatos adicionales a *BRCA1/2*. Este objetivo se ha tratado en dos artículos. El primero de ellos, *“Low prevalence of SLX4 loss-of-function mutations in Non-BRCA1/2 breast and/or ovarian cancer families”*, se encuentra publicado en la revista *European Journal of Human Genetics* y se centra en la búsqueda de variantes patogénicas en el gen *SLX4* implicadas en HBOC. El segundo artículo, *“About 1% of the breast and ovarian Spanish families testing negative for BRCA1 and BRCA2 are carriers of RAD51D pathogenic mutations”* se centra en la búsqueda de variantes patogénicas en el gen *RAD51D* implicadas en HBOC. Este artículo se ha publicado en *International Journal of Cancer*.

## 1. Análisis de la relevancia del estudio del *splicing* en la clasificación de variantes genéticas en *BRCA2*

*Análisis mediante electroforesis capilar de ensayos convencionales de splicing: Clasificación analítica y clínica de 31 variantes en el gen BRCA2 mediante las normas IARC.*

Las variantes de significado incierto en genes de predisposición a enfermedades suponen un gran reto para el consejo genético. Sin embargo, aquellas VUs que producen alteraciones en los procesos de *splicing* pueden ser fácilmente estudiadas en ensayos *in vitro*. Desafortunadamente, la interpretación analítica y clínica de estos ensayos a menudo es compleja. En este trabajo se lleva a cabo el análisis de 31 variantes genéticas en *BRCA2*. Todas las variantes fueron analizadas mediante RT-PCR seguida de electroforesis capilar y secuenciación directa. Cuando los ensayos no ofrecieron resultados claros (Clase 2 o Clase 5 de acuerdo con las guías analíticas de IARC), se realizaron técnicas complementarias como PCR cuantitativa o ensayos de minigenes. Estos últimos fueron realizados como un nuevo vector (pSAD) desarrollado por los autores del presente manuscrito (Patente P201231427 CSIC). Se han identificado tres variantes con relevancia clínica, Clase 5, (c.682-2A>G, c.7617+1G>A, and c.8954-5A>G) y 27 variantes que analíticamente se clasifican como Clase 2 (no inductoras de alteraciones en el *splicing*). Adicionalmente, se demuestra que la variante c.7806-14T>C (rs9534262) es un *splicing quantitative trait locus* (sQTL).



## Capillary Electrophoresis Analysis of Conventional Splicing Assays: IARC Analytical and Clinical Classification of 31 *BRCA2* Genetic Variants

Gorka Ruiz de Garibay,<sup>1</sup> Alberto Acedo,<sup>2</sup> Zaida García-Casado,<sup>3</sup> Sara Gutiérrez-Enríquez,<sup>4</sup> Alicia Tosar,<sup>1</sup> Atocha Romero,<sup>1</sup> Pilar Garre,<sup>1</sup> Gemma Llorc,<sup>5</sup> Mads Thomassen,<sup>6</sup> Orland Díez,<sup>7</sup> Pedro Pérez-Segura,<sup>8</sup> Eduardo Díaz-Rubio,<sup>1,8</sup> Eladio A. Velasco,<sup>2</sup> Trinidad Caldes,<sup>1</sup> and Miguel de la Hoya<sup>1\*</sup>

<sup>1</sup>Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain; <sup>2</sup>Grupo de Splicing y Cáncer, Instituto de Biología y Genética Molecular (CSIC-UVA), Sanz y Forés 3, Valladolid 47003, Spain; <sup>3</sup>Laboratory of Molecular Biology, Fundación Instituto Valenciano de Oncología, Valencia, Spain; <sup>4</sup>Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona, Barcelona Spain; <sup>5</sup>Unidad de Consejo Genético, Servicio de Oncología Médica, Instituto Oncológico del Vallés, Corporación Sanitaria Parc Taulí, Consorci Sanitario de Terrassa, Barcelona, Spain; <sup>6</sup>Department of Clinical Genetics, Odense University Hospital, Soender Boulevard 29, Odense C 5000, Denmark; <sup>7</sup>Oncogenetics Group, University Hospital Vall d'Hebron, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain; <sup>8</sup>Servicio de Oncología Médica, Hospital Clínico San Carlos, Madrid, Spain

Communicated by David E. Goldgar

Received 17 June 2013; accepted revised manuscript 23 September 2013.

Published online 7 October 2013 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22456

**KEY WORDS:** *BRCA2*; splicing; UVs; capillary electrophoresis; minigene; qPCR

**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 and direct sequencing. If assays did not produce clear-cut outputs (Class-2 or Class-5 according to analytical International Agency for Research on Cancer guidelines), 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 three 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 locus.

Hum Mutat 00:1–5, 2013. © 2013 Wiley Periodicals, Inc.

Additional Supporting Information may be found in the online version of this article.

\*Correspondence to: Miguel de la Hoya, Laboratorio de Oncología Molecular, Hospital Clínico San Carlos, C/Martín Lagos s/n, Madrid 28040, Spain. E-mail: mdhoya@hotmail.com

Contract grant sponsors: EAV was supported by research grants PI10/2910 (Instituto de Salud Carlos III), CS1004A10-2 (Consejería de Educación, Regional Government of Castilla y León) and BIO39/VA27/10 (Consejería de Sanidad, Regional Government of Castilla y León). Alberto Acedo is supported by the European Social Fund and Regional Government of Castilla y León (P.O. 2007-2013). ZGC is supported by the Genetic Counseling Program (Regional Government of Valencia). SGE is funded by a Miguel Servet contract (CP10/00617, Instituto de Salud Carlos III). HUVH-VHIO research was supported by a Miguel Servet research grant CP10/00617, Instituto de Salud Carlos III. MH was supported by research grants PI 09/00859 and PI 12/00539 (Instituto de Salud Carlos III). Instituto de Salud Carlos III research grants are partially supported by European Regional Development FEDER Funds.

Germ-line-inactivating mutations in *BRCA2* (MIM #600185) confer markedly increased risks of breast and ovarian cancer. Arguably, *BRCA2* protein-truncating mutations (PTCs) are considered inactivating (clinically relevant) *per se* (PTCs located 3' to the last 50 nucleotides of the penultimate exon representing one exception). However, a high proportion of the genetic variants identified through clinical testing are not PTCs, but rather missense, synonymous, or intronic changes of unknown functional effects [Spurdle et al., 2012]. Several approaches have been proposed to elucidate the clinical relevance (if any) of these so-called unclassified variants (UVs), including association with family history, cosegregation analysis, and co-occurrence with a pathogenic mutation. These and other evidence (provided that they are statistically independent) can be integrated into multifactorial models that allow likelihood classification of variants [Goldgar et al., 2004; Lindor et al., 2012]. Based on posterior probability outputs, the International Agency for Research on Cancer (IARC) working group on UVs proposed a five-tier classification scheme with clinical recommendations [Plon et al., 2008]. According to this scheme, Class-1 variants are non-pathogenic, whereas Class-5 variants are definitely pathogenic.

*In vitro* assays assessing the effect of UVs on splicing (or protein activity) can be highly informative in the classification process, especially if multifactorial likelihood classification is not feasible. In particular, *in vitro* splicing assays are widely used in the clinical setting [Walker et al., 2013], probably reflecting the fact that general approaches to predict and verify splicing alterations are relatively straightforward. Unfortunately, interpretation of results is often more challenging [Spurdle et al., 2008; Walker et al., 2013]. The IARC Working Group on UVs has developed guidelines for the prediction and assessment of splicing alterations [Spurdle et al., 2008]. Their authors adapted the IARC five-tier clinical classification scheme (originally intended for clinical classification of multifactorial model outputs) to the particular case of interpreting *in vitro*

splicing results [Spurdle et al., 2008]. Due to the specific characteristics of these *in vitro* assays (such as absence of numerical outputs), major modifications were introduced, including the fact that this is a four-tier classification scheme according to which Class-2 and Class-5 indicate clear-cut splicing assay outputs, the first indicating that the variant allele produces the same transcript(s) as the wild-type allele, and the latter indicating that the UV allele expresses major PTC transcripts. Class-1 was avoided because splicing analysis outputs *per se* do not allow discarding a pathogenic role (protein function and/or other biological features might be altered). Most relevant, this four-tier is not a clinical, but rather an analytical classification scheme. As such, analytical Class-2, Class-3, and Class-4 outputs do not have specific clinical recommendations. In this regard, analytical Class-5 is exceptional, as *BRCA2* PTCs are considered clinically relevant *per se*. The analytical classification guidelines have been updated recently [Walker et al., 2013]. See Supp. Table S1 for further details.

On practical grounds, quantitative (or semi-quantitative) estimation of aberrant/normal transcripts ratio (as expressed from the UV allele) is mandatory for clear-cut analytical classification [Spurdle et al., 2008; Walker et al., 2013]. Most important, two confounding effects, normal expression from the wild-type allele, and selective degradation of PTC transcripts by the nonsense-mediated mRNA decay (NMD) surveillance pathway, should be mitigated to produce robust outputs. For instance, PTC transcripts can be stabilized by performing cell cultures in the presence of cycloheximide (or other NMD inhibitors). Interferences from the wild-type allele can be avoided by developing reporter minigene assays. Unfortunately, these time-consuming approaches cannot be easily implemented into the routine of a Clinical Diagnostics Laboratory.

Here, we show that incorporating capillary electrophoresis (EP) and direct sequencing of exonic SNPs into the analysis of otherwise conventional splicing assays (RT-PCRs of fresh-blood isolated RNAs) are able to generate clear-cut analytical outputs (so to speak, analytical Class-2 and Class-5 outputs according to the IARC guidelines). However, we show as well that some variants/carriers demand more specific (and time-consuming) analytical approaches (such as dedicated minigenes or qPCR).

As a proof of principle, we analyzed the influence of 31 *BRCA2* variants (Table 1) on splicing (see Supp. Methods and Supp. Table S2 for further details). *In silico* predictions are shown in Table 1 (splice sites) and Supp. Table S3 (enhancer and silencers). However, we performed *in vitro* splicing assays regardless of *in silico* predictions. Furthermore, we analyzed not only *BRCA2* UVs, but also clinical Class-1 and clinical Class-5 variants. We did it because we are interested in finding not only novel *BRCA2* splicing mutations relevant in the clinic, but also non-pathogenic splicing alterations (if any) that might contribute to elucidate the sometimes complex relationship between splicing alterations and disease. Last but not least, we are interested in evaluating whether conventional splicing assays provide clear-cut analytical outputs according to existing guidelines.

As far as we know, *BRCA2* c.682-2A>G is a novel variant not previously reported. Capillary EP analyses in carriers (33 cycles RT-PCR assays) detected (in addition to the full-length) a 264 nt peak representing roughly 10% of the overall signal (Fig. 1A). Interestingly, 45 cycles RT-PCR assays revealed five additional peaks (Fig. 1B). Size callings were compatible with exon 9 skipping (264 nt peak), and activation of five cryptic acceptor sites (CASs) located in intron 8 and exon 9 (Fig. 1B and C). All CASs were supported by *in silico* predictions (Fig. 1C). Direct sequencing demonstrated exon 9 skipping, and CAS4 activation (Fig. 1D). Other CAS activation events were not sequenced due to the low abundance of the corresponding RT-PCR products. These should be considered putative CAS activation

events inferred from capillary EP size calling. Similar peak patterns were observed in four additional variant carriers (Fig. 1E), but not in control samples (data not shown). Importantly, all aberrant splicing events produce PTC transcripts. Segregation analysis together with direct RT-PCR sequencing at informative loci (rs9534262 and rs1799955) demonstrated that full-length *BRCA2* expression was wild-type allele specific (Supp. Fig. S1). Taking together, the data qualify for inclusion into analytical Class-5 (see Supp. Table S1).

*BRCA2* c.7617+1G>A has been reported previously in Spanish [Gutierrez-Enriquez et al., 2009] and Danish [Thomassen et al., 2011] hereditary breast and ovarian cancer families. Both studies detected exon 15 skipping (a PTC event) in variant carriers, but biallelic wild-type expression was not investigated (analytical Class-4 outputs). Yet, cosegregation studies performed in Danish families supported a pathogenic role for this variant (clinical Class-5 variant). We have performed splicing assays in six *BRCA2* c.7617+1G>A carriers. As shown in Supp. Figure S2, capillary EP analysis (33 cycles RT-PCR) detected, in addition to the full-length, a 182 nt shorter peak compatible with exon 15 skipping. Direct sequencing confirmed the skipping event (Supp. Fig. S2). Capillary EP data (semi-quantitative) show roughly identical expression levels of full-length and  $\Delta 15$  transcripts. The data are compatible with the UV allele causing complete (or near complete) exon skipping, but do not prove it. This is so because capillary EP estimates (at best) the ratio between steady-state levels (accumulation) of wild-type and variant transcripts in carriers. This in turn might reflect a combination of multiple underlying processes, including reduced (or unusually higher) transcription rate from the UV allele, instable (or unusually stable) UV transcripts, and wild-type transcripts arising from the UV allele. Yet, segregation analysis together with direct RT-PCR sequencing at one informative locus (rs1799955) revealed that full-length *BRCA2* expression was indeed wild-type allele specific (Supp. Fig. S2). Overall, the data qualify for inclusion into analytical Class-5 (see Supp. Table S1). We have identified eight Spanish families carrying *BRCA2* c.7617+1G>A (Supp. Fig. S3). Genotyping of STRs and SNPs markers detected a common haplotype in carriers from all eight families, suggesting a founder effect (Supp. Table S4).

*BRCA2* c.8954-5A>G has been identified recently in Spanish Hereditary Breast and Ovarian Cancer families [de Juan Jiménez et al., 2013], but splicing defects, if any, were not investigated. The change is reported in the Universal Mutation Database (UMD) as a clinical Class-3 variant (<http://www.umd.be/BRCA2/>). We have performed splicing assays in two *BRCA2* c.8954-5A>G carriers. Agarose gel EP showed a single RT-PCR product both in controls and carriers (data not shown), but capillary EP detected an additional peak (+4 nt) in carriers (Supp. Fig. S4). Direct sequencing revealed a PTC transcript in which the last 4 nt of intron 22 have been incorporated into *BRCA2* exon 23 (Supp. Fig. S4). Measured by capillary EP, the aberrant mRNA represents roughly 30% of the overall signal. The data are compatible with the wild-type expression arising from the UV allele, but do not prove it (as discussed above). Since biallelic wild-type expression was not tested due to the lack of informative loci, it is not possible to classify the output into any specific analytical category (Class-4 or Class-5?) according to the 2008 guidelines. However, the same output qualifies for analytical Class-4 according to the 2013 guidelines (see Supp. Table S1). Interestingly, Human Splicing Finder (HSF) analysis predicted a *de novo* acceptor site at position c.8954-4 without disrupting exon 23 native acceptor site (Table 1), thus suggesting a mechanism through which c.8954-5A>G might indeed contribute to normal splicing. To investigate the wild-type expression arising from the UV allele, we used a reporter minigene assay (see Supp. Methods). The analysis revealed that c.8954-5G does not produce full-length mRNA (Supp. Fig. S5).

**Table 1. Splicing Analysis in 35 BRCA2 Variants**

Location	HGVS description <sup>c</sup>	dbSNP	MAF <sup>d</sup>	In silico splicing analysis <sup>a</sup>		Analytical classification of in vitro splicing assays outputs <sup>b</sup>			
				HSF	MaxEnt	Published data	Routine assay	Minigene	qPCR
Intron 1	c.-39-26G>A	–	.0004	No effect	No effect	–	Class-2	–	–
Intron 4	c.426-89T>C	rs3783265	.022	(c.426-81; 86.04%)	No effect	–	Class-2	–	–
Intron 6	c.516+14C>T	rs182828913	.0016	No effect	No effect	Class-2 <sup>e</sup>	Class-2	–	–
Intron 7	c.631+25C>T	–	.0008	No effect	No effect	–	Class-2	–	–
Exon 8	c.680C>T (p.Ala227Val)	rs149565664	.0004	–0.57%	–3.04%	–	Class-2	–	–
Intron 8	c.681+69C>T <sup>f</sup>	–	.0004	No effect	No effect	–	Class-2	–	–
	c.682-2A>G <sup>g</sup>	–	.0008	–33.39%	–112.46%	–	Class-5	–	–
Exon 10	c.1114A>C (p.His372Asn)	rs9534262	.11	No effect	No effect	Class-2 <sup>g</sup>	–	–	–
	c.1773-1776del (p.Ile591fs)	rs80359305	.0004	No effect	No effect	–	Class-5	–	–
Exon 11	c.3264dupT (p.Gln1089fs)	rs80359380	.0097	(c.3268; 103.25%)	(c.3283; 7.52%)	–	Class-5	–	–
	c.3396A>G (p.Lys1132=)	rs1801406	.25	No effect	No effect	–	–	–	–
	c.3807T>C (p.Val1269=)	rs543304	.14	No effect	No effect	–	–	–	–
	c.6825G>A (p.Gln2275=) <sup>f</sup>	–	.0004	No effect	No effect	–	Class-2	–	–
Exon 12	c.6928A>C (p.Thr2310Pro) <sup>f</sup>	–	.0004	(c.6935; 96.34%)	No effect	–	Class-2	–	–
Exon 14	c.7242A>G (p.Ser2414=)	rs1799955	.20	No effect	No effect	–	–	–	–
Intron 14	c.7435+8A>G <sup>g</sup>	–	.0004	(c.7436+7; 92.39%)	No effect	–	Class-2	–	–
Exon 15	c.7469T>C (p.Le2490Thr)	rs11571707	.0024	No effect	No effect	–	Class-2	–	–
Intron 15	c.7617+1G>A <sup>h</sup>	–	.0016	–29.43%	–83.47%	Nonclassifiable (Class-4) <sup>i,j</sup>	Class-5	–	–
Exon 16	c.7626G>A (p.Thr2542=)	rs61754138	.0004	No effect	No effect	–	Class-2	–	–
	c.7796A>G (p.Gln2599Gly) <sup>f</sup>	–	.0004	(c.7796; 88.03%)	No effect	–	Class-2	–	–
Intron 16	c.7806-14T>C	rs9534262	.40	No effect	–4.68%	–	Nonclassifiable <sup>k</sup>	Class-4 (Class-3)	Class-4 (Class-3)
Intron 17	c.7976+5G>T <sup>f</sup>	–	.0004	–16.83%	–74.19%	–	Class-2	–	–
	c.7976+57G>C	–	.0004	No effect	No effect	–	Class-2	–	–
Exon 18	c.8010_8032del (p.Ala2671fs)	rs80359690	.0004	No effect	No effect	–	Class-5	–	–
	c.8149G>T (p.Ala2717Ser)	rs28897747	.0004	(c.8168; 55.22%)	No effect	–	Class-2	–	–
Exon 20	c.8503T>C (p.Ser2835Pro)	rs11571746	.0024	No effect	No effect	–	Class-2	–	–
	c.8524C>T (p.Arg2842Cys) <sup>l</sup>	rs80359104	.0004	(c.8535; 88.19%)	No effect	–	Class-2	–	–
	c.8592C>T (p.Ala2864=)	–	.0004	(c.8591; 79.82%)	No effect	–	Class-2	–	–
Exon 22	c.8850G>T (p.Lys2950Asn)	rs28897754	.0008	No effect	No effect	–	Class-2	–	–
	c.8851G>A (p.Ala2951Thr)	rs11571769	.0004	No effect	No effect	–	Class-2	–	–
Intron 22	c.8954-5A>G	–	.0008	–0.08%	–43.77%	Nonclassifiable (Class-4) <sup>m,j</sup>	Nonclassifiable (Class-4) <sup>j</sup>	Class-5	–
				(c.8954-4; 103.97%)	(c.8954-4; 98.45%)				
Exon 23	c.9004G>A (p.Glu3002Lys) <sup>l,n</sup>	rs80359152	.0004	(c.9008; 87.90%)	(c.9008; 38.16%)	–	Class-2	–	–
	c.9026_9030del (p.Thr3009fs)	rs80359741	.0020	No effect	No effect	–	Class-5	–	–
	c.9038C>T (p.Thr3013Ile)	rs28897755	.0008	No effect	No effect	Class-2 <sup>g</sup>	Class-2	–	–
Exon 27	c.9976A>T (p.Lys3326Ter)	rs11571833	.009	No effect	No effect	–	Class-5	–	–

<sup>a</sup>De novo splice site predictions are shown in brackets.

<sup>b</sup>Both 2008 [Spurdle et al., 2008] and 2013 [Walker et al., 2013] guidelines have been considered. If discrepant, 2013 Class is shown in brackets. Published data were classified by authors of the present manuscript.

<sup>c</sup>Description according to HGVS guidelines (<http://www.hgvs.org/mutnomen>). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the GenBank reference sequence NM\_000059.3.

<sup>d</sup>MAF in index cases from HBOC families (N = 2558 chromosomes).

<sup>e</sup>[Bonnet et al., 2008].

<sup>f</sup>Novel variants submitted to the Breast Informative Core (BIC) database (<http://research.nhgri.nih.gov/bic/>).

<sup>g</sup>[Sanz et al., 2010].

<sup>h</sup>Segregation analyses supported that this is indeed a clinical Class-5 variant [Thomassen et al., 2011].

<sup>i</sup>[Gutierrez-Enriquez et al., 2009; Thomassen et al., 2011].

<sup>j</sup>UV allele contribution to wild-type expression was not tested.

<sup>k</sup>Poor quality semi-quantitative data do not allow classification.

<sup>l</sup>Inactivating in protein functional assay [Guidugli et al., 2013].

<sup>m</sup>[Menendez et al., 2012].

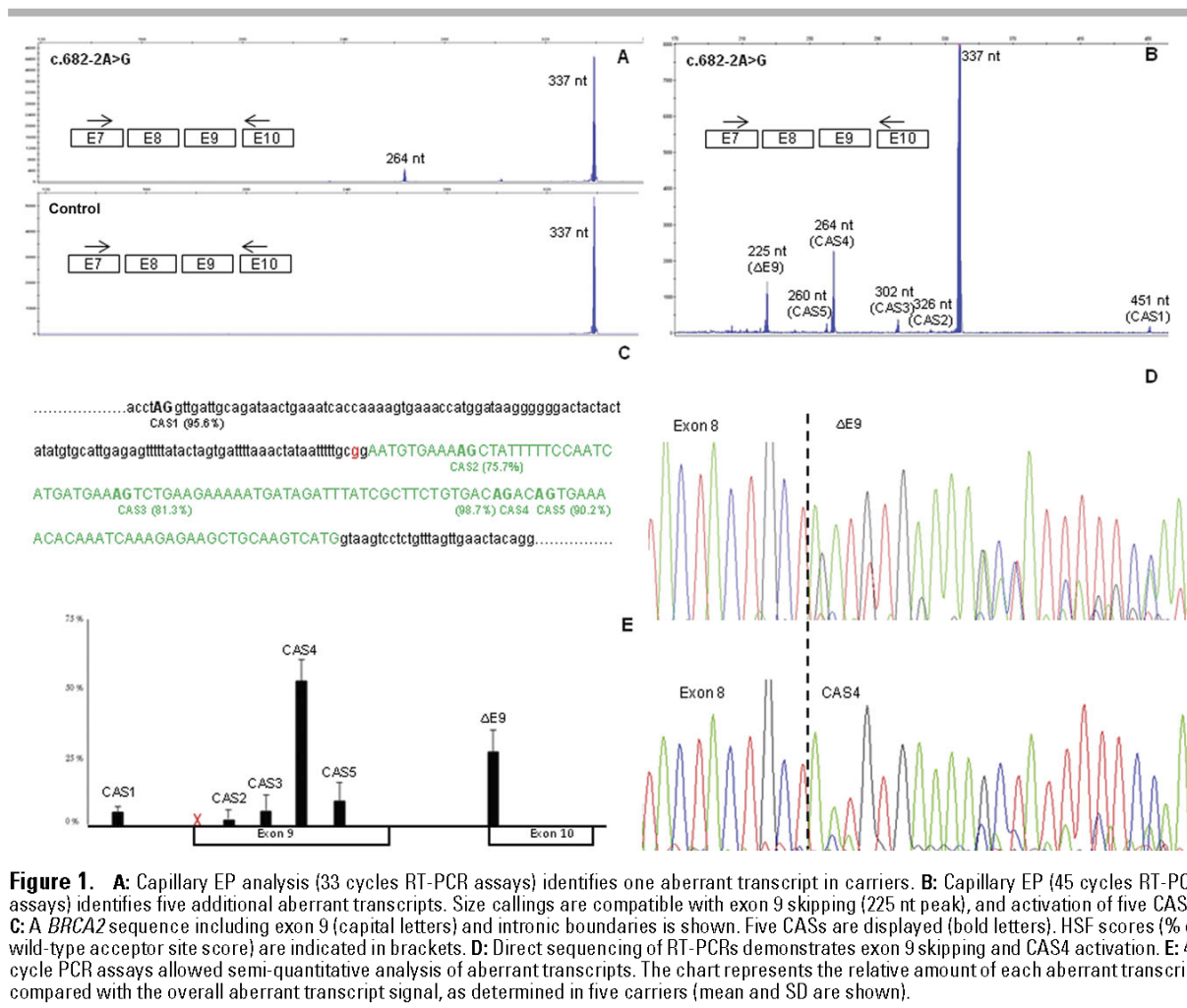
<sup>n</sup>Inactivating in protein functional assays [Biswas et al., 2012].

As a whole, our data qualify for inclusion into analytical Class-5. We have identified four Spanish families carrying this mutation (Supp Fig. S6). Haplotype analysis identified a 1.28-Mb shared haplotype in all carriers, suggesting a founder effect (Supp. Table S4).

Similar assays performed in 23 additional BRCA2 variants did not detect aberrant transcripts, despite the fact that we used high-sensitivity capillary EP analysis. BRCA2 biallelic expression was confirmed in all cases (Supp. Figs. S7–S18). Taking together, all 23 studies qualify for inclusion into analytical Class-2 (see Supp. Table S1). Yet, it is important to remark that analytical Class-2 per se does not have a definite role in the clinical classification of BRCA2 variants.

Not surprisingly, splicing analysis tends to be overlooked in BRCA2 PTC mutation carriers. However, these (as other mutations) might induce splicing alterations (perhaps even regenerating

a full-length open reading frame). This phenomenon is probably unlikely, but certainly real. For instance, it has been reported that c.145G>T (p.Glu49Ter), a PTC mutation located in BRCA2 exon 3, enhances exon 3 skipping (an in-frame splicing isoform), illustrating a complex interplay between RNA and protein effects of potential clinical interest [Sanz et al., 2010]. We have included in our analysis five BRCA2 variants predicted to express PTC transcripts, including four pathogenic (clinical Class-5) mutations (c.1773-1776del, c.3264dupT, c.8010\_8032del22, and c.9026\_9030del) and one benign (clinical Class-1) change (c.9976A>T). However, capillary EP analysis of RT-PCR assays (agarose in the case of c.3264dupT) did not detect splicing alterations. With the single exception of c.9026\_9030del, PTC transcripts were detected at similar levels as wild-type transcripts, suggesting (albeit not proving) that they are



not targeted by the NMD pathway (Supp. Figs. S19 and S20). Since wild-type expression arising from PTC alleles can be discarded *per se*, all five splicing results qualify for analytical Class-5 according to existing guidelines (Supp. Table S1), regardless of the fact that none of these variants affects splicing.

We have identified five common (MAF>0.1) *BRCA2* variants in our cohort (Table 1). *In silico* analyses did not predict splicing alterations, with the single exception that (according to MaxEnt) c.7806-14T>C (rs9534262) produces a slight reduction of the exon 17 acceptor site score (Table 1). Interestingly, two “naturally occurring” alternative splicing events ( $\Delta 17,18$  and  $\Delta 18$ ) have been described in this region [Walker et al., 2010]. Capillary EP did not detect novel transcripts in c.7806-14T>C carriers, but suggested that *BRCA2*  $\Delta 17,18$  was slightly upregulated (data not shown). Hence, we decided to explore the role of c.7806-14T>C in alternative splicing regulation by qPCR (see Supp. Methods). Quantification experiments were performed in 7 T/T, 20 T/C, and 11 C/C carriers. The expression level of *BRCA2*  $\Delta 17,18$  (relative to the overall *BRCA2* expression level) was low ( $\Delta C_T = 5.19 \pm 0.59$ ), reflecting the fact that *BRCA2*  $\Delta 17,18$  is a minor alternative splicing isoform. Yet, the ANOVA test identified differences associated with each genotype ( $P = 0.006$ ). Post-hoc test demonstrated higher levels of *BRCA2*  $\Delta 17,18$  in C/C carriers compared with T/T carriers ( $P = 0.004$ )

(Supp. Fig. S21). To confirm that the C allele is indeed functionally related, and not just in linkage disequilibrium with a true splicing regulatory element, we performed further analysis in a reporter minigene (exons 16–20 of *BRCA2*) based on the new splicing vector pSAD<sup>®</sup> (Consejo Superior de Investigaciones Científicas). In this model, the c.7806-14C reporter minigene showed relative increased levels of both  $\Delta 17,18$  and  $\Delta 18$  isoforms if compared with c.7806-14T reporter minigene (Supp. Fig. S22). Overall, the data indicate that rs9534262 is a bona fide *BRCA2* splicing quantitative trait locus (sQTL). Since *BRCA2*  $\Delta 17,18$  and *BRCA2*  $\Delta 18$  are PTC splicing isoforms, we hypothesized that rs9534262 might be associated with breast cancer risk. To test it, we performed case-control association studies and Kaplan–Meier survival analyses. Overall, these analyses do not support an association with risk (Supp. Fig. S23).

In summary, we have performed routine *in vitro* splicing assays (blood RNA-based RT-PCR experiments) in 31 *BRCA2* genetic variants, producing 29 clear-cut analytical outputs. Of these, two Class-5 (c.682-2A>G and c.7617+1G>A) are clinically relevant findings *per se*. By contrast, neither c.8954-5A>G nor c.7806-14T>C (rs9534262) routine *in vitro* splicing assays resulted in clear-cut analytical classification. The former is of particular clinical relevance since subsequent dedicated assays (a reporter Minigene) demonstrated that *BRCA2* c.8954-5A>G is indeed a clinically relevant Class-5 variant.

Analysis of c.7806-14T>C (rs9534262) with dedicated assays (reported Minigene and qPCR) produced results that qualify for analytical Class 4 (2008 guidelines), or analytical Class-3 (2013 guidelines) (see Supp. Table S1). However, BRCA2 rs9534262 is a clinical Class-1 variant (MAF is close to 0.5 in many populations). Similarly, splicing analysis in c.9976A>T (p.Lys3326Ter) carriers produced an analytical Class-5 output, despite the fact that this is a clinical Class-1 variant based on posterior probability analysis (<http://brca.iarc.fr>). These two examples illustrate the fact that the four-tier classification scheme of splicing assays outputs is more related to the “analytical quality” and “interpretation” of UV splicing assays than to the actual probability of being pathogenic.

Undoubtedly, sequencing informative exonic SNPs (testing biallelic expression) is critical in producing clear-cut routine *in vitro* splicing assays. In addition, we show that incorporating capillary EP into the analysis provides highly-informative analytical data (both high-resolution size-calling and high-sensitivity detecting RT-PCR products). These two features were critical in order to characterize the rather complex splicing outcome observed in BRCA2 c.682-2A>G carriers. This variant induces the expression of several (probably six) different aberrant transcripts. As a result, the “aberrant signal” is split into different components (different RT-PCR products), none of which is detectable by conventional EtBr-stained agarose gel analysis (data not shown). Equally relevant, capillary EP ensures that Class-2 outputs are indeed clear-cut negative results. Capillary EP provides not only high-resolution size-calling and high-sensitivity, but also semi-quantitative information (area under the peaks) that might be useful for analytical purposes. For instance, it allowed us to identify c.7806-14T>C (rs9534262) as a candidate sQTL for further dedicated assays.

We have performed several routine *in vitro* splicing assays in carriers of UV variants with low prior probability (according to *in silico* predictions) of finding an alteration, a conservative approach that we nonetheless recommend in the clinical setting. Undoubtedly, *in silico* analyses might provide high-accuracy predictions, but they do not detect all splicing defects, especially in variants located outside the 5' and 3' consensus sites (Houdayer et al., 2012). Yet, we think that *in silico* analyses have important roles in the clinic. For instance, they assist with prioritizing variants for dedicated splicing assays (such as reporter minigenes) in cases where routine assays do not produce clear-cut outputs, or in cases in which routine assays are not feasible (i.e., lack of fresh blood RNA).

In conclusion, we show that incorporating capillary EP and direct sequencing of exonic-informative SNPs to, otherwise, conventional fresh blood-based RT-PCR splicing assays often provides high-quality results that allow systematic analytical classification according to existing guidelines.

## Acknowledgments

Z.G.C. is a member of the Genetic Counseling Program in the Comunidad Valenciana.

## References

Acedo A, Sanz DJ, Duran M, Infante M, Perez-Cabornero L, Miner C, Velasco EA. 2012. Comprehensive splicing functional analysis of DNA variants of the BRCA2 gene by hybrid minigenes. *Breast Cancer Res* 14:R87.

Biswas K, Das R, Eggington JM, Qiao H, North SL, Stauffer S, Burkett SS, Martin BK, Southon E, Sizemore SC, Pruss D, Bowles KR, et al. 2012. Functional evaluation of BRCA2 variants mapping to the PALB2-binding and C-terminal DNA-binding domains using a mouse ES cell-based assay. *Hum Mol Genet* 21:3993–4006.

Bonnet C, Krieger S, Vezain M, Rousselin A, Tournier I, Martins A, Berthet P, Chevrier A, Dugast C, Layet V, Rossi A, Lidereau R, et al. 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* 45:438–446.

de Garibay GR, Diaz A, Gavina B, Romero A, Garre P, Vega A, Blanco A, Tosar A, Diez O, Perez-Segura P, Diaz-Rubio E, Caldes T, et al. 2013. Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families. *Eur J Hum Genet* 21:883–886.

de Juan Jiménez I, García Casado Z, Palanca Suela S, Esteban Cardenaosa E, López Guerrero JA, Segura Huerta A, Chirivella González I, Sánchez Heras AB, Juan Fita MJ, Tena García I, Guillen Ponce C, et al. 2013. Novel and recurrent BRCA1/BRCA2 mutations in early onset and familial breast and ovarian cancer detected in the Program of Genetic Counseling in Cancer of Valencian Community (eastern Spain). Relationship of family phenotypes with mutation prevalence. *Fam Cancer*. [still ahead of print]

Dosil V, Tosar A, Canadas C, Perez-Segura P, Diaz-Rubio E, Caldes T, de la Hoya M. 2010. Alternative splicing and molecular characterization of splice site variants: BRCA1 c.591C>T as a case study. *Clin Chem* 56:53–61.

Guidugli L, Pankratz VS, Singh N, Thompson J, Erding CA, Engel C, Schmutzler R, Domchek S, Nathanson K, Radice P, Singer C, Tonin PN, et al. 2013. A classification model for BRCA2 DNA binding domain missense variants based on homology-directed repair activity. *Cancer Res* 73:265–275.

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ; Breast Cancer Information Core (BIC) Steering Committee. 2004. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* 75:535–544.

Gutierrez-Enriquez S, Coderch V, Masas M, Balmama J, Diez O. 2009. The variants BRCA1 IVS6-1G>A and BRCA2 IVS15+1G>A lead to aberrant splicing of the transcripts. *Breast Cancer Res Treat* 117:461–465.

Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Léone M, et al. 2012. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined *in silico/in vitro* studies on BRCA1 and BRCA2 variants. *Hum Mutat* 33:1228–1238.

Lindor NM, Guidugli L, Wang X, Vallee MP, Monteiro AN, Tavtigian S, Goldgar DE, Couch FJ. 2012. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Hum Mutat* 33:8–21.

Menendez M, Castellsague J, Mirete M, Pros E, Feliubadalo L, Osorio A, Calaf M, Tornero E, del Valle J, Fernandez-Rodriguez J, Qules F, Salinas M, et al. 2012. Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes. *Breast Cancer Res Treat* 132:979–992.

Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV; IARC Unclassified Genetic Variants Working Group. 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 29:1282–1291.

Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornero L, Esteban-Cardenosa E, Lastra E, Pagani F, Miner C, Velasco EA. 2010. A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clin Cancer Res* 16:1957–1967.

Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM; IARC Unclassified Genetic Variants Working Group. 2008. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat* 29:1304–1313.

Spurdle AB, Healey S, Devereau A, Hogervorst FB, Monteiro AN, Nathanson KL, Radice P, Stoppa-Lyonnet D, Tavtigian S, Wappenschmidt B, Couch FJ, Goldgar DE, et al. 2012. ENIGMA—evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat* 33:2–7.

Thomassen M, Pedersen IS, Vogel I, Hansen TV, Brasch-Andersen C, Brasen CL, Cruger D, Sunde L, Nielsen FC, Jensen UB, Bisgaard ML, Borg A, et al. 2011. A BRCA2 mutation incorrectly mapped in the original BRCA2 reference sequence, is a common West Danish founder mutation disrupting mRNA splicing. *Breast Cancer Res Treat* 128:179–185.

Walker LC, Whiley PJ, Couch FJ, Farrugia DJ, Healey S, Eccles DM, Lin F, Butler SA, Goff SA, Thompson BA, Lakhani SR, Da Silva LM, et al. 2010. Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity. *Hum Mutat* 31:E1484–E1505.

Walker LC, Whiley PJ, Houdayer C, Hansen TV, Vega A, Santamarina M, Blanco A, Fachal L, Southey MC, Lafferty A, Colombo M, De Vecchi G, et al. 2013. Evaluation of a 5-tier scheme proposed for classification of sequence variants using bioinformatic and splicing assay data: inter-reviewer variability and promotion of minimum reporting guidelines. *Hum Mutat* 34:1424–1431.

## **On-line methods**

### Patients and Samples.

All variants under study were first identified through clinical testing in index cases of high risk hereditary breast and ovarian cancer families undergoing genetic counseling according to Spanish National Health System recommendations (N=1279). According to the frequency observed in the index case cohort, we refer throughout the text to rare ( $f < 0.1$ ) and common ( $f > 0.1$ ) *BRCA2* variants. We have performed *in vitro* RNA analysis in index cases (or relatives) carrying the genetic variants of interest. In addition, rs9534262 (c.7806-14T>C) has been genotyped in 376 healthy control individuals, a consecutive series of 173 women diagnosed of breast cancer, and 438 index cases from BRCAX families (high risk families negatives for *BRCA1* and *BRCA2* germ-line mutations after comprehensive genetic testing). The BRCAX cohort has been extensively described elsewhere [de Garibay et al., 2013; Dosil et al., 2010]. All participants signed an informed consent. The study was approved by the Ethics Committee of the participating institutions.

### BRCA2 sequence variants description.

All *BRCA2* genetic variants are described according to HGVS guide-lines (<http://www.hgvs.org/mutnomen>). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the GeneBank reference sequence NM\_000059.3

### In silico splice predictions

Consensus splice sites predictions were performed with HSF and MaxEnt algorithms. Splicing enhancer predictions were performed with ESE finder, RESCUE-ESE, PESE and EIE algorithms. Similarly, splicing silencer predictions were analyzed with ESS, Fas-ESS, PESS and IIE algorithms. All analyses were performed on-line with Human Splicing Finder version 2.4.1 (default settings).

### Blood-based RT-PCR In Vitro Splicing Assays (Routine Assay).

RNAs were extracted from fresh blood using a MagnaPure Compact workstation and Magna Pure RNA kits according to the manufacturer's instructions (Roche Diagnostics). RT was performed with 300ng of RNA and the PrimeScript RT kit (Takara Biotechnology) following manufacturer's protocol. Depending on the interrogated variant, cDNAs were amplified with different combinations of forward and reverse primers (Supp. Table 1). Amplicons were designed taking into consideration both *in silico* predictions and "naturally occurring" alternative splicing events. The later might be relevant for the correct analysis of *BRCA2* splicing variants, as has been previously shown for *BRCA1*

splicing variants [Dosil et al., 2010]. Fluorescent PCRs were performed with FastStart Taq (Roche Diagnostics) in a final volume of 25  $\mu$ L that contained 1  $\mu$ L of cDNA. Samples were denatured at 95 $^{\circ}$  for 10 min, followed by 45 cycles consisting of 95 $^{\circ}$  for 30 sec, 58  $^{\circ}$  for 30 sec, and 72 $^{\circ}$  for 30 sec, and final extension step at 72 $^{\circ}$  for 7 min. For semi-quantitative analysis, the number of PCR cycles was reduced to 33. RT-PCR products were analyzed by capillary electrophoresis on an ABI3130 automated sequencer (Life Technology). Depending on the expected sizes, LIZ-500 or LIZ-1200 (Life Technology) was used as internal size-standard. Size-calling and peak areas (semi-quantitative assays) were analyzed with the Genemapper software v.3.7 (Life Technology). Bi-allelic expression was analyzed by direct sequencing of RT-PCR reactions performed at informative (heterozygous) loci: either the interrogated variant itself (exonic variants) or linked exonic SNPs (in the case of intronic variants). Cycle sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing kit (Life Technology).

#### Minigene assays

Construction of Minigenes:

The construction of the minigene containing *BRCA2* exons 23 and 24 and flanking intronic sequences (MGBR2\_Ex23-24) inserted into the splicing vector pSPL3b was previously described [Acedo et al, 2012]. The minigene containing exons 16 to 20 of the *BRCA2* (MGBR2\_Ex16-20) gene was constructed in three sequential cloning steps in a new splicing vector derived from pSPL3 called pSAD (Patent **P201231427** CSIC, *Priority Patent Application filed*; Acedo et al, manuscript in preparation): exons 17+18 in a single fragment (1,556 bp; intron 17= 485 bp), exon 16 (805 bp) and exons 19+20 in a single fragment (1165 bp; intron 19= 398 bp). These sequences were amplified with the Phusion Hot Start II High Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's instructions, and using the following primers:

for exons 17+18:

mgbr2\_ex17fw, 5'GCTCTAGAACTAGTGGATCCCCCGGTCAGTATGATACTTTGATACATGT3'

mgbr2\_ex18rev, 5'ATAAGCTTGATATCGAATTCCTGCAGTGGAATAGGGATCTGATCAA3',

for exon 16:

mgbr2\_ex16fw, 5'GCTCTAGAACTAGTGGATCCCCCGGCCCTTTTGTTCATCTAAGT3'

mgbr2\_ex16rev, 5'GACATGTATCAAAGTATCATACTGAATAAATGCCTAAGAAAAATGT3'

For exons 19+20:

mgbr2\_ex19fw, 5'CACACACTCGAGATAGCATTAAAGAACTTGTAGCA3'

mgbr2\_ex20rev, 5'CACACAGGTACCATTACAAATGGCTTAGACCTGA3'

For cloning purposes a 5' tail (underlined) was added to each primer. The final structure of the insert of the minigene from the 5' to the 3' end was: ivs15 (333 bp) - exon 16 (188 bp)- ivs16 (234

bp)- // - ivs16 (181 bp) - exon 17 (171 bp) – ivs17 (485 bp) – exon 18 (355 bp) – ivs18 (314 bp)- // - ivs18 (235 bp) – exon 19 (156 bp) – ivs19 (398 bp) – exon 20 (145 bp)– ivs20 (207 bp). Introns 16 (4,595 bp) and 18 (6,853 bp) were too large so they were shortened to facilitate cloning. All the new constructs were sequenced with the Big Dye 3.1 Sequencing Kit (Life Technologies) to confirm the fidelity of cloned sequences.

Site directed mutagenesis:

Mutagenesis was carried out according to the PCR mutagenesis protocol (<http://www.methodbook.net/pcr/pcrmut.html>) with Pfu Turbo DNA polymerase (Agilent, Santa Clara, CA, USA). Wild type (wt) minigenes MGBR2\_Ex23-24 and MGBR2\_Ex16-20 were used as templates to generate DNA variants.

For c.7806-14T>C of intron 16 primers used were

5' TTTATGATAATATTCTACCTTTATTTGTTTCAGGGCTC 3'

5' GAGCCCTGAACAAATAAAGGTAGAATATTATCATAAA 3'

For c.8954-5A>G of intron 22 primers used were

5' TCTTTCTCATCTTTCTCCAGACAGTTATACTGAGTATTTG 3'

5' CAAATACTCAGTATAACTGTCTGGAGAAAGATGAGAAAGA 3'

Transfection of HeLa cells:

Approximately  $10^5$  HeLa cells (human cervical carcinoma) were grown to 90% to 95% confluency in 0.5 mL of growth medium (DMEM, 10% fetal bovine serum, 1% glucose and 1% penicillin/streptomycin) in four-well plates (Nunc, Roskilde, Denmark). Cells were transfected with 1 µg of plasmid minigene and 2 µL of Lipofectamine 2000 (Life Technologies) following the manufacturer's protocol. RNA was purified 24-48 hours after transfection with Nucleospin-RNA-II (Macherey-Nagel, Düren, Germany) that includes on-column rDNAse treatment and quantified in a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific)

RT-PCR of minigenes:

Retrotranscription was carried out with 200 ng of RNA and the transcriptor first strand cDNA synthesis kit (Roche Diagnostics). Fluorescent RT-PCRs were performed in triplicate in a final volume of 20 µL that contained 2 µL of cDNA and the flanking primers of constitutive exons of pSPL3, SD6-PSPL3\_RTFW (FAM-5' TCACCTGGACAACCTCAAAG 3') and SA2-PSPL3\_RTREV (5' TGAGGAGTGAATTGGTCGAA 3'). Samples were denatured at 94°C for two minutes, followed by 25-30 cycles consisting of 94°C for 20 seconds, 58°C for 20 seconds, and 72°C for 30-90 seconds, and a final extension step at 72°C for two minutes. The PCR products were analyzed by capillary electrophoresis. Expected sizes of the RT-PCR products were 488 bp and 1199 bp for minigenes MGBR2\_Ex23-24 and MGBR2\_Ex16-20, respectively

#### Relative quantification of BRCA2 $\Delta$ 17,18 expression level.

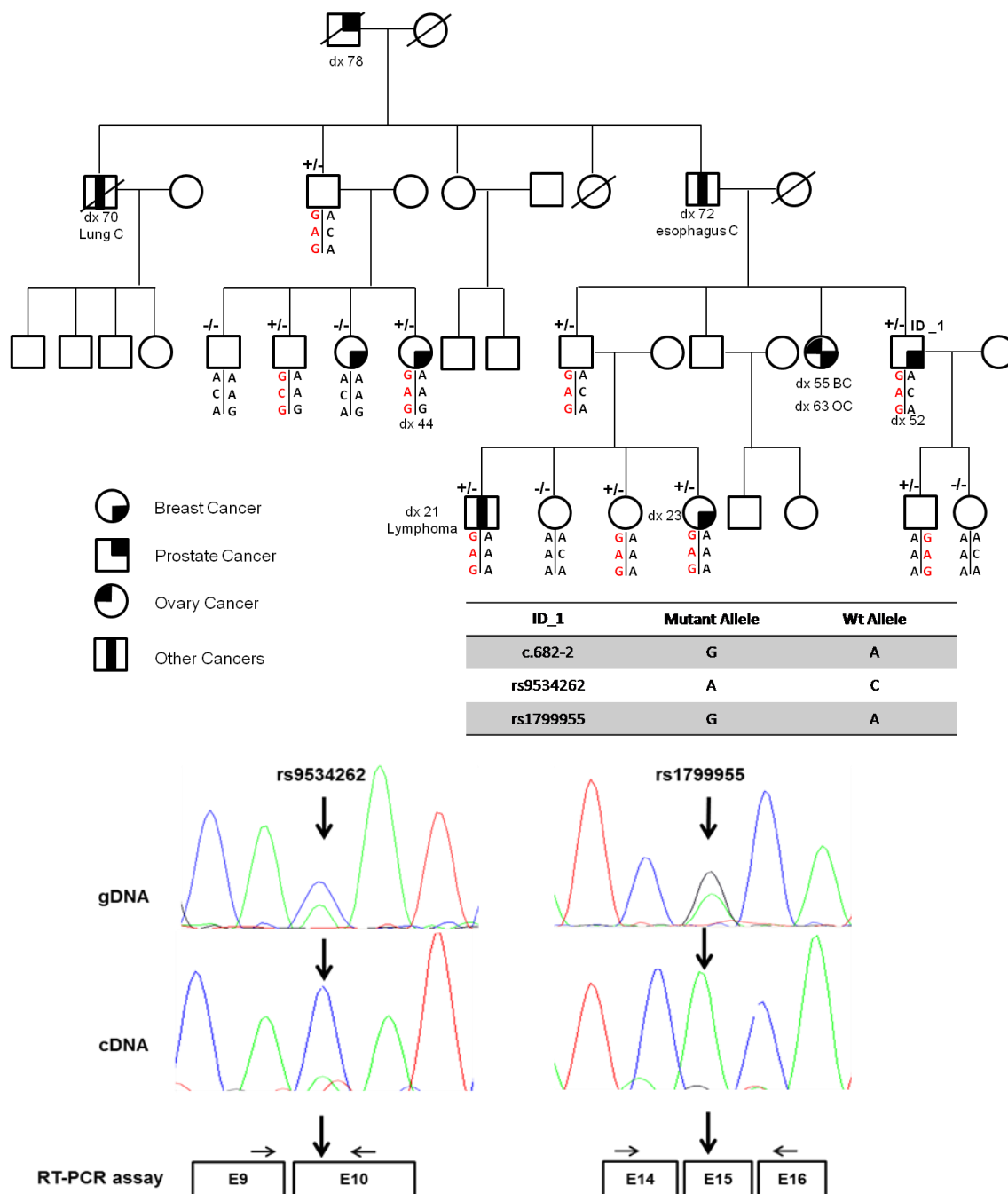
The relative expression level of BRCA2  $\Delta$ 17,18 was measured by real-time PCR with a dedicated FAM-labeled TaqMan gene expression assay. The assay was designed and synthesized using the Assay-by-Design service (Life Technology) that includes proprietary specificity and amplification efficiency QC testing. As endogenous control, we used the VIC-labeled Custom TaqMan gene expression assay Hs01037420\_m (Life Technology) that targets BRCA2 exon 24 (since no BRCA2  $\Delta$ 24 splicing isoform has been reported, BRCA2 exon 24 probably reflects the overall BRCA2 expression level). Both assays were performed in a single tube reaction including 1  $\mu$ L of cDNA, 1X TaqMan Universal PCR MasterMix (Life Technology) and 1X TaqMan assays. Experiments were performed in a 7500 Fast Real-Time PCR system (Life Technology) with default thermal cycling conditions (Life Technology). For quantification purposes, we measured  $\Delta C_T$ , defined as  $C_T$  BRCA2 $\Delta$ 17,18 –  $C_T$  BRCA2e24. Therefore,  $\Delta C_T$  reflects the expression level of BRCA2  $\Delta$ 17,18 relative to the overall expression level from the same locus.  $\Delta C_T$  was analyzed in triplicate in all tested samples.

#### Haplotype analysis of recurrent variants

Haplotype analysis was performed with up to 9 short tandem repeats (STR) covering 14.4Mb of chromosome 13 surrounding the BRCA2 gene. Primer sequences, order of markers and physical distances among D13S260, D13S1699, D13S1698, D13S1697, D13S1701, D13S1715, D13S1493, D13S267, and D13S1694 were obtained from the Ensembl database (<http://www.ensembl.org>). Polymerase chain reaction (PCR) products of STR markers were generated using fluorescently labelled primers. Size-calling was performed by capillary electrophoresis. Additionally, 3 BRCA2 SNPs (rs1799943, rs144848, and rs9534262) were genotyped by HRMA.

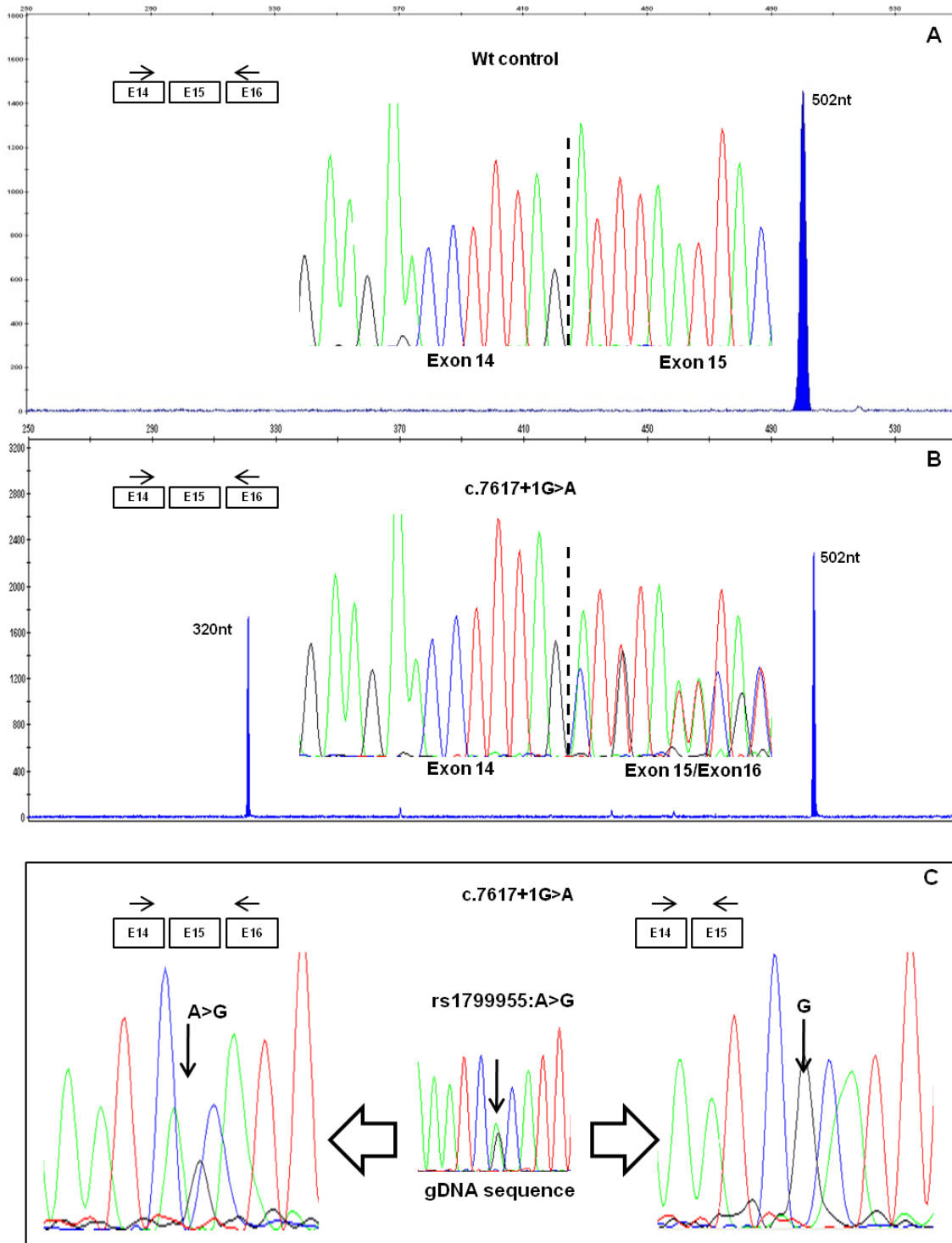
#### Statistical Analyses.

BoxPlots, Kaplan-Meyer survival analysis, chi-square, and ANOVA test were performed with SPSS v13.0 software. Power of the case control studies was calculated using EpiInfo 7. Case control studies were designed to have a power of 80% to detect an O.R. higher than 1.3 with a statistical confidence of 95 %.



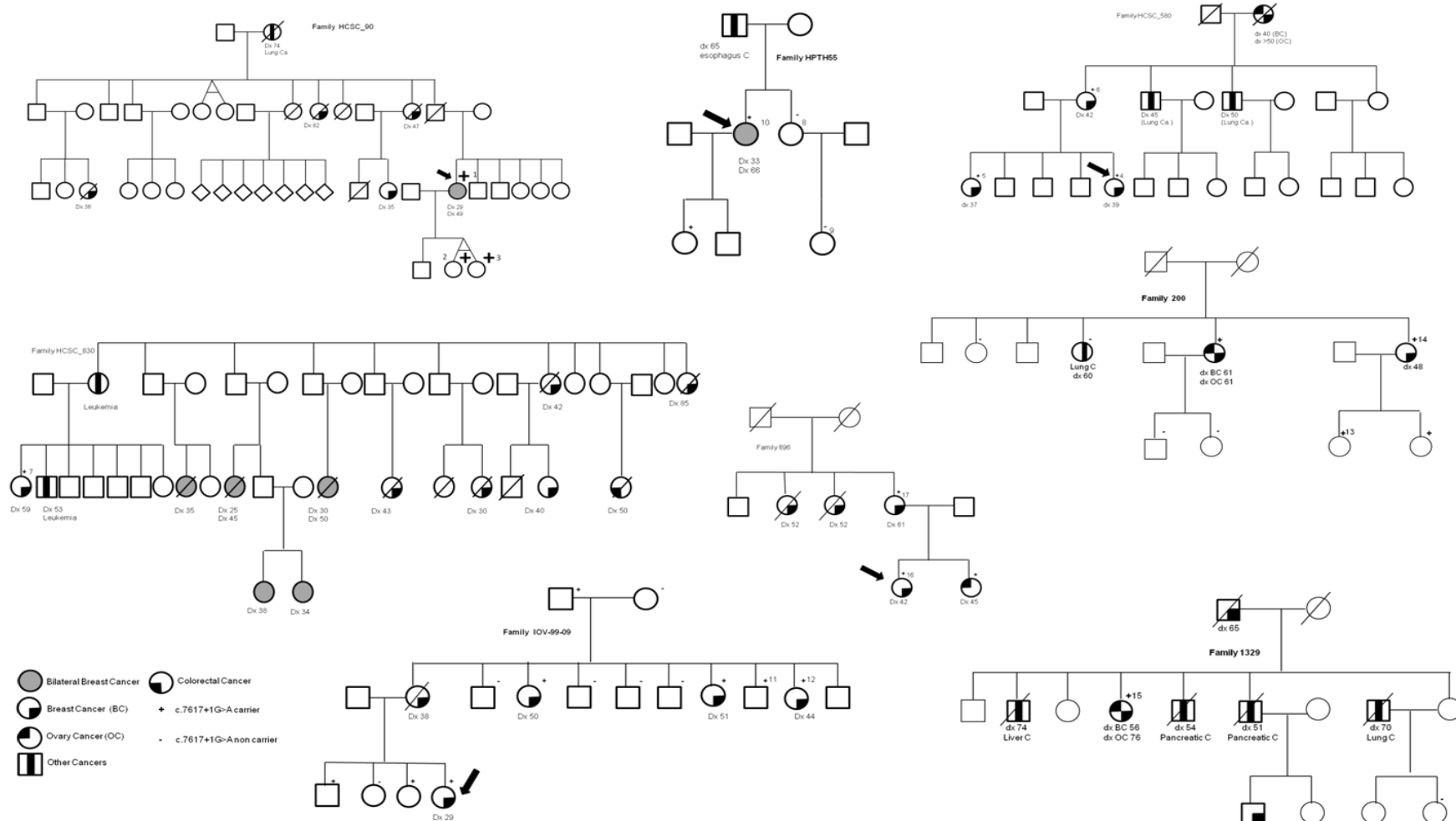
Supp. Figure S 1 A family carrying the BRCA2 variant c.682-2A>G is displayed

Segregation analysis allowed us to determine the phase between the variant allele and two SNPs located in BRCA2 exon 10 (rs9534262) and exon14 (1799955). Direct sequencing of RT-PCR assays (33-cycle PCR) performed with RNA isolated from individual ID\_1 showed that mutant allele contribution to overall BRCA2 expression (if any) is minimal.



Supp. Figure S 2 BRCA2 c.7617+1G>A splicing analysis

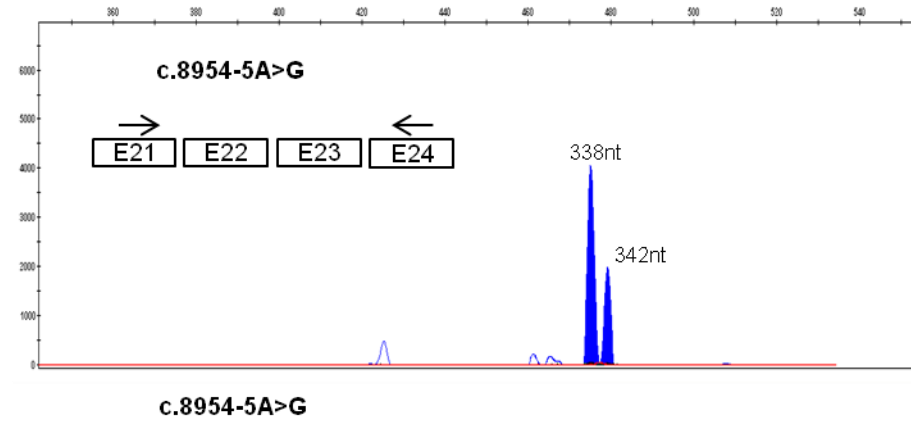
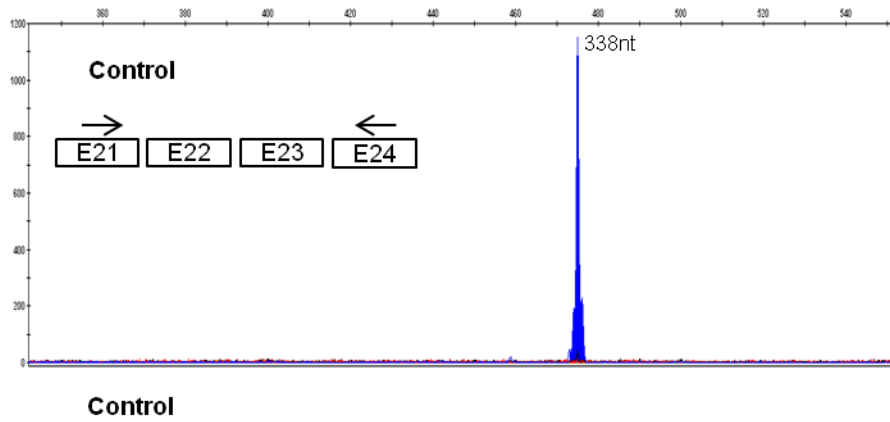
A, B) capillary EP analysis (33 cycles RT-PCR) revealed an aberrant transcript 182nt shorter than wt in c.7617+1G>A carriers, suggesting BRCA2 exon 15 skipping. Direct sequencing confirmed exon 15 skipping. C) An informative locus in exon 14 (rs1799955) allowed us to demonstrate that the variant allele does not contribute to full-length expression. Note that RT-PCR products generated with an exon15 reverse primer (right) are full-length specific. Segregation analysis allowed us to determine the phase between the variant allele and rs1799955 (data not shown).



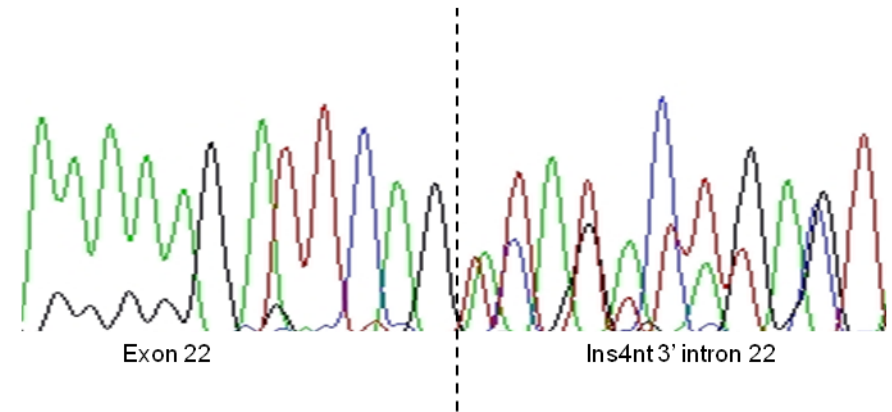
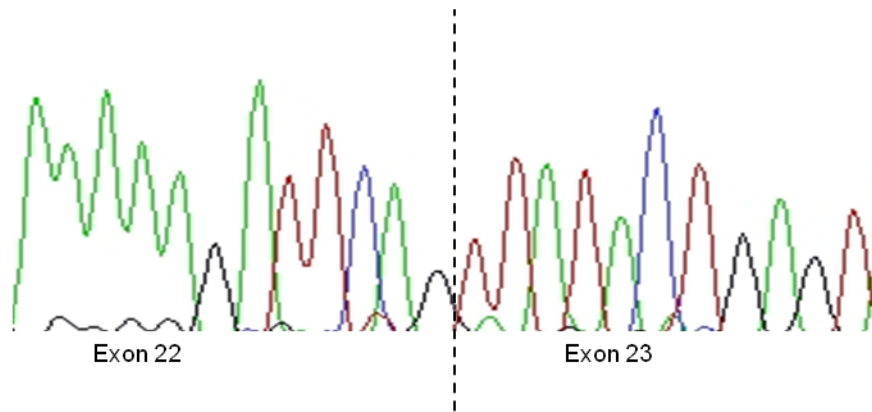
Supp. Figure S 3 Eight Spanish families carrying the BRCA2 genetic variant c.7617+1G>A.

Numbered Individuals were included in the haplotype analysis (See Supp. Table S2).

A

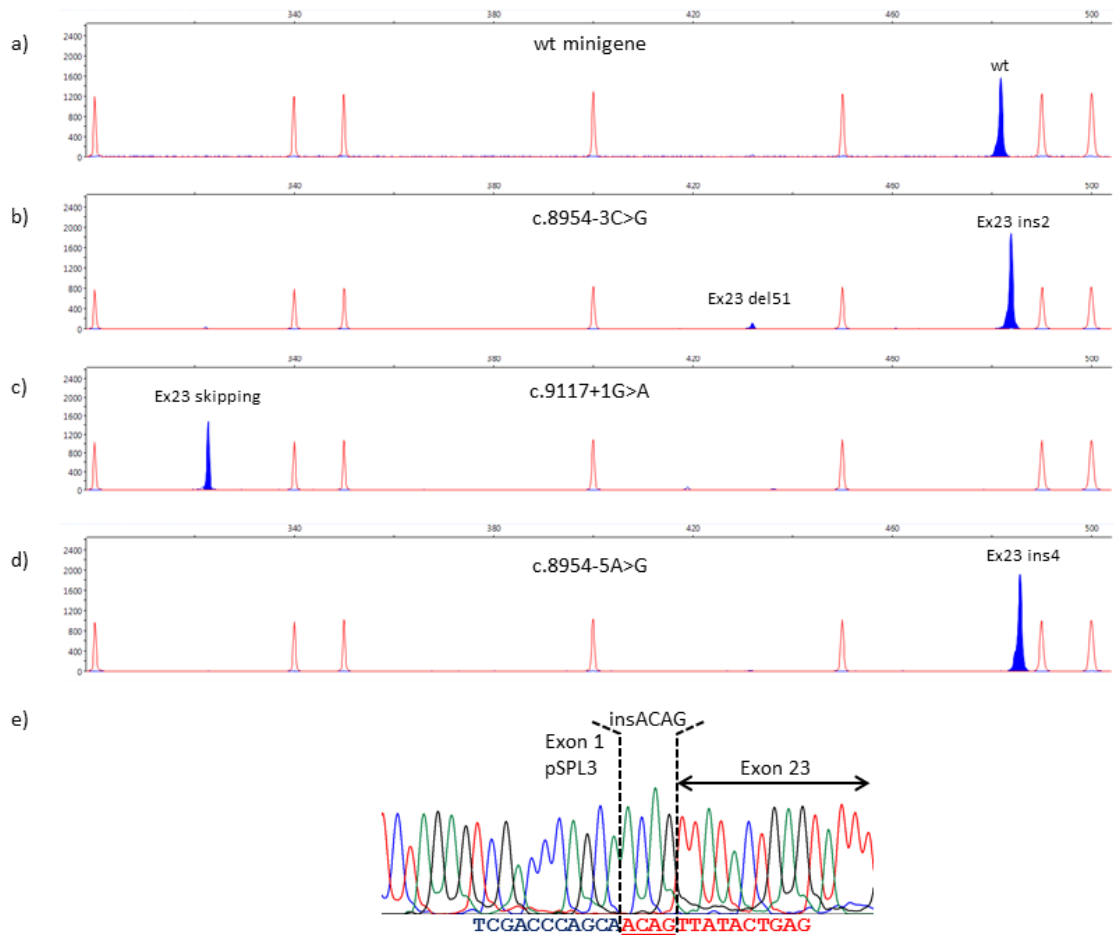


B



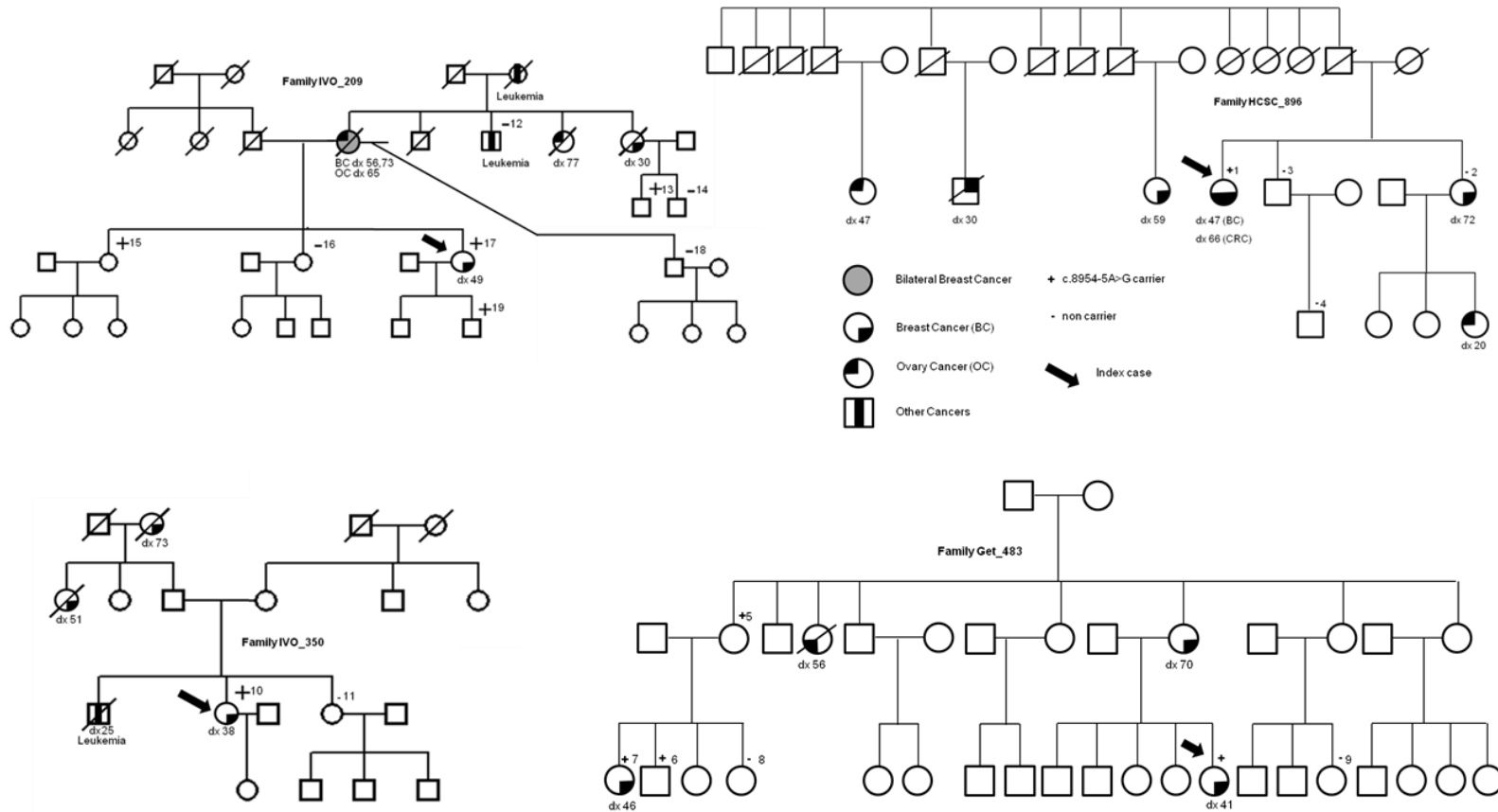
Supp. Figure S 4 Capillary EP analysis of the RT-PCR with primers at exon 21 and 24 in a c.8954-5A>G carrier.

Direct sequencing demonstrates an insertion corresponding to the last 4nt of intron 22.

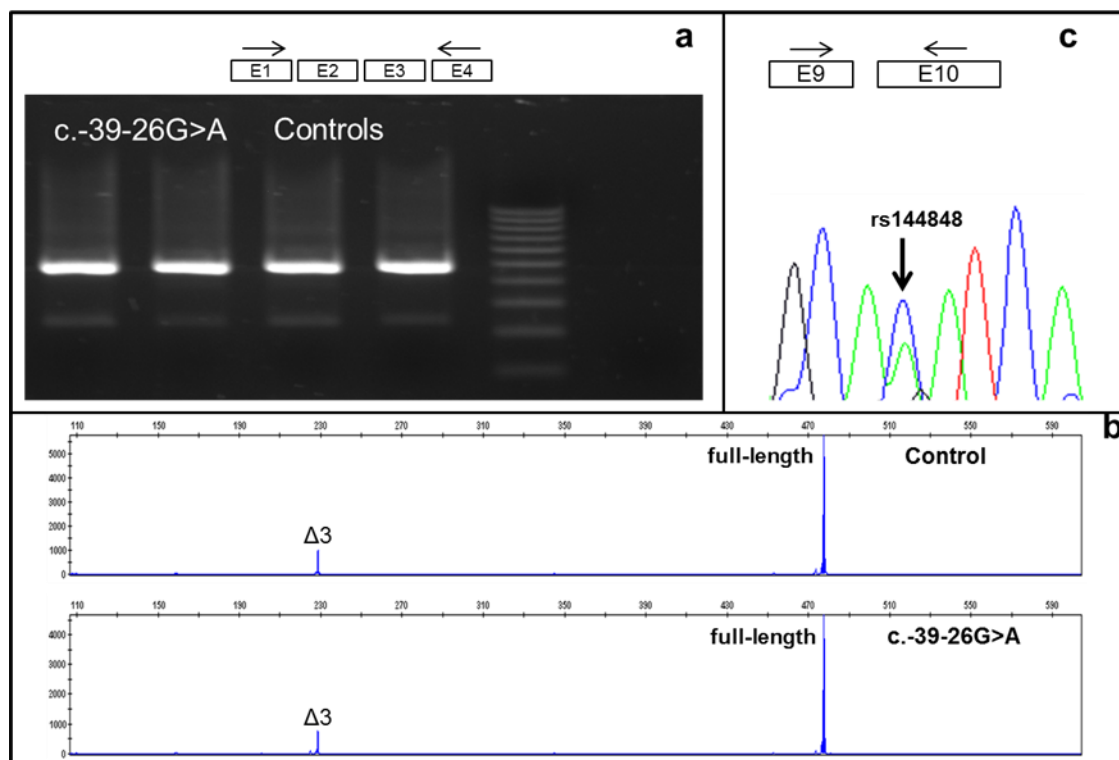


Supp. Figure S 5 Splicing outcomes of variants of exon 23 of BRCA2 in minigene MGBR2\_Ex23-24

RT-PCR products labelled with FAM (blue peaks) of wild type and mutant minigenes were run in an ABI3130 sequencer with Genescan ROX 500 (red peaks) as size standard. The different RNA isoforms are indicated. A) Wild type minigene; B) c.8954-3C>G (intron 22); C) c.9117+1G > A c; D) c.8954-5A>G; E) Sequencing of isoforms produced from c.8954-5A>G.

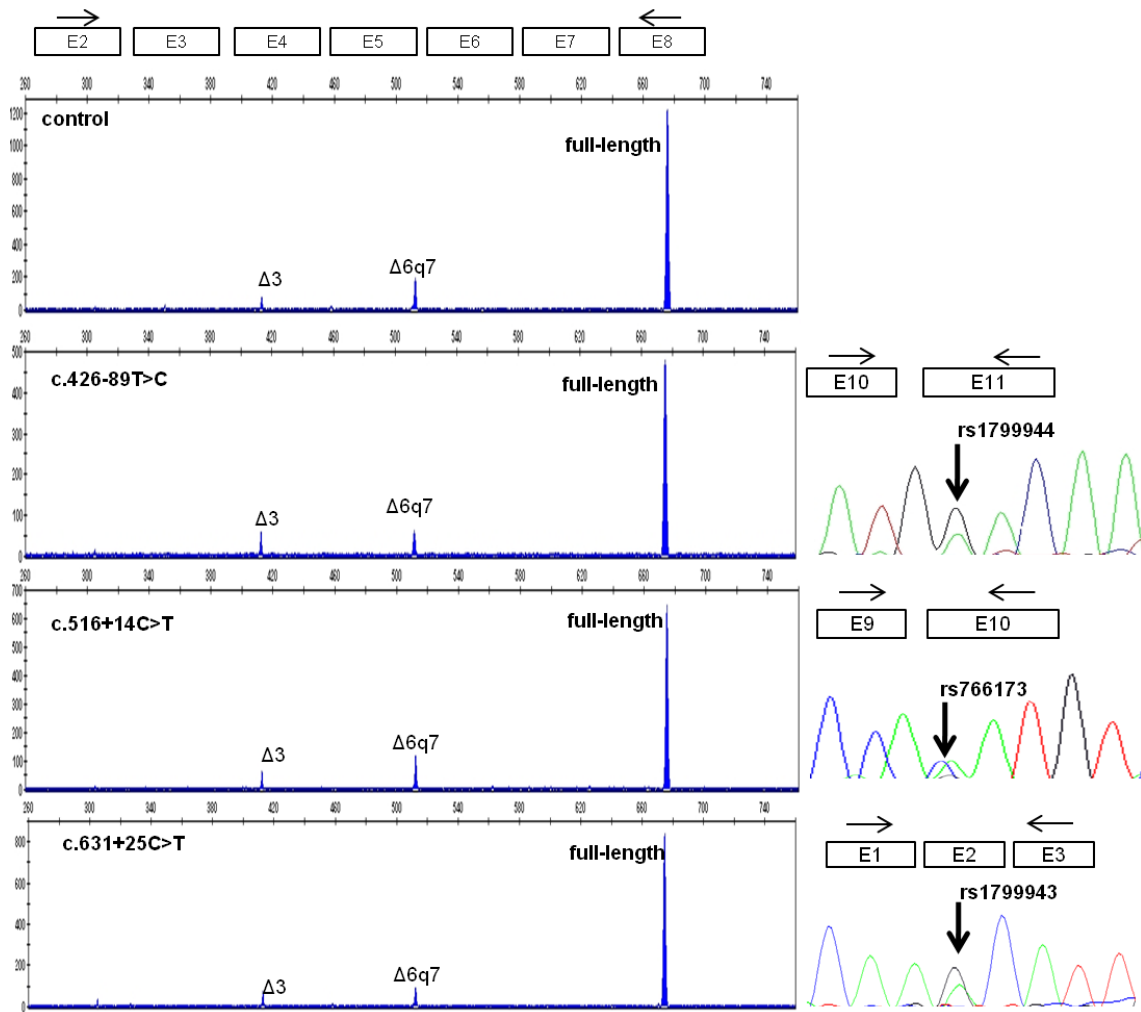


Supp. Figure S 6 Four Spanish families carrying the BRCA2 genetic variant c.8954-5A>G. Numbered Individuals were included in the haplotype analysis (See Supp. Table S3).



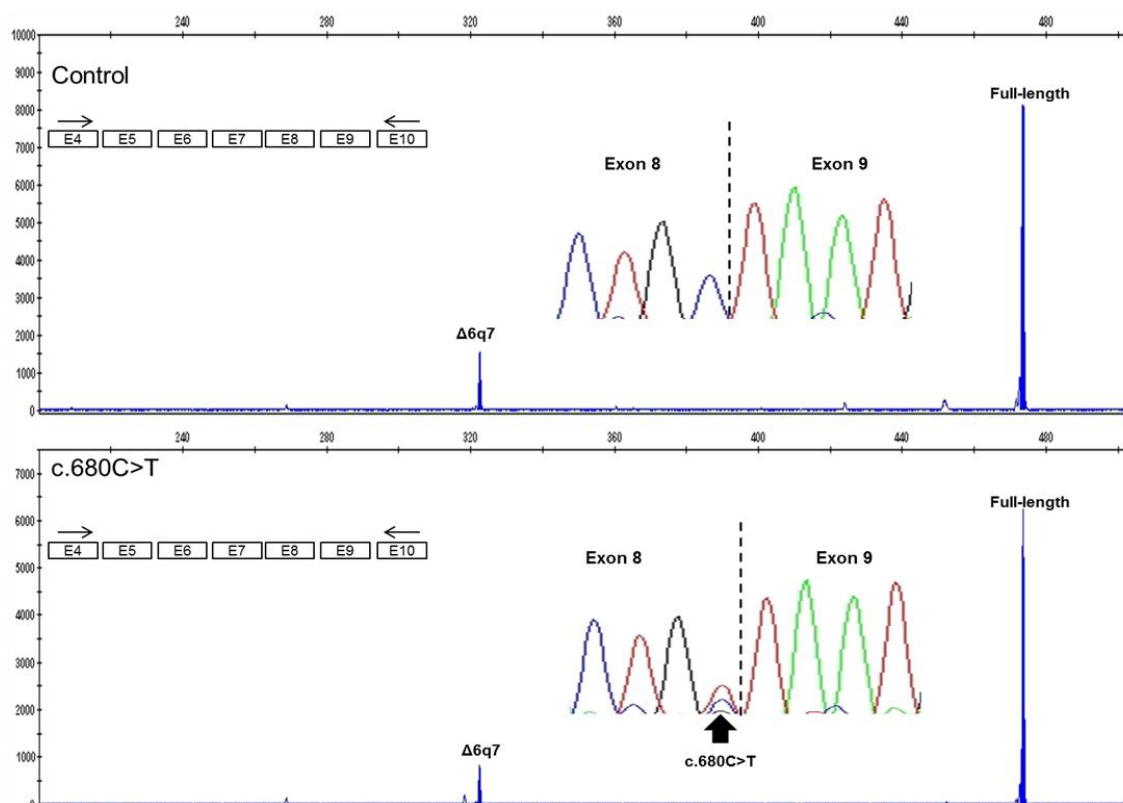
Supp. Figure S 7 Splicing analysis of BRCA2 c.-39-26G>A variant

a) RT-PCR with primers located in exons 1 and 4 shows a similar splicing pattern in controls and c.-39-26G>A carriers, with the presence of full-length and  $\Delta 3$  BRCA2 transcripts. b) Capillary EP analysis of semi-quantitative RT-PCR assays (33 cycle PCR) shows no obvious differences in the amount of  $\Delta 3$  BRCA2 transcripts in c.-39-26G>A carriers. c) Bi-allelic expression was confirmed by direct sequencing of an RT-PCR assay containing an informative locus (rs144848).



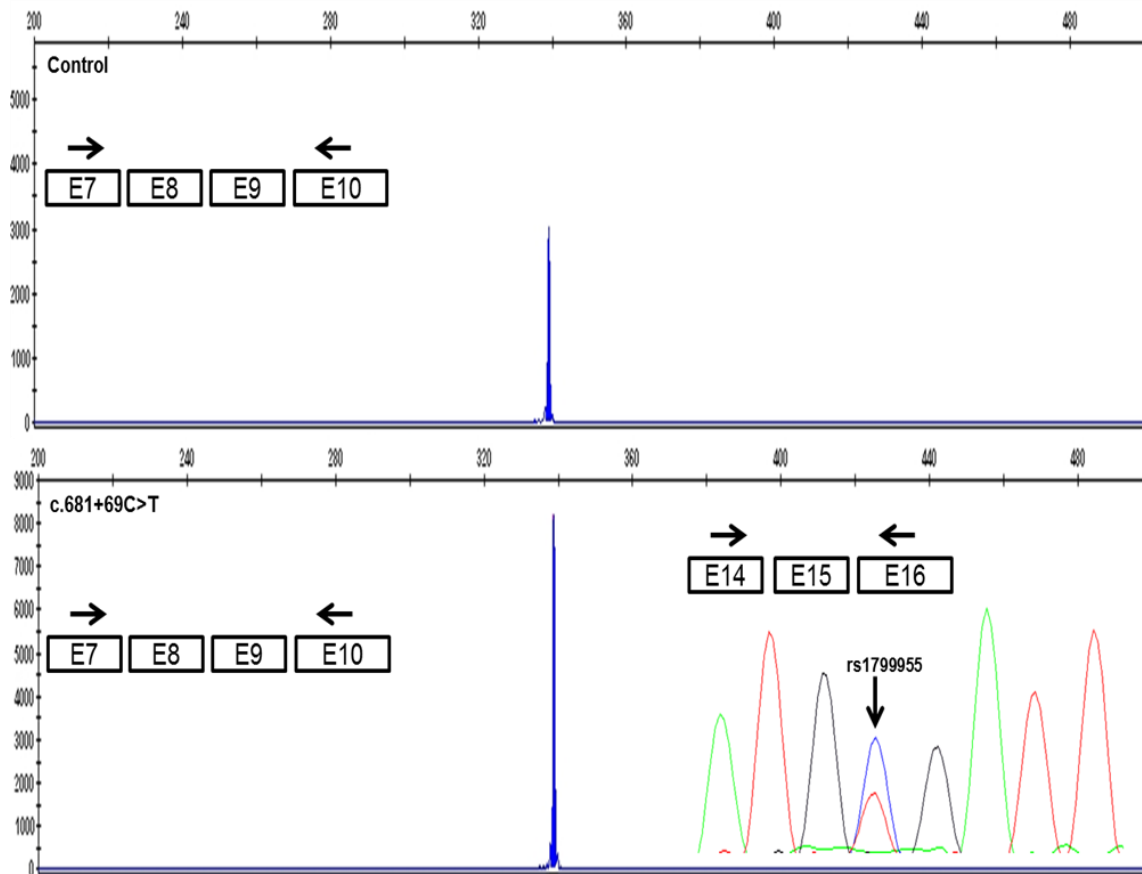
Supp. Figure S 8 Splicing analysis of c.426-89T>C, c.516+14C>T and c.631+25C>T

Capillary EP analysis of semiquantitative RT-PCR with primers located at exons 2 and 8 shows similar patterns of splicing in controls, c.426-89T>C, c.516+14C>T and c.631+25C>T carriers. Full-length, Δ3 and Δ6q7 BRCA2 transcripts were detected in all samples at similar levels. Bi-allelic expression was confirmed by direct sequencing of RT-PCR assays containing informative loci (rs1799944, rs766173, rs1799943).



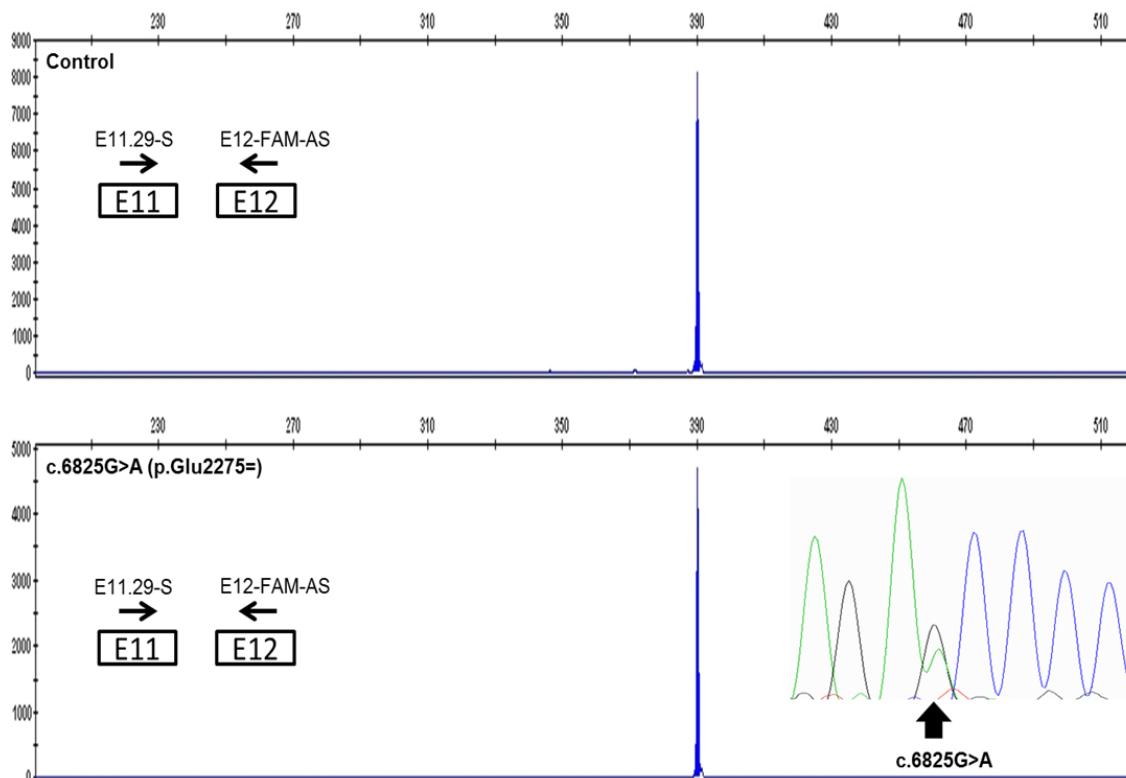
Supp. Figure S 9 Splicing analysis of c.680C>T variant.

Capillary EP analysis of semi-quantitative RT-PCR assays (33 cycle PCR) performed with primers located at exon 4 and 10 shows similar patterns of splicing in wt controls and c.680C>T carriers. Full-length and  $\Delta 6q7$  BRCA2 transcripts were detected in all samples at similar levels. Direct sequencing of RT-PCR assays confirmed bi-allelic expression in c.680C>T carriers.



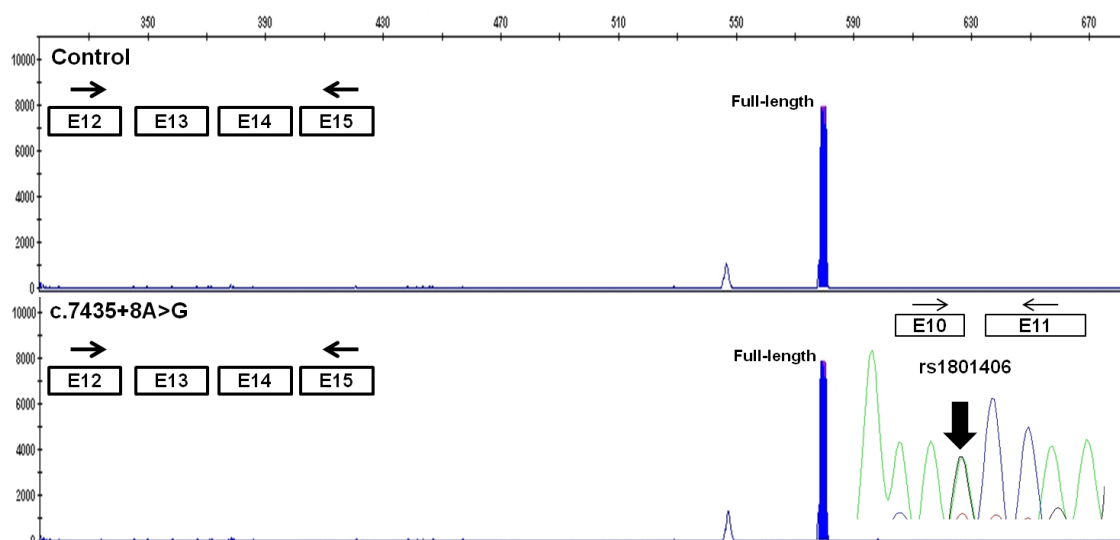
Supp. Figure S 10 Splicing analysis of c.681+69C>T

Capillary EP analysis of RT-PCR assays performed with primers located in exons 7 and 10 shows the presence of a single transcript both in wt controls and one c.681+69C>T carrier. Bi-allelic expression was confirmed by direct sequencing (reverse strand) of RT-PCR assays containing an informative locus in exon 14 (rs1799955).



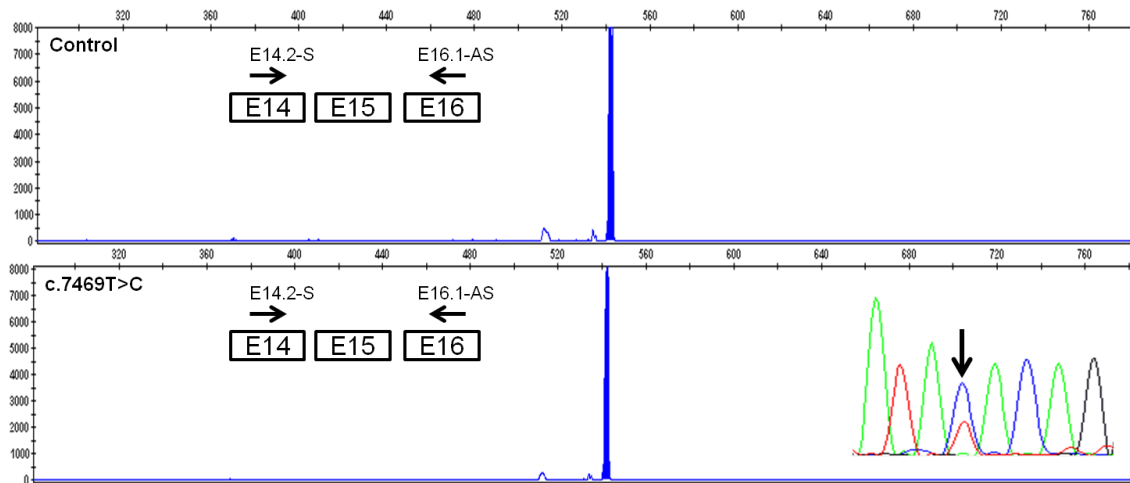
Supp. Figure S 11 Splicing analysis of c.6825G>A (p.Glu2275=) variant

RT-PCR with primers located in exons 11 and 12 (E11.29-S and E12-FAM-AS) shows the presence of a single transcript both in wt controls and one c.6825G>A carrier. Direct sequencing confirmed bi-allelic expression in variant carriers.



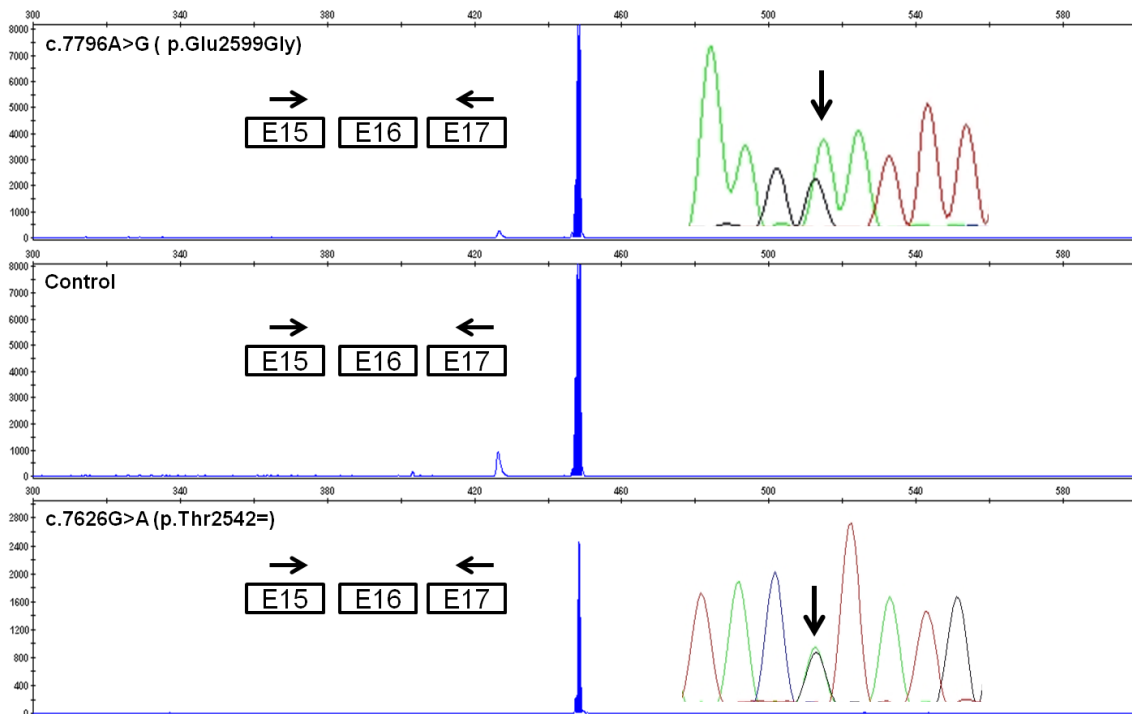
Supp. Figure S 12 Splicing analysis of c.7435+8A>G

RT-PCR with primers located in exons 12 and 15 (E12-S and E15-FAM-AS) shows the presence of a single transcript in wt controls and one c.7435+8A>G carrier. Bi-allelic expression was confirmed by direct sequencing of RT-PCR assays containing an informative locus at exon (rs1801406).



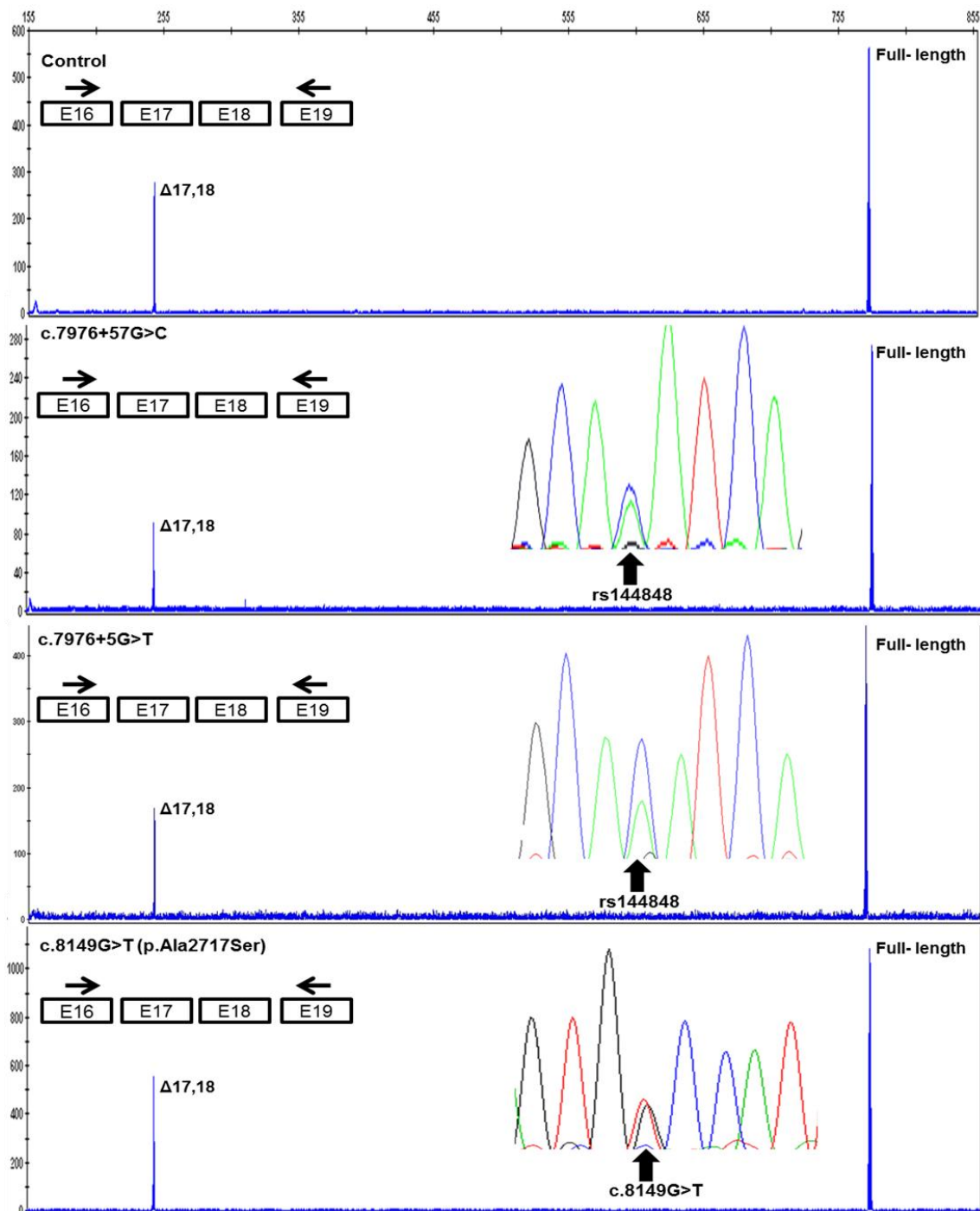
Supp. Figure S 13 Splicing analysis of c.7469T>C (p.Ile2490Thr) variant

a) RT-PCR with primers located at exon 14 and 16 (E14.2-S and E16.1-AS) shows the presence of a single transcript in control and c.7469T>C carrier. b) Sequencing of this product confirms the expression of both alleles.



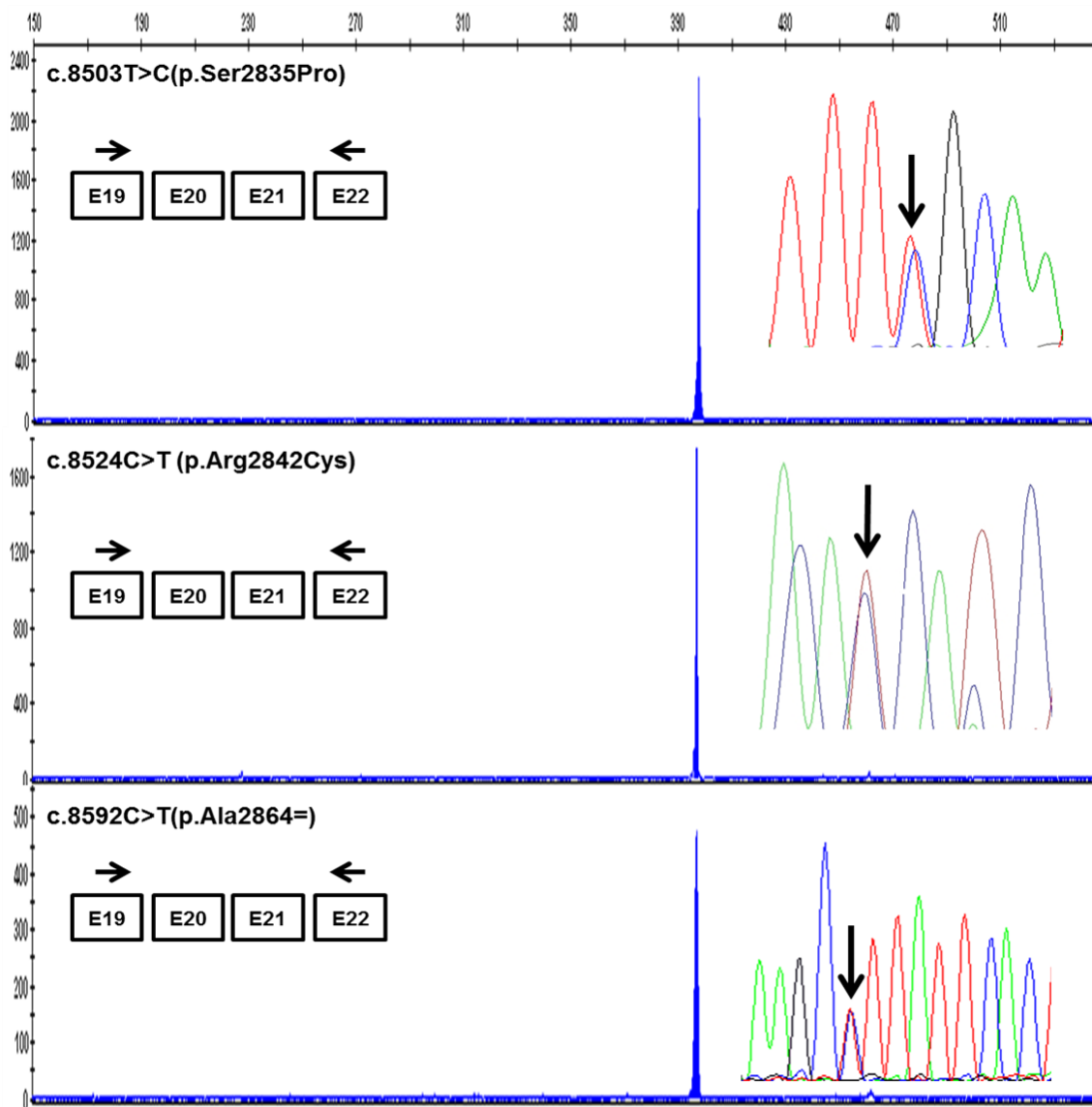
Supp. Figure S 14 Splicing analysis of c.7796A>G (p.Glu2599Gly) and c.7626G>A (p.Thr2542=) variants

RT-PCR with primers located at exon 15 and 17 (E15.2-S and E17.1-AS) shows the presence of a single transcript in controls, c.7796A>G carrier and c.7626G>A carrier. b) Sequencing of these products confirms the expression of both alleles in all cases.



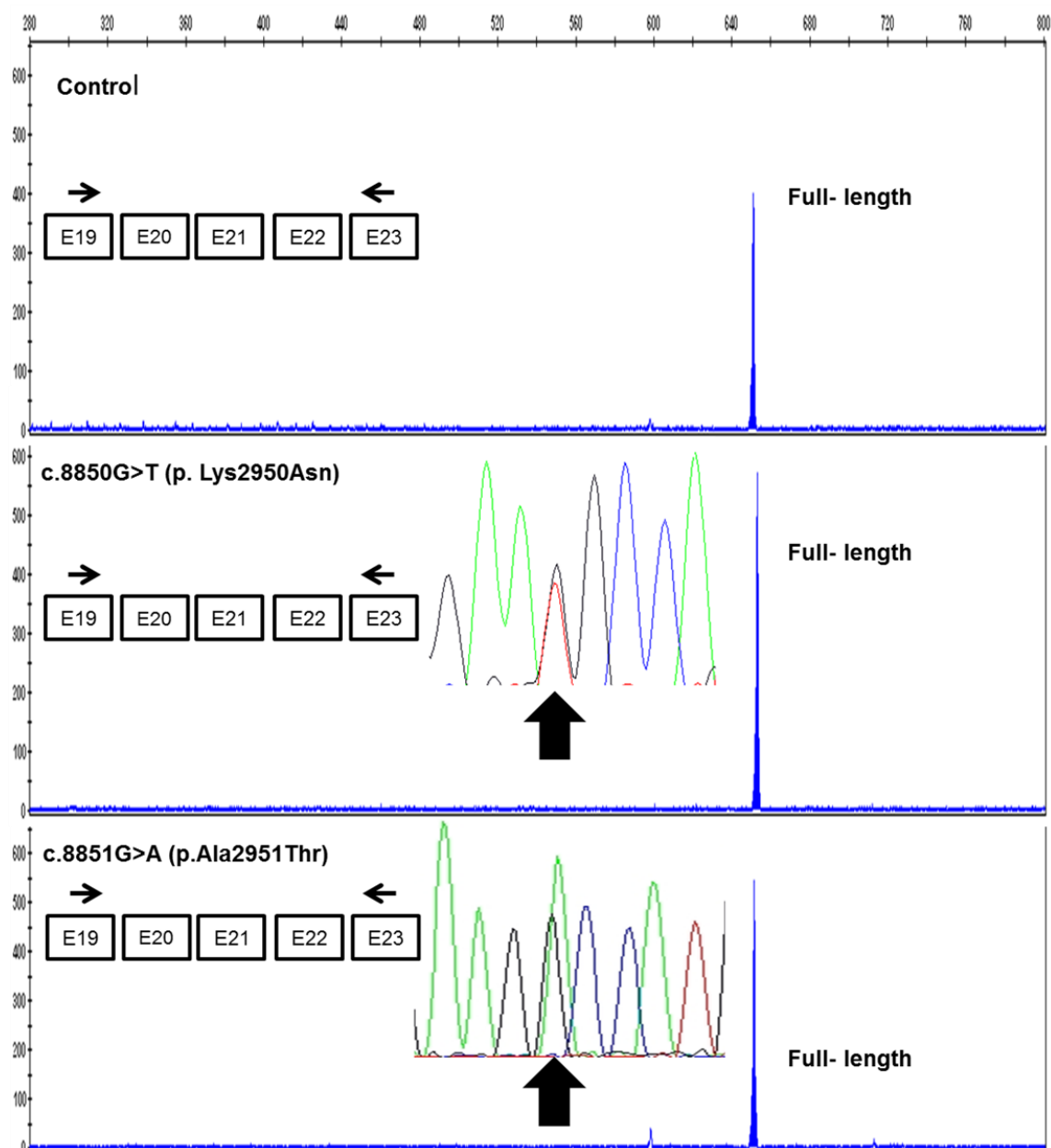
Supp. Figure S 15 splicing analysis of c.7976+57G>C, c.7976+5G>T and c.8149G>T (p.Ala2717Ser) variants.

Semiquantitative RT-PCR with primers located at exon 16 and 19 shows similar splicing patterns in controls, c.7976+57G>C, c.7976+5G>T and c.8149G>T carriers. Full-length and  $\Delta 17,18$  BRCA2 transcripts were detected in all samples at similar levels. Bi-allelic expression was confirmed by direct sequencing of RT-PCRs containing the BRCA2 variant (c.8149G>T), or in the case of intronic variants, an informative exonic locus (rs144848).



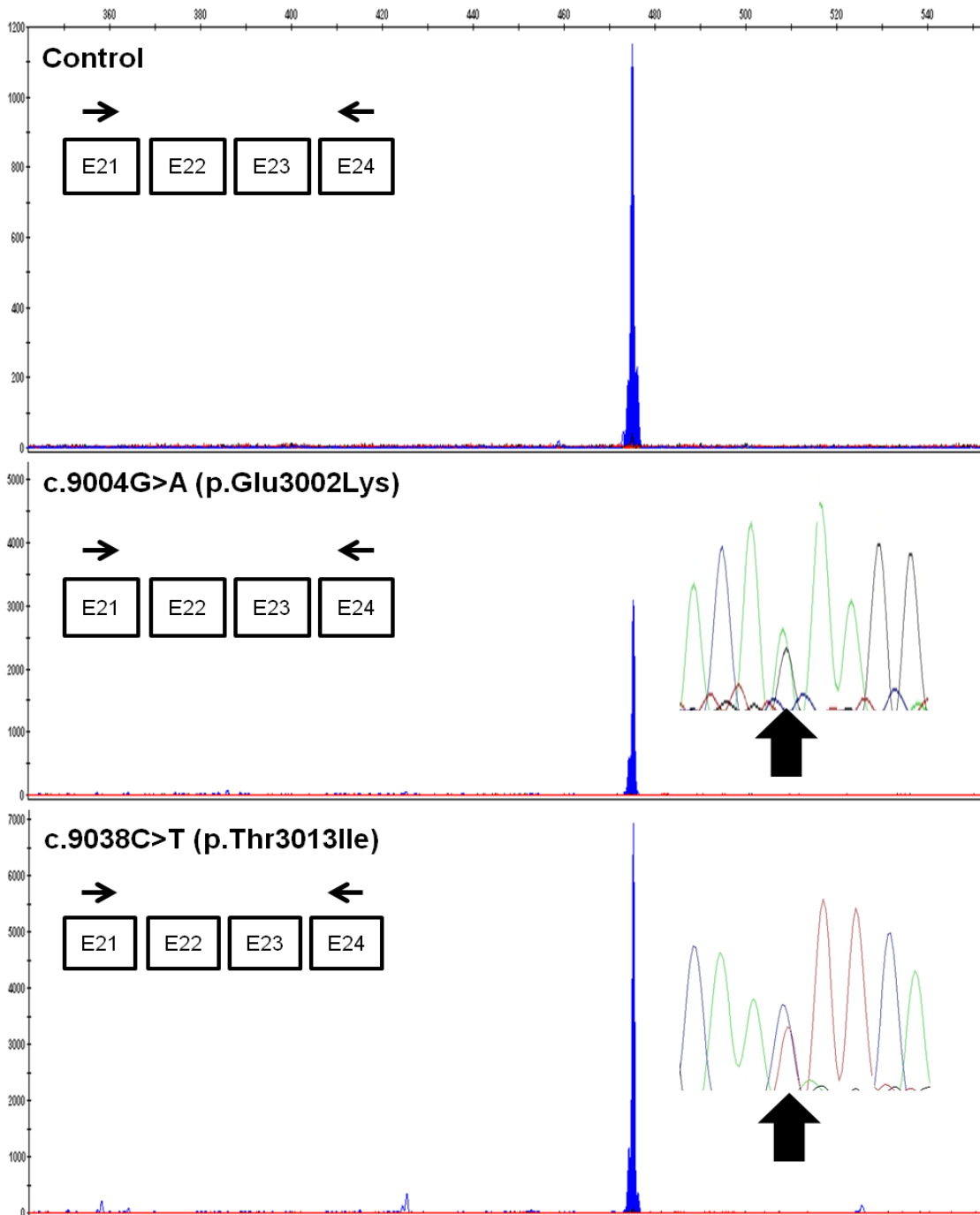
Supp. Figure S 16 Splicing analysis of c.8503T>C (p.Ser2835Pro), c.8524C>T (p.Arg2842Cys), and c.8503T>C (p.Ala2864=) variants

RT-PCR with primers located at exon 19 and 22 (E19-S and E22-FAM-AS) shows the presence of a single transcript in c.8503T>C, c.8524C>T, and c.8503T>C carriers. Direct sequencing of RT-PCR products confirmed the expression of both alleles.



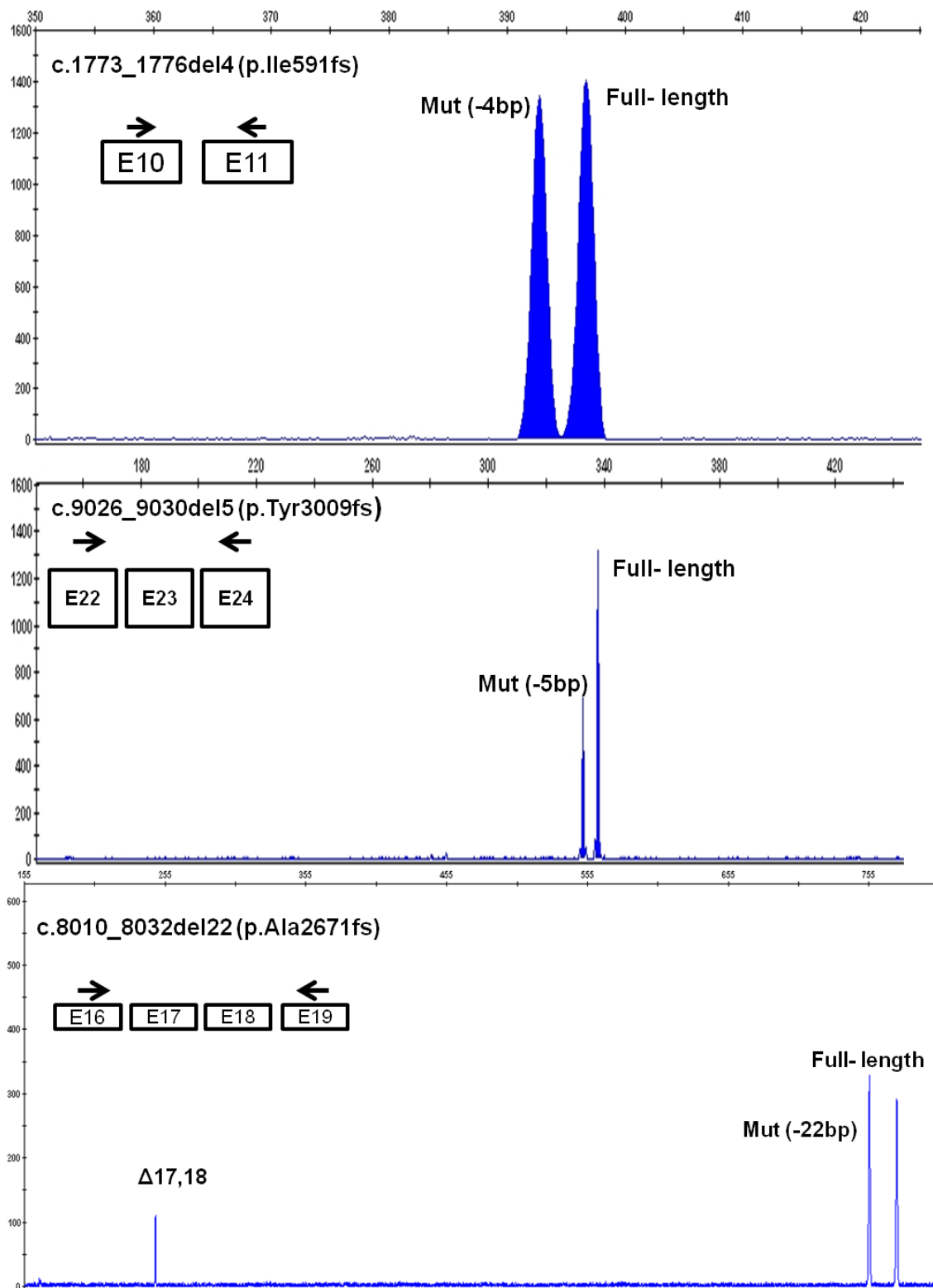
Supp. Figure S 17 Splicing analysis of c.8850G>T (p.Lys1950Asn) and c.8851G>A (p.Ala2951Thr) variants

RT-PCR with primers located at exon 19 and 23 (E19-S and E23-FAM-AS) shows the presence of a single transcript in wt control, c.8850G>T and c.8851G>T carriers. Direct sequencing confirmed bi-allelic expression.

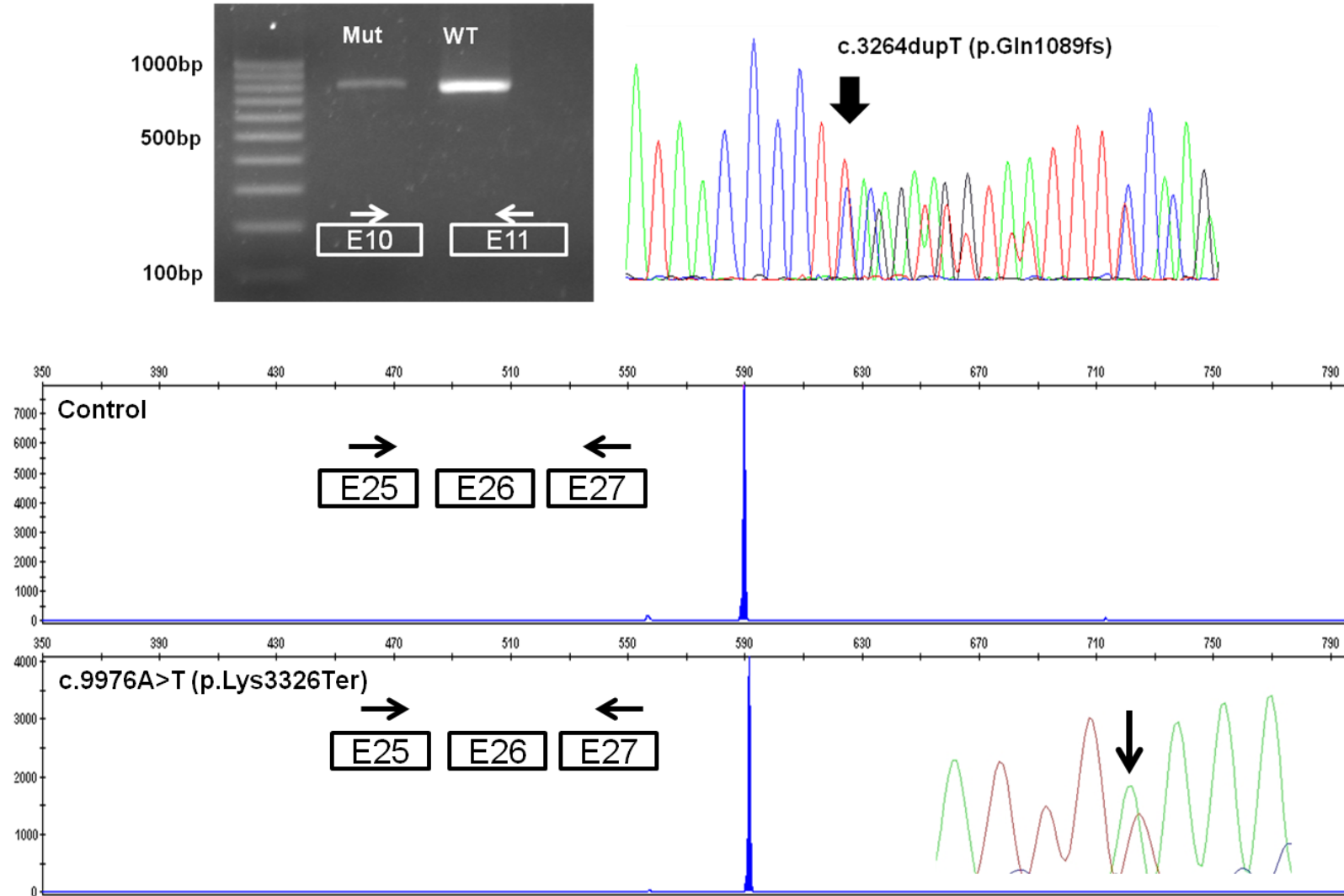


Supp. Figure S 18 Splicing analysis of c.9004G>A (p.Glu3002Lys) and c.9038C>T (p.Thr3013Ile)

RT-PCR with primers located at exon 21 and 24 (E21-S and E24-FAM-AS) shows the presence of a single transcript in control, c.9004G>A and c.9038C>T carriers. Direct sequencing confirmed bi-allelic expression.

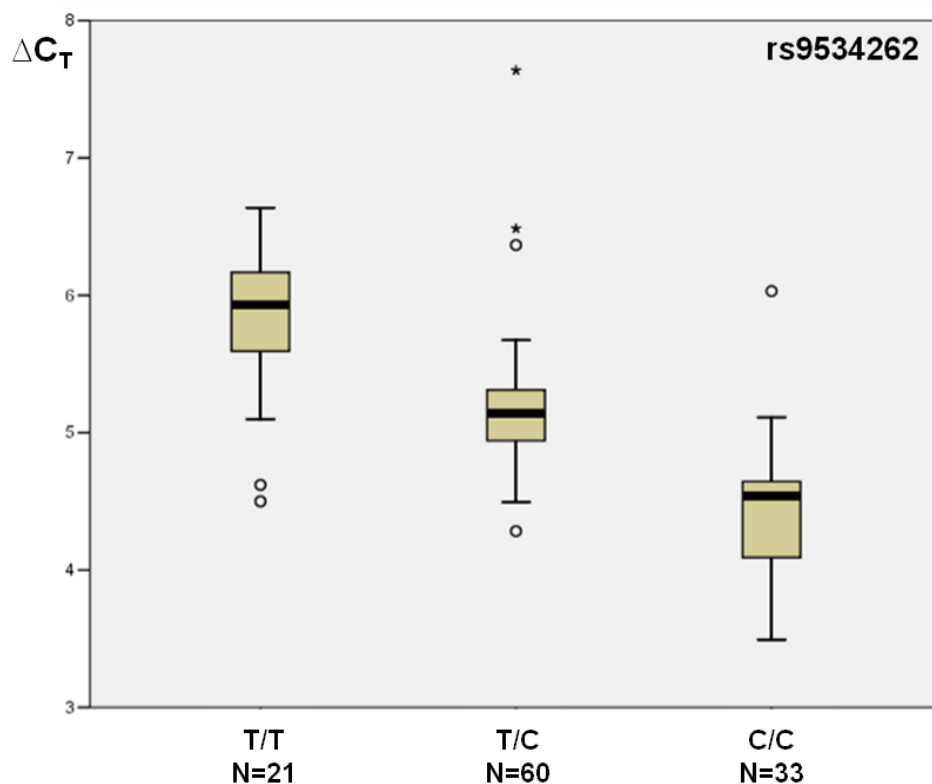


Supp. Figure S 19 Splicing analysis in carriers of three different BRCA2 PTC mutations did not identify splicing alterations  
However, all predicted PTC transcripts were detected. With the single exception of p.Tyr3009fs, PTC transcripts were detected at similar levels than wt transcripts.



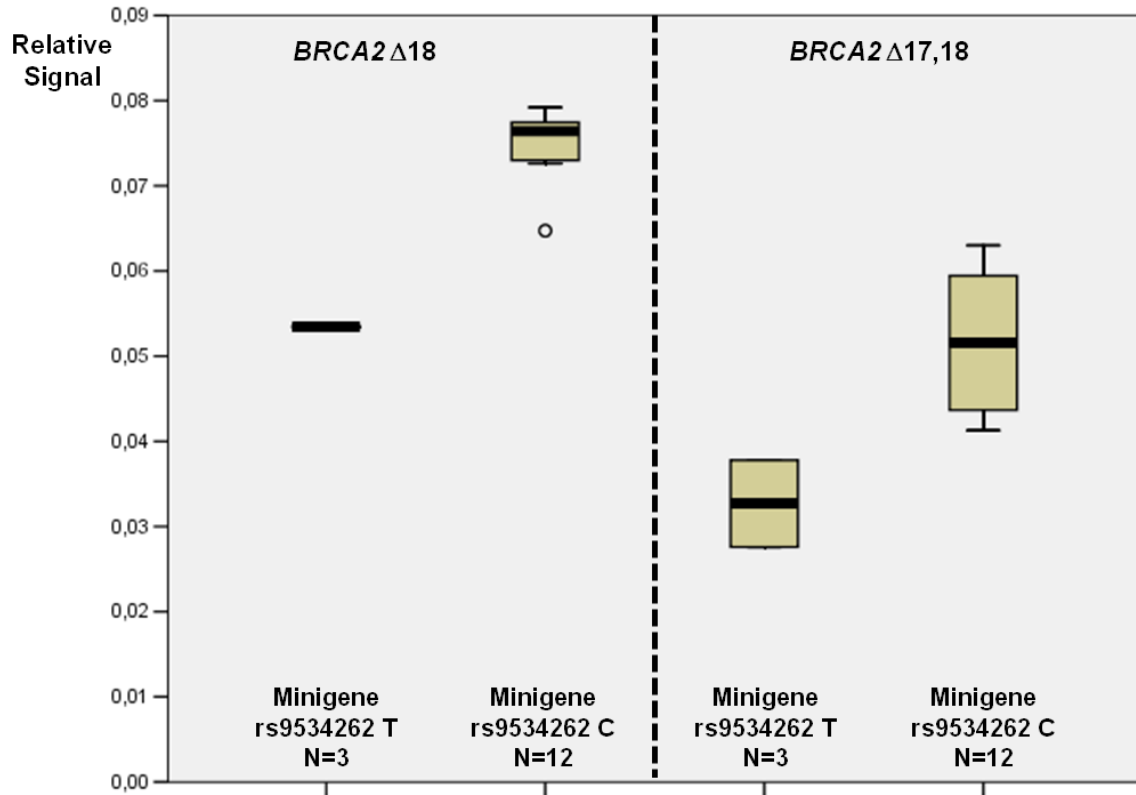
Supp. Figure S 20 Splicing analysis in carriers of two different BRCA2 PTC mutations did not identify splicing alterations.

PTC transcripts were detected at similar levels than wt transcripts (sequencing data).



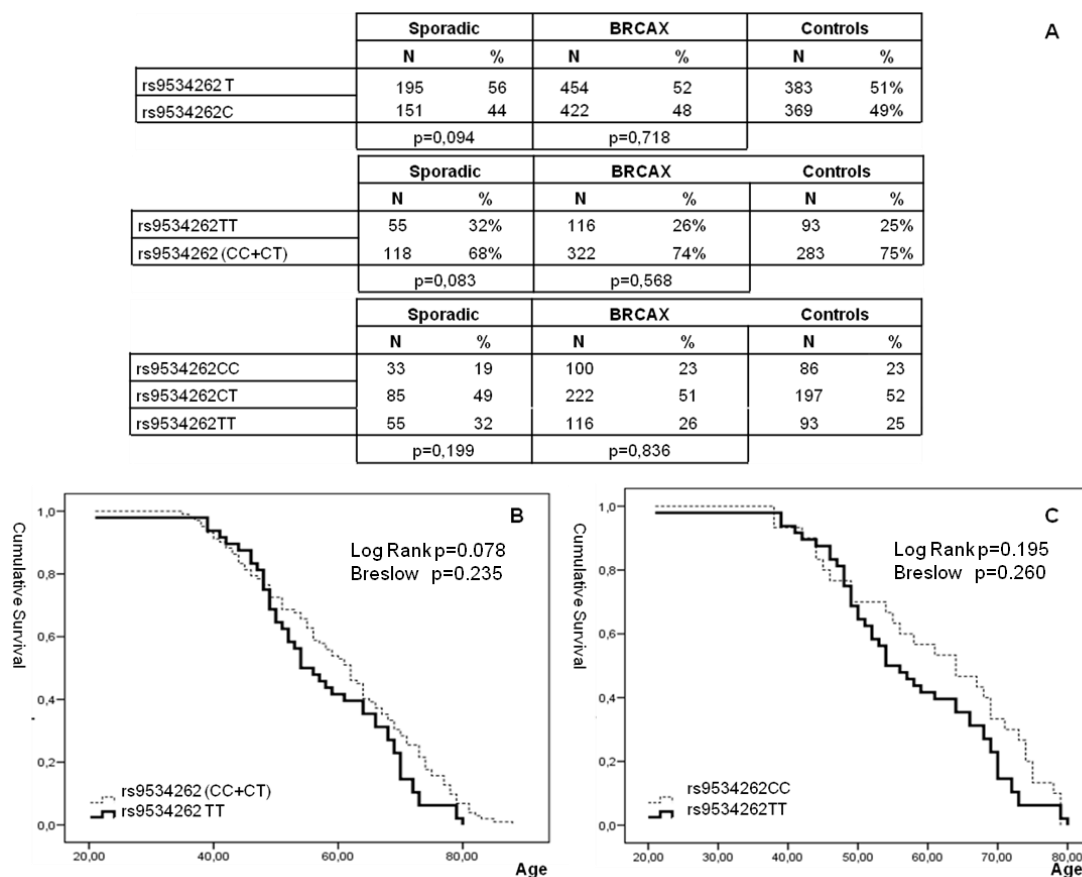
Supp. Figure S 21 qPCR analysis of BRCA2  $\Delta 17,18$  splicing isoform using BRCA2 exon 24 expression (a proxy for overall BRCA2 expression level) as an endogenous control

$\Delta C_T = (C_T \text{ BRCA2 } \Delta 17,18 - C_T \text{ BRCA2 exon24})$ . 7 T/T, 20 T/C, and 11 C/T samples were analyzed in triplicated. Statistically significant differences between C/C and T/T genotypes were observed (BRCA2  $\Delta 17,18$  is upregulated in C/C carriers). Quantification in T/C carriers is compatible with a dosage effect. Overall, the data indicates that rs9534262 (c.7806-14T>C) modulates BRCA2 alternative splicing.



Supp. Figure S 22 Minigene analysis of BRCA2  $\Delta 18$  and  $\Delta 17,18$  expression level using BRCA2 full-length as an internal reference

Relative signal=(peak area splicing isoform/peak area full-length). BRCA2  $\Delta 17,18$  and BRCA2  $\Delta 18$  are upregulated in rs9534262 C minigene.



Supp. Figure S 23 Case control study and Kaplan-Meier analysis of rs9534262

A) Association between rs9534262 genotype and breast cancer risk. Allelic frequencies and two genetic models are tested. Healthy controls, sporadic breast cancer, and BRCAX (index cases from BRCAX families) cohorts are described in on-line method.

B) Kaplan Meier analysis of the age at diagnosis (a dominant model is tested in the sporadic breast cancer cohort). The median age of diagnosis was 62 years for c.7806-14C carriers (CC+CT) and 54 years for non-carriers (TT), but differences were not statistically significant.

C) Similar results were obtained when median age of diagnosis was compared between CC and TT genotypes (64 vs. 54 years).

Supp. Table 1 Analytical classification of in vitro splicing assays outputs

Class	2008 original IARC guide-lines <sup>1</sup>	2013 up-dated guide-lines <sup>2</sup>
5	The variant allele produces <b>major</b> transcript(s) carrying a PTC, or in-frame deletions disrupting known functional domain(s)	The variant allele produces only transcript(s) carrying a PTC, or in-frame deletions disrupting known functional domain(s), determined by semi-quantitative or quantitative methods.
4	The variant allele produces <b>minor</b> transcript(s) carrying a PTC, or in-frame deletions disrupting known functional domain(s), at higher level than similar transcripts corresponding to physiological alternative splicing, or absent in wt controls.	The variant allele produces aberrant transcript(s) carrying a PTC, or in-frame deletions disrupting known functional domain(s), these transcripts are not found in controls assayed in tandem, but further assessment to determine if the UV allele does or does not contribute to production of wildtype transcript.
3	The variant allele produces <b>major</b> or <b>minor</b> transcript(s) carrying an in-frame deletion not disrupting known functional domain(s), at a higher level than similar transcripts corresponding to physiological alternative splicing.	All other transcript patterns e.g. the variant allele produces transcript(s) carrying an in-frame deletion not disrupting known functional domain(s), variant allele upregulates transcripts corresponding to physiological alternative splicing, at least some of which do not carry a PTC or are not an in-frame deletion disrupting known functional domains.
2	The variant allele produces the same transcript(s) as the wildtype allele.	The variant allele produces transcript patterns <b>comparable</b> to that of wildtype controls assessed in tandem.

<sup>1</sup>[Spurdle et al., 2008], <sup>2</sup>[Walker et al., 2013]

In 2008, the IARC proposed a classification system aimed to achieve robustness in the analytical interpretation of in vitro splicing assays. The scheme has been up-dated recently. Somehow inevitably; these guide-lines introduce vague terms such as **major**, **minor**, or **comparable** (highlighted in bold italics by authors of the present report). Note that the 2008 original guide-lines describe splicing assay outputs, while the update is more concerned with methodological issues. By doing so, the meaning of the classification system has been slightly modified. For instance, analytical Class-4 does not reflect any more UV alleles expressing both wild-type and PTC transcripts, but a mere lack of analytical information (the UV contribution to wild-type expression has not been tested). Note that analytical guide-lines do not make a distinction between PTCs located 5' or 3' to the last 50 nucleotides of the penultimate exon, despite the fact that this might have clinical relevance.

*Supp. Table 2 Primers used for Conventional RT-PCR Splicing Assays*

Name	Sequence
<b>Forward primers</b>	
E1	AGCTTACTCCGGCCAAAAA
E2	GGATCCAAAGAGAGGCCAAC
E4	GCAGATGATGTTTCCTGTCC
E7	GGTCGTCAGACACCAAAACA
E9	TGTGACAGACAGTGAAAAGAGAA
E10.6	GCCTCTGAAAGTGGACTGGA
E10.7	AAAAACACAGAAGGAATCGTCA
E10	CTAATTAAGTTCAGCCAG
E11.29	GCAGGCTTCACCTAAAAACG
E11	CCAAGTCATGCCACACATTC
E14	GAGACACTTGATTACTACAG
E15	GAATGCCAGAGATATACAG
E16.2	ATGGCGTTTCTAAACATTGC
E19	CATCGCTTTTCAGTGATGGA
E21	AGTTCGTGCTTTGCAAGATG
E22	AATCATTTTGTGTTAGTAAGGTC
E25	AGGGCCACTTTCAAGAG
<b>Reverse primers (Fam-labeled)</b>	
E3.1	ATTGGTACAGCGGCAGAGTC
E4	GGACAGGAAACATCATCTGCT
E8	ATCATGAGGAAAATACAGTTTCAGAT
E10	TTCCAATGTGGTCTTTGC
E10.3	TCTTTCTTTCTTTTGTCTCTGTGTC
E11	CAGAATTGTCCAAAAGA
E11.6	CGCAGGCTTGCTCAGTTTCTTT
E11.7	TGTAAGTCTGCAATATGT
E12	GCCATCTGGAGTGCTTTTTG
E13	AGTTGTGCGAAAGGGTAC
E15	TTCGCATATCCTGTATATCTCTGG
E16.1	ACCAAAATAATCTTCAGTGTGAAAC
E17.1	CCATAGCTGCCAGTTTCCAT
E19	GGGTCAGGAAAGAATCCAAG
E22	TCATTCAACATTTGCCTGTGA
E23	TTTTAGATTTTGAAGTTGCAAGA
E24	CCGTGGCTGGTAAATCTGA
E27	ATGCAAGTCTTCGTCAGCT

Supp. Table 3 In silico analysis of splicing regulatory elements.

Location	HGVS description <sup>1</sup>	Splicing Enhancer Elements				Splicing Silencer Elements				
		ESE Finder	RESCUE-ESE	PESE	EIE	Silencer motifs	ESS	Fas-ESS	PESS	IIE
Exon 2	c.-39-26G>A	-	-	-	2 new sites	-	-	2 sites broken	-	1 new site
Exon 5	c.426-89T>C	2 new sites	1 new / 1 broken	-	4 sites broken	-	-	-	1 site broken	2 sites broken
Exon 6	c.516+14C>T	1 new site	site broken	1 site broken	3 sites broken	1 site broken	-	-	-	3 new sites
Exon 7	c.631+25C>T	-	3 sites broken	-	4 sites broken	1 site broken	-	-	1 new site	2 new sites
Exon 8	c.680C>T (p.Ala227Val)	2 sites broken	-	1 site broken	-	1 new site	-	-	-	2 new sites
	c.681+69C>T	2 sites broken	-	-	-	-	-	2 new / 1 broken	-	2 new sites
Exon 9	c.682-2A>G	-	3 sites broken	1 new site	3 sites broken	1 new site	-	-	2 sites broken	-
Exon 10	c.1773_1776 delTTAT (p.Ile591fs)	-	-	-	3 new / 4 broken	-	-	-	4 sites broken	4 sites broken
Exon 11	c.3264dupT ( p.Gln1089fs)	1 site broken	2 new sites	new site	2 site broken	1 site broken	-	-	-	-
	c.6825G>A ( p.Glu2275=)	1 site broken	1 new site	5 new sites	1 new site	1 site broken	-	-	-	-
Exon 12	c.6928A>C (p.Thr2310Pro)	1 new/ 2 broken	1 site broken	1 new site	1 new site	1 new site	-	-	-	-
Exon 14	c.7435+8A>G	1 new site	-	-	1 new/1 broken	1 new site	-	-	-	2 new sites
Exon 15	c.7469T>C (p.Ile2490Thr)	-	-	-	1 new site	1 site broken	-	-	-	-
	c.7617+1G>A	2 sites broken	2 new sites	-	4 new sites	1 new site	-	1 site broken	1 new site	1 site broken
Exon 16	c.7626G>A (p.Thr2542=)	-	site broken	-	-	-	-	-	2 new sites	1 site broken
	c.7796A>G (p.Glu2599Gly)	-	-	2 sites broken	1 site broken	1 new site	-	-	1 site broken	-
	c.7976+5G>T	-	-	-	2 sites broken	2 sites broken	-	1 site broken	1 site broken	2 new sites
Exon 18	c.7976+57G>C	1 new site	site broken	-	-	2 new / 1 broken	-	2 sites broken	-	-
	c.8010_8032del (p.Ala2671fs)	2 sites broken	6 sites broken	1 site broken	1 new/14 broken	1 new/4 broken	-	1 new site	1 site broken	2 sites broken
Exon 20	c.8149G>T (p.Ala2717Ser)	new site	1 site broken	1 site broken	3 sites broken	2 sites broken	-	-	-	1 new site
	c.8524C>T ( p.Arg2842Cys)	-	-	-	-	1 site broken	-	-	-	4 new sites
	c.8503T>C ( p.Ser2835Pro)	2 new sites	site broken	1 new site	1 new /2 broken	-	-	-	-	-
Exon 22	c.8592C>T (p.Ala2864=)	-	-	1 new site	3 sites broken	-	-	-	-	2 new sites
	c.8850G>T (p. Lys2950Asn)	1 new site	site broken	-	4 new / 1 broken	1 site broken	-	-	1 new site	-
Exon 23	c.8851G>A (p.Ala2951Thr)	1 new site	2 new sites	2 new sites	6 new sites	1 site broken	-	-	-	-
	c.8954-5A>G	6 new sites	2 sites broken	-	2 new sites	2 sites broken	-	-	-	-
	c.9004G>A (p.Asp3002Lys)	2 sites broken	3 sites broken	4 sites broken	-	1 new/ 2 broken	-	-	-	-
	c.9026_9030del ( p.Tyr3009fs)	1 new site	-	-	2 sites broken	1 new/1 broken	-	-	-	2 new/3 broken
Exon 27	c.9038C>T (p.Thr3013Ile)	1 site broken	4 sites broken	1 site broken	2 sites broken	1 site broken	-	-	-	-
Exon 27	c.9976A>T (p.Lys3326Ter)	-	-	-	1 site broken	-	-	-	2 new sites	3 new sites

<sup>1</sup> Description according to HGVS guide-lines (<http://www.hgvs.org/mutnomen>). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the GeneBank reference sequence NM\_000059.3

Supp. Table 4 Haplotype analysis performed in 18 c.7617+1G>A carriers identified in 8 Spanish and 3 Danish families.

Origen	FAMILY ID	ID	-0,45 Mb	-0,28 Mb	-0,18 Mb	-64 kB	-40 kB	-23 Kb	0	+6 Kb	+0,17Mb	+0,28MB	+1Mb	+1,29 Mb	+14Mb	
			D13S260	D13S1699	D13S1698	D13S1697	rs1799943	rs144848	c.7617+1	rs9534262	D13S1701	D13S1715	D13S1493	D13S267	D13153	
Spain	HCSC_90	1	159/169	145/149	166/178	218/218	G/G	A/A	G/A	T/C	281/285	230/240	227/231	213/219	143/157	
		2	159/161	145/149	158/178	218/218	G/G	A/C	G/A	T/T	281/293	230/240	227/231	219/219	143/159	
		3	159/161	145/149	158/178	218/218	G/G	A/C	G/A	T/T	281/293	230/240	227/231	219/219	143/159	
	HCSC_580	4	159/163	149/149	158/178	218/218	A/G	A/C	G/A	T/T	281/285	234/240	227/231	219/219	151/155	
		5	159/163	149/149	158/178	218/218	A/G	A/C	G/A	T/T	281/285	234/240	227/231	219/219	149/151	
		6	159/165	145/149	162/178	218/218	G/G	A/A	G/A	T/C	269/281	230/240	231/235	207/219	149/155	
	HCSC_830	7	159/161	149/149	158/178	218/218	A/G	A/C	G/A	T/C	281/293	240/244	231/235	207/219	152/155	
	HPTH-55	10	159/159	145/149	172/178	218/218	A/G	A/A	G/A	T/C	281/289	240/244	223/231	207/207	145/153	
	Iov-99-09	11	159/161	149/149	162/178	218/218	G/G	A/C	G/A	T/T	281/289	240/244	223/231	219/219	151/157	
		12	159/161	149/149	162/178	218/218	G/G	A/C	G/A	T/T	281/289	240/244	223/231	219/219	147/157	
	200	13	159/161	149/149	164/178	218/222	A/G	A/C	G/A	T/C	285/293	230/240	231/231	219/219	147/157	
		14	159/161	145/149	166/178	211/218	G/G	A/C	G/A	T/T	281/285	230/240	227/231	207/219	147/149	
	1329	15	159/169	145/149	162/178	214/218	A/G	A/C	G/A	T/C	281/293	234/240	223/235	207/213	157/161	
	696	16	159/165	145/149	174/178	218/222	G/G	A/A	G/A	T/C	281/289	240/244	223/231	207/219	151/157	
		17	159/151	145/149	176/178	218/218	G/G	A/C	G/A	T/C	281/281	240/244	223/231	207/219	145/157	
	Denmark	D1	1	161/167	145/149	160/174	218/218	A/G	A/A	G/A	T/C	289/293	244/244	227/227	207/207	151/157
		D2	2	167/167	145/149	160/166	218/218	G/G	A/A	G/A	T/T	289/289	244/244	227/235	207/209	141/149
D3		3	165/167	145/149	160/162	218/218	G/G	A/A	G/A	T/T	289/297	244/244	223/235	207/221	153/163	

All eight families carrying this disease-causing mutation were resident in the Barcelona or Madrid metropolitan areas. We learnt through personal interviews that 7 families (we lack information on the remaining one) originated in the very same West-Spanish region (Extremadura). The length of the shared haplotype ranged from 0.51Mb to 1.74Mb. Interestingly, BRCA2 c.7617+1G>A has been reported previously as a Danish founder mutation highly prevalent in the Vejle area [Thomassen et al., 2011]. The authors used SNPs arrays to define a shared haplotype of 0.2Mb to 3.3Mb in 19 families carrying this mutation. For that reason, we have included in our analysis three Danish families carrying the BRCA2 c.7617+1G>A data. Our data (STR-based) is compatible with, but do not provide convincing evidence supporting a common origin for the Spanish and Danish mutations.

Supp. Table 5 Haplotype analysis in 10 c.8954-5A>G carriers identified in 4 unrelated Spanish families

		-0,45 Mb	-1.1 Kb	0	+1 Mb	+1.29 Mb	+14 Mb
Family ID	ID	D13S260	rs9534262	c.8954-5	D13S1493	D13S267	D13153
HCSC_896	1	168/168	T/C	A/G	227/235	207/209	151/157
Get483	5	168/168	T/C	A/G	227/243	207/209	157/161
	6	166/168	T/T	A/G	227/243	207/207	151/157
	7	162/168	T/C	A/G	227/235	207/209	145/157
	9	160/164	T/C	A/G	227/235	207/219	141/149
IVO_305	10	166/168	T/T	A/G	223/227	207/213	157/159
IVO_209	13	160/166	T/C	A/G	223/227	207/207	147/155
	15	160/160	T/T	A/G	223/227	207/213	151/153
	17	160/166	T/C	A/G	223/227	207/219	149/153
	19	160/166	T/C	A/G	223/227	207/219	153/153

All four families carrying this disease-causing mutation were residents in the Madrid or Valencia metropolitan areas. Personal interviews did not reveal any obvious common origin for these families. However, haplotype analysis identified a 1.28 Mb shared haplotype in all carriers, suggesting a founder effect. Families IVO\_209 and IVO\_305 have been reported previously [de Juan Jimenez et al., 2013]. BRCA2 c.8954-5A>G has been identified as well in three French families (<http://www.umd.be/BRCA2/>).



## 2. Búsqueda de variantes relacionadas con el síndrome HBOC en familias BRCA1



## 2.A Búsqueda de variantes patogénicas en el gen BRCA2 que no se detectan mediante las técnicas basadas en PCR

*Caracterización de cuatro nuevos grandes reordenamientos genómicos en el locus BRCA2 en familias españolas con cáncer de mama/ovario. Revisión de la literatura y reevaluación de los mecanismos genéticos involucrados en su origen.*

Los grandes reordenamientos genómicos (LGRs, del inglés Large Genomic Rearrangements) en el locus *BRCA2* explican una proporción no despreciable de los síndromes HBOC. El uso de la técnica MLPA ha permitido en los últimos años la identificación de familias portadoras de LGRs en este locus, aunque muy pocas de estas alteraciones han sido caracterizadas a nivel molecular. Sin embargo, la caracterización molecular es esencial para identificar alteraciones recurrentes, para analizar los mecanismos genéticos que subyacen a estas alteraciones, o para investigar posibles relaciones genotipo / fenotipo. Mediante el uso de MLPA se han identificado LGRs en el locus *BRCA2* en 7 de 813 familias española con síndrome HBOC previamente negativas para pequeñas alteraciones genómicas (substituciones e inserciones/delecciones) en *BRCA1/2* y para LGRs en *BRCA1*. Mediante el uso de una combinación de PCR de largo alcance, mapas de restricción y análisis de ADNc, se pudo determinar que Del Exon1-Exon2, Del Exon12-Exon16 and Del Exon22-Exon24 aparecen cada una en una familia, mientras que Del Exon2 parece ser una mutación fundadora española que aparece en cuatro familias independientes. Finalmente, se combinaron datos del presente estudio junto con una revisión exhaustiva de la literatura para reevaluar los mecanismos genéticos que originan los LGRs en el locus *BRCA2*. El presente estudio aumenta de manera substancial el espectro de LGRs en *BRCA2* caracterizados a nivel molecular. Finalmente, los datos sugieren que los mecanismos de recombinación homóloga no alélica (NAHR) podrían estar sobreestimados en cuanto a su involucración en la generación de LGRs en *BRCA2* mientras que los mecanismos mediados por micro-homología estarían infraestimados.



## Characterization of four novel *BRCA2* large genomic rearrangements in Spanish breast/ovarian cancer families: review of the literature, and reevaluation of the genetic mechanisms involved in their origin

Gorka Ruiz de Garibay · Sara Gutiérrez-Enríquez · Pilar Garre · Sandra Bonache · Atocha Romero · Laura Palomo · Ana Sánchez de Abajo · Javier Benítez · Judith Balmaña · Pedro Pérez-Segura · Eduardo Díaz-Rubio · Orland Díez · Trinidad Caldés · Miguel de la Hoya

Received: 28 November 2011 / Accepted: 30 November 2011 / Published online: 21 March 2012  
© Springer Science+Business Media, LLC. 2012

**Abstract** Large genomic rearrangements (LGRs) at the *BRCA2* locus explain a non-negligible proportion of hereditary breast and ovarian cancer (HBOC) syndromes. The multiplex ligation and probe amplification (MLPA) assay has permitted in recent years to identify several families carrying LGRs at this locus, but very few such alterations have been fully characterized at the molecular level. Yet, molecular characterization is essential to identify recurrent alterations, to analyze the genetic mechanisms underlying such alterations, or to investigate potential genotype/phenotype relationships. We have used MLPA to identify *BRCA2* LGRs in 7 out of 813 Spanish HBOC families previously tested negative for *BRCA1* and *BRCA2* small genomic alterations (substitutions and indels) and *BRCA1* LGRs. We used a combination of long-range

PCR, restriction mapping, and cDNA analysis to characterize the alterations at the molecular level. We found that Del Exon1-Exon2, Del Exon12-Exon16 and Del Exon22-Exon24 explain one family each, while Del Exon2 appears to be a Spanish founder mutation explaining four independent families. Finally, we have combined our data with a comprehensive review of the literature to reevaluate the genetic mechanisms underlying LGRs at the *BRCA2* locus. Our study substantially increases the spectrum of *BRCA2* LGRs fully characterized at the molecular level. Further on, we provide data to suggest that non-allelic homologous recombination has been overestimated as a mechanism underlying these alterations, while the opposite might be true for microhomology-mediated events.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10549-011-1909-0) contains supplementary material, which is available to authorized users.

**Keywords** Familial breast cancer · *BRCA2* · MLPA · Genomic rearrangements · Non-allelic homologous recombination · Microhomology

G. Ruiz de Garibay · P. Garre · A. Romero · E. Díaz-Rubio · T. Caldés · M. de la Hoya (✉)  
Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos, C/Martín Lagos s/n, 28040 Madrid, Spain  
e-mail: mdhoya@hotmail.com

J. Benítez  
Grupo de Genética Humana, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain

S. Gutiérrez-Enríquez · O. Díez  
Laboratori d'Oncogenètica, Institut d'Oncologia Vall d'Hebron (VHIO), Barcelona, Spain

J. Balmaña  
Servei d'Oncologia Mèdica, Hospital Universitari Vall d'Hebron, Barcelone, Spain

S. Bonache · L. Palomo  
Laboratori d'Oncogenètica, Institut de Recerca Vall d'Hebron (VHIR), Universitat Autònoma de Barcelona, Barcelone, Spain

P. Pérez-Segura · E. Díaz-Rubio  
Servicio de Oncología Médica, Hospital Clínico San Carlos, Madrid, Spain

A. Sánchez de Abajo  
Servicio de Bioquímica Clínica, Hospitalario Universitario Insular de Gran Canaria, Las Palmas de Gran Canaria, Spain

## Introduction

Germline mutations in *BRCA1* (OMIM 113705) and *BRCA2* (OMIM 600185) cause hereditary breast and/or ovarian cancer (HBOC) syndromes [1, 2]. According to the BIC database (<http://research.nhgri.nih.gov/bic/>), most disease-causing mutations identified so far are either substitutions or indels spread throughout all coding exons and intronic flanking sequences of both genes. Yet, structural alterations explain a non-negligible proportion of HBOC syndromes [3], so that screening of these mutations has been incorporated into the routine of most laboratories. Somehow misleading, the expression “large genomic rearrangements” (LGRs) is often used in the scientific literature to refer collectively to these alterations, despite the fact that its actual size might vary enormously, ranging from less than 500 bp to complete loss of the gene.

Apparently, the contribution of LGRs to the spectrum of disease-causing mutations is consistently lower in *BRCA2* than in *BRCA1*, regardless of the population analyzed. For instance, LGRs account for 2% of all *BRCA2* and 12.5% of all *BRCA1* disease-causing mutations detected in a Danish cohort [4]. Similar findings are observed in Australian (2.3% vs. 14.9%), Malaysian (4% vs. 8%), Italian (0% vs. 10.5%), Czech Republic (0% vs. 12.3%), or USA (6% vs. 22%) cohorts [5–9]. In Spain, previous studies detected *BRCA2* LGRs in 1.5%, and *BRCA1* LGRs in 2.8% of the families analyzed [10, 11]. These differences have been attributed to structural differences between genes. In particular, the higher frequency of Alu repeat sequences in the *BRCA1* gene together with the existence of the  $\psi$ *BRCA1* pseudogene increases the likelihood of intra-locus recombination events [12].

Diverse methodologies such as Southern blot, long-range PCR, and quantitative multiplex PCR of short fragments (QMPSF) have been used (and are still used) to detect and/or analyze LGRs [13]. However, the multiplex ligation and probe amplification (MLPA) assay [14] has emerged in recent years as the most widely used approach to scan for LGRs in these genes. Compared with other methods, the assay is non-expensive, sensitive, and well suited for high-throughput screening in the routine setting. Indeed, the vast majority of *BRCA1* and *BRCA2* LGRs reported so far have been detected with this assay [3].

The MLPA assay is designed to detect suspected LGR carriers, but further analysis are needed to characterize the alteration. Unfortunately, molecular characterization is time-consuming and technically challenging (in particular if the alteration extends upstream or downstream of the target gene), making it difficult to implement in a routine laboratory. Yet, molecular characterization is essential to understand the deleterious effect of LGRs, to identify recurrent alterations, to analyze the genetic mechanisms

underlying such alterations, or to investigate potential genotype/phenotype relationships. Moreover, molecular characterization might be critical to confirm the presence of a LGR, thus discarding MLPA false-positive results that may arise if point mutations are present within the sequence targeted by the probes [14].

We have been able to identify in the literature 19 *BRCA2* LGRs fully characterized at the molecular level [4, 6, 9, 15–24]. Here, we describe the molecular characterization of four additional LGRs identified in Spanish HBOC families.

## Materials and methods

### Patients and samples

MLPA was performed in 813 unrelated individuals ascertained through four Spanish familial cancer clinics: Hospital Clínico San Carlos, Madrid ( $N = 370$ ), Hospital Vall d’Hebrón, Barcelona ( $N = 347$ ), Hospital Universitario Insular, Las Palmas de Gran Canaria ( $N = 50$ ), and Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid ( $N = 46$ ), Spain. Informed written consent was obtained from all patients after appropriate counseling according to the protocols approved by our institutional review boards. After comprehensive *BRCA1* and *BRCA2* mutation scanning and *BRCA1* analysis by MLPA, all 813 individuals were considered negative for the presence of germline mutations. Mutation scanning methods have been described elsewhere [25]. The cohort included the following:

- 137 Index cases from families with both breast and ovarian cancer cases (at least one of each). Most families included 3 or more cases in two consecutive generations.
- 499 Index cases from site-specific breast cancer families (at least two cases). Most families included three or more cases in two consecutive generations.
- 39 Index cases from families selected because they have male breast cancer cases.
- 138 Women diagnosed with breast and/or ovarian cancer with no family history but some other features associated with hereditary syndromes, such as early onset of diagnosis and/or limited family structures [26].

Genomic DNA (gDNA) was extracted from peripheral blood leukocytes using standard protocols. For RNA extraction, fresh peripheral blood was first treated with erythrocytes lysis solution (RBC lysis solution, Qiagen Sciences, Maryland, USA). Total RNA was extracted from leukocytes using TRIzol reagent (Invitrogen Life Technologies, Rockville, MD, USA) according to the manufacturer’s protocol. Isolated RNA was further processed on columns from RNeasy Kit and RNase-Free DNase set to

clean up RNA and remove any DNA traces, according to the protocols provided by the manufacturer (Qiagen Sciences, Maryland, USA). One microgram of total RNA was reverse-transcribed using High Capacity RNA-to-cDNA Master Mix with random primers and oligo(dT) primers (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocols.

#### MLPA

Screening was performed with MLPA commercial kits (MRC-Holland, Amsterdam, The Netherlands) according to the manufacturer's protocol. Depending on the contributing center, P045-B2 or P090-B2 kits were used. *BRCA2* probes are identical in both kits. Each positive result was confirmed in a second independent reaction.

#### PCR amplification

Depending on the expected size of the wt allele, we performed long-range PCR using LA Taq (Takara Bio Inc, Japan)/Expand Long Range kit (Roche Diagnostic GmbH, Mannheim, Germany), or standard PCR using FastStart Taq (Roche Diagnostic GmbH, Mannheim, Germany), according to the manufacturer's protocol. PCR products were separated on 0.8–1% agarose gels and visualized by ethidium bromide staining. Primers are shown in Table S1.

#### Restriction enzyme analysis

Five microliters of long-range PCR product was digested overnight at 37°C with 10 U of HindIII, BglII, and XmaI restriction enzymes (Promega, Madison, WI, USA) in 1X multicore supplemented with 10% BSA. Restriction maps were created with CLC sequence Viewer (CLC bio A/S, Denmark).

#### Sequencing

Cycle sequencing was performed using the BigDye Terminator v1.1 Cycle Sequencing kit. Analysis was performed on the ABI 3130 (Applied Biosystems, Foster City, CA, USA). All LGRs are described at the DNA level [27] using guidelines recently proposed [3]. In all cases, Human Genome Variation Society (HGVS) rules are followed (<http://www.hgvs.org>).

#### Haplotype analysis

Haplotype analysis was performed with four short tandem repeats (STR) encompassing a ~15 Mb of chromosome 13 encompassing *BRCA2* (D13S260, D13S1493, D13S267, and D13S153). Primer sequences were obtained from the

Ensembl database (<http://www.ensembl.org>). One primer of each pair was labeled with 6-carboxyfluorescein. The amplification products were separated on the ABI3130 Genetic Analyzer. Three *BRCA2* intragenic SNPs (rs144848, rs543304, rs9534262) were also included in the analysis.

#### Sequence alignment analysis

The NCBI *BRCA2* RefSeq NG\_012772.1 (<http://www.ncbi.nlm.nih.gov>) was submitted to RepeatMasker (<http://www.repeatmasker.org/>). Analysis with default settings identified 71 SINEs, 25 LINEs, 6 LTRs, and 17 DNA elements. Together, these repeat elements comprise 50.39% of the whole sequence. Simple inspection permitted us to identify breakpoints located at interspersed elements (Table 1). To identify homology-related events, we performed sequence alignments for all 21 *BRCA2* LGRs characterized so far. Alignments were performed using CLUSTAL W2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). In all cases, 61-bp sequences centered on the 5' and 3' breakpoints were aligned. Slow pairwise alignment options with default parameters were used in all cases [28].

## Results

To determine the frequency of *BRCA2* large genomic rearrangements (LGRs) in our population, we examined 813 unrelated breast and/or ovarian cancer patients previously screened negative for small disease-causing mutations in *BRCA1* and *BRCA2*, and for LGRs in *BRCA1*. MLPA analysis identified 10 families with putative deletions targeting exon 1, exons 1–2, exon 2 (four independent families), exon 3 (two independent families), exons 12–16, and exons 22–24.

We performed further sequence analysis in those samples with dosage alterations affecting a single MLPA probe. The analysis revealed that exon 1 and exon 3 alterations were indeed false positives due to previously unnoticed nucleotide changes disturbing probe binding (Table S2). Interestingly, one of these substitutions (c.-227G>C) targets the *BRCA2* transcriptional start point [29]. This change was identified in a woman diagnosed with multifocal breast cancer at age 43. Melanoma and pancreatic cancer were also observed in the family. To investigate the possibility that this alteration was indeed a deleterious mutation abrogating gene expression, we sequenced cDNA at two heterozygous positions: c.3807T>C (p.Val1269=) and c.6471 A>G (p.Gln2157=). Apparently, both alleles were expressed at similar levels, discarding gene expression abrogation. Yet, cDNA sequencing is not quantitative, so we cannot rule out the possibility of some effect of this variant on expression level.

**Table 1** BRCA2 large genomic rearrangements characterized at the molecular level

Exons	Genomic alteration (HGVS nomenclature) <sup>a</sup>	Predicted effect at the protein level (HGVS nomenclature) <sup>b</sup>	Repeat element at 5' breakpoint <sup>c,e</sup>	Repeat element at 3' breakpoint <sup>a,c</sup>	Proposed mechanism <sup>d</sup>	Reference
1–2	g.3793_6539delinv.g.5229_5246	Lack of expression (see text)	-	Alu Sz g.6437_6742+	NHE	This study
1–2	g.4093_6446delins12 <sup>e</sup>	Lack of expression	Alu Sp g.4090_4389+	Alu Sz g.6437_6742+	NAHR (NHE)	[9]
2	g.5670_6155delinsATA	No translation/p.Met1_Lys123del (see text)	MER3 g.5659_5814-	-	NHE	This study
3	g.7674_11736del	p.Asp23_Leu105del	Charlie1a g.7240_7794-	L1MC2 g.11491_11803-	-(MH)	[15]
3	g.8686_8687insAlu	-	-	-	-	[16]
3	g.8807_13871delinsCCAT	p.Asp23AlafsX7	-	-	NHE	[17]
4–11	g.12507_28681dup	p.Gly106_Val1280dup	-	-	NHE	[18]
7–11	g.16099_30416del	p-Ala199GlyfsX17	-	-	NHE (MH)	[19]
12–13	g.31194_40076del	p.Gly2281AlafsX30	AluYa5 g.31146_31403+	AluY g.39989_40303+	NAHR	[20]
12–13	g.31338_39284del	p.Gly2281AlafsX30	AluYa5 g.31146_31403+	AluSx g.39075_39353+	NAHR	[21]
12–13	g.31404_37615delins(A) <sub>~60</sub>	-	AluYa5 g.31146_31403+	-	NHE (MH)	[22]
12–16	g.32051_48841delinsTAG	p.Gly2281ValfsX46	Tigger3a g.32011_32284-	-	NHE	This study
14–18	g.41600_57818del <sup>f</sup>	p.Thr2337PhefsX43	AluSx8 g.41317_41621-	AluSx3 g.57529_57827-	NAHR	[20]
14–16	g.42422_49277del	p.Thr2337_Arg2602del	AGACTG g.42416_42421	AGACTG g.49272_49277	MH	[6]
15–16	g.45679_48141del	p.Asp2479GlyfsX45	AluSp g.45569_45890+	AluIb g.48031_48170+	NAHR (NHE)	[23]
15–16	g.45948_51583delinsTAAATACTGAC <sup>g</sup>	p.Asp2479ValfsX7	-	AluIb g.51540_51849-	NHE	[20]
17–18	g.48897_59728del <sup>h</sup>	p.Arg2602ArgfsX43	TTTTGA g.48892_48896	TTTTGA g.59753_59757	MH	[19]
19–20	g.56115_65818dup	p.Ile2778LysfsX12	AluSz g.56061_56357+	AluSz g.65762_66063+	NAHR	[9]
20	g.60200_65152del	p.Trp2830LysfsX12	AluSq2 g.60132_60413+	AluSp g.65081_65379-	HR	[19]
20	g.60522_61677del <sup>i</sup>	p.Trp2830LysfsX12	-	Tigger1 g.61238_61842	NHE (MH)	[4]
21	g.65942_67442del <sup>k</sup>	p.Glu2878GlyfsX4	AluSz g.65762_66063+	-	NHE	[9]
22	g.68883_68884insAlu	-	-	-	-	[24]
22–24	g.68463_70347del	p.Gly2919AspfsX18	MIR g.68360_68570+	AluY g.70056_70360-	NHE	This study

<sup>a</sup> RefSeq NG\_012772.1

<sup>b</sup> RefSeq\_AAB07223.1

<sup>c</sup> Forward (+) or reverse (-) orientated

<sup>d</sup> NAHR non-allelic homologous recombination, NHE non-homologous event, MH microhomology. Brackets indicate an alternative mechanism proposed in this study (See Supplemental Figures)

<sup>e</sup> The precise sequence was not reported

<sup>f</sup> Described previously as g.37848\_57818 [3]

<sup>g</sup> The 11-bp inserted sequence is not present in RefSeq NG\_012772.1

<sup>h</sup> Described previously as g.48887\_59728del [3]

<sup>i</sup> Described previously as g.60519\_61673 [3]

<sup>j</sup> The Tigger1 repeat is not reported by Hansen et al.

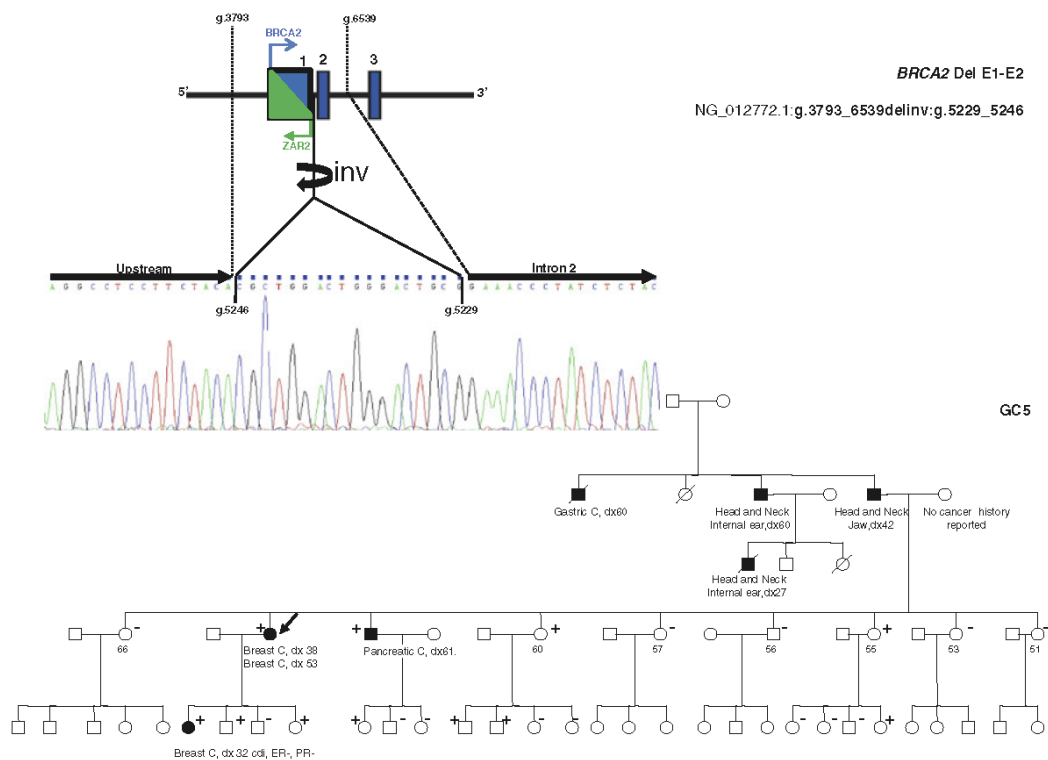
<sup>k</sup> Described previously as g.65941\_67441 [3]

Molecular characterization confirmed that MLPA alterations targeting exons 1–2, exons 12–16, and exons 22–24 represented true LGRs. Prevalence of *BRCA2* LGRs according to clinical criteria is shown in Table S3.

The abnormal MLPA profile suggesting the deletion of *BRCA2* exons 1–2 was first detected in a women diagnosed with bilateral breast cancer at ages 38 and 53 (Fig. 1). MLPA indicated low gene dosage at exon 1 and exon 2 probes. Normal gene dosage was observed at exon 3 and *FRY* probes (data not shown). The later probe is located ~20 kb upstream of *BRCA2*. To characterize the breakpoints, several long-range PCRs were attempted. We used a unique reverse primer located in exon 3 (E3R), but different forward primers randomly located upstream of *BRCA2* exon 1. By using a forward primer located ~15 kb upstream of *BRCA2* exon 1 (–15 kb F), we obtained a PCR product of ~16 kb (expected size, 16,026 bp) in both carriers and control samples. A shorter product of ~13 kb was observed only in carriers, thus suggesting the presence of a ~3-kb deletion (data not shown). Long-range PCR products were digested with *Hind* III, *Bgl* II, and *Xma* I restriction enzymes. Restriction analysis was compatible

with the presence of a ~3-kb deletion encompassing exons 1–2 (data not shown). Based on the later analysis, we designed both a forward primer ~1.3 kb upstream of *BRCA2* exon 1 (–1.3F) and a reverse primer located in intron 2 (I2R). With these primers, we amplified a ~3.5-kb product (expected size, 3,404 bp) in control samples. In carriers, a ~700-bp product was observed, thus confirming the presence of a genomic deletion of ~2.7 kb. The 3.5-kb product was not observed in carriers, probably due to preferential amplification of the mutated (shorter) allele (data not shown). Direct sequencing of the ~700-bp PCR product (Fig. 1) revealed the presence of a 2,748-bp deletion together with a 18-bp insertion into the site of the deletion (Table 1).

A second alteration, an abnormal MLPA profile compatible with a *BRCA2* exon 2 deletion, was detected in four independent families (Fig. 2). We confirm the structural alteration by direct PCR of gDNA in all four families. We first performed a long-range PCR with primers located in exon 1 (E1F) and exon 3 (E3R). A PCR product of ~3.7 kb (expected size, 3,762 bp) was amplified from both carriers and control individuals. A shorter product of



**Fig. 1** Molecular characterization of a genomic rearrangement targeting exons 1 and 2 or *BRCA2*, revealed a 2,747-bp deletion together with a 18-bp insertion at the site of the deletion. The 18-bp sequence is identical to a sequence present in *BRCA2* intron 2, but inserted in the reverse orientation. The deletion targets the

bidirectional *BRCA2/ZAR2* promoter region, including the transcription starting point of both genes. The mutation was identified in a family with several malignancies in men, including rare head and neck tumors. The index case is indicated with an arrow

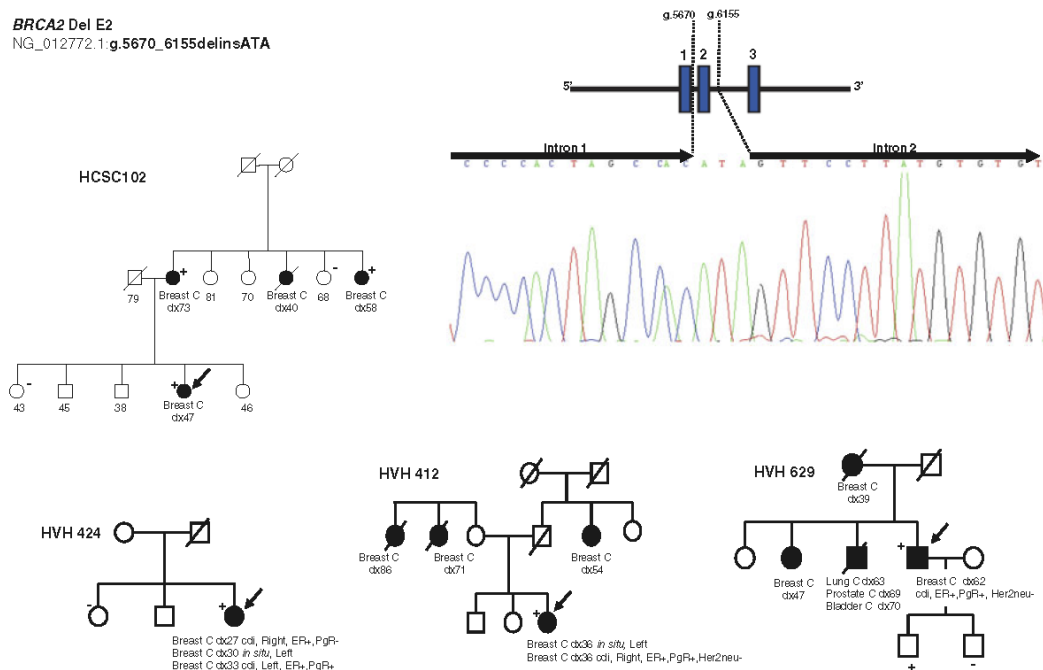
~3.2 Kb was amplified only in MLPA-positive individuals, thus suggesting the presence of a ~500-bp deletion in one allele (Fig. 2). The data were used to design primers located at *BRCA2* intron 1 (I1F) and intron 2 (I2R). By using these primers, we amplified from gDNA a ~900-bp fragment (expected size, 902 bp) in control samples. A shorter fragment of ~400 bp was observed only in MLPA-positive individuals, thus suggesting the presence of a ~500-bp deletion. The ~900-bp fragment was not observed in these samples, indicating preferential amplification of the abnormal (shorter) allele. Direct sequencing of this fragment (Fig. 2) revealed a genomic deletion of 486 bp together with a 3-bp insertion (ATA) at the place of the deletion (Table 1). The alteration was identical in all four families analyzed, thus suggesting a founder effect. Haplotype analysis supported this view (Table S4).

We further characterize the alteration at the cDNA level in one carrier. A PCR with primers in exon 1 (c.1F) and exon 3 (c.3R) produced the expected 380-bp fragment together with an extra fragment of 274 bp with similar intensity (Fig. 2). The data further confirm a heterozygous deletion of *BRCA2* exon 2. At the same time, it reveals that the defective allele is none the less able to produce a stable mRNA.

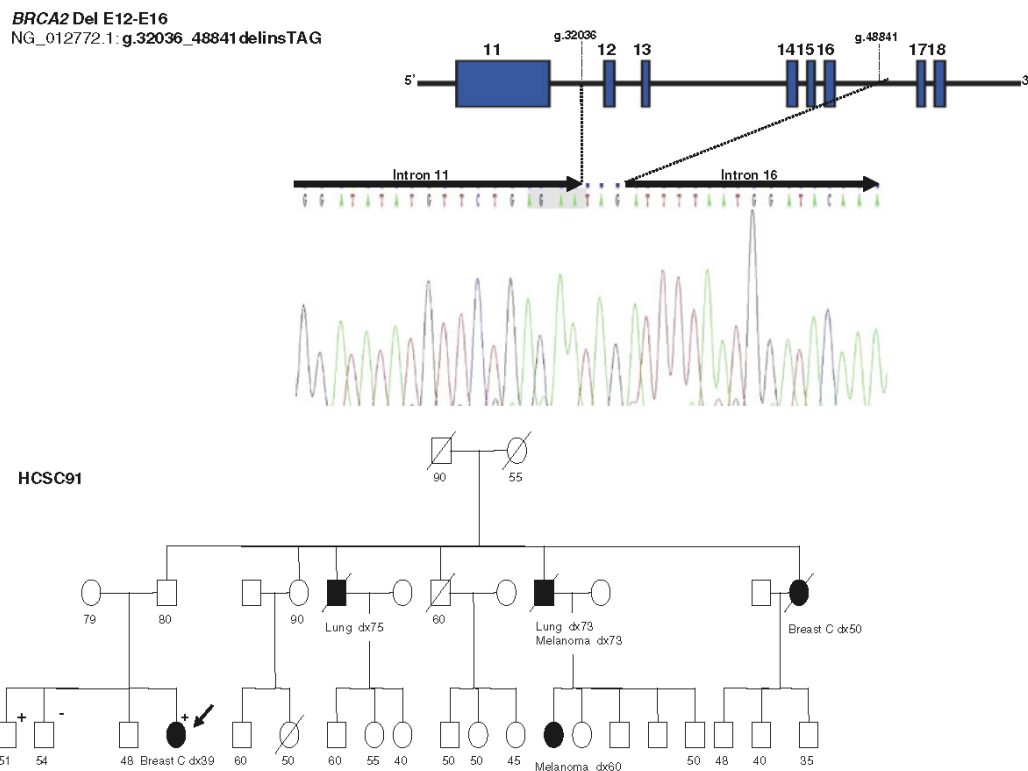
A third alteration, an abnormal MLPA profile suggesting a genomic deletion encompassing *BRCA2* exons 12–16, was observed in a 43-year-old women diagnosed with

breast cancer at age 39 (Fig. 3). We confirmed the suspected structural alteration performing a long-range PCR on gDNA with a forward primer located at the end of exon 11 (e11F) and a reverse primer located at exon 17 (e17R). The expected 21,880-bp product was observed neither in controls nor in positive samples. Most probably, the fragment is too long to be PCR-amplified with our experimental protocol. However, a ~5-kb product was specifically amplified in carriers, thus suggesting the presence of a ~17-kb deletion at the *BRCA2* locus in those samples. Direct sequencing of this PCR product with an internal primer located in intron 11 (i11F) demonstrated the presence of a 16,805-bp deletion of genomic DNA (including the complete sequence of *BRCA2* exons 12–16) plus a 3-bp insertion (TAG) into the site of the deletion (Table 1).

A fourth alteration, an abnormal MLPA profile suggesting a *BRCA2* deletion targeting exons 22–24, was detected in a women diagnosed with breast cancer at age 44 (Fig. 4). We confirmed the presence of a genomic deletion by long-range PCR. By using primers located in exon 21 (e21F) and exon 25 (e25R), we amplified from gDNA a ~18-kb product (expected size, 18,093 bp) only in control samples. A smaller fragment of ~16 kb was observed in carriers, thus suggesting preferential amplification of the abnormal allele containing a ~2-kb deletion in this sample. The data were used to design new primers



**Fig. 2** Molecular characterization of a genomic rearrangement targeting exon 2 of *BRCA2* revealed a 486-bp deletion together with a 3-bp insertion at the site of the deletion. The same alteration was identified in four independent families. Index cases are indicated with an *arrow*



**Fig. 3** Molecular characterization of a genomic rearrangement targeting exons 12–16 of *BRCA2* revealed a 16,790-bp deletion together with a 3-bp insertion into the site of the deletion. The index case is indicated with an arrow

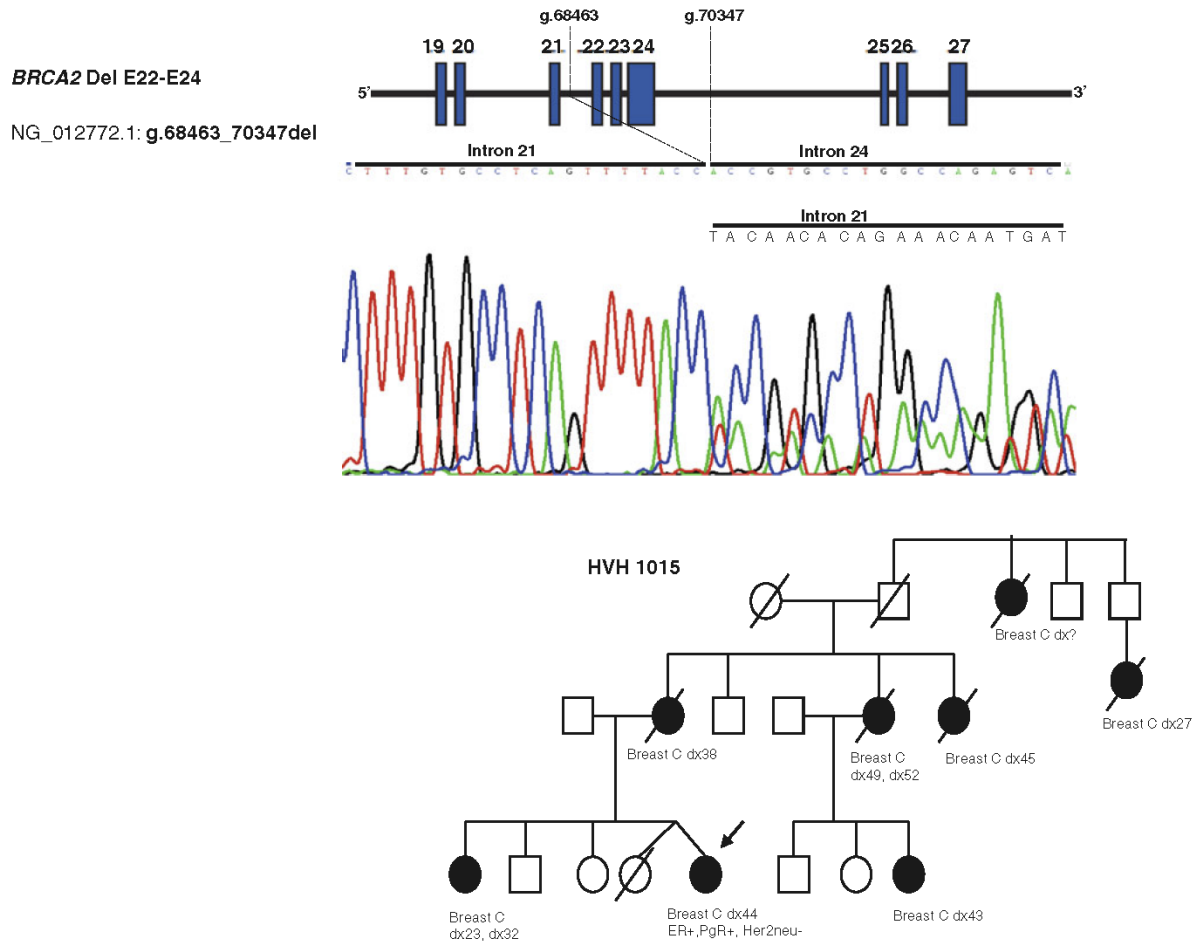
located in intron 21 and intron 24. By using these primers, we amplified from gDNA a ~2.7-kb product in all samples (expected size, 2,696 bp). In carriers, a shorter product of ~800 bp was also amplified. Direct sequencing of these fragments (Fig. 4) confirmed the presence of a genomic deletion of 1,885 bp spanning from intron 21 to intron 24 (Table 1). The alteration was further characterized at the RNA level. By using primers in exon 21 and exon 25, we were able to amplify from control cDNAs a fragment of ~700 bp (expected size, 699 bp). When using cDNA from the deletion carrier, a smaller product of ~200 bp was observed. Direct sequencing revealed the presence of messengers lacking exons 22–24 (data not shown).

Homologous recombination between Alu elements cannot explain any of the four LGRs here characterized (Table 1). Moreover, sequence alignment of the regions surrounding the breakpoints discarded both non-allelic homologous recombination and microhomology-mediated events in all cases (Fig. S2), thus pointing to non-homologous events. The analysis was extended to 19 *BRCA2* LGRs characterized by others, finding evidences that Alu-mediated homologous recombination events might have been overestimated in previous studies (Table 1, Fig. S3).

In particular, we found that this mechanism does not explain well g.4093\_6446delins12 and g.45679\_48141del rearrangements, despite the fact that both 5' and 3' breakpoints are embedded in Alu sequences (Fig. S3). The same analysis revealed that, in addition to previously reported g.42422\_49277del and g.48897\_59728del, microhomology-mediated events might explain four additional LGRs at the *BRCA2* locus (Table 1, Fig. S4).

**Discussion**

In this study, we have identified four novel structural alterations at the *BRCA2* locus, and we have characterized them at the molecular level, thus expanding the spectrum of *BRCA2* LGRs already characterized (Table 1). Interestingly, g.5670\_6155delinsATA appears to be a founder mutation explaining a significant fraction of all the *BRCA2* LGR families identified so far in Spain (Table S5). Overall, we have detected *BRCA2* LGRs in 7 out of 836 HBOC families analyzed, supporting a low proportion of these alterations in our population. Indeed, a low proportion (ranging from 0 to 2.5%) has been reported in most populations analyzed [3, 5–9, 30].



**Fig. 4** Molecular characterization of a genomic rearrangement targeting exons 22–24 of *BRCA2* revealed a 1,885-bp deletion. The index case is indicated with an *arrow*

Some authors have reported male breast cancer patients in most families harboring *BRCA2* LGRs, thus suggesting a strong association with this malignancy [4, 5, 21]. Our data do not support any particular association distinguishing *BRCA2* LGR families from those carrying other mutations (Table S3). Indeed, only one out of the seven families here reported had a male breast cancer patient. Somehow surprisingly, ovarian cancer was not observed in any of these families, but ovarian cancer has been reported previously in several families with *BRCA2* LGRs [17, 23, 31].

Interestingly, Del E1-E2 extends upstream of the *BRCA2* transcriptional start site, targeting the minimal promoter region defined from g.4814 to g.5310, and including the E-box (g.4981\_4985) essential for gene expression [29]. It is therefore expected that this alteration abrogates gene expression. We have not been able to demonstrate the latter, since we did not have access to RNA material. However, it has been reported that another LGR targeting the minimal

promoter (g.4093\_6446delins12) impairs *BRCA2* expression [9], supporting that this is indeed the most probable outcome of g.3793\_6539delinv:g.5229\_5246. Recently, the *BRCA2* promoter has been reported to be bidirectional [32]. The reverse activity regulates *ZAR2* expression. Thus, LGRs targeting the *BRCA2* promoter are predicted to abrogate *ZAR2* expression, as well. Interestingly, this gene codifies for a transcription factor likely regulating many other genes, including *BRCA2* itself [32]. Whether this may be reflected in a distinctive phenotype is not known. The family is remarkable in that several malignancies in men are observed, including pancreatic, gastric, and rare head and neck tumors. While pancreatic cancer is a well-established *BRCA2*-related tumor, the same is not true for gastric and head and neck malignancies [33, 34]. Unfortunately, we could not genotype gastric or head and neck cancer patients, so we cannot support a link between *ZAR2/BRCA2* promoter deletion and these malignancies. Moreover, we have

identified in the literature a minimum of seven HB(O)C families carrying LGRs targeting the bidirectional promoter, including one family carrying a Del E1-E13 alteration [31], 3 families carrying Del E1-E2 alterations [5, 9], and 3 families carrying whole-gene deletions [21, 23, 35]. Head and neck malignancies were not reported in any of these families, thus arguing against any specific genotype/phenotype relationship.

Del E2 is an internal gene deletion targeting the first coding exon. Interestingly, the mutated allele is able to express a stable transcript at similar levels than the wt allele, thus indicating that the deleted region (g.5670\_6155) does not contain promoter elements critical for *BRCA2* expression. Probably, the translation of this transcript is completely abolished due to the lack of the canonical start codon plus 39 bp of 5' UTR. However, we cannot rule out the possibility that translation proceeds, if an alternative start codon becomes available. The first in-frame ATG is located in exon 4 (codon 124). On the assumption that this ATG serves as a translation initiation codon, the translation of the mutant mRNA will produce a *BRCA2* protein lacking the first 123 aa (p.Met1\_Lys123del). This alteration has been detected in four independent families. As far as we know, these families are unrelated. Another Spanish family with a *BRCA2* exon 2 deletion has been reported previously [10], but this alteration was not fully characterized at the molecular level. Our data support that g.5670\_6155delinsATA is indeed a founder mutation explaining a substantial fraction of all *BRCA2* LGR-related Spanish HB(O)C families identified so far (Table S4). Several recurrent *BRCA1* LGRs have been described in the literature [3], but as far as we know, there is only one previous example of a recurrent *BRCA2* LGRs [36].

LGRs may arise through different genetic mechanisms including non-allelic homologous recombination (NAHR), microhomology-mediated events involving very short (2–15 bp) homologous sequences (MH), and poorly characterized mechanisms not involving sequence homology and hence collectively referred as non-homologous events (NHE) [37]. Frequently, NHE result in deletions plus short insertions at the site of the deletion [37]. Three of the four LGRs here reported fall in this category. The 18-bp insertion observed in Del E1-E2 is identical to a *BRCA2* intron 1 sequence, albeit inserted in the reverse orientation, suggesting that this short intron 1 sequence was indeed involved in the genetic event leading to this structural alteration (Fig. 1). The origin of the insertions observed in Del E2 and Del E12-E16 is much uncertain. These insertions are too short (3 bp) to perform meaningful sequence alignments. Yet, in both cases a 3-bp sequence identical to the insertion is located in the deleted sequence close to the breakpoints (Fig. S2), suggesting that these sequences might be involved in the non-homologous events.

NHE and Alu-mediated NAHR together explain most *BRCA2* LGRs described so far, while MH-related events are rare (Table 1). We propose that NAHR has been overestimated, while the opposite may be true for MH events (Table 1). For instance, Del E1-E2 (g.4093\_6446delins12) and Del E15-E16 (g.45679\_48141del) have been considered NAHR events because both breakpoints are located at Alu sequences [3]. However, alignment analysis fails to identify any evidence of homologous recombination (Fig. S3), indicating that location of breakpoints at Alu sequences is not a proxy for NAHR.

**Acknowledgments** We thank the individuals whose participation made this project possible. We thank Paula Diaque, Anna Tenés, Miriam Masas, and Alicia Barroso for technical assistance. This study was supported by Instituto de Salud Carlos III, Fondo de Investigación Sanitaria (FIS) Research Grants 09/00859 and 08/1120 to MdH and JB, respectively. Fundación Canaria de Investigación y Salud (FUNCIS) Research Grant 08/08 to ASdA. Fundación Mutua Madrileña (FMM) Research Grant FMM-08 to MdH. Asociación Española contra el Cáncer (AECC) Research Grant (2009) to JB. Red Temática de Investigación Cooperativa en Cáncer; Instituto de Salud Carlos III, Fondo Europeo de Desarrollo Regional (RETICC 06/0020/0021) supported PG, AR, PPS, EDR, TC, and MdH. Centro de Investigación en Red de Enfermedades Raras (CIBERER) supported JB. SGE is supported by a Miguel Servet contract of the Instituto de Salud Carlos III.

**Conflicts of interest** None.

## References

1. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266:66–71
2. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G (1995) Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* 378:789–792
3. Sluiter MD, van Rensburg EJ (2011) Large genomic rearrangements of the *BRCA1*, *BRCA2* genes: review of the literature, report of a novel *BRCA1* mutation. *Breast Cancer Res Treat* 125:325–349
4. Hansen TO, Jønson L, Albrechtsen A, Andersen MK, Ejler Jensen B, Nielsen FC (2009) Large *BRCA1* and *BRCA2* genomic rearrangements in Danish high risk breast-ovarian cancer families. *Breast Cancer Res Treat* 115:315–323
5. Woodward AM, Davis TA, Silva AG, Kirk JA, Leary JA, kConFab Investigators (2005) Large genomic rearrangements of both *BRCA2* and *BRCA1* are a feature of the inherited breast/ovarian cancer phenotype in selected families. *J Med Genet* 42:e31
6. Kang P, Mariapun S, Phuah SY, Lim LS, Liu J, Yoon SY, Thong MK, Mohd Taib NA, Yip CH, Teo SH (2010) Large *BRCA1* and *BRCA2* genomic rearrangements in Malaysian high risk breast-ovarian cancer families. *Breast Cancer Res Treat* 124:579–584
7. Buffone A, Capalbo C, Ricevuto E, Sidoni T, Ottini L, Falchetti M, Cortesi E, Marchetti P, Scambia G, Tomao S, Rinaldi C, Zani

- M, Ferraro S, Frati L, Screpanti I, Gulino A, Giannini G (2007) Prevalence of BRCA1 and BRCA2 genomic rearrangements in a cohort of consecutive Italian breast and/or ovarian cancer families. *Breast Cancer Res Treat* 106:289–296
8. Ticha I, Kleibl Z, Stribrna J, Kotlas J, Zimovjanova M, Mateju M, Zikan M, Pohlreich P (2010) Screening for genomic rearrangements in BRCA1 and BRCA2 genes in Czech high-risk breast/ovarian cancer patients: high proportion of population specific alterations in BRCA1 gene. *Breast Cancer Res Treat* 124:337–347
  9. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, Roach KC, Mandell J, Lee MK, Ciermnikova S, Foretova L, Soucek P, King MC (2006) Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 295:1379–1388
  10. Gutiérrez-Enríquez S, de la Hoya M, Martínez-Bouzas C, Sanchez de Abajo A, Ramón y Cajal T, Llorc G, Blanco I, Beristain E, Díaz-Rubio E, Alonso C, Tejada ML, Caldés T, Díez O (2007) Screening for large rearrangements of the BRCA2 gene in Spanish families with breast/ovarian cancer. *Breast Cancer Res Treat* 103:103–107
  11. de la Hoya M, Gutiérrez-Enríquez S, Velasco E, Osorio A, Sanchez de Abajo A, Vega A, Salazar R, Esteban E, Llorc G, Gonzalez-Sarmiento R, Carracedo A, Benítez J, Miner C, Díez O, Díaz-Rubio E, Caldes T (2006) Genomic rearrangements at the BRCA1 locus in Spanish families with breast/ovarian cancer. *Clin Chem* 52:1480–1485
  12. Mazoyer S (2005) Genomic rearrangements in the BRCA1 and BRCA2 genes. *Hum Mutat* 25:415–422
  13. Ewald IP, Ribeiro PL, Palmero EI, Cossio SL, Giugliani R, Ashton-Prolla P (2009) Genomic rearrangements in BRCA1 and BRCA2: a literature review. *Genet Mol Biol* 32:437–446
  14. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 30:e57
  15. Muller D, Rouleau E, Schultz I, Caputo S, Lefol C, Bieche I, Caron O, Nogues C, Limacher JM, Demange L, Lidereau R, Fricker JP, Abecassis J (2011) An entire exon 3 germ-line rearrangement in the BRCA2 gene: pathogenic relevance of exon 3 deletion in breast cancer predisposition. *BMC Med Genet* 12:121
  16. Teugels E, De Brakeleer S, Goelen G, Lissens W, Sermijn E, De Grève J (2005) De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum Mutat* 26:284
  17. Nordling M, Karlsson P, Wahlström J, Engwall Y, Wallgren A, Martinsson T (1998) A large deletion disrupts the exon 3 transcription activation domain of the BRCA2 gene in a breast/ovarian cancer family. *Cancer Res* 58:1372–1375
  18. Lim YK, Lau PT, Ali AB, Lee SC, Wong JE, Putti TC, Sng JH (2007) Identification of novel BRCA large genomic rearrangements in Singapore Asian breast and ovarian patients with cancer. *Clin Genet* 71:331–334
  19. Agata S, Dalla Palma M, Callegaro M, Scaini MC, Menin C, Ghiotto C, Nicoletto O, Zavagno G, Chieco-Bianchi L, D'Andrea E, Montagna M (2005) Large genomic deletions inactivate the BRCA2 gene in breast cancer families. *J Med Genet* 42:e64
  20. Casilli F, Tournier I, Sinilnikova OM, Coulet F, Soubrier F, Houdayer C, Hardouin A, Berthet P, Sobol H, Bourdon V, Muller D, Fricker JP, Capoulade-Metay C, Chompret A, Nogues C, Mazoyer S, Chappuis P, Maillet P, Philippe C, Lortholary A, Gesta P, Bézieau S, Toulas C, Gladieff L, Maugard CM, Provencher DM, Dugast C, Delvincourt C, Nguyen TD, Faivre L, Bonadona V, Frébourg T, Lidereau R, Stoppa-Lyonnet D, Tosi M (2006) The contribution of germline rearrangements to the spectrum of BRCA2 mutations. *J Med Genet* 43:e49
  21. Tournier I, Paillerets BB, Sobol H, Stoppa-Lyonnet D, Lidereau R, Barrois M, Mazoyer S, Coulet F, Hardouin A, Chompret A, Lortholary A, Chappuis P, Bourdon V, Bonadona V, Maugard C, Gilbert B, Nogues C, Frébourg T, Tosi M (2004) Significant contribution of germline BRCA2 rearrangements in male breast cancer families. *Cancer Res* 64:8143–8147
  22. Wang T, Lerer I, Gueta Z, Sagi M, Kadouri L, Peretz T, Abeliovich D (2001) A deletion/insertion mutation in the BRCA2 gene in a breast cancer family: a possible role of the Alu-polyA tail in the evolution of the deletion. *Genes Chromosom Cancer* 31:91–95
  23. del Valle J, Feliubadaló L, Nadal M, Teulé A, Miró R, Cuesta R, Tornero E, Menéndez M, Darder E, Brunet J, Capellà G, Blanco I, Lázaro C (2010) Identification and comprehensive characterization of large genomic rearrangements in the BRCA1 and BRCA2 genes. *Breast Cancer Res Treat* 122:733–743
  24. Miki Y, Katagiri T, Kasumi F, Yoshimoto T, Nakamura Y (1996) Mutation analysis in the BRCA2 gene in primary breast cancers. *Nat Genet* 13:245–247
  25. Díez O, Osorio A, Durán M, Martínez-Ferrandis JI, de la Hoya M, Salazar R, Vega A, Campos B, Rodríguez-López R, Velasco E, Chaves J, Díaz-Rubio E, Jesús Cruz J, Torres M, Esteban E, Cervantes A, Alonso C, San Román JM, González-Sarmiento R, Miner C, Carracedo A, Eugenia Armengod M, Caldés T, Benítez J, Baiget M (2003) Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: a high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat* 22:301–312
  26. Iversen ES Jr, Katki HA, Chen S, Berry DA, Parmigiani G (2007) Limited family structure and breast cancer risk. *JAMA* 298:2007
  27. den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 15:7–12
  28. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948
  29. Davis PL, Miron A, Andersen LM, Iglehart JD, Marks JR (1999) Isolation and initial characterization of the BRCA2 promoter. *Oncogene* 18:6000–6012
  30. Preisler-Adams S, Schönbuchner I, Fiebig B, Welling B, Dworniczak B, Weber BH (2006) Gross rearrangements in BRCA1 but not BRCA2 play a notable role in predisposition to breast and ovarian cancer in high-risk families of German origin. *Cancer Genet Cytogenet* 168:44–49
  31. Palma MD, Domchek SM, Stopfer J, Erlichman J, Siegfried JD, Tigges-Cardwell J, Mason BA, Rebbeck TR, Nathanson KL (2008) The relative contribution of point mutations and genomic rearrangements in BRCA1 and BRCA2 in high-risk breast cancer families. *Cancer Res* 68:7006–7014
  32. Misra S, Sharma S, Agarwal A, Khedkar SV, Tripathi MK, Mittal MK, Chaudhuri G (2010) Cell cycle-dependent regulation of the bi-directional overlapping promoter of human BRCA2/ZAR2 genes in breast cancer cells. *Mol Cancer* 9:50
  33. van Asperen CJ, Brohet RM, Meijers-Heijboer EJ, Hoogerbrugge N, Verhoef S, Vasen HF, Aoustens MG, Menko FH, Gomez Garcia EB, Klijn JG, Hogervorst FB, van Houtwelingen JC, van't Veer LJ, Rookus MA, van Leeuwen FE, Netherlands Collaborative Group on Hereditary Breast Cancer (HEBON) (2005) Cancer risks in BRCA2 families: estimates for sites other than breast and ovary. *J Med Genet* 42:711–719
  34. Levy-Lahad E, Friedman E (2007) Cancer risks among BRCA1 and BRCA2 mutation carriers. *Br J Cancer* 96:11–15
  35. Rodríguez M, Torres A, Borràs J, Salvat M, Gumà J (2010) Large genomic rearrangements in mutation-negative BRCA families: a population-based study. *Clin Genet* 78:405–407

36. Peixoto A, Santos C, Rocha P, Pinheiro M, Príncipe S, Pereira D, Rodrigues H, Castro F, Abreu J, Gusmão L, Amorim A, Teixeira MR (2009) The c.156\_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. *Breast Cancer Res Treat* 114:31–38
37. Hastings PJ, Lupski JR, Rosenberg SM, Ira G (2009) Mechanisms of change in gene copy number. *Nat Rev Genet* 10:551–564

Supplementary Figures

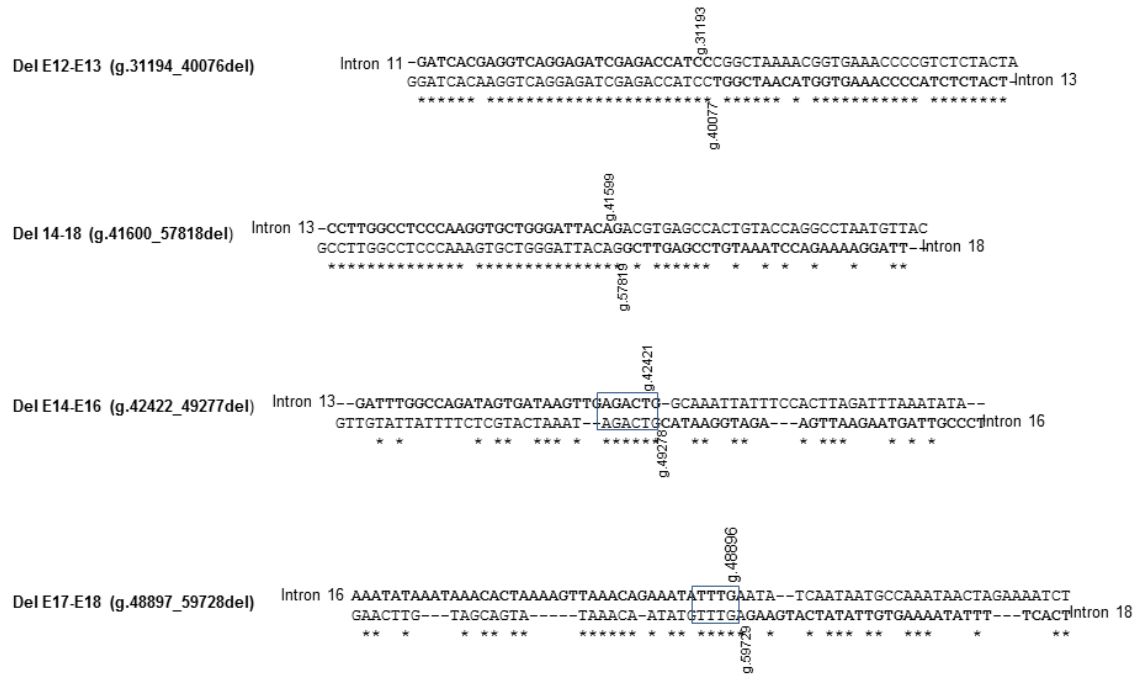


Fig S 1 Alignment analysis

Representative examples of BRCA2 LGR explained by Non Allelic-Homologous Recombination (top) or microhomology (bottom) events. Boxes indicate microhomology regions. The sequence of the mutated alleles is indicated in bold

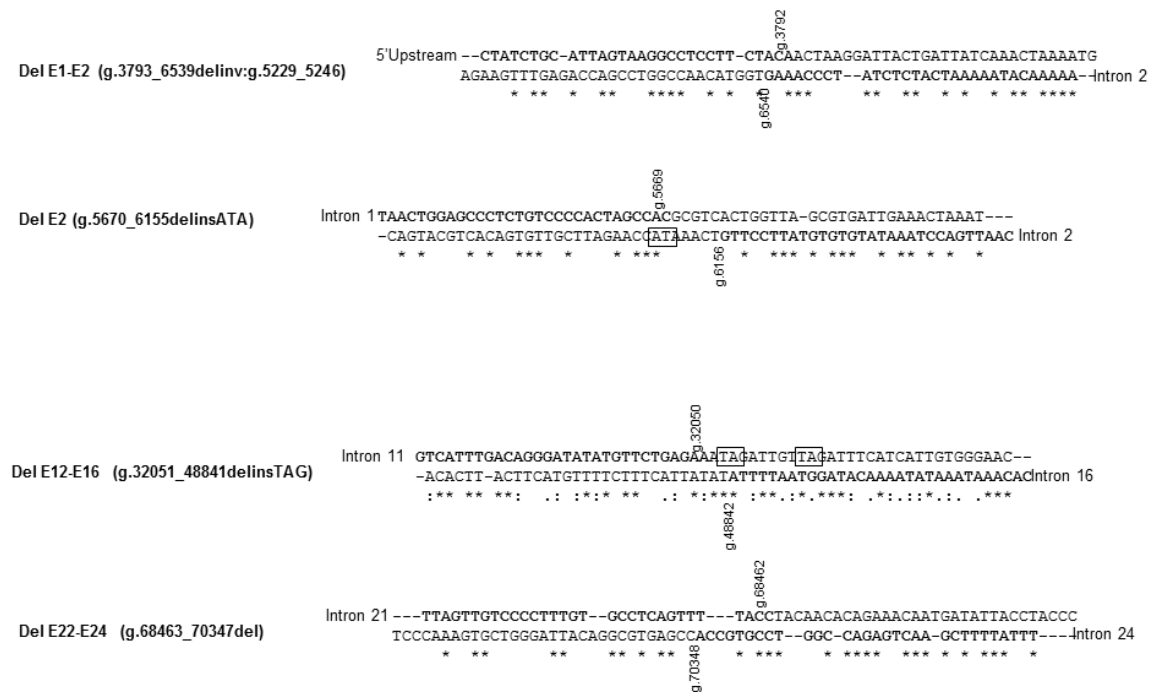


Fig S 2 Alignment analysis

None of the four BRCA2 LGR characterized in the present study can be explained by nonn allelic homologous recombination or micro-homology events. Boxes indicate nucleotides coincident with insertions observed at the breakpoints. The role of these particular nucleotides (if any) in the genomic event leading to the rearrangement is unknown. The sequence of the mutated alleles is indicated in bold.

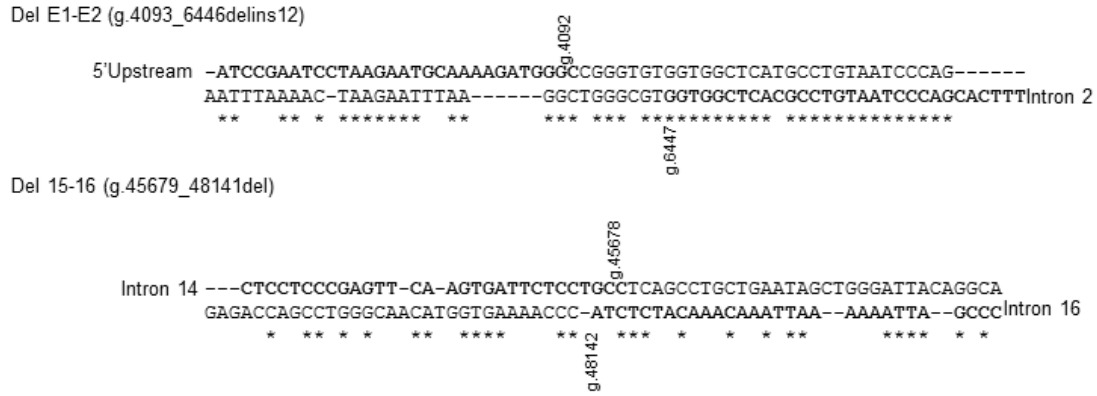


Fig S 3 Aligment analysis.

Non allelic homologous recombination is not supported by alignment analysis, despite the fact that breakpoints are located at Alu sequences in both cases. The sequence of the mutated alleles is indicated in bold. In the case of Del E1-E2 (g.4093\_6446delins12), the 12-bp insertion was indeed supporting a non homologous event.

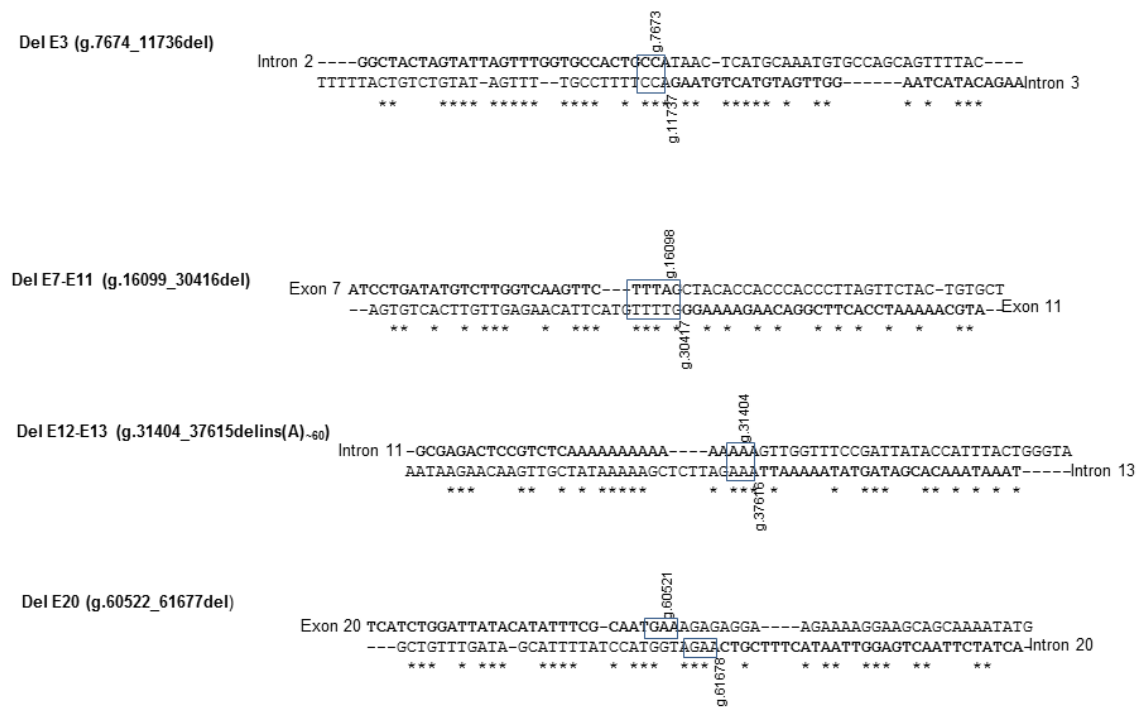


Fig S 4 Aligment analysis

Four BRCA2 LGR likely explained by micro-homology. The evidence is weaker in the case of DelE7-E11. However, a hypothetical variant present in the ancestral allele (g.30415T>A), will make the micro-homology event more plausible. The sequence of the mutated alleles is indicated in bold.

Table S 1: PCR primers used to characterize BRCA2 LGRs.

LGR	Location	Name	Nucleotide sequence	Nucleotide position <sup>A</sup>
BRCA2 del 1-2	-15 kb	-15 kb F	gcaccctatttccttcatgg	g.13692_13711 <sup>B</sup>
	-1.3 kb	-1.3 kb F	TTTGGTAAGTGGCAGCTGTT	g.3682_3701
	Intron 2	I2R	GATGTTGCGACATGGGAAAG	g.7066_7085
	Exon 3	E3R	tcagcccttgctcttgaat	g.8770_8789
BRCA2 del 2	Exon 1	E1F	AGCTTACTCCGGCCAAAAA	g.5149_5167
	Intron 1	I1F	TAAGCCCCTGGTAATTGCTG	g.5574_5593
	Intron 2	I2R	GGCTGGTCTCAAACCTTCTGG	g.6507_6526
	Exon3	E3R	AGTCAGCCCTTGCTCTTTGA	g.8772_8791
BRCA2 del12-16	Exon 11	E11F	CAGGCTTCACCTAAAAACG	g.30427_30444
	Intron 11	I11F	GGCAGAGAGCATTTCAGTAAGTGTT	g.31850_31873
	Intron 16	I16R	AGGAGGGTGGATCCTGAGG	g.50111_50129
	Exon 17	E17R	TGTGCAACATTTTGACATGG	g.52287-52306
BRCA2 del 22-24	Exon 21	E21F	AGTTCGTGCTTTGCAAGATG	g.66240_66259
	Intron 21	I21F	TGAGGAGAACGTAGGCATCC	g.67908_67927
	Intron 24	I24R	CATGGGAACAATGTGGCTTA	g.70585_70604
	Exon 25	E25R	CACTGGAGGTTGCTTGCAG	g.84315_84333

<sup>A</sup> RefSeq NG\_012772.1 <sup>B</sup> RefSeq NG\_017006

Table S 2: Sequence Variants interfering with MLPA analysis

MLPA probe ID <sup>A</sup>	Location	probe sequence <sup>B</sup>	Variant <sup>C</sup>	dbSNP130
02283-L12281 (right)	Exon1	- <b>G</b> TGGCGCGAGCTTCTGAACTAGGCGGCAGAGGCGGAG	c.-227G> <b>C</b>	-
01599-L10642 (left)	Exon3	GTCAGCTGGCTT <b>C</b> AACTCCAATAATATTCAA-	c.227C> <b>G</b> (p.Ser76X)	rs80358498
08898-L09587 (left)	Intron3	GTACATTCATTAATTGTGTCATGCTGGGCAAAT <b>C</b> A-	c.316+108A> <b>G</b>	rs115376548

<sup>A</sup> MLPA Salsa Kits P045/P090. <sup>B</sup> The mismatch is indicated by bold. <sup>C</sup> RefSeq NM\_000059.3

Table S 3: Incidence of BRCA2 LGRs according to clinical criteria

Clinical Criteria	Families	LGRs (%)
≥2 BC/OC	137	0
≥2 BC	499	5 (1%)
≥1 MBC	39	1(2.6%)
1BC/1OC	138	1 (0.7%)
Combined	813	7 (0.9%)

BC breast cancer, OC ovarian cancer, MBC Male Breast Cancer

Table S 4: Haplotype analysis in four BRCA2 DeIE2 (g.5670\_6155delinsATA) families

	Genetic Markers <sup>A</sup>							
	D13S260	BRCA2 LGRs	rs144848	rs543304	rs9534262	D13S1493	D13S267	D13S153
	-0.45Mb	0	+22kb	+22kb	+46kb	+1Mb	+1.29Mb	+14.8Mb
Family ID								
HVH424	<b>161/169</b>	g.5670_6155delinsATA	C/A	C/T	T/T	<b>223/231</b>	213/ <b>218</b>	<b>223/231</b>
HVH412	<b>161/161</b>	g.5670_6155delinsATA	C/A	C/T	T/T	<b>223/235</b>	213/ <b>218</b>	<b>223/235</b>
HVH629	<b>161/167</b>	g.5670_6155delinsATA	C/C	C/C	T/T	<b>223/235</b>	215/ <b>218</b>	<b>223/235</b>
HVH629	<b>161/161</b>	g.5670_6155delinsATA	C/C	C/C	T/C	<b>223/235</b>	213/ <b>218</b>	<b>223/235</b>
HCSC102	<b>161/161</b>	g.5670_6155delinsATA	C/C	C/C	T/T	<b>223/227</b>	213/ <b>218</b>	<b>223/227</b>

<sup>A</sup> Physical Distance according to Ensembl Human (GRCh37)

Table S 5: BRCA2 LGRs detected in Spanish HB(O)C families.

Exons	N <sup>A</sup>	HGVS nomenclature <sup>B</sup>	Reference
Del E1-E24	1	?	[23]
Del E1-E24	1	?	[35]
Del E1-E2	1	g.3793_6539delinv:g.5229_5246	This study
Del E2	1	?	[10]
Del E2	4	g.5670_6155delinsATA	This study
Del E10-E12	1	?	[10]
Del E12-E16	1	g.32036_48841delinsTAG	This study
Del E15-E16	1	?	[10]
Del E15-E16	1	g.45679_48141del	[23]
Dup20	1	?	[10]
Dup20	1	?	[10]

<sup>A</sup> Unrelated Families <sup>B</sup> RefSeq NG\_012772.1

*Análisis de la mutación c.6937+594T>G en el gen BRCA2 en una cohorte de 302 familias HBOC españolas negativas para mutación en BRCA1/2*

El resultado del test genético para un 80% de las familias con múltiples casos de cáncer de mama y ovario que acuden a las unidades clínicas de cáncer familiar es la ausencia de detección de mutación en los genes *BRCA1/2*. Una proporción de estos casos podría estar explicada por mutaciones en genes adicionales, como demuestra la identificación de *BRIP1*, *PALB2*, *RAD51C*, y *RAD51D* como nuevos genes de susceptibilidad. No obstante, el desarrollo de nuevas técnicas de cribado, como por ejemplo el análisis de grandes reordenamientos genómicos, ha demostrado que un porcentaje pequeño aunque significativo de estos casos se explican por mutaciones en *BRCA1/2* no detectables mediante técnicas convencionales de screening. Recientemente se ha descrito por primera vez la presencia de una variante deletérea intrónica localizada alejada de los exones (c.6937+594T>G) que altera el splicing normal de *BRCA2* y explica nueve familias HBOC francesas. Este tipo de mutaciones no están incluidas en los métodos de cribado rutinarios que se utilizan en la mayoría de los laboratorios clínicos, y por ello se deberían introducir nuevas técnicas para su detección. Sin embargo, en un primer lugar es relevante establecer la incidencia de mutaciones intrónicas profundas en los genes *BRCA1/2*. En este trabajo se analizó la presencia de c.6937+594T>G en 302 casos índices con historia personal y familiar de cáncer de mama y/o ovario en los que previamente no se habían detectado mutaciones en *BRCA1/2*.

***BRCA2* c.6937+594T>G screening in a cohort of 302 *BRCA1/2* mutation negative Spanish HBOC families.**

Gorka Ruiz de Garibay<sup>1</sup>, Pedro Pérez-Segura<sup>2</sup>, Eduardo Díaz-Rubio<sup>1,2</sup>, Trinidad Caldés<sup>1</sup>, Miguel de la Hoya<sup>1</sup>.

<sup>1</sup>Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain. <sup>2</sup>Servicio de Oncología Médica, Hospital Clínico San Carlos. Madrid. Spain.

**Abstract**

*BRCA1/2* germ-line mutations are not detected in approximately 80 % of the multiple-case breast and/or ovarian cancer families attending familial cancer clinics. A proportion of these cases can be explained by the presence of mutations in additional HBOC genes such as *BRIP1*, *PALB2*, *RAD51C*, and *RAD51D*. The improvement of the BRCA genetic screening techniques, i.e. screening of large genomic rearrangements, confirms that another proportion accounts for false negative BRCA tests. Recently, the first *BRCA2* deep intronic mutation (c.6937+594T>G) has been described in nine French families. *BRCA1/2* deep intronic mutations are not screened systematically in clinical laboratories. Nevertheless, the incidence of *BRCA1/2* deep intronic mutations has not been established yet. We analyse the presence of c.6937+594T>G in 302 index cases with personal and family history of breast and/or ovarian cancer that tested negative for *BRCA1/2* mutations.

## **Introduction**

5% of the breast and ovary cancer cases occur in the context of a hereditary cancer susceptibility syndrome. *BRCA1* and *BRCA2* genes were discovered as the main susceptibility genes thanks to the use of linkage studies [1, 2]. The discovery of these genes allowed the application of genetic counseling to familial breast and ovary cancer. Nevertheless, only 44% of patients predicted by BRCAPRO algorithm to carry a *BRCA1* or *BRCA2* mutation test positive for recognized mutations in either gene [3]. In addition, in some of the familial breast cancer families that map to the *BRCA1/2* through linkage analysis sequencing of genomic DNA was not able to identify any mutation [5, 6]. The latter suggests that some mutations might be systematically overlooked. Indeed, the application of screening techniques for large genomic rearrangement –which are not detected by the PCR-based techniques–, is able to detect mutations that, depending on the population, contribute substantially to the burden of mutations in these genes (up to 30% in the most extreme cases) [7].

Diagnostic screening of the *BRCA1/2* genes in breast cancer families is mostly performed on genomic DNA. Mutation screening techniques target exons and flanking regions, while most of the intronic sequences remain unanalyzed. Many functional elements lay within introns, such as the canonical branchpoint and splicing enhancers and silencers. In human genetics, deep intronic mutations generally appear to comprise less than 1% of known splicing mutations, but this figure is very likely to be underestimated owing to detection bias [8]. Growing evidences support the existence of such mutations in cancer related genes such as *CDKN2A* [9], *RB1* [10], *ATM* [11], *PTEN* [12] and *MSH2* [13]. The *NF1* gene was methodically screened for mutations that altered splicing. 44% of mutations detected affected the correct splicing, and approximately 5% of the identified mutations that altered splicing were deep intronic mutations [14].

The first *BRCA2* deep intronic mutation was recently identified as part of a global screening of *BRCA2* transcripts in families with a very strong family history and no mutation identified in the coding sequences or exon–intron boundaries [15]. c.6937+594T>G was found to activate the splicing of a cryptic exon introducing a premature stop codon. This mutation was present in 9 of 2000 French families. In order to evaluate the presence of such deep intronic mutations in our Spanish cohort, we screened the *BRCA1/2* negative individuals for c.6937+594T>G mutation.

## **Methods**

### *Study population and DNA samples*

1537 unrelated individuals attending the familial cancer clinic at Hospital Clinico San Carlos, Madrid, Spain from 1998 to 2013 were referred for *BRCA1* and *BRCA2* genetic testing due to a personal and family history of breast and/or ovarian cancer compatible with a hereditary syndrome.

After comprehensive *BRCA1* and *BRCA2* genetic testing (full coding sequence, intron/exon boundaries, and genomic rearrangements), 1041 individuals were considered negative for the presence of germ-line mutations. A subset of these individuals, with the highest risk, were selected for the analysis of the presence of *BRCA2* c.6937+594T>G deep intronic mutation. In total 302 index cases with personal and family history of breast and/or ovarian cancer were selected for this study. The inclusion criteria were as follow:

- 216 index cases from families with three or more cases of female only breast cancer
- 39 index cases from families with three or more cases of female only breast cancer and/or ovary cancer
- 47 index cases from families with presence of male breast cancer regardless the number of affected individuals.

#### *BRCA2 c.6937+594T>G screening*

*BRCA2* intron12 was analyzed as published previously [15]. Briefly, genomic DNA was amplified with forward primer located at c.6937+495-c.6937+522 position and reverse primer located at c.6937+683-c.6937+710 position and analyzed by HRM. Samples showing an aberrant profile were cycle sequenced using the BigDye Terminator v3.1 Cycle Sequencing kit and the ABI 3130 (Applied Biosystems, Foster City, CA, USA) capillary electrophoresis system.

#### *RT-PCR screening of additional pseudoexons transcribed from intron 12*

Twelve index cases negative for *BRCA1/2* mutations were screened for additional pseudoexons transcribed from intron 12. RNA was obtained from blood leukocytes with MagNA Pure LC Total Nucleic Acid Extraction kit and a MagNA Pure LC instrument (Roche Diagnostics, Penzberg, Germany). RT-PCR design was as follow: 300 ng of RNA were reverse transcribed into cDNA using Takara Kit following instructions. 0.3 µl of cDNA were PCR amplified in a final volume of 25. PCR protocol consisted in: initial denaturation step at 95° for 5 min, 40 cycles (95° 30 sec, 58° 30 sec, 72° 3 minutes) and final elongation at 72° for 15 min. Forward primer located at exon 11 (CCAAGTCATGCCACACATTC) and reverse primer located exon 14 (ACGGTTCAACTCTGTGAAAATGTG) were used. Long elongation times were used to allow the amplification of putatively long pseudoexons.

## **Results**

Four cases showed an aberrant melting pattern, an example is shown in fig. 1a. Sequencing of these samples discarded the presence of c.6937+594T>G. The variant of unknown clinical significance, c.6937+565T>C, was found instead in all these cases (fig 1 B). The implication of

c.6937+565T>C in aberrant splicing is very unlikely since splicing prediction algorithms do not predict any effect at c.6937+565 position (table 1). Unfortunately RNA from carrier individuals was not available to test this possibility.

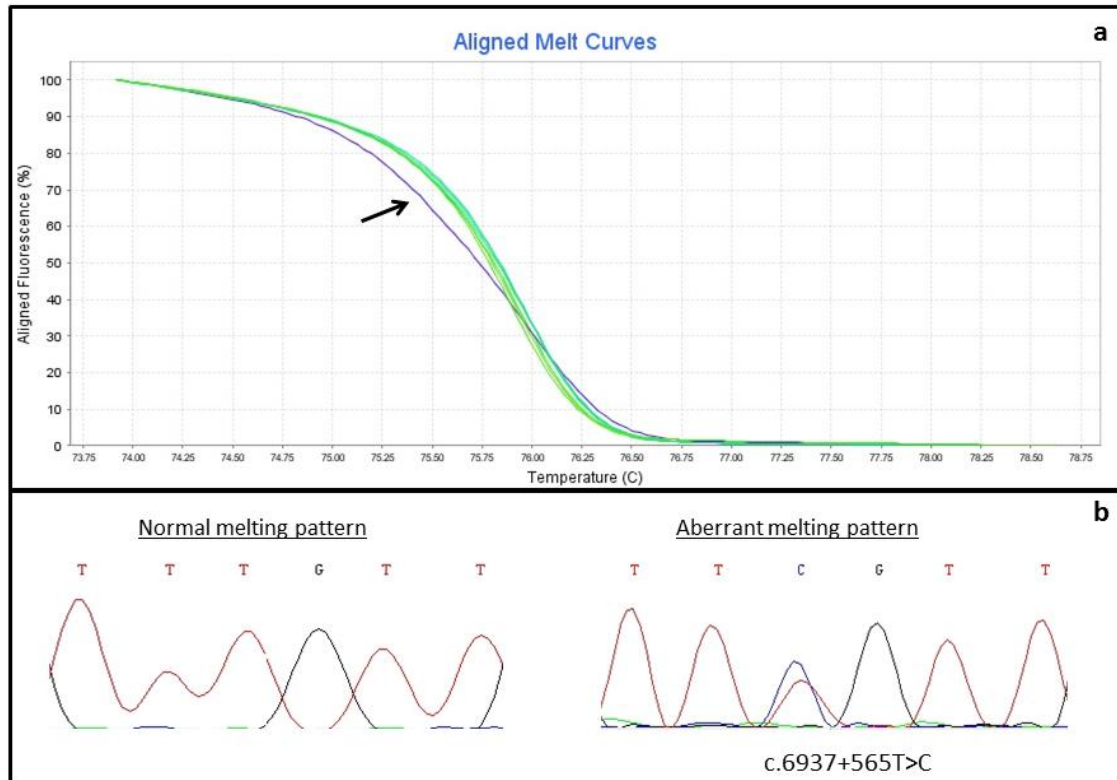


Figure. 1 Identification of c.6937+565 variant

The discovery of a pseudoexon transcribed from *BRCA2* intron 12 in the French study confirms the existence of a particularly strong cryptic acceptor site at position c.6937+495. This site might be coupled not only with c.6937+589 donor site but with additional new donor sites created by deep intronic mutations. The description of c.6937+565T>C in our population pointed at the possibility of the existence of additional genetic variability along the 2196 nucleotides of intron 12. To analyze this variability, all genetic variants from intron 12 located downstream of c.6937+495 were retrieved from ensemble database. In total 20 additional variants were identified at these positions. A prediction of the effect on splicing is shown in table 1. Four variants c.6938-948A>G, c.6938-729A>G, c.6938-561\_6938-560insA and c.6938-467T>G are predicted to increase the strength of cryptic donor sites. The sizes of the putative pseudoexons produced by the theoretical usage of this newly created donor sites range from 731 to 1212 nucleotides. Twelve index cases included in this study were screened for the presence of these putative pseudoexons using an RT-PCR based strategy. None of the samples analyzed presented any splicing alteration (data not shown).

Tab. 1 Genetic variability of BRCA2 intron 13 according to Ensemble database.

Prediction of putative splicing alterations was performed on all BRCA2 intron 13 genetic variants downstream of the cryptic c.6937+495acceptor site

ID	Position	HSF score		MaxEnt Score		size of putative pseudoexon	Hz in controls (#)*
		original	mutant	original	mutant		
rs182608844	c.6937+537G>A	-	-	-	-		no
rs11571675	c.6937+565T>C	69.04	68.89	-	-		no
rs191253965	c.6937+594T>G	77.92	90.24	7.64	9.79	95	yes (1)
rs4942443	c.6937+736A>T	-	-	-	-		yes (53)
rs142165468	c.6937+789T>G	-	-	-	-		no
rs151074481	c.6937+810A>G	-	-	-	-		no
rs9567576	c.6937+847T>G	-	-	-	-		yes(105)
rs11571676	c.6937+978A>T	-	-	-	-		no
rs183216047	c.6937+1041T>C	-	-	-	-		no
rs206078	c.6938-948A>G	59.34	86.17	-0.73	7.44	731	yes (2)
rs115057418	c.6938-843C>T	-	-	-	-		no
rs188491685	c.6938-754T>C	-	-	-	-		no
rs202156906	c.6938-729A>G	39.51	66.35	-	-	950	-
rs199876527	c.6938-643_6938-642insT	-	-	-	-		yes (1)
rs74834625	c.6938-624G>T	-	-	-	-		yes (7)
rs35497963	c.6938-561_6938-560insA	62.75	69.07	-	-	1117	-
rs57263247	c.6938-515G>C	-	-	-	-		no
rs79240676	c.6938-467T>G	53.14	82.08	-	-	1212*	-
rs206079	c.6938-346G>A	-	-	-	-		yes (320)
rs192145481	c.6938-336T>C	-	-	-	-		no
rs185087806	c.6938-313C>T	-	-	-	-		no
rs187941100	c.6938-261C>A	-	-	-	-		no

\*Controls from hapmap and 1000 genomes \*\* Would not alter the reading frame.

## Discussion

Historical reasons suggest a shared genetic background between Spanish and French populations, the profile of BRCA1/2 recurrent mutations in France [16] and Spain [17] supports this

shared background. Of the 16 most frequent French and the 9 most frequent Spanish mutations four are shared (*BRCA1*: c.68\_69delAG and c.5266dupT both of Jewish origin and *BRCA2*: c.2808\_2811del4 and c.9026\_9030del5). In addition, two newly discovered Spanish founder mutations, c.8954-5A>G and c.7617+1G>A (Ruiz de Garibay et al. Human Mutation) are present in seven French families. While c.6937+594T>G mutation was present in 9 of 2000 French families, was not detected in 302 index cases negative for the presence of germ-line *BRCA1/2* mutations. These differences might be explained by the smaller size of the Spanish population included in this study or by differential distribution of c.6937+594T>G mutation within different populations. Since c.6937+594T>G has only been detected in one population and at relatively high frequency, a specific French founder effect cannot be discarded.

Regardless of the presence of this specific mutation in the Spanish population, *BRCA2* deep intronic mutations still undiscovered might be present in the Spanish and in other populations.. Certainly, the screening of deep intronic regions (excluded from many current protocols), will led to the discovery of variants of uncertain significance, as is the case of c.6937+565T>C, found in four index cases in our population.

The availability of RNA will be the limiting factor to perform both functional studies for this newly discovered variants and RT-PCR based screening strategies. Indeed, our retrospective collection of RNA allowed us to collect only twelve samples out of 302 index cases included in this study. The advantages of *BRCA1/2* DNA/RNA based screening strategies should be considered. For other cancer susceptibility genes such as NF1, which has been methodically screened for mutations that alter splicing, 44% of the mutations detected affected the correct splicing, and approximately 5% of the identified mutations that altered splicing were deep intronic mutations [14]. Considering the fact that previous modifications of screening algorithms lead to an increase in the *BRCA1/2* mutations identified, as was the introduction of LGR screening[7, 18], it should be analyzed wether the implementation of DNA/RNA based *BRCA1/2* screening strategies would be cost-effective.

## References

1. Miki, Y., et al., A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 1994. **266**(5182): p. 66-71.
2. Wooster, R., et al., Identification of the breast cancer susceptibility gene BRCA2. *Nature*, 1995. **378**(6559): p. 789-92.
3. Nanda, R., et al., Genetic testing in an ethnically diverse cohort of high-risk women: a comparative analysis of BRCA1 and BRCA2 mutations in American families of European and African ancestry. *JAMA*, 2005. **294**(15): p. 1925-33.
4. Tarabeux, J., et al., Streamlined ion torrent PGM-based diagnostics: BRCA1 and BRCA2 genes as a model. *Eur J Hum Genet*, 2013.
5. Ford, D., et al., Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet*, 1998. **62**(3): p. 676-89.
6. Tavtigian, S.V., et al., The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet*, 1996. **12**(3): p. 333-7.
7. Sluiter, M.D. and E.J. van Rensburg, Large genomic rearrangements of the BRCA1 and BRCA2 genes: review of the literature and report of a novel BRCA1 mutation. *Breast Cancer Res Treat*, 2011. **125**(2): p. 325-49.
8. Cooper, D.N., et al., Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum Mutat*, 2010. **31**(6): p. 631-55.
9. Harland, M., et al., A deep intronic mutation in CDKN2A is associated with disease in a subset of melanoma pedigrees. *Hum Mol Genet*, 2001. **10**(23): p. 2679-86.
10. Dehainault, C., et al., A deep intronic mutation in the RB1 gene leads to intronic sequence exonisation. *Eur J Hum Genet*, 2007. **15**(4): p. 473-7.
11. Cavalieri, S., et al., Deep-intronic ATM mutation detected by genomic resequencing and corrected in vitro by antisense morpholino oligonucleotide (AMO). *Eur J Hum Genet*, 2012.
12. Clendenning, M., et al., *Mutation deep within an intron of MSH2 causes Lynch syndrome*. *Fam Cancer*, 2011. **10**(2): p. 297-301.
13. Agrawal, S., R. Pilarski, and C. Eng, Different splicing defects lead to differential effects downstream of the lipid and protein phosphatase activities of PTEN. *Hum Mol Genet*, 2005. **14**(16): p. 2459-68.
14. Pros, E., et al., Nature and mRNA effect of 282 different NF1 point mutations: focus on splicing alterations. *Hum Mutat*, 2008. **29**(9): p. E173-93.
15. Anczukow, O., et al., BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy. *Clin Cancer Res*, 2012. **18**(18): p. 4903-9.

16. Caputo, S., et al., Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. *Nucleic Acids Res*, 2012. **40**(Database issue): p. D992-1002.
17. Diez, O., et al., Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: a high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat*, 2003. **22**(4): p. 301-12.
18. Casilli, F., et al., The contribution of germline rearrangements to the spectrum of BRCA2 mutations. *J Med Genet*, 2006. **43**(9): p. e49.

## 2.B Búsqueda de variantes patogénicas implicadas en HBOC en genes candidatos adicionales a BRCA1/2

*Baja prevalencia de mutaciones inactivantes de SLX4 en familias con cáncer de mama y/o ovario no portadoras de mutación en BRCA1/2.*

La anemia de Fanconi (FA) es una enfermedad autosómica recesiva genéticamente heterogénea. Se caracteriza por defectos en el desarrollo, fallo en la médula ósea e incidencia de cánceres en la infancia. Numerosas evidencias apuntan hacia la existencia de una base genética común para la FA y la susceptibilidad al cáncer de mama/ovario. Recientemente se ha demostrado que mutaciones germinales bialelicas en *SLX4* son la causa de un subtipo de FA desconocido previamente (FA-P). En este trabajo se ha analizado el papel de *SLX4/FANCP* en la susceptibilidad al cáncer de mama/ovario realizando una búsqueda exhaustiva de mutaciones en 486 casos índice de familias españolas con múltiples casos de cáncer de mama y/o ovario no portadoras de mutación en *BRCA1/2* (Familias No BRCA1/2). Se detectó la presencia de una mutación claramente inactivante (p.Glu1517X). Adicionalmente, se observó que la mutación no sinónima (p.Arg372Trp), para la cual se predice un efecto patogénico mediante análisis *in silico*, cosegrega con la enfermedad en una de las familias. En general, el estudio sugiere que el análisis de mutaciones en *SLX4* tendrá muy poco impacto en el consejo genético de familias No BRCA1/2





## SHORT REPORT

Low prevalence of *SLX4* loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer familiesGorka Ruiz de Garibay<sup>1</sup>, Avellaneda Díaz<sup>1</sup>, Belén Gaviña<sup>1</sup>, Atocha Romero<sup>1</sup>, Pilar Garre<sup>1</sup>, Ana Vega<sup>2</sup>, Ana Blanco<sup>2</sup>, Alicia Tosar<sup>1</sup>, Orland Díez<sup>3</sup>, Pedro Pérez-Segura<sup>4</sup>, Eduardo Díaz-Rubio<sup>1,4</sup>, Trinidad Caldés<sup>1</sup> and Miguel de la Hoya<sup>\*1</sup>

Fanconi anemia is a genetically heterogeneous autosomal recessive disorder characterized by development abnormalities, bone marrow failure, and childhood cancers. Compelling evidence indicates a common genetic basis for FA and breast/ovarian cancer susceptibility. Recently, biallelic germ-line mutations in *SLX4* have been demonstrated to cause a previously unknown FA subtype (FA-P). We address the role of *SLX4/FANCP* in breast/ovarian cancer susceptibility by conducting a comprehensive mutation scanning in 486 index cases from non-BRCA1/BRCA2 multiple-case breast and/or ovarian cancer families (non-BRCA1/2 families) from Spain. We detected one unequivocal loss-of-function mutation (p.Glu1517X). In addition, one missense change (p.Arg372Trp) predicted to be pathogenic by *in silico* analysis co-segregates with disease in one family. Overall, the study indicates that *SLX4* mutation screening will have a very low impact (if any) in the genetic counseling of non-BRCA1/2 families.

European Journal of Human Genetics (2013) 21, 883–886; doi:10.1038/ejhg.2012.268; published online 5 December 2012

**Keywords:** hereditary breast cancer; FA; *SLX4*; *FANCP*; genetic testing

## INTRODUCTION

*BRCA1* and *BRCA2* germ-line mutations explain roughly 20% of the multiple-case breast and/or ovarian cancer families attending familial cancer clinics. Non-BRCA1/2 families have been the focus of several genetic studies aimed to identify additional causative loci, albeit with little success.<sup>1</sup>

Fanconi Anemia (FA) is a genetically heterogeneous rare autosomal recessive disorder characterized by developmental abnormalities, bone marrow failure, and childhood cancers.<sup>2</sup> Somehow unexpectedly, a common genetic basis for breast/ovarian cancer susceptibility and FA emerged when biallelic *BRCA2* loss-of-function mutations were identified in FA type D1 patients.<sup>3</sup> Later, mutations in *FANCF/BRIP1*, *FANCG/PALB2*, *FANCO/RAD51C*, and *XRCC2* have been found to lead both to FA (when biallelic) and to increased breast cancer risk (when monoallelic).<sup>4–7</sup>

Recently, two studies have demonstrated that *SLX4* loss-of-function biallelic mutations explain a novel FA subgroup (FANCP).<sup>8,9</sup> Not surprisingly, both reports suggest that *SLX4* might be a breast and/or ovarian cancer predisposing gene. So far, this hypothesis has been tested by analyzing *SLX4* in 52 German/Byelorussian, 94 Spanish and 526 Italian non-BRCA1/2 families, but obvious loss-of-function germ-line mutations (or any other evidence supporting a causative role) have not been reported.<sup>10–12</sup> To further investigate the role of *SLX4* in breast/ovarian cancer susceptibility, we have performed a comprehensive scanning of germ-line mutations in a cohort of 486 index cases from Spanish non-BRCA1/2 families.

## PATIENTS AND METHODS

486 unrelated women (465 breast and 21 ovarian cancer patients) who were referred for *BRCA1* and *BRCA2* genetic testing due to a family history of breast and/or ovarian cancer compatible with a hereditary syndrome (a minimum of two additional breast and/or ovarian cancer cases diagnosed at any age in two generations of the same parental branch). After comprehensive genetic testing (full coding sequence, intron/exon boundaries and genomic rearrangements), all 486 women were considered non-BRCA1/2 patients, that is, genetic testing discarded the presence of *BRCA1/2* truncating mutations or any other sequence variant described as a disease-causing mutation in the Breast Informative Core (BIC) database (<http://research.nhgri.nih.gov/bic/>). The MLPA KIT P045 (MRC-Holland) detected the *CHEK2*\*1100delC allelic variant in one woman (later confirmed by direct sequencing). As this is a low-penetrance allele,<sup>13</sup> the carrier was not excluded from further analysis. Families were ascertained through three participating centers: Hospital Clínico San Carlos, Madrid, (*N*=348), Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, (*N*=98), and Hospital Vall d'Hebrón, Barcelona (*N*=40). The study was approved by the Ethics Committee of each participating center. All patients signed an informed consent to participate in the study. DNA samples were obtained from peripheral blood leukocytes using the standard procedures.

A total of 29 *SLX4* amplicons were designed (Supplementary Table 1). PCR amplification and high-resolution melting analysis were carried out in a 7500 Fast Real-Time PCR Instrument (Applied Biosystems, Foster City, CA, USA) with Melt-Doctor HRM Master Mix (Applied Biosystems) and default settings. Samples were sequenced using the Big Dye Terminator v1.1 Cycle Sequencing Reaction kit (Applied Biosystems), according to the manufacturer's protocol. All sequence variants are named according to the Human Genome Variation Society (HGVS) nomenclature, using Ensemble transcript ID

<sup>1</sup>Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain; <sup>2</sup>Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica-USC, CIBERER-ER, IDIS, Santiago de Compostela, Spain; <sup>3</sup>Laboratorio de Oncogenética, Hospital Universitario Vall d'Hebrón & Vall d'Hebrón Instituto Oncológico (VHIO), Barcelona, Spain; <sup>4</sup>Servicio de Oncología Médica, Hospital Clínico San Carlos, Madrid, Spain

\*Correspondence: Dr M de la Hoya, Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), planta baja sur, Hospital Clínico San Carlos, c/Martín Lagos s/n, Madrid 28040, Spain. Tel: +34 91 3303348; E-mail: mdhoya@hotmail.com

Received 28 June 2012; revised 12 October 2012; accepted 2 November 2012; published online 5 December 2012



Table 1 9 SLX4 germ-line variants predicted to be deleterious

exon	Nucleotide <sup>a</sup>	Protein <sup>a</sup>	dbSNP135 <sup>b</sup>	Polyphen-2 <sup>c</sup>	Panther <sup>c</sup>	SIFT <sup>d</sup>	GVGD <sup>e</sup>	BRCA2 (N = 972) <sup>f</sup>	Controls (N = 960) <sup>f</sup>	Previously detected in BRCA2 families?
2	c.421G>T	p.Gly141Trp	rs137976282	0.964	undefined <sup>g</sup>	0	undefined <sup>g</sup>	1	2	(refs. 11,12)
4	c.832C>T	p.Arg278Trp	rs141597706	0.997	0.771	0.01	C0	1	—	no
5	c.1114C>T	p.Arg372Trp	—	1	0.862	0	C65	1	0	no
7	c.1637A>G	p.Tyr546Cys	rs150547487	1	0.771	0	C65	1	—	no
12	c.2359G>A	p.Glu787Lys	rs140600202	0.995	0.397	0.01	C55	2	1	(ref. 12)
12	c.4333C>T	p.Arg1445Trp	—	1	0.837	0.02	C15	1	—	no
12	c.4409C>T	p.Pro1470Leu	rs72778139	0.7	0.15	0.15	C55	1	—	(ref. 11)
12	c.4549G>T	p.Glu1517X	—	—	—	—	—	1	0	no
13	c.4648C>T	p.Arg1550Trp	rs77021998	0.06	0.847	0.01	C25	3	2	(ref. 12)

<sup>a</sup>Ensembl transcript ID ENST0000029400 as a reference.

<sup>b</sup>dbSNP135 (accessed on 12 June 2012).

<sup>c</sup>Polyphen and Panther scores range from 0 (less likely pathogenic) to 1 (most likely pathogenic).

<sup>d</sup>SIFT scores range from 1 (less likely damaging) to 0 (most likely damaging), strongly supporting a damaging effect for scores  $\leq 0.05$ .

<sup>e</sup>GVGD classifies missense changes in seven categories C0, C15, C25, C35, C45, C55, and C65 (from less to most likely deleterious).

<sup>f</sup>N = chromosomes.

<sup>g</sup>Weak evolutionary conservation of the SLX4 N-terminal domain prevents accurate predictions.

ENST00000294008 as a reference sequence. All variants were investigated in dbSNP135 (<http://www.ncbi.nlm.nih.gov/SNP/>) and Exome Variant Server (<http://evs.gs.washington.edu/EVS/>), Exome Sequencing Project (ESP), Seattle, WA (Table 1, and Supplementary Table 2).

Splicing predictions were performed with Human Splicing Finder and MaxEnt splice site matrices (<http://www.umhd.be/HSF/>). Discrepant results were analyzed with NNSPLICE ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)), and GeneSplicer (<http://www.cbc.umd.edu/software/genesplicer/>). Bioinformatic analyses of missense substitutions were performed using PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>), SIFT (<http://sift.jcvi.org/>), Align-GVGD (<http://agvgd.iarc.fr/>), and Panther (<http://www.pantherdb.org/tools/csnpl/>). In all cases, default settings were applied. For Align-GVGD analysis, we performed Clustal-W multiple-sequence alignment of Homo, Chimpanzee, Cow, Dog, Guinea, Mouse, Opossum, and Zebrafish SLX4 sequences. Statistical analyses were performed with JavaStat (<http://statpages.org/>).

RESULTS

We used high-resolution melting analysis to investigate the whole coding region and flanking sequences of the SLX4 gene in 486 index cases from non-BRCA1/2 families. Fifty-seven different sequence variants (11 not previously annotated in dbSNP) were identified, including missense (N = 31), silent (N = 18), intronic (N = 6), 5'UTR (N = 1) and nonsense (N = 1) changes (Supplementary Table 2).

The c.2854G>A (p.Ala952Thr) and c.2855C>T (p.Ala952Val) substitutions are recorded in dbSNP135 as two individual common variants (rs59939128, and rs78637028, respectively). However, we noted that all c.2855C>T carriers were c.2854G>A carriers as well. The most parsimonious model indicates that the double-nucleotide (doublet) substitution c.2854\_55GC>AT (p.Ala952Met) is the actual common variant. We identified one true c.2854G>A (p.Ala952Thr) carrier, indicating that this is a rare variant (at least in our population). Similar observations have been made previously.<sup>10,12</sup>

Variants with MAF > 0.01, and/or variants in which homozygous carriers have been reported (Supplementary Table 2) were considered likely neutral polymorphisms (not clinically relevant) and discarded for further analysis (N = 23). The truncating mutation p.Glu1517X (not annotated in dbSNP135) was the only unequivocal loss-of-function SLX4 mutation detected. The mutated allele codifies for a truncated protein lacking 318 residues at the C-terminal end. The deleted region includes SAP (SAF-A/B, Acinus and PIAS) and SBD (SLX1-binding domain) domains critical for proper SLX4 function.<sup>14</sup>

Unfortunately, samples from other family members were not available, so it was not possible to perform co-segregation analysis (Figure 1). The mutation was not detected in 480 healthy control samples (Table 1). The remaining variants (N = 36) were analyzed *in silico* to evaluate their potential impact on mRNA splicing and/or protein function. Neither HSF nor MaxEntScan predicted the splicing effects for any variant, with the single exception of c.4115G>A (p.Arg1372Gln). This variant was predicted by HSF to create a new donor site. MaxEntScan, NNSPLICE, and GeneSplicer did not predict any effect, so it was considered unlikely that this variant affects normal splicing.

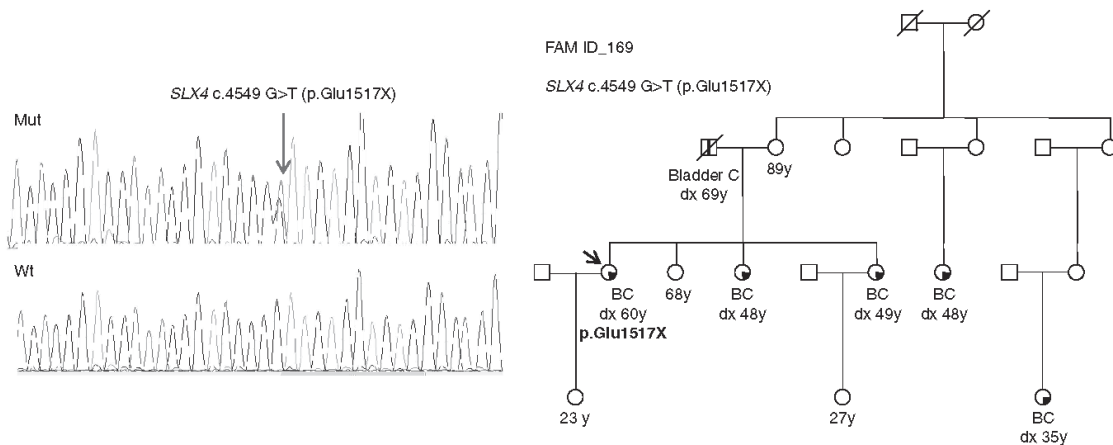
Eight missense variants were predicted to be functionally relevant by two or more algorithms (Table 1). Because of lack of additional family members, co-segregation analysis was restricted to p.Gly141Trp and p.Arg372Trp families (Figure 2). The p.Gly141Trp variant was detected in one family carrying the CHEK2\*1100delC variant as well, thereby complicating co-segregation analysis. The p.Arg372Trp variant co-segregates with disease in two affected family members other than the proband. Four missense variants were tested in a control group and two of these, including p.Arg372Trp, were only found in the patient group (Table 1).

DISCUSSION

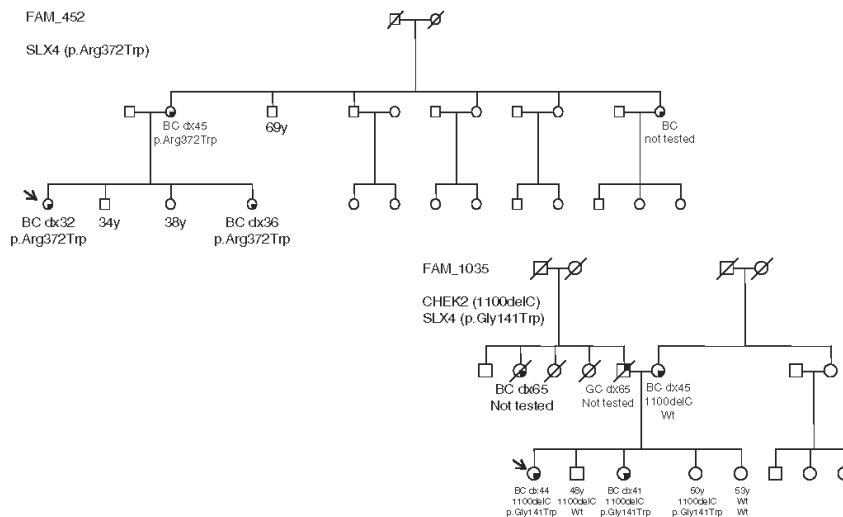
We have detected one unequivocal loss-of-function SLX4 germ-line mutation (p.Glu1517X) in 486 non-BRCA1/2 families tested, confirming that these alterations are rare, but do exist. Previous studies conducted in non-BRCA1/2 families failed to detect such alterations.<sup>10-12</sup> Interestingly, not a single SLX4 truncating (or consensus splice) variant has been observed in > 12,000 exomes from the ESP project, supporting the relevance of our finding.

Yet, our analysis provides further evidence supporting that the proportion of non-BRCA1/2 families explained by SLX4 germ-line defects (if any) is probably very low. Indeed, the existence of genes able to explain (individually) a substantial fraction of non-BRCA1/2 families seems unlikely, as genome-wide linkage analysis failed to identify such hypothetical loci.<sup>1</sup> Consistently with that, RAD51C and XRCC2 germ-line mutations have been reported in a very low proportion (< 1%) of non-BRCA1/2 families.<sup>6,7</sup>

Certainly, the actual prevalence of germ-line mutations in those genes might be underestimated due to the fact that non-truncating disease-causing mutations are difficult to distinguish from clinically



**Figure 1** Left panel displays a representative electropherogram demonstrating the *SLX4* germ-line mutation c.4549G>T (p.Glu1517X). This truncating mutation was identified in a pedigree (right panel) with multiple cases of breast cancer, but not ovarian cancer. The mutation was identified in the index case (arrow). Unfortunately, it was not possible to perform co-segregation analysis. BC, breast cancer. dx, age at diagnosis.



**Figure 2** The *SLX4* variants p.Arg372Trp and p.Gly141Trp are predicted to be pathogenic by *in silico* analysis. In the case of p.Arg372Trp, co-segregation analysis is compatible with a pathogenic role (family ID\_452). In the case of p.Gly141Trp, co-segregation analysis is complicated by the fact that *CHEK2*\*1100delC is also detected in this family (family ID\_1035). Index cases are indicated with arrows. BC, breast cancer. GC, gastric cancer. dx, age at diagnosis.

irrelevant rare variants.<sup>15</sup> For instance, we suspected that p.Arg372Trp is a true pathogenic mutation. This is because this variant is predicted to be pathogenic by four algorithms with very high scores, co-segregates with disease, and has not been detected in 480 healthy control individuals (nor in 12,000 exomes from the ESP project, Supplementary Table 2). Moreover, this variant targets the UBZ4 domain (Ubiquiting-Binding Zinc finger motif 4) essential for *SLX4* interaction with multiple-partner proteins,<sup>14</sup> and recruitment to sites of DNA damage.<sup>16</sup> Indeed, four FA patients harbor *SLX4* splicing mutations causing in-frame exon 5 skipping,<sup>8,9</sup> supporting that the UBZ4 domain (encoded by exon 5) is clinically relevant. Other variants do not appear to represent disease-causing mutations, in particular those detected in control samples, but some might

represent low- or moderate-risk alleles contributing to the overall cancer risk. Interestingly, our cohort included one family harboring both the *SLX4* variant p.Gly141Trp, and *CHEK2*\*1100delC. It has been postulated that *CHEK2*\*1100delC might interact with another rare gene (or genes) to cause breast cancer risk comparable with those conferred by *BRCA1* or *BRCA2*.<sup>17</sup> Our data suggest that *SLX4* might be one of such rare-interacting genes.

We think that routine analysis of *SLX4* in non-BRCA1/2 families by classical technologies such as Sanger sequencing is not warranted. Yet, this and other FA-related genes can be incorporated easily into novel testing strategies who analyze large panels of breast and ovarian cancer susceptibility genes with massively parallel sequencing.<sup>18</sup> In this scenario, the identification of people whose breast cancer is



related to homologous recombination DNA-repair defects seems clinically relevant, as those patients might benefit from Poly (ADP-ribose) polymerase (PARP) inhibitor therapies.

In conclusion, the prevalence of *SLX4* loss-of-function mutations among non-BRCA1/2 families is probably very low. Yet, to definitely demonstrate (or discard) a role for *SLX4* in breast and/or ovarian cancer genetic susceptibility is of outstanding scientific interest, so that further studies in very large cohorts are warranted.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank the patients whose participation made this project possible. We thank Paula Diaque for technical assistance. This study was supported by Instituto de Salud Carlos III, Fondo de Investigación Sanitaria (FIS) Research Grants 09/00859, and Fundación Mutua Madrileña (FMM) Research Grant FMM-08, to MH. Xunta de Galicia (10PXIB 9101297PR) and Fundación Mutua Madrileña (FMM-10) to AV. Red Temática de Investigación Cooperativa en Cáncer; Instituto de Salud Carlos III, Fondo Europeo de Desarrollo Regional (RETICC 06/0020/0021) supported PG, AR, PPS, EDR, TC, and MdH. None of these study sponsors had any role in study design, in the collection, analysis, and interpretation of data; in the writing of the report; and/or in the decision to submit the paper for publication.

- 1 Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P: Genetic susceptibility for breast cancer: how many more genes to be found? *Crit Rev Oncol Hematol* 2007; **63**: 125–149.
- 2 Moldovan GL, D'Andrea AD: How the fanconi anemia pathway guards the genome. *Annu Rev Genet* 2009; **43**: 223–249.

- 3 Howlett NG, Taniguchi T, Olson S et al: Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 2002; **297**: 606–609.
- 4 Seal S, Thompson D, Renwick A et al: Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 2006; **38**: 1239–1241.
- 5 Rahman N, Seal S, Thompson D et al: PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 2007; **39**: 165–167.
- 6 Meindl A, Hellebrand H, Wiek C et al: Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet* 2010; **42**: 410–414.
- 7 Park DJ, Lesueur F, Nguyen-Dumont T et al: Rare mutations in XRCC2 increase the risk of breast cancer. *Am J Hum Genet* 2012; **90**: 734–739.
- 8 Stoopker C, Hain K, Schuster B et al: SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet* 2011; **43**: 138–141.
- 9 Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, Smogorzewska A: Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet* 2011; **43**: 142–146.
- 10 Landwehr R, Bogdanova NV, Antonenkova N et al: Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer. *Breast Cancer Res Treat* 2011; **130**: 1021–1028.
- 11 Fernández-Rodríguez J, Quiles F, Blanco I et al: Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. *BMC Cancer* 2012; **12**: 84.
- 12 Catucci I, Colombo M, Verderio P et al: Sequencing analysis of SLX4/FANCP gene in Italian familial breast cancer cases. *PLoS One* 2012; **7**: e31038.
- 13 Meijers-Heijboer H, van den Ouweland A, Kljin J et al: CHEK2-Breast Cancer Consortium, low-penetrance susceptibility to breast cancer due to CHEK2\*(\*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 2002; **31**: 55–59.
- 14 Svendsen JM, Smogorzewska A, Sowa ME et al: Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* 2009; **138**: 63–77.
- 15 Osorio A, Endt D, Fernández F et al: Predominance of pathogenic missense variants in the RAD51C gene occurring in breast and ovarian cancer families. *Hum Mol Genet* 2012; **21**: 2889–2898.
- 16 Yamamoto KN, Kobayashi S, Tsuda M et al: Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proc Natl Acad Sci USA* 2011; **108**: 6492–6496.
- 17 Oldenburg RA, Kroeze-Jansema K, Kraan J et al: The CHEK2\*(\*)1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer Res* 2003; **63**: 8153–8157.
- 18 Walsh T, Lee MK, Casadei S et al: Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. *Proc Natl Acad Sci USA* 2010; **107**: 12629–12633.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)

**SUPPLEMENTARY TABLES**

Supplemental table 1 Primers used for HRM mutation scanning of the SLX4 gene.

Exon	Amplicon	Forward	Reverse	Size (bp)	Genomic location <sup>A</sup>
2	2A	GCCAGTTGTGGACTTTAACA	GCTTTCAGGCTGGTCTTCAG	271	7461-7731
	2B	TCATCTGTCTGCCTGTCTG	TGGTCCTTTCAATTTGCTTC	226	7674-7900
	2C	ACACAAAAGGCTGCCTCAA	GGGTGTTTTGTGCTGTTCC	256	7846-8101
	2D	AAATGGCAAGCAAGTGAACC	GCCCTTCCAGGAAGTTTC	207	7999-8205
3	3	CAGCTACAACAGATGGTTGTGAA	CTGTGTGGAAGACAGAAACACA	298	9846-10143
4	4	GGGTGAGCTTGTCTTTGGAG	CTTGGGGATTCTGGTCACAT	298	14224-14521
5	5	GGGAGGGTGGTTCTCCTGTA	CACAGCATTATTTGGCTTGG	295	15356-15650
6	6	GAAGCGTGCACTCCTCTAGC	CTCCAGGGTCACTCTTCTG	286	18534-18819
7	7	ACTCCACCGTTGATGGTGTG	AGCCTATCCATGTGCCTGAG	388	18852-19239
8	8	ATGTGATGGCTTCTGCAGTG	AGTCCATGGCAGCCTCAG	393	20119-20511
9	9	TCTTCTGGTGGTGGATGTG	ACATGGCAAGTGGGATGG	207	20829-21035
10	10	CCCAGTTAGCAGAGCAGAGG	GAGAGGAAAGGAGGGCACAC	242	21929-22170
11	11	CAGTAAACCTCTGTGTTGCAT	CTGCCAGTCTCGTCAGTTA	247	23687-23934
12	12A	CGATGACTGTCGATGTTTCG	TCTTCGTGGTCTTGGATT	223	25247-25469
	12B	GAGGAGGAAGCGGAGACTTT	CTCCACCTGTCCCACTGTT	234	25426-25659
	12C	CTCCTAGCAGGTGCCAGGT	GTAATCCCCGCATCATCT	282	25615-25896
	12D	GCCAGGCAGAGAGAAAAGAA	ACGACCCACTTGTGTGATGA	270	25841-26110
	12E	CTCCCTGGCAGGCATCTC	TTTTTCTGGGAACATCAGGA	293	26015-26307
	12F	CGTGCTGGAGTGCAGAAATA	GCTCACAGGACCTAGGGCTAA	257	26262-26518
	12G	GATCCTCTGGAAGAAAAGAAAGC	GAACAGTCACGGCTTCTGCT	290	26467-26756
	12H	CCAGCACACAGACACCTC	CTTCTGTGGCCTTGCCTTCT	283	26690-26972
	12I	ACTCCCGTCTCTCCAGGAAC	TGCTCATCGTCACTGTCTCC	230	26923-27152
	12J	CAAGCTTCTGAACCAGACC	TCCAGGTTCCAGTGGTCAAT	190	27089-27278
	12K	CCAATTGACGACTGCTGCT	CAGAGCCGAATTCAGAAAGC	267	27211-27477
	12L	AGCTGCACTACCCAGAGGAA	GACTGATCTTCCCACCCTCTT	270	27388-27658
13	13	TGGGATCAGATGTGGGAAAT	ATGGGAGGGAAAAGACCACT	216	31649-31865
14	14A	GGGTCTGAGATCCACCTGT	GCCCTTGGTCTTAGCAGGTC	283	33044-33327
	14B	AGGAGGCCACCACAGGAC	CACTGACCCCATCGCATC	249	33269-33517
15	15	CCTGGTCTGGCTCTTTTCAG	CTGCTGATGGCAGGTTGG	411	33872-342802

<sup>A</sup> Refseq NG\_028123



Supplemental Table 2. SLX4 genetic variants identified in 486 Spanish Non-BRCA1/2 families

486 Non-BRCA1/2 families				Controls	In Silico Analysis							dbSNP135			EVS (ESP)		Previous studies in Non-BRCA1/2 families		
EXON/INTRON	Variant	HET	HZ	MAF (N=972)	MAF (N=960)	Polyphen-2	SIFT	AGVGD	Panther	HSF	MaxEnt	rs.	MAF	HZ	MAF	HZ	(ref. 10)	(ref. 12)	(ref. 11)
2 (UTR)	c.-31G>T	1	0	0.001	-	-	-	-	-	neutral	neutral	-	-	-	<0.0001	no	-	-	-
2	c.88 G>A (p.Arg29His)	1	0	0.001	-	Benign	Tolerated	undefined	undefined	neutral	neutral	rs149117119	<0.001	no	0.0002	no	-	-	-
2	c.90C>T (p.Ser30=)	20	0	0.02	-	-	-	-	-	neutral	neutral	rs118089506	0.006	no	0.005	no	-	0.014	-
2	c.376A>G (p.Thr126Ala)	1	0	0.001	-	Benign	Tolerated	undefined	undefined	neutral	neutral	-	-	-	-	-	-	-	-
2	c.421G>T (p.Gly141Trp)	1	0	0.001	0.001	0.964	0 (low confidence)	undefined	undefined	neutral	neutral	rs137976282	<0.0001	no	0.0007	no	-	0.002	0.005
2	c.465A>G (p.Glu155=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs138512851	<0.0001	no	0.0001	no	-	-	-
3	c.555C>T (p.Asp185=)	42	0	0.043	-	-	-	-	-	-	-	rs74640850	0.046	yes	0.049	yes	-	-	0.048
3	c.590T>C (p.Val197Ala)	3	0	0.003	-	Benign	Tolerated	undefined	0.093	neutral	neutral	rs147826749	0.003	no	0.002	no	0.010	0.003	-
3	c.610 C>T (p.Arg204Cys)	42	0	0.043	-	-	-	-	-	-	-	rs79842542	0.061	yes	0.071	yes	0.096	0.040	0.053
3	c.678C>T (p.His226=)	26	0	0.027	-	-	-	-	-	-	-	rs28516461	0.11	yes	0.105	yes	0.010	0.028	0.016
3	c.753G>A (p.Ala251=)	52	1	0.055	-	-	-	-	-	-	-	rs8061528	0.26	yes	0.218	yes	0.173	0.211	0.149
3	c.760+10C>G	7	0	0.007	-	-	-	-	-	neutral	neutral	-	-	-	-	-	-	-	-
4	c.832 C>T (p.Arg278Trp)	1	0	0.001	-	0.997	0.01	C0	0.771	neutral	neutral	rs141597706	<0.0001	no	0.0002	no	-	-	-
5	c.1114C>T (p.Arg372Trp)	1	0	0.001	0	1.0	0	C65	0.862	neutral	neutral	-	-	-	-	-	-	-	-
5	c.1152A>G (p.Pro384=)	44	0	0.045	-	-	-	-	-	-	-	rs112511042	0.05	yes	0.069	yes	0.096	0.041	0.059
5	c.1153 C>A (p.Pro385Thr)	3	0	0.003	-	Benign	Tolerated	C0	0.179	neutral	neutral	rs115694169	0.008	no	0.008	no	-	0.006	0.005
5	c.1156A>G (p.Met386Val)	44	0	0.045	-	-	-	-	-	-	-	rs113490934	0.05	yes	0.069	yes	0.096	0.041	0.058
5	c.1163+10C>T	44	0	0.045	-	-	-	-	-	-	-	rs80116508	0.059	no	0.074	yes	0.096	0.041	0.058
6	c.1366+11T>C	1	0	0.001	-	-	-	-	-	-	-	rs76350200	0.074	no	0.095	yes	0.096	0.048	0.064
7	c.1637A>G (p.Tyr546Cys)	1	0	0.001	-	1.0	0	C65	0.771	neutral	neutral	rs150547487	0.005	no	0.005	no	-	-	-
9	c.1991 C>T (p.Pro664Leu)	1	0	0.001	-	Benign	Tolerated	C0	0.173	neutral	neutral	-	-	-	<0.0001	no	-	-	-
9	c.2012T>C (p.Leu671=)	46	0	0.047	-	-	-	-	-	-	-	rs77985244	0.09	yes	0.082	yes	0.096	0.042	0.058
9	c.2013+21G>T	1	0	0.001	-	-	-	-	-	neutral	neutral	rs139099529	-	-	0.007	no	-	-	-
9	c.2013+23G>A	45	0	0.046	-	-	-	-	-	-	-	rs112226642	0.09	yes	0.069	yes	0.096	0.041	0.058
11	c.2235C>T (p.Thr745=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs75184268	-	-	-	-	-	0.002	-
11	c.2305G>C (p.Glu769Gln)	3	0	0.003	-	Benign	Tolerated	C0	0.130	neutral	neutral	rs150712805	0.001	no	0.001	no	-	-	-
12	c.2359 G>A (p.Glu787Lys)	2	0	0.002	0.001	0.995	0.01	C55	0.397	neutral	neutral	rs140600202	<0.0001	no	0.001	no	-	0.008	-
12	c.2654C>T (p.Pro885Leu)	1	0	0.001	-	0.930	Tolerated	C0	0.366	neutral	neutral	rs147492092	0.001	no	<0.0001	no	-	-	-
12	c.2667 A>G (p.Gln889=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs146790631	0.001	no	<0.0001	no	-	-	-
12	c.2824G>C (p.Glu942Gln)	8	0	0.008	-	-	-	-	-	-	-	rs114014006	0.014	yes	0.018	yes	-	-	-
12	c.2854 G>A (p.Ala952Thr)	1	0	0.001	-	0.0152	Tolerated	C0	0.551	neutral	neutral	rs59939128	0.1	yes	0.032	yes	-	0.002	-
12	c.2854 55G>C>AT (p.Ala952Met)	47	0	0.048	-	-	-	-	-	-	-	-	-	-	0.047	yes	0.096	0.040	0.042
12	c.2924C>T (p.Pro975Leu)	8	0	0.008	-	Benign	Tolerated	C0	0.223	neutral	neutral	rs114472821	0.006	no	0.005	no	-	0.007	-
12	c.2975G>A (p.Gly992Glu)	1	0	0.001	-	Benign	Tolerated	C0	0.173	neutral	neutral	rs139287784	0.001	no	0.0003	no	-	0.001	-
12	c.3162G>A (p.Ser1054=)	23	0	0.024	-	-	-	-	-	-	-	rs76488917	0.04	yes	0.033	yes	0.010	0.026	0.005
12	c.3308G>A (p.Arg1103His)	1	0	0.001	-	Benign	Tolerated	C0	0.169	neutral	neutral	-	-	-	-	-	-	0.001	-
12	c.3365C>T (p.Pro1122Leu)	80	1	0.08	-	-	-	-	-	-	-	rs714181	0.21	yes	0.212	yes	0.096	0.062	0.063
12	c.3662C>T (p.Ala1221Val)	47	0	0.048	-	-	-	-	-	-	-	rs3827530	0.14	yes	0.049	yes	0.096	0.033	0.053
12	c.3783G>A (p.Pro1261=)	1	0	0.001	-	-	-	-	-	neutral	neutral	-	-	-	0.021	yes	-	0.006	-
12	c.3812C>T (p.Ser1271Phe)	46	0	0.046	-	-	-	-	-	-	-	rs3810813	0.09	yes	0.040	yes	0.058	0.049	0.021
12	c.3868C>A (p.His1290Asn)	1	0	0.001	-	Benign	Tolerated	C0	0.156	neutral	neutral	rs112596894	0.001	no	0.001	no	-	-	-
12	c.3939C>A (p.Pro1313=)	1	0	0.001	-	-	-	-	-	neutral	neutral	-	-	-	<0.0001	no	-	-	-
12	c.3963G>A (p.Pro1321=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs116781836	0.006	no	0.006	no	-	0.010	-
12	c.4115G>A (p.Arg1372Gln)	2	0	0.002	-	Benign	Tolerated	C0	0.117	new acceptor site	neutral	rs79174372	0.008	no	0.010	yes	-	-	-
12	c.4241C>T (p.Pro1414Leu)	1	0	0.001	-	Benign	Tolerated	C0	0.061	neutral	neutral	rs140892471	0.002	no	0.002	no	-	-	-
12	c.4333C>T (p.Arg1445Trp)	1	0	0.001	-	1.0	0.02	C15	0.837	neutral	neutral	-	-	-	0.0001	no	-	-	-
12	c.4344C>T (p.Pro1448=)	1	0	0.001	-	-	-	-	-	neutral	neutral	-	-	-	<0.0001	no	-	-	-
12	c.4409C>T (p.Pro1470Leu)	1	0	0.001	-	0.7	Tolerated	C65	0.563	neutral	neutral	rs72778139	<0.0001	no	<0.0001	no	-	-	0.005
12	c.4494 G>A (p.Leu1498=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs146532299	<0.0001	no	<0.0001	no	-	-	-
12	c.4500T>C (p.Asn1500=)	239	114	0.48	-	-	-	-	-	-	-	rs3810812	0.45	yes	0.37	yes	0.510	0.510	0.447
12	c.4530 G>T (p.Leu1510=)	1	0	0.001	-	-	-	-	-	neutral	neutral	rs139254595	<0.0001	no	0.0001	no	-	-	0.005
12	c.4549G>T (p.Glu1517X)	1	0	0.001	0	-	-	-	-	neutral	neutral	-	-	-	-	-	-	-	-
12	c.4581G>A (p.Pro1527=)	4	0	0.004	-	-	-	-	-	-	-	rs78635099	0.03	yes	0.021	yes	-	0.009	-
12	c.4600G>T (p.Gly1534Ser)	1	0	0.001	-	-	-	-	-	-	-	rs78770603	0.05	yes	0.016	yes	-	0.001	-
13	c.4648C>T (p.Arg1550Trp)	3	0	0.003	0.002	Benign	0.01	C25	0.847	neutral	neutral	rs77021998	0.002	no	0.002	no	-	0.001	-
13	c.4739+24G>T	93	7	0.11	-	-	-	-	-	-	-	rs12933120	0.17	yes	0.098	yes	0.115	0.159	0.127

HET, heterozygous. HZ, homozygous. MAF, minor allele frequency. dbSNP135 accessed 12/06/2012. EVS (ESP) accessed 10/09/2012



### Relevancia del estudio del *splicing* en la clasificación de variantes genéticas en BRCA2

En el año 1995 se identificó a *BRCA2* como gen de susceptibilidad al cáncer de mama y ovario (Wooster et al., 1995). En ese trabajo, de las seis variantes patogénicas identificadas, una de ellas consistía en la delección de las cinco últimas bases del exón 13 y la primera base del intrón, sugiriéndose como efecto de esta variante una posible alteración del *splicing*. Un trabajo posterior, del año 1996, describió la inserción de una secuencia *alu* en el exón 22 con el resultado de la exclusión de dicho exón en el ARNm (Miki et al., 1996). En la actualidad se han descrito alteraciones en el proceso de *splicing* en un total de 95 variantes de *BRCA2*, según una reciente revisión realizada por el grupo de trabajo de *splicing* del consorcio ENIGMA. (Walker et al., 2013)

#### Aspectos Metodológicos

Uno de los aspectos relevantes a la hora de la realización de un análisis del *splicing* es el diseño de la metodología. Para poder establecer una clasificación clínicamente relevante en función del análisis del *splicing* es importante que el análisis se realice mediante una metodología estandarizada y validada. El análisis de los estudios publicados hasta el momento sugiere la existencia de una gran variabilidad metodológica. Esta variabilidad se refleja en varios aspectos del proceso analítico como por ejemplo:

- Diseño del estudio: el origen de las muestras de ARN puede ser una recolección *in vivo* (de muestras originarias del paciente) o bien *ex vivo* (mediante el uso de minigenes).
- Tejido de la muestra de RNA: En los estudios que utilizan muestras derivadas de pacientes se utiliza sangre periférica fresca (Bonatti et al., 2006), cultivos primarios de linfocitos (Machackova et al., 2008), cultivos de fibroblastos (Vreeswijk et al., 2009) o cultivos de líneas linfoblastoides (Tesoriero et al., 2005). En los estudios basados en el uso de minigenes se han utilizado diversas líneas celulares como HeLa, COS7 y MCF7.
- Algunos estudios incluyen tratamiento con cicloheximida o puromicina para inhibir la posible degradación de transcritos portadores de codones de parada prematuros, mientras que otros estudios no lo incluyen.
- Detección de transcritos: la detección de los transcritos se ha realizado mediante electroforesis en gel o bien en capilar y la identificación de su identidad mediante

secuenciación directa, secuenciación tras la purificación de la banda o secuenciación tras la clonación del producto amplificado.

El estudio presentado en el esta tesis es un esfuerzo colaborativo internacional del consorcio ENIGMA, con el objetivo de establecer una serie de criterios de estandarización. Como conclusiones se determina una serie de aspectos que son claves a la hora del correcto análisis e interpretación de los defectos del *splicing*, que son:

- Transcriptasa reversa: Se detectó que el uso de ciertas transcriptasas reversas (Superscript II) y de la combinación oligodT/random hexámeros favoreció la síntesis de ADNc particularmente largos.
- El diseño de la PCR: la localización de los cebadores fue fundamental a la hora de detectar de manera correcta todos los transcritos. Resulta muy difícil el proponer unos cebadores de referencia debido a que el diseño del análisis del *splicing* se realiza de manera individualizada para cada una de las variantes que se estudia. Sin embargo, el conocimiento de los posibles transcritos alternativos que suceden de manera natural en el gen podría ayudar a un diseño de cebadores adecuado. Así se ha demostrado, por ejemplo, en el caso de la variante *BRCA1* c.591C>T (Dasil et al., 2010). Además, el uso de herramientas *in silico* que orienten sobre los posibles transcritos aberrantes producidos como consecuencia de una variante podrían facilitar un diseño adecuado de cebadores. Por otra parte, el tiempo de elongación (adaptable para transcritos largos) y el número de ciclos (mínimo sugerido de 30) también fueron esenciales para la detección de transcritos particularmente largos y/o con bajos niveles de expresión.
- Métodos de detección de los transcritos: La detección de transcritos mediante electroforesis capilar ofrece una mejora en la sensibilidad frente al uso de en gel de agarosa, lo cual permite la detección de transcritos minoritarios o sometidos a degradación.
- Métodos de secuenciación: La clonación previa a la secuenciación demostró una mayor sensibilidad frente a la secuenciación directa o el aislamiento de bandas de gel de agarosa.

Cabe destacar que, debido al diseño del estudio basado en el análisis de muestras procedentes de pacientes, el análisis basado en minigenes queda fuera de la validación. Sería interesante en un futuro el diseño de un estudio de validación específico del análisis de minigenes y de un estudio comparativo de ambas aproximaciones.

### Clasificación de variantes mediante el análisis de *splicing*.

El protocolo de análisis del *splicing* utilizado de manera rutinaria en esta tesis cumple la mayoría de los estándares sugeridos como consenso por el grupo de trabajo de *splicing* del consorcio ENIGMA y por ello pensamos que desde un punto de vista metodológico los resultados obtenidos son robustos. En nuestra serie se analizaron 31 UVs utilizando como único criterio de selección de UVs la disponibilidad de ARN. El resultado del análisis del *splicing* permitió clasificar como patogénicas únicamente a tres; c.682-2A>G, c.7617+1G>A y c.8954-5A>G. La mayoría de las bases de datos del gen *BRCA2* que contienen resultados de análisis del *splicing* recogen un porcentaje mayor de variantes que producen alteraciones en el *splicing*. La base de datos LOVD (último acceso Noviembre 2013) contiene 63 UVs analizadas a nivel de *splicing*, demostrándose alteraciones que permitieron clasificar como patogénicas al 50%. Por su parte, la base de datos UMD-*BRCA2* (Caputo et al., 2012) contiene 57 variantes analizadas a nivel de *splicing*, de las cuales un 35% demostraron alteraciones en el *splicing* que permitieron clasificar a la variante como patogénica.

En el año 2011, Mucaki y colaboradores (Mucaki et al., 2011) aplicaron el método bioinformático de predicción de *splicing* Automated Splice Site Analysis (ASSA) al total de variantes depositadas en la base de datos BIC (1240 variantes en *BRCA2*). Para 89 de las variantes se predijo el debilitamiento de los sitios naturales de *splicing*. Además, para 46 de las variantes se predijo la potenciación de sitios alternativos de *splicing*. En total un 11% de las variantes en *BRCA2* presenta predicciones positivas para algún tipo de alteración en el *splicing*. La predicción realizada mediante métodos bioinformáticos ha demostrado una buena correlación con los resultados experimentales del análisis del *splicing*, habiéndose establecido mediante el análisis de curvas ROC valores de corte para los distintos algoritmos disponibles (Houdayer et al., 2012). Sin embargo un trabajo reciente en el gen *BRCA1* detectó la existencia de variantes exónicas que alteran el *splicing* y que escapan a las predicciones bioinformáticas (Wappenschmidt et al., 2012).

En principio, cualquier variante es susceptible de alterar el *splicing* independientemente de su localización y significado a nivel proteico, por lo que un análisis de *splicing* nunca debe descartarse *a priori*. A lo largo de multitud de trabajos realizados estudiando el impacto en el *splicing* se han descrito una gran variedad de variantes genéticas en *BRCA2* produciendo diversos tipos de alteraciones. Ejemplos de estas alteraciones son:

- Variantes que alteran los sitios de *splicing*, produciendo la eliminación del exón del ARNm maduro. Es el caso por ejemplo de c.9117+2T>C y de c.9117G>A (p.=), que producen la exclusión del exón 23 del ARNm maduro (Bonatti et al., 2006; Houdayer et al., 2012).

- Variantes que alteran los sitios de *splicing*, produciendo la utilización de un sitio críptico de *splicing* pre-existente. Es el caso por ejemplo de c.8754+4A>G, que induce la utilización del sitio de *splicing* situado en posición c.8754+46 (Bonatti et al., 2006)
- Variantes potenciando un sitio de *splicing* no utilizado previamente, exónico o intrónico como son los casos por ejemplo de c.1763A>G (p.Asn588Ser) (Menendez et al., 2012; Sanz et al., 2010) y de c.8954-15T>G (Menendez et al., 2012).
- Variantes afectando sitios reguladores de *splicing* como por ejemplo c.520C>T (p.Arg174Cys) que induce la exclusión del exón 7 en el transcrito maduro en un 66% de los transcritos portadores de la variante (Gaildrat et al., 2012) o c.145G>T (p.Glu49\*) que induce la exclusión del exón 3 (Sanz et al., 2010).

Estos ejemplos muestran como variantes de cualquier tipo, bien sean intrónicas o exónicas (incluyendo sinónimas, cambios de aminoácido o codones de parada prematuros), pueden producir una alteración de los procesos normales de *splicing*. Es bastante común que, debido a la laboriosidad que conllevan algunos métodos de análisis del *splicing*, las variantes se sometan previamente a un proceso de priorización mediante métodos *in silico*. Este hecho, junto con el sesgo de publicación de resultados negativos produce una aparente sobreestimación de las variantes en *BRCA2* que alteran el *splicing*.

Independientemente de que las alteraciones de *splicing* representen un porcentaje más o menos elevado del conjunto de mutaciones patogénicas en *BRCA2* (dato difícil de estimar en estos momentos), lo cierto es que la demostración de una alteración en el *splicing* permite, en algunos casos, clasificar a una variante directamente como clase 5 (claramente patogénica). Este hecho, convierte al análisis del *splicing* en una herramienta muy relevante desde el punto de vista del consejo genético. Trabajos recientes sugieren para algunas variantes de *BRCA2* relaciones muy complejas entre las alteraciones funcionales del *splicing* y las alteraciones funcionales de la proteína (Acedo et al., 2012). De manera ideal, se debería estudiar la posibilidad de alteraciones en el *splicing* para todas las variantes descritas en *BRCA2*. Según los resultados presentados en esta tesis, la disponibilidad de un método sencillo y adaptado al entorno de trabajo de un laboratorio de diagnóstico molecular permite realizar un análisis de *splicing* en un número elevado de variantes y de manera coste-efectiva, presentándose como única limitación la disponibilidad de muestras. La mayor parte de las variantes analizadas de esta manera podrán ser clasificadas analíticamente siguiendo el sistema propuesto por el grupo de trabajo de *splicing* de IARC. De esta forma, sólo aquellas variantes que requieran métodos más complejos de análisis (menos coste-efectivos) serían seleccionadas según predicciones bioinformáticas. En nuestra experiencia, ese porcentaje de variantes es muy bajo.

---

## Búsqueda de variantes relacionadas con el síndrome HBOC en familias BRCA

*Búsqueda de variantes patogénicas en el gen BRCA2 que no se detectan mediante las técnicas basadas en PCR*

### **Grandes reordenamientos genómicos en el gen BRCA2**

Los resultados presentados en esta tesis demuestran la presencia de reordenamientos genómicos en el gen *BRCA2* en población española en aproximadamente 1% de los casos estudiados. La frecuencia de reordenamientos en el gen *BRCA2* es significativamente inferior a los encontrados en el gen *BRCA1* en familias HBOC, que se aproxima al 10% (de la Hoya et al., 2006). Un resultado interesante presentado en esta tesis es la descripción por primera vez en población española de un reordenamiento en el gen *BRCA2* con efecto fundador. Se trata del reordenamiento g.5670\_6155delinsATA, que ha sido encontrado en más de la mitad de las familias portadoras de reordenamiento en *BRCA2*. El número de familias analizadas en el presente trabajo, únicamente 800, no permite estimar de una manera precisa la frecuencia de este reordenamiento en población española. Convendría realizar el análisis específico de este reordenamiento en una cohorte mayor, para obtener una estimación de su frecuencia más precisa. A modo de ejemplo, el reordenamiento en *BRCA2* c.156\_157insAlu fue descrito por primera vez en una familia de una serie compuesta por 50 familias (Teugels et al., 2005). Tras un siguiente estudio extensivo, a fecha de hoy se ha establecido que c.156\_157insAlu es una variante deletérea fundadora de origen portugués que aparece en más del 25% de los casos portadores de mutación *BRCA1/2* de la población del norte y centro de Portugal (Peixoto et al., 2009). De manera interesante, recientemente se ha descrito en población española también el primer reordenamiento con efecto fundador en el gen *BRCA1* (g.8097\_22733del), que se encuentra presente en el 20% de las familias portadoras de reordenamientos del este de España (Palanca et al., 2013).

Ante la baja prevalencia de reordenamientos en *BRCA2* observados en población española, cabría plantearse si su análisis es coste-efectivo. Los recientes descubrimientos de reordenamientos con efecto fundador en población española en ambos genes plantean la posibilidad de que se analice la presencia específica de los reordenamientos g.5670\_6155delinsATA de *BRCA2* y g.8097\_22733del de *BRCA1*. Los datos publicados hasta el momento sobre individuos portadores de reordenamientos en población española se encuentran resumidos en la página 123 (*BRCA2*) y en la Tabla 3 (*BRCA1*). En la literatura científica se han descrito 54 familias españolas portadoras de reordenamientos, 39 en el gen *BRCA1* y 15 en el gen *BRCA2*. De ellas, son portadoras de reordenamiento con efecto fundador, 4 familias en *BRCA1* y 5 en *BRCA2* (una de ellas pendiente de

confirmar los puntos de corte del reordenamiento). Ambos reordenamientos representan un porcentaje muy bajo del total de reordenamientos descritos en población española que se aproxima al 15 %, por lo que en principio, la búsqueda específica de ambos reordenamientos no parece una estrategia efectiva. Sin embargo, g.5670\_6155delinsATA representa un porcentaje muy elevado de todos los reordenamientos encontrados en el gen *BRCA2*, superior al 30%. Debido a la baja prevalencia de reordenamientos en *BRCA2* observados en población española y a la alta prevalencia de g.5670\_6155delinsATA dentro de los reordenamientos detectados en este gen, se sugiere como estrategia coste-efectiva la búsqueda específica de este reordenamiento. Tras haberse descartado la presencia de mutaciones puntuales en los genes *BRCA1/2* y la presencia de reordenamientos en *BRCA1* mediante el uso de MLPA, se comprobaría la presencia de g.5670\_6155delinsATA mediante el uso de PCR convencional y en aquellas muestras negativas se realizaría el MLPA de *BRCA2*. La valoración de la efectividad de esta estrategia depende de estudios que analicen en una población de mayor tamaño la frecuencia de cada uno de los dos reordenamientos con efecto fundador descritos en población española y de la posible confirmación de un efecto fundador en ciertos LGR que aparecen de manera recurrente como Del E1-E2 y Del E8-E13 en el gen *BRCA1*.

**Tabla 3:** Reordenamientos genómicos en el gen *BRCA1* descritos en población española.

Exones	N	nomenclatura HGVS <sup>A</sup>	Referencias
Del E1-E2	5	g.85341_99155del	(de Juan Jimenez et al., 2013; Palanca Suela et al., 2008)
Del E1-E2	1	-	(del Valle et al., 2010)
Dup E1-E2	1	-	(del Valle et al., 2010)
Del E1-E24	1	-	(de la Hoya et al., 2006)
Del E1-E24	1	-	(Blay et al., 2013)
Del E1-E24	1	-	(de Juan Jimenez et al., 2013)
Del E1-E13	1	-	(del Valle et al., 2010)
Del E1-E13	1	-	(Blay et al., 2013)
Del E3	1	-	(de Juan Jimenez et al., 2013)
Del E3-E5	4	g.97346_111983del	(Palanca et al., 2013)
Del E5-E7	1	g.110447_115470del	(Palanca Suela et al., 2008)
Del E8	1	-	(de Juan Jimenez et al., 2013)
Del E8-E13	4	g.116216_139982del	(de Juan Jimenez et al., 2013; de la Hoya et al., 2006; Palanca Suela et al., 2008)
Del E8-E13	1	-	(Miramar et al., 2008)
Dup E9-E24	1	-	(del Valle et al., 2010)
Del E11-E15	1	g.122702_145814del	(de la Hoya et al., 2006)
Del E14	2	g.137093_142042del	(de la Hoya et al., 2006)
Del E16-E23	1	g.145024_171817del	(del Valle et al., 2010)
Dup E19-E20	1	-	(de la Hoya et al., 2006)
Del E20	1	g.160599_164583del	(Palanca Suela et al., 2008)
Del E20	1	-	(Blay et al., 2013)
Del E20	1	g.[160897_163720del; 163847_163878delinsGG]	(Del Valle et al., 2011)
Dup E20	2	-	(de la Hoya et al., 2006)
Del E22	1	g.168592_168844del	(del Valle et al., 2010)
Amplificación E21-E22	2	-	(de la Hoya et al., 2006)
Del E23-E24	1	-	(de la Hoya et al., 2006)

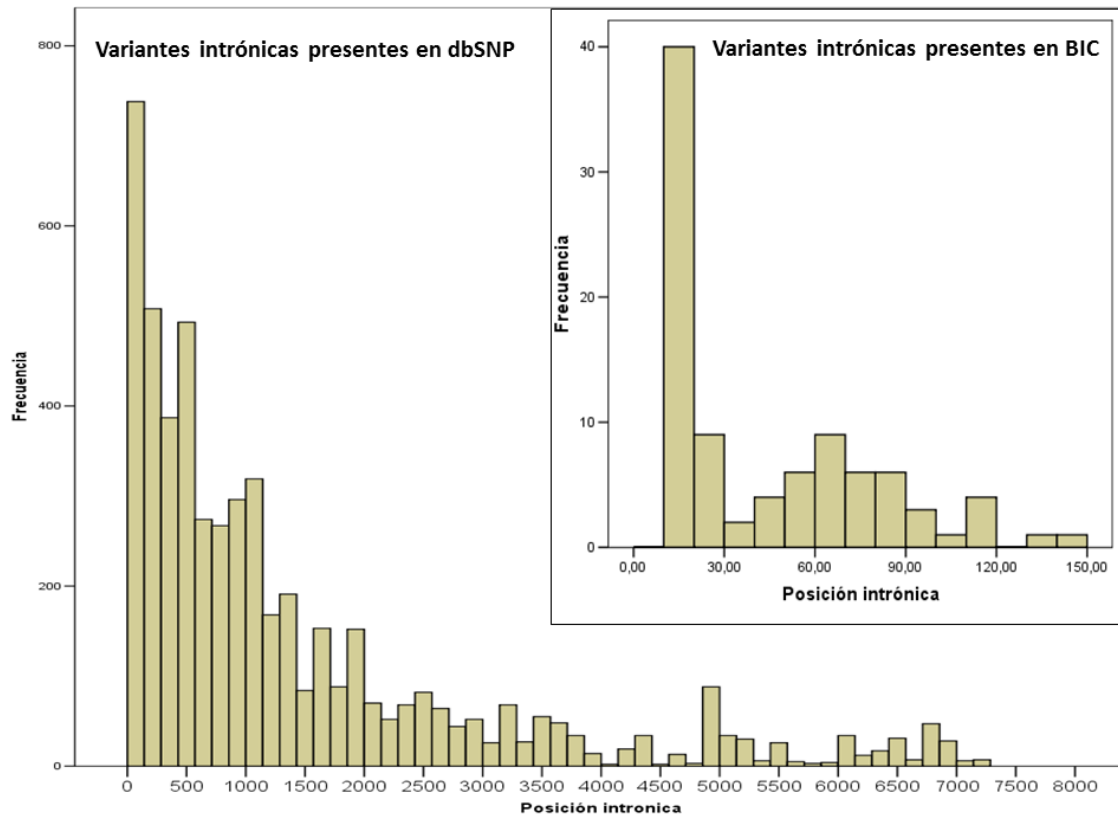
### Variantes intrónicas alejadas del exón

Parte del intrón12 de *BRCA2* fue analizado en 302 casos índices de familias BRCA. No se detectó la presencia de la variante patogénica c.6937+594T>G, en su lugar se detectó la variante de significado incierto c.6937+565T>C en un 1 % de los individuos. La detección de esta variante de significado incierto plantea las siguientes preguntas: ¿Cuál es la variabilidad genética en los intrones de *BRCA2*?, ¿Qué parte de esta variabilidad se encuentra pendiente de descubrir en individuos con HBOC?, ¿Cuál es la mejor metodología para analizar esta variabilidad?

Recientemente se ha publicado una revisión (Romano et al., 2013) que analiza el papel de los pseudoexones en cáncer. Se ha descrito la activación de pseudoexones por mutaciones en genes relacionados con síndromes de cáncer hereditario como *APC*, *ATM*, *BRC-ABL*, *ESR1*, *MSH2*, *NF1*, *NF2*, *RB1*, *TSC2* y *BRCA2*. Por lo general, la implicación de pseudoexones y la presencia de mutaciones *deep intronic* no se analizan de manera sistemática en la mayoría de estos genes, estos hallazgos se tratan de descubrimientos esporádicos. Como excepción, el gen *NF1* se ha analizado de manera sistemática a nivel de ADN y de ARN (Pros et al., 2008). En este gen se han encontrado un porcentaje elevado (45%) de variantes que producen alteraciones del *splicing*, sin embargo únicamente un 5 % de las alteraciones de *splicing* son debidas a variantes *deep intronic*.

En la base de datos BIC (último acceso Septiembre de 2013) se han detectado 92 variantes intrónicas alejadas del exón (> de 10 pares de bases). En la figura 12 se muestra la distribución de estas variantes ordenadas por su posición en el intrón y se compara con la distribución de las variantes intrónicas en el mismo gen descritas en la base de datos dbSNP (último acceso Septiembre de 2013). Las posiciones intrónicas incluidas en la base de datos BIC cubren únicamente un 15% de la variabilidad genética presente en dbSNP. Este dato demuestra que la mayor parte de la variabilidad genética de los intrones de *BRCA2* permanece inexplorada en el contexto del síndrome HBOC.

El desarrollo de nuevos métodos de ultrasecuenciación y su popularización progresiva en los laboratorios de diagnóstico genético está originando un cambio en los métodos de cribado. Se espera que la aplicación de técnicas de ultrasecuenciación produzca una reducción de costes en el cribado genético de *BRCA1/2* así como una reducción en la carga de trabajo relacionada con la simplificación y automatización de los procesos (Tarabeux et al., 2013). La reducción en costes y en carga de trabajo, sin embargo, puede ser desbordada por la necesidad de dotar de interpretación clínica a las nuevas variantes descritas. Alguno de los métodos basados en la ultrasecuenciación propuestos para el análisis de *BRCA1/2* incluyen la secuenciación de ambos loci completos, incluyendo la totalidad de las regiones intrónicas (Hernan et al., 2012; Walsh et al., 2010).



**Figura 12:** Localización de las variantes intrónicas en el gen BRCA2 presentes en las base de datos BIC y dbSNP.

La aplicación de este tipo de técnicas podría ayudar a conocer la auténtica variabilidad genética de los intrones de *BRCA1/2*. Sin embargo, la tendencia generalizada es optar por aquellos métodos de ultrasecuenciación que se basan en el enriquecimiento y secuenciación de los exones y regiones limítrofes de *BRCA1/2*. Esto es debido a que el análisis completo de los intrones produciría una mayor carga a la hora de analizar los resultados y posiblemente un incremento en el número de UVs reportadas, que no se verían compensados con un incremento significativo en la detección de variantes patogénicas.

Parece mucho más factible analizar la presencia de mutaciones *deep intronic* mediante el análisis sistemático y en paralelo de muestras de ADN y ARN. Esta posibilidad debería ser analizada desde un punto de vista coste-efectivo, ya que permitiría la identificación de mutaciones *deep intronic* así como el análisis sistemático de alteraciones en el *splicing* de cualquier UV detectada. Sin duda, aumentará la posibilidad de descubrir nuevas variantes intrónicas alejadas de los exones, siendo posible que alguna de estas nuevas variantes produzca alteraciones en el *splicing* de manera similar a lo observado para c.6937+594T>G (Anczukow et al., 2012).

---

*Sobre la búsqueda de variantes patogénicas implicadas en HBOC en genes candidatos adicionales a BRCA1/2*

#### **SLX4**

En el año 2009 cuatro estudios centraron la atención en SLX4 y demostraron su papel en la recombinación homóloga mediando la resolución de los intermediarios de Holliday (Andersen et al., 2009; Fekairi et al., 2009; Munoz et al., 2009; Svendsen et al., 2009). Dos estudios del año 2011 encontraron un total de 6 individuos con anemia de Fanconi portadores bialelicos de mutaciones inactivantes de *SLX4* (*FANCP*) (Kim et al., 2011; Stoepker et al., 2011). Un estudio posterior aplicó la secuenciación del exoma a un paciente de anemia de Fanconi, identificando también mutaciones bialelicas en *SLX4* (Schuster et al., 2013).

En la actualidad seis estudios (incluido el presentado en esta tesis) han analizado *SLX4* como gen candidato en la susceptibilidad al cáncer de mama y ovario familiares (Bakker et al., 2013; Catucci et al., 2012; de Garibay et al., 2012; Fernandez-Rodriguez et al., 2012; Landwehr et al., 2011; Shah et al., 2013). En conjunto se han analizado 2625 casos HBOC no explicados por mutaciones en *BRCA1/2* y se han encontrado tres mutaciones truncantes en *SLX4*. Dos mutaciones introducen un codón de parada prematuro, c.2469G>A p.W823\* y c.4549G>T p.Glu1517\*); por su parte c.2013+2T>A produce una alteración del *splicing* que elimina el exón 8 del ARNm maduro, induciendo un cambio en la fase de lectura. En el caso de p.W823\* se ha demostrado que no es capaz de complementar la sensibilidad a agentes que producen daño a ADN Mitomicina C, Camptotecina y Olaparib de la línea RA3331 de fibroblastos humanos *SLX4*<sup>-/-</sup> (Shah et al., 2013), evidenciándose de esta manera un defecto funcional asociado a p.W823\*.

Además de las tres mutaciones truncantes en el gen *SLX4*, se han identificado otras 197 VUs, (ver anexo II). 88 de estas variantes han sido descritas por primera vez en el contexto de familias con síndrome HBOC, incluyendo 45 cambios de aminoácido. De estos, 17 se predicen como deletéreos para la estructura de la proteína según los algoritmos SIFT y Polyphen (Tabla 4). No se ha observado defectos funcionales para ninguna de las siete variantes testadas en ensayos de complementación (Tabla 4).

Para establecer una relación causal entre las variantes del gen *SLX4* y el síndrome HBOC se requieren evidencias adicionales como estudios de asociación o estudios de co-segregación. Cuatro de los seis estudios realizados hasta el momento han analizado un total de 16 variantes raras en 2421 controles. Cada estudio ha analizado en controles únicamente una selección de variantes, lo cual hace que el número de individuos analizados para cada variante sea de media aproximada 400

controles y no sobrepase en ningún caso los 1000. Debido a estos números tan bajos, es muy difícil realizar un estudio caso control clásico.

El proyecto ESP6500, ha analizado un número superior de controles que incluye 6503 individuos. Se ha encontrado una frecuencia de individuos portadores de cambios que alteran la fase de lectura del gen *SLX4* de 0,15%. Esta frecuencia no difiere estadísticamente de la frecuencia de mutaciones inactivantes en el gen *SLX4* encontrada en los casos HBOC, 0,11%. Sin embargo el análisis detallado de los resultados de ESP6500 muestra que nueve de las diez variantes encontradas en controles colocalizan con regiones de homopolímeros, pudiendo tratarse de falsos positivos. Existen un total de 174 variantes raras (aquellas con un frecuencia inferior al 1%) en la población de casos HBOC, de ellas 74 están presentes en la población ESP6500. El estudio de la frecuencia de portadores de variantes raras en casos y controles sugiere una frecuencia inferior de variantes raras en la población HBOC (12%) que en la población control (21%) (O.R=0,27 IC=0,24-0,30  $p<0,0001$ ). La frecuencia de portadores de variantes raras con predicción bioinformática de alterar la estructura de la proteína sigue la misma tendencia (O.R=0,39 IC=0,29-0,51  $p<0,0001$ ). Ambos datos apuntan a que en general, no existe un exceso de variantes raras en el gen *SLX4* en la población de casos HBOC. Existe la posibilidad de que sin embargo, algunas de las variantes descritas en el gen *SLX4* sí que afecten a la función de la proteína *SLX4* y se encuentren sobrerrepresentadas dentro de los casos HBOC. De hecho, Shah et al demostraron que los individuos portadores de cambios de aminoácido en *SLX4* presentaban un adelanto en la edad media de diagnóstico del cáncer de mama de dos años en comparación con los individuos no portadores, de 49 a 47 años (Shah et al., 2013). Diferenciar qué variantes en concreto, alteran la función de *SLX4* es complejo, ya que los resultados obtenidos hasta el momento sugieren que la predicción mediante algoritmos bioinformáticos no presenta correlación con los resultados experimentales. ,

La descripción de variantes inactivantes del gen *SLX4* en familias HBOC apunta hacia la implicación del *SLX4* en el síndrome, sin embargo, los datos disponibles hasta el momento no son suficientes para establecer un relación causal. Por otro lado, el estudio de segregación presentado por Bakker et al para c.2013+2T>A sugiere una penetrancia intermedia (Bakker et al., 2013). Se requerirá el análisis de una población de un tamaño superior, así como la realización de ensayos complementarios para poder dilucidar el papel de las mutaciones en *SLX4* en el síndrome HBOC.

**Tabla 4:** Predicciones y estudios funcionales en las variantes de SLX4 descritas por primera vez en casos HBOC. La tabla muestra las predicciones bioinformáticas de cambios en la estructura y en el proceso de splicing de las 59 variantes de SLX4 descritas por primera vez en casos HBOC, así mismo se muestra el resultado de ensayos funcionales en el caso que éstos hayan sido realizados. En la columna de predicción de splicing, los datos en negrita se corresponden con variantes que podrían alterar los sitios originales de splicing, mientras que los demás datos se corresponden con aquellas variantes que podrían generar nuevos sitios críticos de splicing. Las predicciones del algoritmo Polyphen-2 se dividen en tres clases. En orden de menor a mayor probabilidad de alteración son benigna, posible deletéreo y probable deletéreo (ambas fueron consideradas como predicción positiva de alteración). Aquellas variantes claramente deletéreas aparecen subrayadas.

Exón	Cambio de nucleótido	Cambio de aminoácido	Predicciones Bioinformáticas				Estudios Funcionales	
			Predicciones estructurales		Predicción de <i>splicing</i>		Ensayos de complementación	Análisis del <i>Splicing</i>
			SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)		
2	c.-31G>T							
2	c.248G>C	p.Gly83Ala	Deletéreo	Benigno				
2	c.376A>G	p.Thr126Ala	Tolerada	Benigno	75%			
2	c.497A>C	p.Gln166Pro	Tolerado	Benigno				
3	c.760+10C>G							
4	c.851A>G	p.His284Arg	Tolerado	Benigno	83%	33%		
4	c.950+29C>T							
5	c.1114C>T	p.Arg372Trp	Deletéreo	Probable Deletéreo	88%	44%		
5	c.1132C>A	p.Pro378Arg	Tolerado	Posible Deletéreo			Normal	
5	c.1163+194C>T							
6	c.1164-16T>C							
6	c.1164-40C>A							
6	c.1271C>T	p.Ala424Val	Deletéreo	Probable Deletéreo	86%			
7	c.1419C>T	p. =			88%	48%		
7	c.1513G>A	p.Ala505Thr	Tolerado	Benigno		73%		
7	c.1517G>A	p.Ser506Asn	Deletéreo	Probable Deletéreo		94%		
8	c.1736T>C	p.Leu579Pro	Tolerado	Benigno				
8	c.1845C>T	p. =						
8	c.1924+5G>A				<b>87%</b>	<b>54%</b>		
8	c.1924+59C>T							
8	c.1924+127G>A							
9	c.1925-122C>T							
9	c.1925-106T>G							
9	c.1925G>T	p.Gly642Val	Deletéreo	Probable Deletéreo	<b>95%</b>	<b>85%</b>	Normal	
9	<u>c.2013+2T&gt;A</u>				<b>70%</b>	<b>7%</b>		Exclusión Exon 9
9	c.2013+130C>G							
10	c.2098G>A	p.Gly700Arg	Deletéreo	Probable Deletéreo	89%			
12	c.2328-35del13							
12	c.2346C>T	p. =				88%		
12	c.2437G>A	p.Glu813Lys	Deletéreo	Probable Deletéreo				
12	c.2469G>C	p.Trp823Cys	Deletéreo	Probable Deletéreo	87%	25%		

Exón	Cambio de nucleótido	Cambio de aminoácido	Predicciones Bioinformáticas				Estudios Funcionales	
			Predicciones estructurales		Predicción de <i>splicing</i>		Ensayos de complementación	Análisis del <i>Splicing</i>
			SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)		
12	<u>c.2469G&gt;A</u>	<u>p.Trp823*</u>			101%	62%	Aberrante	
12	c.3034A>G	p.Arg1012Gly	Tolerado	Benigno				
12	c.3337G>C	p.Gly1113Arg	Deletéreo	Benigno	80%	59%	Normal	
12	c.3860T>G	p.Val1287Gly	Tolerado	Benigno				
12	c.3872C>T	p.Thr1291Met	Deletéreo	Benigno				
12	c.3913G>A	p.Ala1305Thr	Tolerado	Benigno	81%			
12	c.4272C>T	p. =						
12	c.4333C>T	p.Arg1445Trp	Deletéreo	Probable Deletéreo				
12	c.4403G>A	p.Arg1468His	Tolerado	Benigno	79%			
12	c.4410G>A	p. =			99%	113%		
12	<u>c.4549G&gt;T</u>	<u>p.Glu1517*</u>			110%	118%		
12	c.4597GC>TA	p.Ala1533Thr	Tolerado	Benigno				
13	c.4637-125C>T							
13	C.4637-210A>					41%		
13	C.4637-234G>A							
13	c.4637-267C>T							
13	c.4739+10C>T							
13	c.4739+31G>A							
13	c.4739+90C>G							
14	c.4797C>T	p. =						
14	c.4887C>T	p. =						
14	c.5072A>G	p.Asn1691Ser	Tolerado	Benigno	88%			
14	c.5081C>T	p.Ala1694Val	Deletéreo	poss. Deletéreo	86%	13%		
14	c.5150A>C	p.Gln1717Pro	Deletéreo	Probable Deletéreo			Normal	
15	c.5189C>T	p.Ser1730Phe	Deletéreo	poss. Deletéreo		15%		
15	c.5389C>T	p. =				100%		
15	c.5440C>T	p.Arg1814Cys	Deletéreo	Probable Deletéreo			Normal	
15	c.*102G>A							

**Tabla 5:** Descripción de las variantes raras en SLX4 presentes en las series de HBOC y en ESP6500.

La tabla muestra la frecuencia y las predicciones bioinformáticas de cambios en la estructura y en el proceso de splicing de las 74 variantes raras de SLX4 presentes en ESP6500. En la columna de predicción de splicing los datos en **negrita** se corresponden con variantes que podrían alterar los sitios originales de splicing mientras que los demás datos se corresponden con aquellas variantes que podrían generar nuevos sitios críticos de splicing. Fueron consideradas predicciones positivas aquellas en las que los algoritmos SIFT y POLYPHEN de manera simultánea apuntan a un efecto deletéreo. También se consideró predicción positiva el marcado incremento del score del sitio crítico de splicing detectado por los algoritmos MaxEnt y HSF para c.833G>A.

Exón	Cambio de nucleótido	Cambio de aminoácido	Total N= 4289		ESP6500 n= 6503		Predicciones Bioinformáticas				Estudios Funcionales
			Portadores	Frecuencia	Portadores	Frecuencia	Predicciones estructurales		Predicción de <i>splicing</i>		
							SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)	
2	c.85C>T	p.Arg29Cys	1	0,0004	2	0,0003	Deletéreo	Benigno			
2	c.86G>A	p.Arg29His	6	0,0023	3	0,0005	Tolerado	Benigno			
2	c.112C>T	p.Leu38Phe	2	0,0008	1	0,0002	Deletéreo	Benigno			
2	c.145G>A	p.Glu49Lys	1	0,0004	1	0,0002	Deletéreo	Posible Deletéreo	77%	16%	
2	c.237T>C	p. =	2	0,0008	1	0,0002					
2	c.339T>C	p. =	1	0,0004	14	0,0022					
2	c.421G>T	p.Gly141Trp	9	0,0034	10	0,0015	Deletéreo	Probable Deletéreo	83%	27%	
2	c.465A>G	p. =	2	0,0008	2	0,0003					
3	c.590T>C	p.Val197Ala	12	0,0046	27	0,0042	Tolerado	Benigno	77%	31%	
3	c.710G>A	p.Arg237Gln	17	0,0065	120	0,0185	Tolerado	Benigno	92%		
3	c.742G>A	p.Glu248Lys	1	0,0004	6	0,0009	Deletéreo	Posible Deletéreo	80%		
4	c.832C>T	p.Arg278Trp	1	0,0004	3	0,0005	Deletéreo	Probable Deletéreo			
4	c.833G>A	p.Arg278Gln	1	0,0004	1	0,0002	Tolerado	Benigno	101%	141%	
5	c.999C>T	p. =	2	0,0008	116	0,0179					
5	c.1065G>A	p. =	1	0,0004	1	0,0002					
5	c.1077G>T	p.Lys359Asn	1	0,0004	1	0,0002	Deletéreo	Benigno	76%	7%	

Exón	Cambio de nucleótido	Cambio de aminoácido	Total N= 4289		ESP6500 n= 6503		Predicciones Bioinformáticas				Estudios Funcionales
			Portadores	Frecuencia	Portadores	Frecuencia	Predicciones estructurales		Predicción de <i>splicing</i>		Ensayos de complementación
							SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)	
5	c.1153C>A	p.Pro385Thr	18	0,0069	100	0,0154	Tolerado	Benigno			
6	c.1243G>A	p.Glu415Lys	4	0,0015	1	0,0002	Deletéreo	Posible Deletéreo	84%	25%	
7	c.1372A>G	p.Lys458Glu	3	0,0011	1	0,0002	Deletéreo	Probable Deletéreo		4%	
7	c.1637A>G	p.Tyr546Cys	1	0,0004	66	0,0102	Deletéreo	Probable Deletéreo			
8	c.1702G>A	p.Val568Met	1	0,0004	1	0,0002	Tolerado	Benigno			
8	c.1755C>T	p. =	8	0,0030	55	0,0085					
8	c.1755C>A	p. =	1	0,0004	55	0,0085			89%		
8	c.1803G>A	p. =	8	0,0030	56	0,0086			66%		
8	c.1898G>A	p.Gly633Asp	3	0,0011	88	0,0135	Tolerado	Benigno			
9	c.1925-30G>A		2	0,0008	62	0,0095					
9	c.1991C>T	p.Pro664Leu	1	0,0004	1	0,0002	Tolerado	Benigno			
9	c.2013+21G>T		1	0,0004	96	0,0148					
10	c.2014-28G>A		1	0,0004	3	0,0005					
11	c.2305G>C	p.Glu769Gln	5	0,0019	19	0,0029	Tolerado	Benigno	90%	9%	
12	c.2328-9G>C		1	0,0004	33	0,0051					
12	c.2359G>A	p.Glu787Lys	16	0,0061	15	0,0023	Tolerado	Posible Deletéreo		86%	Normal
12	c.2609C>T	p.Ala870Val	4	0,0015	8	0,0012	Tolerado	Posible Deletéreo	86%	15%	
12	c.2654C>T	p.Pro885Leu	1	0,0004	6	0,0009	Deletéreo	Posible Deletéreo			
12	c.2667A>G	p. =	2	0,0008	9	0,0014			92%	70%	
12	c.2681T>G	p.Val894Gly	2	0,0008	4	0,0006	Tolerado	Benigno			
12	c.2824G>C	p.Glu942Gln	16	0,0061	224	0,0345	Tolerado	Probable Deletéreo	85%		
12	c.2854G>A	p.Ala952Thr	6	0,0023	316	0,0486	Tolerado	Benigno			
12	c.2924C>T	p.Pro975Leu	23	0,0088	67	0,0103	Tolerado	Benigno			

Exón	Cambio de nucleótido	Cambio de aminoácido	Predicciones Bioinformáticas								Estudios Funcionales
			Total N= 4289		ESP6500 n= 6503		Predicciones estructurales		Predicción de <i>splicing</i>		
			Portadores	Frecuencia	Portadores	Frecuencia	SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)	
12	c.2975G>A	p.Gly992Glu	3	0,0011	4	0,0006	Tolerado	Posible Deletéreo			
12	c.3019C>A	p.Gln1007Lys	4	0,0015	4	0,0006	Tolerado	Benigno			
12	c.3101T>C	p.Arg1034His	1	0,0004	2	0,0003	Tolerado	Benigno			
12	c.3109T>C	p. =	2	0,0008	2	0,0003					
12	c.3178C>T	p.Arg1060Trp	10	0,0038	9	0,0014	Deletéreo	Benigno			
12	c.3189C>T	p. =	12 (3)	0,0057	3	0,0005			83%	49%	
12	c.3315C>T	p. =	1	0,0004	1	0,0002			82%	41%	
12	c.3316G>A	p.Val1106Met	1	0,0004	1	0,0002	Tolerado	Benigno			
12	c.3368C>A	p.Ser1123Tyr	2	0,0008	11	0,0017	Deletéreo	Posible Deletéreo			
12	c.3676C>T	p.Arg1226Trp	2	0,0008	20	0,0031	Deletéreo	Probable Deletéreo			
12	c.3783G>A	p. =	8	0,0030	258	0,0397					
12	c.3857C>T	p.Ala1286Val	6	0,0023	29	0,0045	Tolerado	Benigno	91%	14%	
12	c.3868C>A	p.His1290Asn	1	0,0004	20	0,0031	Tolerado	Benigno			
12	c.3939C>A	p. =	1	0,0004	1	0,0002					
12	c.3963G>A	p. =	4	0,0015	84	0,0129					
12	c.4024A>G	p.Ser1342Gly	3	0,0011	1	0,0002	Deletéreo	Probable Deletéreo	77%		
12	c.4057C>A	p.His1353Asn	1	0,0004	6	0,0009	Tolerado	Benigno			
12	c.4068G>A	p. =	1	0,0004	51	0,0078					
12	c.4241C>T	p.Pro1414Leu	1	0,0004	27	0,0042	Tolerado	Benigno			
12	c.4261A>T	p.Ile1421Phe	5	0,0019	8	0,0012	Deletéreo	Posible Deletéreo			
12	c.4338C>T	p.Thr1446Ser	1	0,0004	112	0,0172	Deletéreo	Posible Deletéreo			
12	c.4344C>T	p. =	1	0,0004	1	0,0002					
12	c.4409C>T	p.Pro1470Leu	4	0,0015	2	0,0003	Tolerado	Posible Deletéreo		27%	

Exón	Cambio de nucleótido	Cambio de aminoácido	Total N= 4289		ESP6500 n= 6503		Predicciones Bioinformáticas				Estudios Funcionales
			Portadores	Frecuencia	Portadores	Frecuencia	Predicciones estructurales		Predicción de <i>splicing</i>		Ensayos de complementación
							SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)	
12	c.4427C>G	p.Thr1476Ser	1	0,0004	1	0,0002	Tolerado	Benigno	87%		
12	c.4494G>A	p. =	1	0,0004	1	0,0002			86%	59%	
12	c.4530G>T	p. =	3	0,0011	2	0,0003					
12	c.4580C>T	p.Pro1527Leu	1	0,0004	1	0,0002	Tolerado	Benigno		60%	
12	c.4581G>A	p. =	14	0,0053	261	0,0402			97%	16%	
12	c.4600G>A	p.Gly1534Ser	5	0,0019	193	0,0278	Tolerado	Benigno	76		
13	c.4648C>T	p.Arg1550Trp	10	0,0038	28	0,0040	Deletéreo	Benigno			Normal
13	c.4649G>A	p.Arg1550Gln	1	0,0004	1	0,0002	Tolerado	Benigno	89%		
14	c.5029C>T	p.Pro1677Ser	2	0,0008	51	0,0078	Tolerado	Benigno	93%	38%	
14	c.5040G>C	p.Arg1680Ser	1	0,0004	1	0,0001	Tolerado	Benigno			
14	c.5146T>A	p.Ser1716Thr	3	0,0011	114	0,0164	Deletéreo	Probable Deletéreo			
15	c.5501A>G	p.Asn1834Ser	24	0,0091	74	0,0077	Tolerado	Benigno	70%		
Total de variantes raras			121	0,126	2980	0,452					
Variantes con predicción bioinformática positiva			38	0,014	361	0,054					

## RAD51D

En la actualidad siete estudios (incluido el presentado en el anexo de esta tesis) han analizado *RAD51D* como gen candidato en la susceptibilidad a cáncer de mama y ovario tanto familiares como esporádicos (Golmard et al., 2013; Loveday et al., 2011; Osher et al., 2012; Pelttari et al., 2012; Thompson et al., 2013; Wickramanyake et al., 2012). Se han analizado un total de 8034 casos, de los cuales 5026 han sido analizados mediante estudios de resecuenciación y 3008 han sido analizados para la presencia de la variante deletérea fundadora de origen finlandés c.576+1G>A. Se han identificado un total de 59 variantes en *RAD51D*, de las cuales 14 son variantes deletéreas (ver Tabla 10). Se han descrito un total de 25 portadores de variantes deletéreas (0,31%), 17 en familias HBOC y 8 en cáncer esporádicos. La variante fundadora finlandesa c.576+1G>A representa siete de estos casos, además, otras tres variantes aparecen de manera recurrente: c.556C>T p.Arg186\* (n=4), c.694C>T p.Arg232\* (n=3) y c.803G>A p.Trp268\* (n=2). Se desconoce si se trata también de mutaciones fundadoras o si bien son casos de puntos calientes mutacionales.

En estos estudios se han analizado además un total de 3628 controles, de los cuales 1060 han sido analizados mediante resecuenciación, 466 mediante resecuenciación sólo de aquellos exones con presencia de nuevas variantes no sinónimas en casos HBOC y 2102 han sido analizados para la presencia de la mutación fundadora de origen finlandés. En total se han descrito tres individuos portadores de variantes deletéreas en población control (0,08). La variante c.748delC se ha detectado únicamente en controles mientras que c.576+1G>A y c.270\_271dup aparecen también en casos. Debido a que en la mayoría de los controles no se ha resecuenciado la totalidad del gen, no es posible realizar un estudio caso control de variantes deletéreas. Como alternativa, se ha realizado un estudio caso control comparando la frecuencia de variantes deletéreas en los 5026 casos estudiados mediante resecuenciación en comparación con los 6503 controles de la población de ESP6500. En los casos se identificaron 20 variantes deletéreas (0,39%) y en controles 3 (0,04%),. La diferencia en frecuencia es significativa estadísticamente y confiere una OR=8,6 (IC 2,5-29,1 p=0,005). De las 59 variantes identificadas en *RAD51D* (Anexo II), 30 han sido detectadas tanto en casos como en controles de los propios estudios o de las poblaciones ESP6500 y 1000 genomas. 29 variantes se han descrito únicamente en casos, de ellas, se presenta en la Tabla 6 un resumen de las predicciones bioinformáticas funcionales y de *splicing* y de los resultados experimentales.

**Tabla 6 :** Predicciones y estudios funcionales en las variantes de RAD51D descritas por primera vez en casos HBOC y esporádicos

La tabla muestra las predicciones bioinformáticas de cambios en la estructura y en el proceso de splicing de las 29 variantes de RAD51D descritas por primera vez en casos HBOC y esporádico, así mismo se muestra el resultado de los estudios experimentales de splicing en el caso que éstos hayan sido realizados. En la columna de predicción de splicing los datos en negrita se corresponden con variantes que podrían alterar los sitios originales de splicing, mientras que los demás datos se corresponden con aquellas variantes que podrían generar nuevos sitios crípticos de splicing. Las predicciones del algoritmo Polyphen-2 se dividen en tres clases, en orden de menor a mayor probabilidad de alteración son benigno, posible deletéreo y probable deletéreo (ambas fueron consideradas como predicción positiva de alteración). Aquellas variantes claramente deletéreas aparecen subrayadas.

Exón	Cambio de nucleótido	Cambio de aminoácido	Predicciones Bioinformáticas				Análisis del Splicing
			Predicciones estructurales		Predicción de splicing		
			SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)	
5'UTR	c.-111G>A						
5'UTR	c.-60C>T						
5'UTR	c.-39C>T						
1	<u>c.1A&gt;T</u>	<u>p.Met1?</u>					
2	c.83-4T>C					<b>94,8%</b>	
2	c.117A>T	p.=			84,5%		
2	<u>c.131_144+24del38</u>	<u>p.Val28Glyfs*21</u>			<b>7,6%</b>	<b>0%</b>	Delección exón 2
2	c.137C>G	p.Ser46Cys	Deletéreo	Probable deletéreo			
3	c.211C>T	p.Leu71Phe	Deletéreo	Probable deletéreo	89,2%		
4	c.264-6C>T				<b>89,2%</b>		
4	c.264-50_264-49del					6,10%	
4	c.286G>T	p.Gly96Cys	Deletéreo	Probable deletéreo		11,50%	
4	c.292G>A	p.Asp98Asn					
5	<u>c.345G&gt;C</u>	<u>p.Gln115His</u>	Deletéreo	Probable deletéreo	<b>86%</b>	<b>0%</b>	Delección exones 3 y 4
5	<u>c.363delA</u>	<u>p.Ala122Glnfs*14</u>					
5	<u>c.480+1G&gt;A</u>				<b>72,60%</b>	<b>0%</b>	
6	<u>c.556C&gt;T</u>	<u>p.Arg186*</u>					
7	<u>c.580delA</u>	<u>p.Thr194Leufs*6</u>			99,7%	48%	
7	c.629C>T	p.Ala210Val	Deletéreo	Probable deletéreo			
7	c.630G>A	p.=			92,80%	22,10%	
7	<u>c.667+2_667+23del</u>	<u>p.Val193Alafs*4</u>			<b>48,90%</b>	<b>0%</b>	Delección del exón 7
8	<u>c.694C&gt;T</u>	<u>p.Arg232*</u>			74,70%		
8	c.715C>T	p.Arg239Trp	Deletéreo	Probable deletéreo			
8	c.738+54T>C					6%	
9	c.753A>G	p.Ile251Met	Benigno	Benigno	85,70%		
9	<u>c.757C&gt;T</u>	<u>p.Arg253*</u>					
9	c.796C>T	p.Arg266Cys	Deletéreo	Probable deletéreo	80,70%		
9	<u>c.803G&gt;A</u>	<u>p.Trp268*</u>			77,10%	43%	
9	c.911G>A	p.Gly304Asp	Deletéreo	Benigno			

**Tabla 7:** Descripción de las variantes raras en RAD51D presentes en las series de HBOC/casos esporádicos y en ESP6500.

La tabla muestra la frecuencia y las predicciones bioinformáticas de cambios en la estructura y en el proceso de splicing de las 12 variantes raras de RAD51D presentes en ESP6500. En la columna de predicción de splicing los datos en **negrita** se corresponden con variantes que podrían alterar los sitios originales de splicing mientras que los demás datos se corresponden con aquellas variantes que podrían generar nuevos sitios críticos de splicing. Fueron consideradas predicciones positivas aquellas en las que los algoritmos SIFT y POLYPHEN de manera simultánea apuntan a un efecto deletéreo. También se consideró predicción positiva el marcado descenso del score del sitio aceptor de splicing detectado por el algoritmo MaxEnt para c.904-11T>A.

Exón	Cambio de nucleótido	Cambio de aminoácido	Predicciones Bioinformáticas							
			Total N= 4289		ESP6500 n= 6503		Predicciones estructurales		Predicción de <i>splicing</i>	
			Portadores	Frecuencia	Portadores	Frecuencia	SIFT	Polyphen-2	HSF (% sitio original)	MaxEnt (% sitio original)
1	c.26G>C	p.Cys9Ser	5	0,0012	9	0,001	Deletéreo	Posible Deletéreo	-	-
3	c.185C>T	p.Ser62Leu	1	0,0002	1	0,0002	Deletéreo	Posible Deletéreo	-	-
5	c.355T>C	p.Cys119Arg	2	0,0005	1	0,0002	Tolerado	Posible Deletéreo	-	14,70%
5	c.394G>A	p.Val132Ile	1	0,0002	1	0,0002	Deletéreo	Benigno	-	-
7	c.620C>T	p.Ser207Leu	3	0,0007	2	0,0003	Deletéreo	Posible Deletéreo	-	-
8	c.695G>A	p.Arg232Gln	7	0,0016	59	0,009	Tolerado	Posible Deletéreo	77,10%	-
9	c.793G>A	p.Gly265Arg	2	0,0005	2	0,0003	Deletéreo	Posible Deletéreo	90,50%	53,80%
9	c.872G>A	p.Arg291His	1	0,0002	1	0,0002	Tolerado	Benigno	-	-
9	c.879G>A	p.=	2	0,0005	1	0,0002			75,90%	6,25%
10	c.904-11T>A	p. =	1	0,0002	2	0,0003			<b>95,70%</b>	<b>64,20%</b>
10	c.904-3C>T	p. =	1	0,0002	3	0,0005			<b>91,70%</b>	<b>95,22%</b>
10	c.932T>A	p.Ile311Asn	1	0,0002	3	0,0005	Deletéreo	Posible Deletéreo	-	-
Total de variantes raras			27	0,006	85	0,013				
Variantes con predicción bioinformática positiva			13	0,003	17	0,003				

Se han descrito un total de 54 variantes raras (aquellas con un frecuencia inferior al 1%) en la población de casos HBOC y esporádico, de ellas 12 están presentes en la población ESP6500 (ver Tabla 7). El estudio de la frecuencia de portadores de variantes raras en casos y controles sugiere un frecuencia inferior de variantes raras en la población HBOC y esporádicos (0,62%) que en la población control (1,3%) (O.R=0,47 IC=0,31-0,74 p=0,0009). La frecuencia de portadores de variantes raras con predicción bioinformática de alterar la estructura de la proteína o el *splicing* no es significativamente diferente (0,30% vs 0,26% p=0,687). Ambos datos apuntan a que en general no existe un exceso de variantes raras en el gen *RAD51D* en los casos. La Tabla 8 presenta un resumen de la frecuencia de las variantes claramente deletéreas encontradas en estos estudios, estratificadas por el tipo de tumor e historia familiar. El análisis global de todos los estudios confirma a *RAD51D* como gen de susceptibilidad al cáncer de ovario, tanto en el contexto familiar como en el esporádico. De manera similar a lo observado en el primer estudio (*Loveday et al., 2011*), la presencia de mutaciones en *RAD51D* no se asocia a familias con incidencia exclusivamente de cáncer de mama ni a cáncer de mama esporádico. Llama la atención la asociación con familias que presentan casos tanto de cáncer de ovario como de cáncer de mama. Se desconoce si esta asociación es específica o se debe a un enriquecimiento en la selección de aquellas familias que presentan casos de ambos tipos. En conjunto, la edad media de diagnóstico de todos los individuos afectados de cáncer de mama portadores de mutaciones en *RAD51D* es de 44.77 años, la cual es inferior a la edad media de diagnóstico de cáncer esporádico de mama que se aproxima a los 50 años (*Haffty et al., 2009*) y muy cercana a la media de edad de diagnóstico de portadores de variante deletérea en *BRCA1/2* reportada en las poblaciones de origen de estas familias (*Diez et al., 2003; Mavaddat et al., 2010; Panchal et al., 2010*). El número tan bajo de familias que presentan únicamente casos de ovario analizadas hasta el momento (n=81) no permite discriminar la implicación de *RAD51D* en familias con cáncer de ovario familiar frente a su implicación en familias que presentan cáncer de mama y ovario.

**Tabla 8:** Frecuencia de las mutaciones encontradas en el gen *RAD51D*

		Portadores (%)	Total de casos	OR (95% CI)	p-Valor
<b>Historia Familiar</b>	Mama	1 (0.04)	2577	0.47 (0.05-4.5)	0.51
	Ovario	0	81		
	Mama y ovario	16 (0.8)	1985	9.82 (2.86-33.75)	0.0003*
<b>Casos no seleccionados</b>	Mama	1 (0.04)	2171	0.55 (0.06-5.36)	0.61
	Ovario	7 (0.6)	1188	7.16 (1.84-27.74)	0.004*
	Mama y ovario	0	32		
<b>Controles</b>	-	3 (0.08)	3628		

## Perspectivas futuras

La revisión de la literatura científica sugiere una relación entre la genética de la anemia de Fanconi y la del síndrome HBOC. De hecho un total de cuatro genes están involucrados de forma clara en ambos procesos, *BRCA2 (FANCD1)*, *RAD51C (FANCO)*, *BRIP1 (FANCI)*, *PALB2 (FANCN)*. Adicionalmente *SLX4 (FANCP)* y *XRCC2* (Hilbers et al., 2012; Park et al., 2012; Shamseldin et al., 2012) podrían estar involucrados también en ambos procesos. Un reciente estudio bioinformático, centrado en la búsqueda de nuevos genes candidatos para la anemia de Fanconi, sitúa a *RAD51D* en un listado de 150 genes donde se localizan los 16 genes conocidos involucrado en anemia de Fanconi junto con otros genes relacionados funcional y estructuralmente (Haitjema et al., 2013). Se ha sugerido que la susceptibilidad común al cáncer de mama/ovario y la anemia de Fanconi estaría relacionada únicamente con mutaciones en las proteínas efectoras de la ruta. Por su parte, mutaciones en el resto de las proteínas estarían relacionadas únicamente con la anemia de Fanconi (Bakker et al., 2013; Green and Kupfer, 2009). La implicación en la susceptibilidad al cáncer de mama de todas las proteínas de la ruta de anemia de Fanconi se ha evaluado en una serie de estudios. Por una parte, un estudio en una cohorte de 944 familiares de individuos con anemia de Fanconi demostró una mayor incidencia de cáncer de mama, tanto en el grupo de portadores heterocigotos de mutaciones en *BRCA2 (FANCD1)* como en el de portadores heterocigotos de *FANCC* (Solyom et al., 2011). Por otra parte, el total de los genes de anemia de Fanconi han sido resequeenciados en familias HBOC (negativos para variante deletérea en *BRCA1/2* (Ameziane et al., 2009; Barroso et al., 2006; Garcia et al., 2009a; Garcia et al., 2009b; Lewis et al., 2005; Seal et al., 2003; Solyom et al., 2011; Thompson et al., 2012b), aunque la mayoría de estudios son de un tamaño muestral muy pequeño. Se han descrito 4 individuos portadores de mutaciones inactivantes de *FANCC* de una serie de 957 casos HBOC (Thompson et al., 2012b) y un individuo portador de un reordenamiento genómico en el gen *FANCA* de una cohorte de 100 casos HBOC (Solyom et al., 2011). El conjunto de estos resultados sugiere la posible implicación de todas las proteínas de la anemia de Fanconi, y no únicamente de las proteínas efectoras de la ruta de anemia de Fanconi, en el síndrome HBOC. Sin embargo, la detección de estas mutaciones, presumiblemente muy raras, es muy dependiente del tamaño muestral (Ellis and Offit, 2012).

La resequeenciación de genes candidatos ha demostrado ser una estrategia efectiva para el descubrimiento de nuevos genes implicados en el síndrome HBOC, sin embargo, individualmente cada uno de los nuevos genes descritos explica un porcentaje muy bajo de casos, cercano al 1%. De esta forma, un porcentaje significativo del componente genético del síndrome HBOC permanece aún inexplicado. La implicación de rutas adicionales a la respuesta al daño al ADN en el riesgo a cáncer de mama detectada mediante alelos de bajo riesgo en los ensayos GWAS (ver página 48), sugiere la

posible existencia de alelos de alto-moderado riesgo involucrados en las mismas rutas. Futuros ensayos de genes candidatos implicados en estas rutas podrían contribuir al descubrimiento de nuevos genes de susceptibilidad al cáncer familiar.

La popularización en los últimos años de técnicas de ultrasecuenciación ha permitido llevar los estudios de resecuenciación a escala genómica. Actualmente, se pueden realizar estudios de secuenciación de genomas completos o de exomas completos. De estos estudios se espera que contribuyan a descubrir nuevos genes y nuevas rutas relacionadas con el síndrome HBOC .

La limitación inherente a estas técnicas es la dificultad de establecer de manera robusta una asociación causal entre una variante recién descubierta y el riesgo a cáncer de mama. Esta asociación causal estará, a menudo, limitada por los conocimientos científicos disponibles, convirtiendo un método agnóstico (por ejemplo, secuenciación del genoma) en un método sesgado por hipótesis *a priori* que se aplican a la interpretación de los resultados. Con el objetivo de establecer una asociación causal, se aplican criterios como la selección de variantes localizadas en regiones codificantes, predilección por variantes con un efecto sobre el marco de lectura de la proteína, priorización de variantes localizadas en genes de rutas previamente implicadas en susceptibilidad al cáncer de mama o filtrado de variantes por frecuencia. Feng et al (Feng et al., 2011) modelizan en su artículo distintas estrategias para la interpretación de variantes encontradas tras la resecuenciación de exomas de individuos afectados por enfermedades poligénicas, como el cáncer de mama hereditario.

El análisis de exomas ha generado grandes expectativas en el campo de la genética del HBOC y de hecho ha demostrado ser una estrategia efectiva, como sugieren la identificación de individuos portadores de mutaciones en los genes *XRCC2*, *FANCC* y *BML* (Park et al., 2012; Thompson et al., 2012b), que hasta el momento, no se habían asociado con el síndrome HBOC. El hecho de que los únicos genes identificados, hasta el momento, mediante análisis de exomas estén implicados en distintos pasos de la respuesta al daño al ADN, sugiere o bien que únicamente esta ruta está implicada en el síndrome HBOC o bien una limitación en la interpretación de los resultados provenientes de exomas. A modo de guía, el artículo de Snape et al. (Snape et al., 2012) describe la metodología llevada a cabo en el diseño e interpretación del análisis del exoma de 50 individuos afectados de cáncer de mama familiar. La complejidad en la interpretación de los resultados la ilustra el total de 1296 variantes truncantes identificadas en los 50 casos. La efectividad de la estrategia fue demostrada mediante la identificación de mutaciones en genes conocidos (*BRCA2*, *CHEK2* y *ATM*) que habían pasado inadvertidos durante el proceso rutinario de screening. Sin embargo, tras el análisis de los resultados se llega a la conclusión de que, si estos genes no hubiesen sido conocidos previamente habría sido imposible distinguir estas variantes del resto de variantes truncantes y se

sugiere la necesidad de realizar análisis complementarios junto con el análisis de exomas. Otros trabajos publicados han complementado el análisis de exomas con estudios de co-segregación, estudios caso control y estudios de validación mediante resecuenciación de genes concretos. Por el momento, no se ha logrado identificar de forma clara, ninguna variante en genes nuevos (Gracia-Aznarez et al., 2013; Hilbers et al., 2013; Lynch et al., 2013).

Tras la finalización del Proyecto Genoma Humano, aproximadamente el 2% del genoma se encontró asociado a genes codificantes de proteínas. El restante ADN, durante tiempo, se consideró un mero contenedor, llegando a acuñarse el término de “ADN basura” por carecer de una función conocida. Los datos aportados por el proyecto ENCODE (2012) expanden el concepto de funcionalidad en el genoma mucho más allá que el de codificar proteínas. Aproximadamente el 80 % del genoma humano se ha encontrado participando en funciones bioquímicas, fundamentalmente relacionadas con la regulación de la expresión génica. El diseño de los estudios de resecuenciación genómica y la interpretación de la asociación causal de muchas de las variantes descubiertas deberá considerar la relevancia de estas funcionalidades reguladoras del genoma.



## Conclusiones

A continuación se indican las conclusiones de la presente Tesis Doctoral:

1. Se ha desarrollado una metodología de análisis de *splicing* que produce, en la mayoría de los casos, resultados analíticos inequívocos de acuerdo a las guías de la IARC. La metodología es de fácil implementación en la rutina de un laboratorio de diagnóstico genético.
2. Las variantes *BRCA2* c.7617+1G>A, c.682-2A>G y c.8954-5A>G provocan alteraciones de *splicing* inequívocamente patogénicas.
3. El SNP (rs9534262) es un sQTL que regula los niveles de los transcritos alternativos *BRCA2Δ17* y *BRCA2Δ17-18*. Los ensayos caso-control y de supervivencia no detectaron implicación de rs9534262 en el riesgo a cáncer de mama.
4. Los reordenamientos génicos en el locus *BRCA2* explican menos del 1% de las familias españolas con sospecha de padecer un síndrome hereditario de cáncer de mama y ovario.
5. En lo que se refiere a los mecanismos moleculares responsables de reordenamientos en el locus *BRCA2*, la relevancia de la recombinación homóloga ha sido probablemente sobreestimada, mientras que la contribución de mecanismos mediados por micro-homología han sido infraestimados.
6. La mutación patogénica *BRCA2* c.6937+594T>G no explica una fracción significativa de las familias españolas con sospecha de padecer un síndrome hereditario de cáncer de mama y ovario.
7. Se han identificado efectos fundadores asociados a las mutaciones patogénicas *BRCA2* g.5670\_6155delinsATA, c.7617+1G>A y c.8954-5A>G.
8. El análisis de mutaciones en el gen *SLX4* en 486 familias BRCAX encontró una variante claramente inactivante de la proteína (p.Glu1517\*) y un cambio de aminoácido (p.Arg372Trp) que presentó cosegregación con los casos de cáncer. Los resultados no son concluyentes respecto a la implicación de *SLX4* en el síndrome HBOC.

9. El análisis de mutaciones en el gen *RAD51D* en 842 familias BRCA1 encontró dos variantes claramente inactivante de la proteína p.Arg231\* y p.Val193Alafs\*4, en un total de tres familias y en una familia la variante c.1A>T p.Met1? probablemente inactivante. Los resultados confirman la implicación de *RAD51D* en el riesgo a cáncer de ovario hereditario, pero no permiten descartar su posible implicación en el riesgo a cáncer de mama hereditario.

En este anexo se presentan aquellos los dos trabajos realizados por el doctorando como autor no principal y relevantes para el contenido de esta tesis.

*Comparación de protocolos de análisis del splicing del ARNm entre múltiples laboratorios: Guía de buenas prácticas para el análisis clínico estandarizado.*

La evaluación de UVs en genes de predisposición al cáncer es esencial para el manejo clínico y depende de un análisis multifactorial de datos clínicos, genéticos, anatomo-patológicos, bioinformáticos y funcionales. La precisión de los ensayos funcionales depende de un diseño apropiado y una correcta interpretación. Se ha llevado a cabo un estudio multicentrico para comparar los protocolos de análisis del splicing del ARNm utilizados por los miembros del consorcio ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles). En concreto, se compararon los resultados derivados del análisis de un panel de variantes en *BRCA1* y *BRCA2* con un efecto alterador del *splicing* ya conocido (*BRCA1*:c.135-1 G>T, c.591 C>T, c.594-2 A>C, c.671-2 A>G, c.5467+5G>C and *BRCA2*:c.426-12\_8delGTTTT, c.7988 A>T, c.8632+1G>A, c.9501+3A>T).

Posteriormente, se analizaron las diferencias en los protocolos con el objetivo de determinar cuáles elementos son críticos para un diseño robusto del análisis. Se observó que un conocimiento previo de los posibles transcritos alternativos junto con las estrategias de diseño de primers, las condiciones de PCR y los métodos para detectar amplicones fueron los factores claves para un resultado preciso del análisis de *splicing*. Se observó mayor variabilidad en los resultados en relación con la detección de transcritos de baja abundancia y esto se atribuyó en parte a la naturaleza estocástica de la PCR cuando la cantidad de molde a amplificar es baja. Este estudio hace unas recomendaciones de buenas prácticas y sugiere una serie de aspectos claves a tener en cuenta para el diseño de en análisis del ARNm para la evaluación de UVs.



## Comparison of mRNA Splicing Assay Protocols across Multiple Laboratories: Recommendations for Best Practice in Standardized Clinical Testing

Phillip Whiley,<sup>1,2</sup> Miguel de la Hoya,<sup>3</sup> Mads Thomassen,<sup>4</sup> Alexandra Becker,<sup>5,6</sup> Rita Brandão,<sup>7</sup> Inge Sokilde Pedersen,<sup>8</sup> Marco Montagna,<sup>9</sup> Mireia Menéndez,<sup>10</sup> Francisco Quiles,<sup>10</sup> Sara Gutiérrez Enríquez,<sup>11</sup> Kim De Leeneer,<sup>12</sup> A. Tenés,<sup>11</sup> G. Montalban,<sup>11</sup> D. Tserpelis,<sup>7</sup> T. Yoshimatsu,<sup>13</sup> Carole Tirapo,<sup>14</sup> M. Raponi,<sup>15</sup> T. Caldes,<sup>3</sup> Ana Blanco,<sup>16</sup> M. Santamariña,<sup>17</sup> Lucia Guidugli,<sup>18</sup> G. Ruiz de Garibay,<sup>3</sup> Ming Wong,<sup>19</sup> M. Tancredi,<sup>20</sup> L. Fachal,<sup>16</sup> Y. Ding,<sup>21</sup> T. Kruse,<sup>4</sup> Vanessa Lattimore,<sup>22</sup> Ava Kwong,<sup>23</sup> Tsun Chan,<sup>23</sup> Mara Colombo,<sup>24</sup> G. De Vecchi,<sup>24</sup> Maria Caligo,<sup>19</sup> Diana Baralle,<sup>15</sup> Conxi Lázaro,<sup>10</sup> Fergus Couch,<sup>17</sup> Paolo Radice,<sup>24</sup> Melissa Southey,<sup>18</sup> Susan Neuhausen,<sup>21</sup> Claude Houdayer,<sup>14</sup> Jim Fackenthal,<sup>13</sup> Thomas Hansen,<sup>25</sup> Ana Vega,<sup>16</sup> Orland Diez,<sup>11</sup> Rien Blok,<sup>7</sup> Kathleen Claes,<sup>12</sup> Barbara Wappenschmidt,<sup>5,6</sup> Logan Walker,<sup>22</sup> Amanda Spurdle,<sup>1</sup> and Melissa Brown<sup>2</sup> on behalf of the ENIGMA consortium

**BACKGROUND:** Accurate evaluation of unclassified sequence variants in cancer predisposition genes is essential for clinical management and depends on a multifactorial analysis of clinical, genetic, pathologic, and bioinformatic variables and assays of transcript length and abundance. The integrity of assay data in turn relies on appropriate assay design, interpretation, and reporting.

**METHODS:** We conducted a multicenter investigation to compare mRNA splicing assay protocols used by members of the ENIGMA (Evidence-Based Network for the Interpretation of Germline Mutant Alleles) consortium. We compared similarities and differences in results derived from analysis of a panel of breast cancer 1,

early onset (*BRCA1*) and breast cancer 2, early onset (*BRCA2*) gene variants known to alter splicing (*BRCA1*: c.135-1G>T, c.591C>T, c.594-2A>C, c.671-2A>G, and c.5467+5G>C and *BRCA2*: c.426-12\_8delGTTTT, c.7988A>T, c.8632+1G>A, and c.9501+3A>T). Differences in protocols were then assessed to determine which elements were critical in reliable assay design.

**RESULTS:** PCR primer design strategies, PCR conditions, and product detection methods, combined with a prior knowledge of expected alternative transcripts, were the key factors for accurate splicing assay results. For example, because of the position of primers and PCR extension times, several isoforms associated with

<sup>1</sup> Genetics & Computational Biology Division, Queensland Institute of Medical Research, Brisbane, Queensland, Australia; <sup>2</sup> School of Chemistry and Molecular Biosciences, the University of Queensland, Queensland, Australia; <sup>3</sup> Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain; <sup>4</sup> Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark; <sup>5</sup> Center for Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany; <sup>6</sup> Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany; <sup>7</sup> Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, the Netherlands; <sup>8</sup> Section of Molecular Diagnostics, Department of Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark; <sup>9</sup> Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV—IRCCS, Padua, Italy; <sup>10</sup> Genetic Diagnosis Unit, Hereditary Cancer Program, Institut Català d'Oncologia, Barcelona, Spain; <sup>11</sup> Oncogenetics Laboratory, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Institut de Recerca, University Hospital Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain; <sup>12</sup> Department of Medicine, the University of Chicago Medical Center, Chicago, IL; <sup>13</sup> Service de Génétique et U830, Institut Curie, et Université Paris Descartes, Sorbonne Paris Cité, Paris, France; <sup>14</sup> Human Development and Health Academic Unit, Faculty of Medicine, University of Southampton, Southampton General Hospital, Southampton, UK; <sup>15</sup> Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica-USC, CIBERER, IDIS, Santiago

de Compostela, Spain; <sup>17</sup> Grupo de Medicina Xenómica—USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain; <sup>18</sup> Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, NY; <sup>19</sup> Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia; <sup>20</sup> Istituto di Anatomia Patologica, Università di Pisa, Pisa, Italy; <sup>21</sup> Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA; <sup>22</sup> Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; <sup>23</sup> Center for Genomic Medicine, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; <sup>24</sup> Department of Pathology, University of Otago, Christchurch, New Zealand; <sup>25</sup> Center for Medical Genetics, Ghent University Hospital, De Pintelaan, Ghent, Belgium; <sup>26</sup> Division of Molecular Pathology, Department of Pathology, Hong Kong Sanatorium & Hospital, Hong Kong, China.

\* Address correspondence to this author at: The University of Queensland, Cooper Rd., St. Lucia Brisbane, QI, Australia 4072. Fax 61-07-33654622; e-mail melissa.brown@uq.edu.au.

The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute or NIH. Received June 25, 2013; accepted October 21, 2013.

Previously published online at DOI: 10.1373/dinchem.2013.210658

*BRCA1*, c.594-2A>C and c.671-2A>G, were not detected by many sites. Variation was most evident for the detection of low-abundance transcripts (e.g., *BRCA2* c.8632+1G>A  $\Delta$ 19,20 and *BRCA1* c.135-1G>T  $\Delta$ 5q and  $\Delta$ 3). Detection of low-abundance transcripts was sometimes addressed by using more analytically sensitive detection methods (e.g., *BRCA2* c.426-12\_8delGTTTT ins18bp).

**CONCLUSIONS:** We provide recommendations for best practice and raise key issues to consider when designing mRNA assays for evaluation of unclassified sequence variants.

© 2013 American Association for Clinical Chemistry

---

Germline mutations in the breast cancer susceptibility genes breast cancer 1, early onset (*BRCA1*)<sup>26</sup> and breast cancer 2, early onset (*BRCA2*) (OMIM #113705 and #600185, respectively) are associated with a significantly increased risk of breast and other cancers (1). Although many thousands of disease-associated mutations have been identified in these genes, many DNA sequence changes found during genetic screening fall into the category of unclassified variants because their functional and clinical significance is not immediately clear. Such unclassified variants pose a challenge for clinical management of variant carriers.

Unclassified variants have the potential to alter protein function by changing the coding sequence of a transcript, or the level or structure of the gene transcript, and by disrupting regulatory regions in promoters, untranslated regions, exons, or introns (2–5). Such regulatory variants include those affecting normal splicing of *BRCA1* and *BRCA2*, many of which have been shown to be clinically significant by use of cDNA studies and multifactorial likelihood analysis methods that combine bioinformatic, pathologic, and clinical information (6–8). These variants include those that affect splicing by disrupting or weakening the motifs at intron–exon boundaries, introducing de novo splice acceptor or donor sites, activating cryptic splice sites, or disrupting enhancer and silencer sequences. Several studies have shown that bioinformatic prediction tools can be used to prioritize variants for splicing assays (9–14).

To date, a total of 82 studies have reported findings related to splicing in *BRCA1* or *BRCA2* (15). The majority of these used reverse transcriptase PCR (RT-

PCR)<sup>27</sup> analysis of RNA extracted from blood of variant carriers or alternatively, minigene constructs containing the variant and assayed in non-patient-derived cell lines. The interpretation of splicing results for variant carriers can be complicated by the detection of normal alternatively spliced transcripts that occur in healthy individuals—an issue that has yet to be extensively addressed in the literature. The effect of the range of variables found in protocols used in research and clinical testing laboratories, including the PCR assay design, reagents used, and tools for visualizing and characterizing transcripts identified by PCR on assay result interpretation, is also unclear.

There are 4 instances of inconsistent or conflicting splicing results (6, 8, 14, 16–19). These include *BRCA1* c.212+3A>G, c.670+8C>T, and c.736T>G and *BRCA2* c.517-19C>T (4, 19–25). Reports of splicing results from a further 7 variants differed in the number of aberrant bands found in each study. The potential clinical implications of such inconsistencies highlight the need to establish the advantages and limitations of the various techniques in practice.

Guidelines for clinical interpretation and reporting of unclassified variants analyzed using splicing assays are available in the UK and Netherlands via the UK Clinical Molecular Genetics Society ([http://www.cmsgs.org/BPGs/Best\\_Practice\\_Guidelines.htm](http://www.cmsgs.org/BPGs/Best_Practice_Guidelines.htm)) and Dutch Society of Clinical Genetic Laboratory Specialists (<http://www.vkgl.nl/>). In addition, a range of in silico approaches have been compared with one another, and with transcript analysis, by the splice network of the French *BRCA* diagnostic testing laboratories, recently reported by Houdayer et al. (11). In this study (11), Houdayer et al. investigated the value of combining Splice-site Finder and MaxEntScan prediction tools and showed that major splice defects were consistently identified across a number of different laboratories. The authors did find some discrepancies with results previously reported in the literature and recommended a large cross-validation study as a future priority.

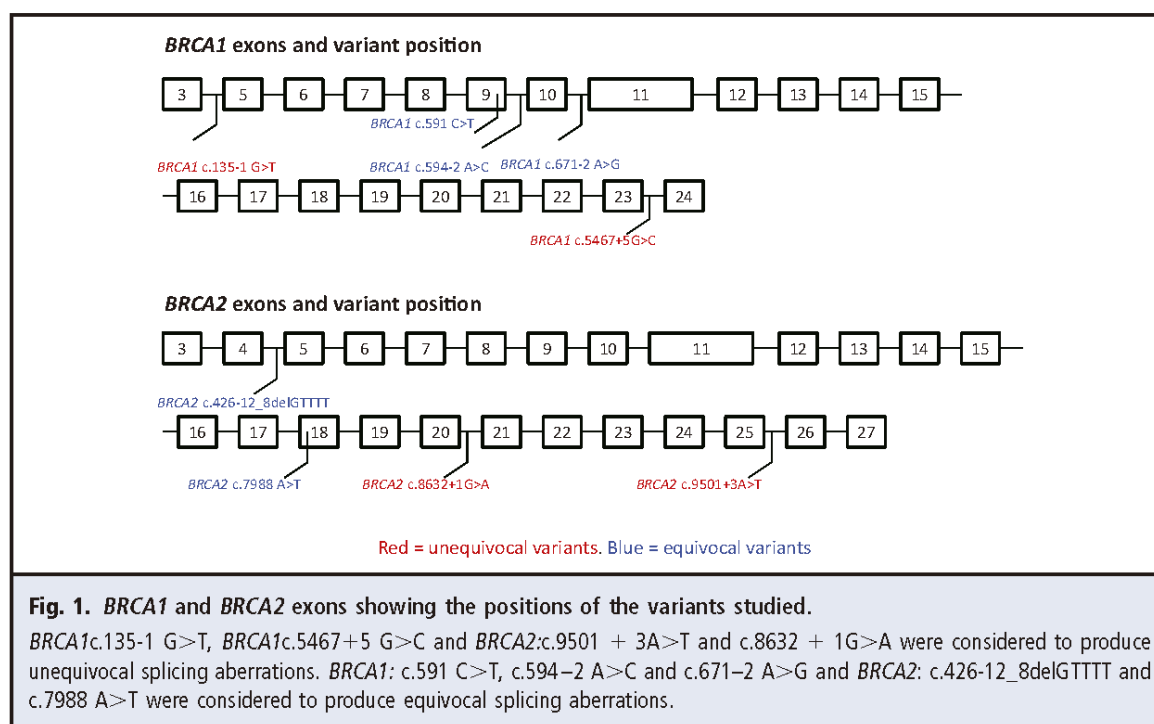
The Evidence-Based Network for the Interpretation of Germline Mutant Alleles (ENIGMA) consortium was established in 2009 with the purpose of sharing data, methods, and resources to facilitate classification of unclassified variants (21). To date, a total of 3286 unique *BRCA1* and *BRCA2* variants considered to be of uncertain clinical significance have been submitted to ENIGMA from more than 43 sites in 19

---

<sup>26</sup> Human genes: *BRCA1*, breast cancer 1, early onset; *BRCA2*, breast cancer 2, early onset.

<sup>27</sup> Nonstandard abbreviations: RT-PCR, reverse transcriptase PCR; ENIGMA, Evidence-Based Network for the Interpretation of Germline Mutant Alleles; NMD, nonsense-mediated decay; LCL, lymphoblastoid cell lines; kConFab, Kathleen Cunningham Consortium for Research into Familial Breast Cancer; CE, capillary electrophoresis.

## Guidelines for mRNA Splicing Assays



countries. The consortium has established several working groups, including one dedicated to examining variants that potentially alter RNA splicing.

Here we describe the outcome of an ENIGMA Splicing Working Group study to assess the importance of various mRNA assay components on consistency of results. We identified a variety of differences in protocols from 23 laboratories, the majority of which conduct routine clinical assays (See Table 1 in the Data Supplement that accompanies the online version of this article at <http://www.clinchem.org/content/vol60/issue2>). We report the critical elements on assay design that should be considered in the analysis of variants that may impact RNA splicing.

### Materials and Methods

Each participating laboratory submitted information about the mRNA splicing protocol in use at their site. These protocols were then compared on the basis of the source of biological material; the use of a nonsense mediated decay (NMD) inhibitor, RNA extraction, or removal of contaminating genomic DNA; the choice of cDNA synthesis primer, reverse transcriptase, and DNA *Taq* polymerase; the method of PCR product detection; and whether products were isolated, subcloned or sequenced (see online Supplemental Table 1).

To compare the assays used by laboratories within the ENIGMA consortium, 23 sites were sent aliquots of

samples from the same lymphoblastoid cell lines (LCLs) that had been generated by the Kathleen Cunningham Consortium for Research into Familial Breast Cancer (kConFab) from 9 carriers of *BRCA1* or *BRCA2* variants known to be associated with splicing defects (Fig. 1) and from 11 controls. Four LCLs carried variants that produce unequivocal splicing aberrations resulting in a clear exon-skipping event. Five LCLs carried variants considered to produce equivocal splicing aberrations, based on the observation that they confer more subtle and variable effects, such as altering the availability of naturally occurring isoforms to a variable extent and/or producing a large and variable number of uncharacterized splicing products.

The project was conducted in 2 phases. In the initial phase (phase 1), 16 sites used an mRNA splicing protocol they routinely use in their laboratory (see online Supplemental Table 1), summarized their results, and submitted these to PJW and MAB. for collation. Following an analysis of phase 1 results, phase 2, informed by the phase 1 findings, was initiated, during which some sites repeated each assay using a standard set of PCR primers and cycling conditions (see online Supplemental Table 2). All other components of the protocol were per phase 1, apart from site 8, which used a Bioanalyzer in phase 1 and capillary electrophoresis (CE) in phase 2. Seven sites that participated in phase 1 repeated the assays under the controlled conditions of phase 2. An addi-

tional 3 sites joined the study to assay all variants for phase 2. A further 3 sites joined phase 2 to specifically assay *BRCA1* c.671-2A>G, following the finding that this equivocal variant gave rise to the greatest range of alternatively spliced transcripts.

## Results

The initial comparison of protocols used by participating laboratories revealed that cycloheximide or puromycin was sometimes used for NMD treatment, with incubation times between 4 and 8 h and concentrations between 100 and 250  $\mu\text{g/mL}$ , the use of 8 cDNA synthesis kits, 12 different DNA polymerases, and transcript isolation strategies that included band excision, subcloning, and sequencing. The majority of laboratories used agarose gel electrophoresis for visualizing transcripts, but several used digital visualization strategies.

In phase 1 of the study, all sites detected the full-length transcript for each of the 4 unequivocal variants (Table 1). All sites also detected the most prominent single-exon skipping events not seen in controls for each of the unequivocal variants, apart from site 4 and 14, which did not detect the  $\Delta 20$  transcript for the unequivocal variant *BRCA2* c.8632+1G>A. Not all sites detected all of the less abundant transcripts from this variant, however, with only 3/16 sites detecting the  $\Delta 19\&20$  transcript and only 6/16 sites detecting the ins i20 transcript. For the unequivocal variant *BRCA1* c.135-1G>T, which has been associated with multiple splice isoforms (22), only 3/16 detected the  $\Delta 5q$  transcript, and only 1 site detected the  $\Delta 3$  transcript.

Detailed analysis of each of the protocols and resulting data revealed that the range of PCR design strategies contributed to the variation in detection of transcripts, in particular PCR primer design and PCR cycling conditions. For example, 11 out of 16 sites that analyzed equivocal *BRCA1* c.671-2A>G were unable to detect all of the transcripts because primer position did not allow some, clearly unanticipated, fragments to be amplified (Table 1). Forward primers positioned in exon 9 or 10 were unable to amplify isoforms lacking those exons, including  $\Delta 9/10$ ,  $\Delta 9/10/11$ , or  $\Delta 9/10/11q$  isoforms seen in controls, or the  $\Delta 9/11$  or  $\Delta 10/11$  variant-associated isoforms detected by other sites.

The length of extension time during PCR amplification was also found to be a contributing factor, with several protocols using times that were likely to be too short to detect the longer PCR products amplified from some splice isoforms. For example, 5 sites used elongation times of 3 min or less and were unable to amplify full-length transcripts or transcripts containing exon 11 ( $\Delta 9$  or  $\Delta 9/10$ ) for *BRCA1* c.671-2A>G, which are longer than 3 kb. As for the results observed for un-

equivocal variants, an additional explanation for variation in detection of transcripts was the low abundance of some transcripts, including those identified in the variant carrier only (e.g.,  $\Delta 9/11$ ,  $\Delta 10/11$ , and  $\Delta >3\text{kb}$  exon 11 transcripts), which is known to lead to variable PCR amplification. PCR cycle number was also important, with site 23 detecting only a limited number of transcripts (Table 1), likely reflecting the use of only 25 cycles (see online Supplemental Table 1).

Given that phase 1 showed that many transcripts were not observable due to the positioning of primers or elongation time, phase 2 of the study was initiated. Phase 2 included assays conducted by 10–12 sites (depending on the variant analyzed) using a standard set of primers and elongation times appropriate for the expected lengths of the transcripts (see online Supplemental Table 2). The outcome was a much greater analytical sensitivity and consistency of results (Table 2). For example, all sites were now able to detect relatively high-abundance isoforms or variant-associated transcripts reported in previous studies, but not consistently reported in phase 1 [ $\Delta 17,18$  for *BRCA2* c.7988A>T,  $\Delta 20$  for *BRCA2* c.8632+1G>A,  $\Delta 5$  for *BRCA2* c.426-12\_8delGTTTT, and  $\Delta 10$  for *BRCA1* c.594-2A>C (5, 7, 8)]. Importantly, unlike phase 1, in phase 2 all study sites were able to detect at least 1 aberrant band (cf. controls) and thus may have been able to better classify the variant using the IARC (International Agency for Research on Cancer) 5-tier classification scheme.

There remained some inconsistencies in the phase 2 data. Further comparison of protocols suggested that the method of PCR product detection was likely to be a contributing factor. Sites 2 and 8 in phase 2 were the only sites to use CE exclusively for detection of transcripts. Site 2 had higher overall detection compared to the other sites. Indeed, 10 of the 23 transcripts (43.5%) identified across all sites in the phase 1 analysis of unequivocal variants were detected only by CE, demonstrating it to be a comparatively more analytically sensitive detection method. This trend continued for equivocal variants analyzed in phase 2, with 12 of the 49 (24.5%) transcripts detected only by capillary CE. The sites employing a Qiaxcel visualization, Bioanalyzer, or MultiNA systems demonstrated that these systems were often more analytically sensitive than gel electrophoresis. For example, sequencing of transcripts identified by Qiaxcel analysis of *BRCA2* c.426-12\_8delGTTTT (site 9, phase 2) showed that it was the only system to discriminate the small insertion of 18 nucleotides from the full-length transcript (Table 2; also see online Supplemental Fig. 1).

Analysis of *BRCA1* c.594-2A>C in phase 2 identified 11 different transcripts. Excising bands from agarose gel or sequencing PCR products directly enabled

Guidelines for mRNA Splicing Assays

Table 1. Phase 1 results submitted by 17 sites. <sup>a</sup>			Site (n = 17)																
Variant	Transcript		1	2	3	4	5	6 <sup>b</sup>	7	8	9	10	11	12	13	14	15	16	17
Unequivocal																			
BRCA2 c.9501+3A>T	Full length		+	+	+	+	+	+	NA	+	+	+	+	+	+	+	+	+	+
	Δ25 <sup>d</sup>		+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+
BRCA2 c.8632+1G>A	Full length		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Δ20 <sup>d</sup>		+	+	+	-	+	+	+	+	+	+	+	+	-	+	+	-	
	Δ19&20 <sup>d</sup>		-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+	-
	ins i20		+	+	-	+	+	+	-	-	-	+	-	-	-	-	-	-	-
BRCA1 c.135-1G>T	Full length		+	+	+	+	+	+	NA	+	+	+	+	+	+	+	+	+	+
	Δ5		+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+
	Δ5q		-	+	-	-	-	+		-	-	+	-	-	-	-	-	-	-
	Δ3		-	+	-	-	-	-		-	-	-	-	-	-	-	-	-	-
BRCA1 c.5467+5G>C	Full length		+	+	+	+	+	+	NA	+	+	+	+	+	+	+	+	+	+
	Δ23 <sup>d</sup>		+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	-
Equivocal																			
BRCA2 c.7988A>T	Full length		+	+	+	+	+	+	NA	+	+	+	+	+	+	+	+	+	+
	Δ18		+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	-
	Δ17&18		+	+	-	-	-	+		+	+	+	-	+	+	-	-	+	+
BRCA2 c.426-12_8delGTTT	Full length		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Δ6q,7		+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	Δ3,5 <sup>d</sup>		-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
	Δ3,4,5 <sup>d</sup>		-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	ins 18bp <sup>d</sup>		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
	Δ4,5,6,7 <sup>d</sup>		-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
	Δ5		+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	-
	Δ5,6,7 <sup>d</sup>		-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
	Δ6,7 <sup>d</sup>		-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
	BRCA1 c.591C>T	Full length		+	+	+	+	+	+	NA	+	+	+	+	+	NA	+	+	+
Δ8&9			-	-	-	-	+	-		-	-	-	-	-	-	-	-	-	-
Δ8,9,10			-	-	-	-	+	-		-	-	-	-	-	-	-	-	-	+
Δ9			-	+	+	+	-	+		+	-	+	+	-		+	-	+	-
Δ9,10			+	+	+	-	-	+		+	+	+	-	+		-	+	+	+
Δ9,10,11q			+	+	-	-	-	-		-	-	+	-	-		-	-	+	-
Δ11q			+	-	-	-	-	-		-	-	+	-	-		-	-	+	-
Δ9,11q <sup>d</sup>			+	+	-	-	-	-		-	-	+	-	-		-	-	+	-
Δ10 <sup>d</sup>			-	-	-	-	-	-		+	-	-	-	-		-	-	-	-
BRCA1 c.594-2A>C		Full length		+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
	Δ9		-	+	-	-	-	+	-	-	-	+	-	-	-	-	+	-	-
	Δ9,10		+	+	+	+	+	+	+	+	+		-	-	-	+	-	+	-
	Δ9,10,11q		+	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+
	Δ9,10,11		+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Δ11q		+	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+
	Ins 21bp <sup>d</sup> Intron 9 <sup>d</sup>		+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Δ9,11q <sup>d</sup>		-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+
	Δ10,11q <sup>d</sup>		-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+
Δ10 <sup>d</sup>		+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	

Continued on page XX

**Table 1.** Phase 1 results submitted by 17 sites.<sup>a</sup> (Continued from page XX)

Variant	Transcript	Site (n = 17)																
		1	2	3	4	5	6 <sup>b</sup>	7	8	9	10	11	12	13	14	15	16	17
BRCA1 c.671-2A>G	Full length	+	+	+	-	-	+	-	+	NA	+	-	-	+	+	-	+	NA
	Δ9,11 <sup>d</sup>	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Δ10,11 <sup>d</sup>	+	+	+	-	-	-	-	+	-	+	+	+	-	-	+	+	-
	Δ11&12 <sup>d</sup>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Δ~3.2 kb exon 11 <sup>d</sup>	+	-	+	-	-	-	-	+	-	+	-	-	-	-	-	-	-
	Δ9,10,11,12 <sup>d</sup>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
	Δ9	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-
	Δ9,10	+	+	-	-	-	-	-	+	-	+	-	-	-	-	+	-	+
	Δ9,10,11	+	+	-	-	+	-	-	+	-	+	-	-	-	-	-	+	+
	Δ9,10,11q	+	+	-	-	-	-	+	+	-	+	-	-	-	-	-	-	-
	Δ11	+	+	+	+	+	+	-	+	-	+	+	-	+	-	+	+	+
	Δ11q	-	+	-	+	+	+	-	+	-	+	+	-	+	-	-	-	-

<sup>a</sup> +, detection of the transcript; -, no detection.  
<sup>b</sup> Site 6 reported Δ1q and Δ5,1q (BRCA1 c.135-1G>T) and Δ3 (BRCA2 c.426-12\_8delGTTT) following a review of the data after results were initially compiled. HGVS (Human Genome Variation Society) nomenclature descriptions are available in online Supplemental Table 3. The Δ6,7 transcript identified in the analysis of BRCA2 c.426-12\_8delGTTT is not sequence confirmed, and it is possible that this transcript is actually Δ6q,7, which differs by only 2 bp. Additional minor and uncharacterized peaks were detected by some sites but have not been included in the table. For example, site 17 also reported Δ8p and Δ14p splice slipping events.  
<sup>c</sup> NA, not assessed.  
<sup>d</sup> Transcripts detected in the variant carrier only. Other transcripts were identified in at least 1 control.

detection of 3–6 transcripts (sites 3, 4, 17, and 18). Cloning PCR products followed by sequencing detected 6–7 transcripts (sites 1 and 16), and CE detected 10–11 transcripts (sites 2 and 8). This showed that cloning PCR products improved analytical sensitivity, and visualization by the Qiaxcel system or capillary CE together with sequence analysis is optimal to identify and characterize transcripts. The number of clones sequenced also appeared to improve analytical sensitivity; screening 40 clones (site 16) in comparison to 24 clones (site 1) enabled the detection of 1 additional transcript.

Finally, we examined the effect of using different reverse transcriptase enzymes with the same RNA, cDNA synthesis primers, and PCR primers, enzymes, and conditions. As shown in Fig. 2, the amplification of the longest transcripts was not possible with GoScript; with M-MuLV we missed in the patient with the c.671-2A>G variant the wild-type transcript; only Superscript II allowed amplification of the longest transcript in both controls and variant carriers.

It is important to note that all transcripts shown in Tables 1 and 2 were the outcome of results by scorers who were blind to the transcripts identified by other participants, to avoid biasing the interpretation and thus the value of each approach. Once the full range of transcripts was known, however, it was possible to find

some missing transcripts, demonstrating the importance of prior knowledge in both the design of the assays and the interpretation of results.

There was no clear evidence of any differences as a result of using (a) cycloheximide vs puromycin treatment for NMD inhibition; (b) differing RNA extraction methods; (c) oligo d(T) and random hexamers vs gene specific primers; (e) various methods of DNase treatment; and (f) a particular type or brand of *Taq* polymerase.

A summary of the recommendations arising from this study is provided in Table 3.

## Discussion

RNA splicing assays are commonly used in diagnostic and research settings to assess the potential effects of unclassified variants in multiple genes, including *BRCA1* and *BRCA2*. There are a multitude of differing protocols used in clinical and research laboratories, including those within the ENIGMA consortium, and this prompted a study aimed at establishing assay guidelines.

This study shows that prior knowledge of the expected transcripts, including naturally occurring isoforms and aberrant transcripts predicted to occur in variant-carrying samples, is important for assay design.

## Guidelines for mRNA Splicing Assays

<b>Table 2. Summary of results from phase 2 of the study, in which PCR primers and conditions were controlled.<sup>a</sup></b>												
Variant	Transcripts <sup>b</sup>	Phase 2 results (10 sites) <sup>c</sup>										
		1*	2 <sup>^</sup>	3*	4*	8 <sup>^</sup>	9 <sup>^</sup>	16*	17*	18*	19*	
<b>Unequivocal</b>												
BRCA2 c.9501+3A>T	Full length	+	+	+	+	+	+	+	+	+	+	
	Δ25	+	+	+	+	+	+	+	+	+	+	
	ins i23	-	+	-	-	-	-	-	-	-	-	
BRCA2 c.8632+1G>A	Full length	+	+	+	+	+	+	NA	+	+	+	
	Δ19	-	-	-	-	+	-	-	-	-	-	
	Δ20	+	+	+	+	+	+	-	+	+	+	
	Δ19&20	+	+	-	-	+	+	-	+	-	-	
	ins i20	-	+	-	+	-	-	-	-	-	-	
BRCA1 c.135-1G>T	Extra peaks	-	+	-	-	-	-	-	-	-	-	
	Full length	+	+	+	+	+	+	NA	+	+	+	
	Δ5q	+	+	-	-	+	-	-	+	-	-	
	Δ5	+	+	+	+	+	+	-	+	+	+	
	Δ3	-	+	-	-	+	-	-	-	-	-	
	Δ3,5	-	+	-	-	+	-	-	-	-	-	
	Δ5,6	-	+	-	-	-	-	-	-	-	-	
	ins i3	-	+	-	-	+	-	-	-	-	-	
BRCA1 c.5467+5G>C	ins i3+Δ5	-	+	-	-	-	-	-	-	-	-	
	Full length	+	+	+	+	+	+	+	+	+	+	
	Δ21	-	-	-	-	+	-	-	-	-	-	
	Δ21,23	-	-	-	-	+	-	-	-	-	-	
	Δ22,23	-	+	-	-	+	-	-	+	-	-	
	Δ23	+	+	+	+	+	+	+	+	+	+	
<b>Equivocal</b>	Δ22	-	+	-	+	+	-	-	-	-	-	
	BRCA2 c.7988A>T	Full length	+	+	+	+	+	+	+	+	+	+
		Δ18	+	+	+	+	+	+	+	+	+	+
		Δ17&18	+	+	+	+	+	+	+	+	+	+
	BRCA2 c.426-12_8delGTTT*	FL	+	+	+	+	+	+	NA	+	+	+
		Δ6q,7	+	+	-	+	+	-	-	+	-	-
		ins 18bp	-	-	-	-	-	+	-	-	-	-
		Δ5	+	+	+	+	+	+	-	+	+	+
		Δ5,6,7	-	+	-	-	+	-	-	+	-	-
	BRCA1 c.591C>T	Δ6,7	-	-	-	-	-	+	-	-	-	+
		Full length	+	+	+	+	+	+	+	+	+	NA
Δ9		-	+	+	+	+	-	+	+	-	-	
Δ9,10		+	+	+	+	+	+	+	+	+	+	
Δ9,10,11q		+	+	+	-	+	+	+	+	-	-	
Δ9,10,11		-	+	-	+	+	-	-	-	-	-	
Δ11q		+	+	+	-	+	+	+	+	-	-	
Δ9,11q		+	+	-	-	+	-	+	-	-	-	
Δ10,11q		-	-	-	-	+	-	-	-	-	-	
Δ10		-	+	-	-	+	-	-	-	-	-	
Δ11q+insi13		-	+	-	-	+	-	-	-	-	-	
ins i21 intron 9		-	-	-	-	+	-	-	-	-	-	
extra peaks	-	+	-	-	-	-	-	-	-	-		

Continued on page XX

**Table 2.** Summary of results from phase 2 of the study, in which PCR primers and conditions were controlled.<sup>a</sup> (Continued from page XX)

Variant	Transcripts <sup>b</sup>	Phase 2 results (10 sites) <sup>c</sup>											
		1*	2 <sup>^</sup>	3*	4*	8 <sup>^</sup>	9 <sup>^</sup>	16*	17*	18*	19*		
BRCA1 c.594-2A>C	Full length	+	+	+	+	+	+	+	+	+	+	NA	
	Δ9	-	+	-	-	+	+	-	-	-	-		
	Δ9,10	+	+	+	+	+	+	+	+	+	+		
	Δ9,10,11q	+	+	+	-	+	+	+	+	-	-		
	Δ9,10,11	+	+	-	-	-	+	-	-	-	-		
	Δ10	+	+	+	+	+	+	+	+	+	+		
	Δ11q	+	+	+	+	+	+	+	+	-	-		
	ins i21	-	+	-	-	+	-	-	-	-	-		
	Δ9,11q	-	+	-	-	+	+	+	-	-	-		
	Δ10,11q	-	+	+	-	+	+	+	+	-	-		
	Δ11q+insi13	-	+	-	-	+	-	-	-	-	-		
		Phase 2 BRCA1 c0.671-2A>G results (12 sites)											
		1*	2 <sup>^</sup>	3*	4*	8 <sup>^</sup>	9 <sup>^</sup>	16*	17*	18*	20 <sup>^</sup>	21*	22 <sup>^</sup>
BRCA1 c0.671-2A>G	Full length	+	+	+	+	+	+	+	+	+	+	+	+
	Δ9	-	+	+	-	+	-	-	-	-	+	-	-
	Δ9,10	+	+	+	+	+	+	+	+	+	+	+	-
	Δ9,10,11	+	+	+	+	+	+	+	+	-	+	+	-
	Δ9,10,11q	+	+	-	-	+	+	-	-	+	-	-	-
	Δ10	-	+	-	-	+	-	-	-	-	-	-	-
	Δ11	+	+	+	+	-	+	+	+	+	+	-	+
	Δ11q	-	+	-	+	+	+	-	-	+	+	-	+
	Δ9,10,11,12	-	-	-	-	-	-	+	-	-	-	-	-
	Δ9,11	-	+	-	-	-	-	-	-	-	-	-	-
	Δ10,11	+	+	+	+	+	+	+	+	-	-	-	+
	Δ10,11q	-	-	-	-	+	-	-	-	-	-	-	-
	Δ11&12	-	+	-	-	-	-	-	-	-	+	-	-
	Δ~3.2kb exon 11	+	-	-	-	-	-	-	-	-	-	-	+
	Δ9,11q	-	+	-	-	+	-	-	-	-	-	-	-
	ins i21	-	-	-	-	+	-	+	-	-	-	-	-
Δ11q+insi13	-	+	-	-	+	-	-	-	-	-	-	-	

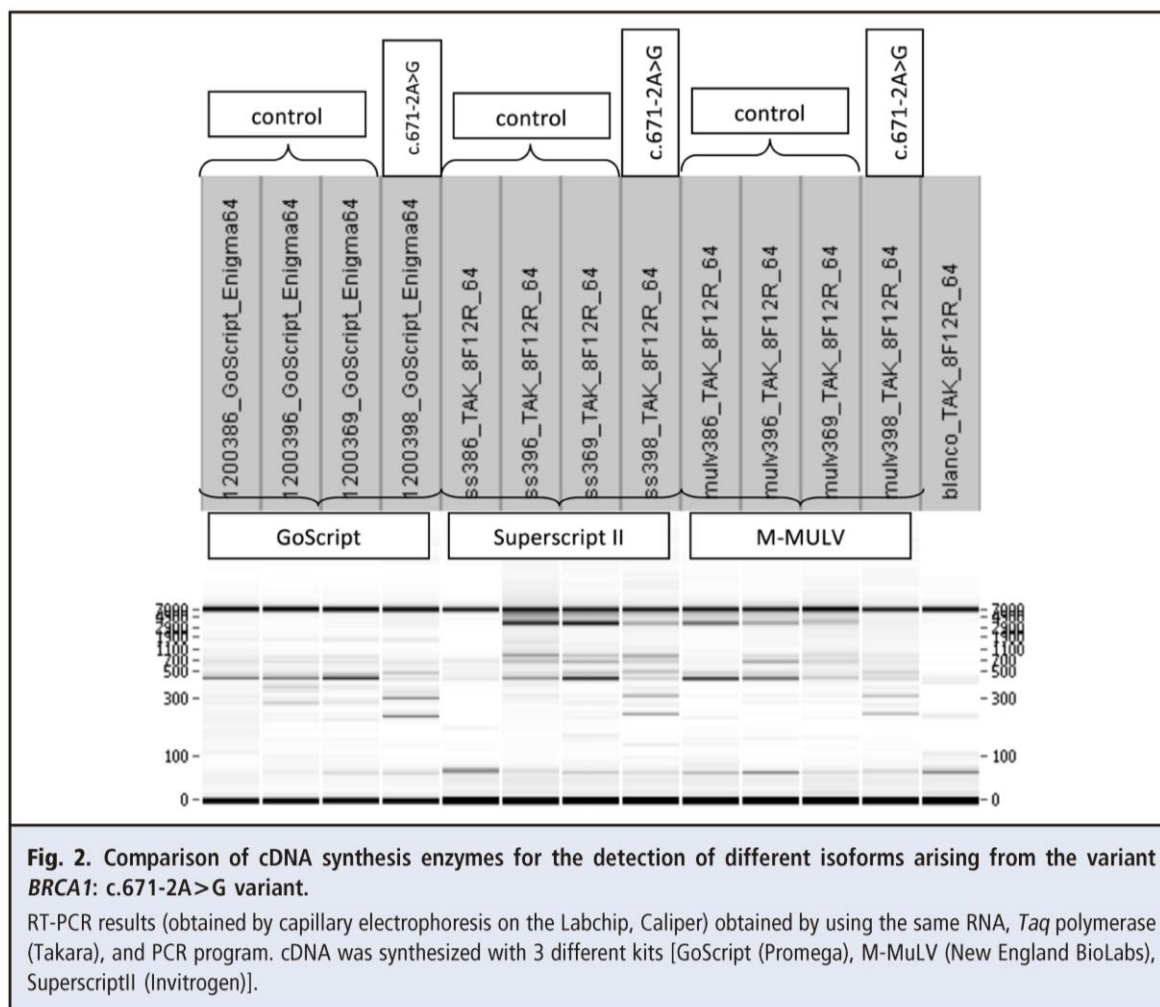
<sup>a</sup> +, detection of the transcript; -, no detection.  
<sup>b</sup> Transcripts were identified in at least one control. The Δ6,7 transcript identified in the analysis of BRCA2:c.426-12\_8delGTTT is not sequence confirmed and it is possible that this transcript is actually Δ6q,7, which differs by only 2 bp. Splice slipping events were also reported by several sites, for example 442\_444del3 from BRCA1 c.591C>T by site 11. Δ8p and Δ14p splice slipping events were also reported by site 11 and 17 following a review of the data after results were initially compiled. Additional minor and uncharacterized peaks were detected by some sites but have not been included in the table. HGVS (Human Genome Variation Society) nomenclature descriptions are available in online Supplemental Table 3.  
<sup>c</sup> \*, Agarose gel detection; ^, other detection; NA, not assessed. Sites 2 and 8 used capillary EP, 9 used Qiaxcel, 20 used MultiNA, 22 used Labchip GX.

Phase 1, followed by phase 2, demonstrated that the selection of primers used to amplify exons and the design of cycling conditions appropriate for that primer design explain the vast majority of the differential success of detecting some isoforms. In phase 2 of the project, during which primer design and extension time were controlled, all sites detected the full-length transcript and the predominant alternative transcripts, suggesting that high-abundance aberrant transcripts

will be detectable regardless of assay protocol, which is consistent with the conclusions of Houdayer et al. (12).

Variability in overall detection increased as the apparent abundance of individual transcripts in a sample decreased, and thus detection became more dependent on the sensitivity of the method of analysis. This variability is also likely to occur between replicates done in a single laboratory, in addition to that between different laboratories. A controlled comparison of different

## Guidelines for mRNA Splicing Assays



reverse transcriptases showed that Superscript is much better able to copy longer transcripts (Fig. 2). It is also possible that the maximum span length of some PCR polymerases contributed to the ability of some groups to detect longer transcripts. Furthermore, primer pairs that selectively amplify disease-associated isoforms rather than naturally occurring isoforms could increase assay sensitivity.

Sites that used gel electrophoresis visualization alone were unable to detect some bands because of the inherent insensitivity of this technique, combined with the stochastic nature of PCR when analyzing low levels of target (26). An example of this is site 1, which when analyzing the unequivocal variant *BRCA2* c.8632+1G>A detected the ins 21bp intron 9 in phase 1 but not in phase 2, despite using the same primers and PCR conditions.

Some sites sequenced PCR products. Sites that directly sequenced the products of PCR reactions expe-

rienced some challenges in determining the sequence of low-level transcripts. An accurate assessment of transcript sequence was also confounded by the presence of multiple (3 or more) PCR products of similar lengths. In these instances, adjustments to the concentration of agarose and running times of electrophoresis may improve analytical sensitivity. However, it appears that this may be less relevant if CE systems are adopted (see below). Cloning single PCR products into a vector system is a useful alternative for isolating transcripts and appears to improve sensitivity over band excision and sequencing alone. Furthermore, by increasing the number of clones screened it is possible to marginally increase the number of transcripts detected. However, to identify low-abundance transcripts, analysis of very large numbers of clones (100s or 1000s) or next generation sequencing would be necessary.

Of all the detection methods used, CE was shown to be the most analytically sensitive. For example, site 8

Table 3. Protocol recommendations.	
Protocol	Recommendation
NMD inhibitor	Need for this depends on sensitivity of detection method. For agarose gel detection, either CHX (cycloheximide) or puromycin are recommended.
RNA extraction	RNA extraction protocols were indistinguishable.
Dnase treatment	Recommended to avoid genomic DNA contamination.
cDNA synthesis primer(s)	Gene-specific OR oligodT + random hexamers are recommended.
cDNA synthesis	SuperscriptII reverse transcriptase is better for longer transcripts.
PCR primers	Forward and reverse primers must be at least 1 whole exon 5' or 3' of variant, respectively.
PCR polymerase	Different PCR polymerases were indistinguishable.
PCR conditions	Extension time long enough to copy amplicon (see enzyme manufacturer's instructions). At least 30 cycles.
Detection	CE was the most sensitive, followed by Qiaxcel, and then sequencing and agarose gel electrophoresis.
Sequencing	Cloning and sequencing is more sensitive than direct sequencing, but need to sequence at least 40 clones.

showed an increase in sensitivity from phase 1 to phase 2 after switching from using a Bioanalyzer to using CE. In addition to analytical sensitivity, the CE system has the added advantage of a greater resolution (1–2 bp) compared to Qiaxcel (3–5 bp). However, the limitation with both the Bioanalyzer and CE is the inability to harvest and thus perform sequence analysis of the PCR product. Also, CE relies on a prediction of the splicing event based on the length of the product observed, which can be limited by the inaccuracy of size standards, so a secondary set of primers may be required. It is also worth noting that very long full-length (or alternative) transcripts (like those involving *BRCA1* exon11 and *BRCA2* exons 10 and 11) cannot be analyzed by CE.

The results presented here represent each laboratory's initial assessment of each variant. Each site had the opportunity to reassess their results after the data from all sites were released to the group and several sites reported that they detected additional transcripts in addition to (and thus not shown) the initial conclusions reported in Tables 1 and 2. This finding suggests that a prior knowledge of all potential splice transcripts related to variant carriers, from studies such as these, as well as those that occur as naturally occurring isoforms in healthy controls, is essential not only to design detection strategies (see above) but to interpret results.

The use of analytically sensitive PCR product detection (CE and Qiaxcel in phases 1 and 2, Bioanalyzer in phase 1) enabled the identification of several novel low-abundance transcripts, in both normal controls and variant carriers. This raises the question of which detectable transcripts are functional and thus relevant for determining the pathogenicity of clinically identified unclassified variants, and whether or not low abundance transcripts are of biological or pathological

significance in vivo. It is generally accepted that variants resulting in single major transcripts that lack an open reading frame will be deleterious (27). However, it is much less clear whether changes in the levels of low-abundance alternative splicing events will have an impact either directly or through altering the function or levels of endogenous transcripts including full-length mRNA.

It is possible, for example, that a reduction in the full-length expression will have a deleterious effect on known *BRCA1* functions (DNA repair, cell cycle control) (28). A quantitative analysis of the range of naturally occurring isoforms relative to full-length expression and relative to other *BRCA1* or *BRCA2* isoforms is required, as is a comprehensive analysis of the functional role of each of these isoforms in both the healthy functioning of *BRCA* genes and the consequences of sequence variation on this process (29). It will also be important to extend this investigation to breast and ovarian tissue, to gain a broader understanding of the tissue-specific nature of splice-isoform regulation. Importantly, this information will be essential to determine whether knowing the full complement of transcripts has the potential to have an impact on the final classification of the variant as pathogenic or otherwise. For example, does the expression profile of the 16 alternately spliced transcripts detected in *BRCA1* c.671-2A>G carriers change at different tissue sites, and will this new information influence the classification of the variant?

In summary, we have shown that primer design, PCR conditions, and PCR product detection methodology, together with prior knowledge of potential transcripts, are important contributors to the analytical sensitivity of PCR-based assays for detecting alternatively spliced RNA transcripts from variant carriers and

## Guidelines for mRNA Splicing Assays

from wild-type sequences. These factors must be considered when designing assays, particularly when they form the basis of clinical decision-making. Furthermore, the formulation of standard assay design and detection methods is indicated for all variants, but particularly for those that may impact on isoform expression.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

**Authors' Disclosures or Potential Conflicts of Interest:** Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

**Employment or Leadership:** S. Neuhausen, the Morris and Horowitz Families Endowed Professor;

**Consultant or Advisory Role:** None declared.

**Stock Ownership:** None declared.

**Honoraria:** None declared.

**Research Funding:** The Clinical Follow Up Study was funded 2001–2009 by NHMRC and currently by the National Breast Cancer Foundation and Cancer Australia #628333. kConFab is supported by grants from the National Breast Cancer Foundation, the National Health and Medical Research Council (NHMRC) and by the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. The Beckman Research Institute of City of Hope (BRICOH) study was supported by Grant Number P30 CA033572 from the National Cancer Institute. R. Brandao, Fundação para a Ciência e Tecnologia; A.B. Spurdle, Australian NHMRC Project grant ID #1010719, The University of Queensland, NHMRC

Senior Research Fellowship; M.A. Brown, Australian NHMRC Project grant ID #1010719, The University of Queensland; C. Lazaro, the Spanish Health Research Fund, Carlos III Health Institute, Catalan Health Institute, Government of Catalonia and the Spanish Association Against Cancer (Contract grant numbers: ISCIIRETIC: RD06/0020/1051, 2009SGR290, PI10/01422); C. Houdayer, INCA DHOS 2010 Recherche Translationnelle sur le Cancer, FASDEC grant; L.C. Walker, Health Research Council NZ Sir Charles Hercus Health Research Fellowship; D. Baralle, CRUK Project grant; M. Colombo, the Italian Association for Cancer Research (AIRC; grant number 11897); G. De Vecchi, the Italian Association for Cancer Research (AIRC; grant number 11897); P. Radice, the Italian Association for Cancer Research (AIRC; grant number 11897); M. de la Hoya, the NEYE Foundation, the Xunta de Galicia (10PXIB 9101297PR) and FMM Foundation and a Fondo de Investigación Sanitaria (FIS) research grant PI 12/00539, an initiative of the Instituto de Salud Carlos III (Spain), partially supported by European Regional; S. Gutierrez Enriquez, Miguel Servet contract from the Spanish Carlos III Health Institute - Miguel Servet Project grant CP10/00617; A. Tenes, Miguel Servet contract from the Spanish Carlos III Health Institute; A. Vega, Xunta de Galicia, FMM Foundation; O. Diez, Fondo de Investigación Sanitaria (FIS) research grant PI 12/02585 from the Spanish Carlos III Health Institute.

**Expert Testimony:** None declared.

**Patents:** None declared.

**Role of Sponsor:** The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

**Acknowledgments:** We thank all members of the ENIGMA consortium Splicing Working Group for useful suggestions relating to study execution and interpretation. We also wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow-Up Study for their contributions to this resource, and the many families who contribute to kConFab. K. Claes and K. De Leeener would like to thank Anneke Grunewald for her contributions to the project.

## References

- Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, et al. A comprehensive model for familial breast cancer incorporating *BRCA1*, *BRCA2* and other genes. *Br J Cancer* 2002;86:76–83.
- Chatterjee S, Pal JK. Role of 5'- and 3'- untranslated regions of mRNAs in human diseases. *Biol Cell* 2009;101:251–62.
- Brewster BL, Rossiello F, French JD, Edwards SL, Wong M, Wronski A, et al. Identification of fifteen novel germline variants in the *BRCA1* 3'UTR reveals a variant in a breast cancer case that introduces a functional miR-103 target site. *Hum Mutat* 2012;33:1665–75.
- Chen X, Truong T-TN, Weaver J, Bove BA, Cattie K, Armstrong BA, et al. Intronic alterations in *BRCA1* and *BRCA2*: effect on mRNA splicing fidelity and expression. *Hum Mutat* 2006;27:427–35.
- Lovelock PK, Healey S, Au W, Sum EY, Tesoriero A, Wong EM, et al. Genetic, functional, and histopathological evaluation of two C-terminal *BRCA1* missense variants. *J Med Genet* 2006;43:74–83.
- Farrugia DJ, Agarwal MK, Pankratz VS, Deffenbaugh AM, Pruss D, Frye C, et al. Functional assays for classification of *BRCA2* variants of uncertain significance. *Cancer Res* 2008;68:3523–31.
- Thomassen M, Blanco A, Montagna M, Hansen TV, Pedersen IS, Gutierrez-Enriquez S, et al. Characterization of *BRCA1* and *BRCA2* splicing variants: a collaborative report by ENIGMA consortium members. *Breast Cancer Res Treat* 2012;132:1009–23.
- Walker LC, Whiley PJ, Couch FJ, Farrugia DJ, Healey S, Eccles DM, et al. Detection of splicing aberrations caused by *BRCA1* and *BRCA2* sequence variants encoding missense substitutions: implications for prediction of pathogenicity. *Hum Mutat* 2010;31:E1484–505.
- Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, et al. Intronic variants in *BRCA1* and *BRCA2* that affect RNA splicing can be reliably selected by splice-site prediction programs. *Hum Mutat* 2009;30:107–14.
- Mucaki EJ, Ainsworth P, Rogan PK. Comprehensive prediction of mRNA splicing effects of *BRCA1* and *BRCA2* variants. *Hum Mutat* 2011;32:735–42.
- Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on *BRCA1* and *BRCA2* variants. *Hum Mutat* 2012;33:1228–38.
- Houdayer C, Dehainault C, Mattler C, Michaux D, Caux-Moncoutier V, Pages-Berhouet S, et al. Evaluation of in silico splice tools for decision-making in molecular diagnosis. *Hum Mutat* 2008;29:975–82.
- Radice P, De Summa S, Caleca L, Tommasi S. Unclassified variants in *BRCA* genes: guidelines for interpretation. *Ann Oncol* 2011;22(Suppl 1):i18–23.
- Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabomero I, Esteban-Cardenas E, et al. A high proportion of DNA variants of *BRCA1* and *BRCA2* is associated with aberrant splicing in breast/ovarian cancer patients. *Clin Cancer Res* 2010;16:1957–67.
- Walker LC, Whiley PJ, Houdayer C, Hansen TO, Vega A, Santamarina M, et al. Evaluation of a 5-tier scheme proposed for classification of sequence variants using bioinformatic and splicing assay data: inter-reviewer variability and promotion of minimum reporting guidelines. *Hum Mutat* 2013;34:1424–31.

16. Claes K, Poppe B, Madhockova E, Coene I, Foretova L, De Paepe A, Messiaen L. Differentiating pathogenic mutations from polymorphic alterations in the splice sites of *BRCA1* and *BRCA2* Genes. *Chrom Cancer* 2003;37:314–20.
17. Gayther SA, Warren W, Mazoyer S, Russell PA, Harrington PA, Chiano M, et al. Germline mutations of the *BRCA1* gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation. *Nat Genet* 1995;11:428–33.
18. Madhockova E, Foretova L, Lukesova M, Vasickova P, Navratilova M, Coene I, et al. Spectrum and characterisation of *BRCA1* and *BRCA2* deleterious mutations in high-risk Czech patients with breast and/or ovarian cancer. *BMC Cancer* 2008; 8:140.
19. Menendez M, Castellsague J, Mirete M, Pros E, Feliubadalo L, Osorio A, et al. Assessing the RNA effect of 26 DNA variants in the *BRCA1* and *BRCA2* genes. *Breast Cancer Res Treat* 2012;132: 979–92.
20. Vega A, Campos B, Bressac-De-Paillerets B, Bond PM, Janin N, Douglas FS, et al. The R71G *BRCA1* is a founder Spanish mutation and leads to aberrant splicing of the transcript. *Hum Mutat* 2001; 17:520–1.
21. Spurdle AB, Healey S, Devereau A, Hogervorst FB, Monteiro AN, Nathanson KL, et al. ENIGMA-evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in *BRCA1* and *BRCA2* genes. *Hum Mutat* 2012;33:2–7.
22. Tesoriero AA, Wong EM, Jenkins MA, Hopper JL, Brown MA, Chenevix-Trench G, et al. Molecular characterization and cancer risk associated with *BRCA1* and *BRCA2* splice site variants identified in multiple-case breast cancer families. *Hum Mutat* 2005;26:495.
23. Bonnet C, Krieger S, Vezain M, Rousselin A, Tournier I, Martins A, Berthet P, et al. 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* 2008;45: 438–46.
24. Thery JC, Krieger S, Gaildrat P, Revillon F, Buisine MP, Killian A, et al. Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the *BRCA* genes. *Eur J Hum Genet* 2011;19:1052–8.
25. Colombo M, De Vecchi G, Caleca L, Foglia C, Ripamonti CB, Ficarazzi F, et al. Comparative in vitro and in silico analyses of variants in splicing regions of *BRCA1* and *BRCA2* genes and characterization of novel pathogenic mutations. *PLoS One* 2013;8:e57173.
26. Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, et al. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res* 1996;24:3189–94.
27. Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Similnikova OM. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat* 2008;29:1304–13.
28. Orban TI, Olah E. Expression profiles of *BRCA1* splice variants in asynchronous and in G1/S synchronized tumor cell lines. *Biochem Biophys Res Commun* 2001;280:32–8.
29. Sevcik J, Falk M, Macurek L, Kleiblova P, Lhota F, Hojny J, et al. Expression of human *BRCA1* Delta17–19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response. *Cell Signal* 2013;25:1186–93.

## Supplementary Material

Supplementary Table 1 PCR primers and conditions used in Phase 2 of the study.

Variant	Forward Primer (1)	Reverse Primer (1)	full-length size (bp)	Recommended extension	Recommended annealing temp °C	Forward Primer (2)	Reverse Primer (2)	full-length size (bp)	Recommended extension	Recommended annealing temp °C
BRCA2c.7988 A>T	GAGTCTTTTCAGTTTCACACTG	CAAGTTTGGTATACCAGCGAG	706	45s to 1min30s	55					
BRCA2c.426-12_8delGTTTT	GCAGATGATGTTTCCTGTCC	TGACTTGCAGCTTCTCTTGA	413	45s to 1min15s	55					
BRCA1c.591 C>T	CCAACTCTTAACCTTGGAAGTGTG	CTCCAGCCCATCTGTTATGTTG	510	1min to 2min	56	CCAACTCTTAACCTTGGAAGTGTG	GATGACCTTCCACTCCTGGTTC	719 (del11q)	1min to 2min	56
BRCA1c.594-2 A>C	CCAACTCTTAACCTTGGAAGTGTG	CTCCAGCCCATCTGTTATGTTG	510	1min to 2min	56	CCAACTCTTAACCTTGGAAGTGTG	GATGACCTTCCACTCCTGGTTC	719 (del11q)	1min to 2min	56
BRCA1c.671-2 A>G	CCAACTCTTAACCTTGGAAGTGTG	CTCCAGCCCATCTGTTATGTTG	510	1min to 2min	56	CCAACTCTTAACCTTGGAAGTGTG	GATGACCTTCCACTCCTGGTTC	719 (del11q)	1min to 2min	56
BRCA2c.9501+3A>T	TAGGAAGGCCATGGAATCTG	TCCTTTGGCCATACAAAGTG	869	45s to 1min30s	60					
BRCA2c.8632+1G>A	GAAGCCCCAGAATCTCTTATG	CTCTTAAGTCTTCACTG	457	45s to 1min15s	54					
BRCA1c.135-1 G>T	ATGGATTATCTGCTCTTCGC	CCAATCAATGTAGACAGACG	456	45s to 1min15s	55					
BRCA1c.5467+5G>C	GGAGATGTGGTCAATGGAAGAAAC	TGCTACACTGTCCAACCCCACTC	318	45s to 1min30s	62					

Supplementary Table 2. HGVS nomenclature descriptions of variant transcripts.

Phase 1

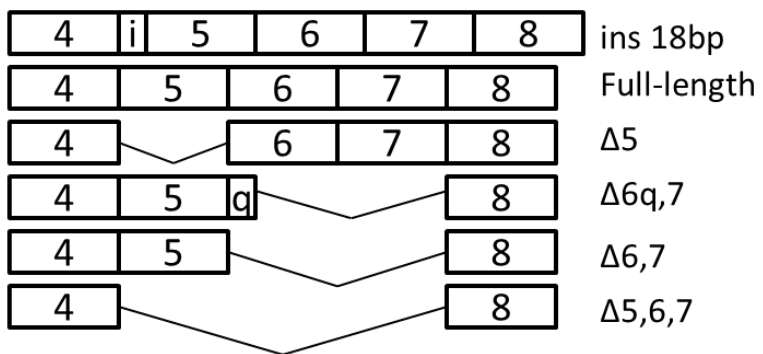
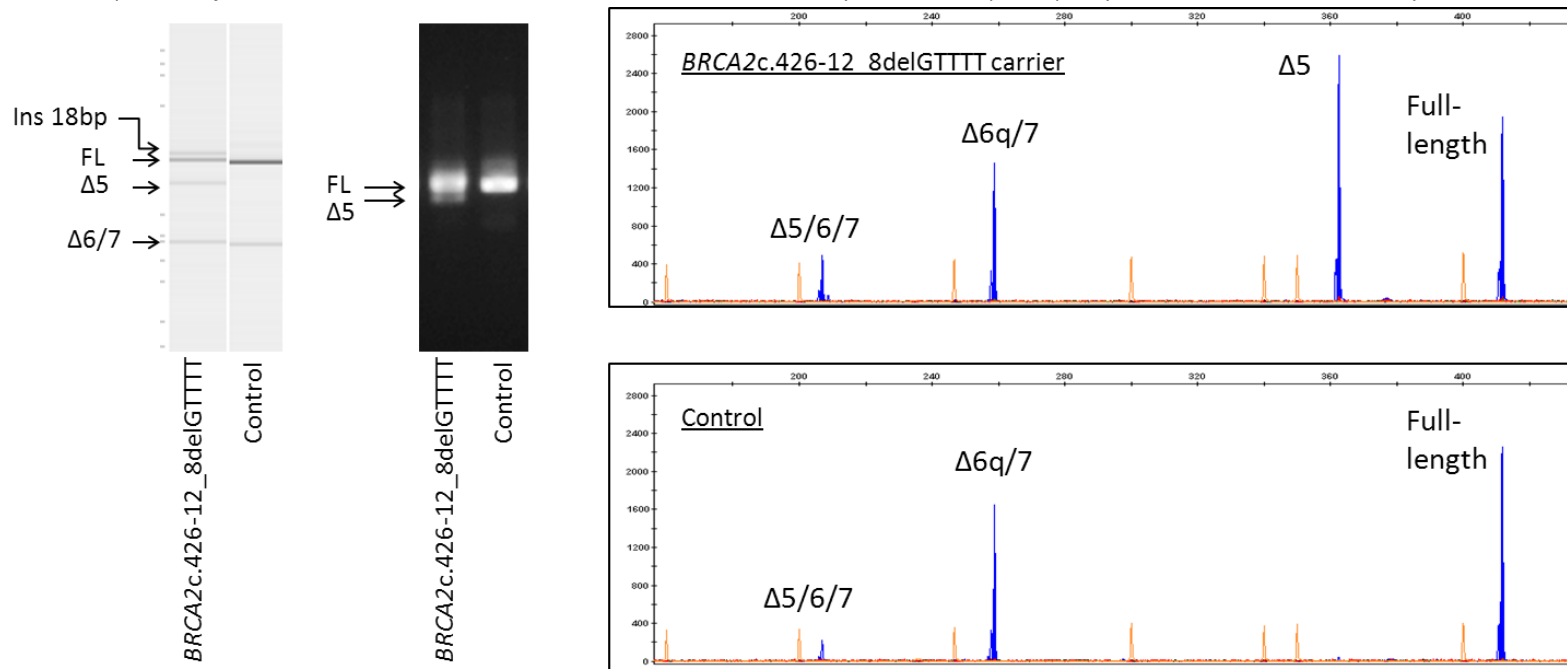
Variant	Transcript	HGVS
<b>BRCA2c.9501+3A&gt;T</b>	Full-length	
	<b>Δ25</b>	<b>c.9257_9501del245</b>
<b>BRCA2c.8632+1G&gt;A</b>	Full-length	
	<b>Δ20</b>	<b>c.8488_8632del145</b>
	<b>Δ19&amp;20</b>	<b>c.8332_8632del301</b>
	ins i20	c.8633-1327_c.8633-1264ins64
<b>BRCA1c.135-1 G&gt;T</b>	Full-length	
	<b>Δ5</b>	c.135_212del78
	<b>Δ5q</b>	c.191_212del22
	<b>Δ3</b>	c.81_134del54
<b>BRCA1c.5467+5G&gt;C</b>	Full-length	
	<b>Δ23</b>	<b>c.5407_5467del61</b>
<b>BRCA2c.7988 A&gt;T</b>	Full-length	
	<b>Δ18</b>	c.7977_8331del355
	<b>Δ17&amp;18</b>	c.7806_8331del526
<b>BRCA2c.426-12_8delGTTTT</b>	Full-length	
	<b>Δ6q,7</b>	c.478_631del154
	<b>Δ3,5</b>	<b>c.68_316del249_c.426_475del50</b>
	<b>Δ3,4,5</b>	<b>c.68_475del408</b>
	ins 18bp	<b>c.426-18_c.426-1ins18</b>
	<b>Δ 4,5,6,7</b>	<b>c.317_631del315</b>
	<b>Δ5</b>	c.426_475del50
	<b>Δ5,6,7</b>	c.426_631del206
	<b>Δ6,7</b>	c.476_631del156
<b>BRCA1c.591 C&gt;T</b>	Full-length	
	<b>Δ 8&amp;9</b>	c.442_593del152
	<b>Δ 8,9,10</b>	c.442_670del229
	<b>Δ9</b>	c.548_593del46
	<b>Δ9,10</b>	c.548_670del123
	<b>Δ9,10,11q</b>	c.548_670del123_c.788_4096del3309
	<b>Δ11q</b>	c.788_4096del3309
	<b>Δ9,11q</b>	<b>c.548_593del46_c.788_4096del3309</b>
	<b>Δ10</b>	<b>c.594_670del77</b>
<b>BRCA1c.594-2 A&gt;C</b>	Full-length	
	<b>Δ9</b>	c.548_593del46
	<b>Δ9,10</b>	c.548_670del123
	<b>Δ9,10,11q</b>	c.548_670del123_c.788_4096del3309
	<b>Δ9,10,11</b>	c.548_4096del3549
	<b>Δ11q</b>	c.788_4096del3309
	ins 21bp intron 9	<b>c.592-20_c.592-1ins21</b>
	<b>Δ9,11q</b>	<b>c.548_593del46_c.788_4096del3309</b>
	<b>Δ10,11q</b>	<b>c.594_670del77_c.788_4096del3309</b>
		<b>Δ10</b>

Variant	Transcript	HGVS
<b>BRCA1c.671-2 A&gt;G</b>	Full-length	
	<b>Δ9,11</b>	<b>c.548_593del46_c.671_4096del3426</b>
	<b>Δ10,11</b>	<b>c.594_4096del3503</b>
	<b>Δ11&amp;12</b>	<b>c.671_4185del3515</b>
	<b>Δ~3.2kb exon 11</b>	<b>multiple minor transcripts</b>
	<b>Δ9,10,11,12</b>	<b>c.548_4185del3638</b>
	<b>Δ9</b>	<b>c.548_593del46</b>
	<b>Δ9,10</b>	<b>c.548_670del123</b>
	<b>Δ9,10,11</b>	<b>c.548_4096del3549</b>
	<b>Δ9,10,11q</b>	<b>c.548_670del123_c.788_4096del3309</b>
	<b>Δ11</b>	<b>c.671_4096del3426</b>
	<b>Δ11q</b>	<b>c.788_4096del3309</b>

## Phase 2

Variant	Transcripts	HGVS
<b>BRCA2c.9501+3A&gt;T</b>	Full-length	
	<b>Δ25</b>	<b>c.9257_9501del245</b>
	<b>ins i23</b>	<b>c.9117+1_c.9118-1ins93</b>
<b>BRCA2c.8632+1G&gt;A</b>	Full-length	
	<b>Δ19</b>	<b>c.8332_8487del156</b>
	<b>Δ20</b>	<b>c.8488_8632del145</b>
	<b>Δ19&amp;20</b>	<b>c.8332_8632del301</b>
	<b>ins i20</b>	<b>c.8633-1327_c.8633-1264ins64</b>
<b>extra peaks</b>	<b>multiple minor transcripts</b>	
<b>BRCA1c.135-1 G&gt;T</b>	Full-length	
	<b>Δ5q</b>	<b>c.191_212del22</b>
	<b>Δ5</b>	<b>c.135_212del78</b>
	<b>Δ3</b>	<b>c.81_134del54</b>
	<b>Δ3,5</b>	<b>c.81_212del132</b>
	<b>Δ5,6</b>	<b>c.135_301del167</b>
	<b>ins i3</b> <b>ins i3+Δ5</b>	<b>c.135-4047_135-3932ins116</b> <b>c.135-4047_135-3932ins116_c.135_212del78</b>
<b>BRCA1c.5467+5G&gt;C</b>	Full-length	
	<b>Δ21</b>	<b>c.5278_5332del55_c.</b>
	<b>Δ21,23</b>	<b>c.5278_5332del56_c.5407_5467del61</b>
	<b>Δ22,23</b>	<b>c.5333_5467del135</b>
	<b>Δ23</b> <b>Δ22</b>	<b>c.5407_5467del61</b> <b>c.5333_5406del74</b>
<b>BRCA2c.7988 A&gt;T</b>	Full-length	
	<b>Δ18</b>	<b>c.7977_8331del355</b>
	<b>Δ17&amp;18</b>	<b>c.7806_8331del526</b>
<b>BRCA2c.426-12_8delGTTT</b>	FL	
	<b>Δ6q,7</b>	<b>c.478_631del154</b>
	<b>ins 18bp</b>	<b>c.426-18_c.426-1ins18</b>

Supplementary Figure 1. Qiaxcel visualisation (left), gel electrophoresis (centre) and capillary EP (right) comparison for analysis of BRCA2: c.426-12\_8delGTTTT. Increased separation of bands enabled the Qiaxcel tool to discriminate between similarly sized transcripts. Capillary EP demonstrates the sensitivity to detect the low abundance isoform  $\Delta 5/6/7$



*Aproximadamente el 1% de las familias españolas con cáncer de mama y ovario negativas para BRCA1 y BRCA2 son portadoras de variantes patogénicas en RAD51D*

Recientemente se han identificado mutaciones en *RAD51D* en familias con cáncer de mama y ovario. Aunque parece obvio un rol etiológico en el cáncer de ovario, la asociación de mutaciones en *RAD51D* y el cáncer de mama no es tan claro. Nuestro objetivo ha sido determinar la prevalencia de mutaciones germinales en *RAD51D* en familias HBOC españolas. Se analizaron 842 casos índice: 491 de familias con cáncer de mama y ovario, 171 de familias con cáncer de mama, 51 de familias con cáncer de ovario y 129 pacientes sin historia familiar pero con cáncer de mama o ovario de diagnóstico temprano o bien cáncer de mama y ovario metacrónico. La detección de mutaciones se realizó mediante HRM, DHPLC o secuenciación sanger.

Se identificaron tres mutaciones en cuatro familias con casos de cáncer de mama y ovario (0,82%): las mutaciones nuevas c.1A>T (p.Met1?) y c.667+2\_667+23del (produce la exclusión del exon 7 del ARNm maduro) y c.674C>T (p.Arg232\*), mutación ya descrita previamente. c.667+2\_667+23del mostró co-segregación en la familia con un caso de cáncer de mama de diagnóstico temprano y dos casos de cáncer de mama bilateral. Además se identificó la variante c.629C>T (p.Ala210Val), potencialmente patogénica según las predicciones *in silico*.

Aproximadamente el 1% de las familias BRCA1/2 españolas son portadoras de mutaciones en *RAD51D*. La presencia de varios portadores de mutación afectos de cáncer de mama sugiere un incremento del riesgo de cáncer de mama, aunque en el contexto de cáncer de ovario familiar. Este riesgo debería ser considerado para el desarrollo de medidas de seguimiento y diagnóstico temprano en portadores. El análisis de *RAD51D* debería ofrecerse en el contexto clínico a familias con cáncer de mama y ovario, independientemente del número de casos de cáncer de ovario presentes en la familia.





IJC

International Journal of Cancer

## About 1% of the breast and ovarian Spanish families testing negative for *BRCA1* and *BRCA2* are carriers of *RAD51D* pathogenic variants

Sara Gutiérrez-Enríquez<sup>1</sup>, Sandra Bonache<sup>1,2</sup>, Gorka Ruíz de Garibay<sup>3</sup>, Ana Osorio<sup>4,5</sup>, Marta Santamaría<sup>6</sup>, Teresa Ramón y Cajal<sup>7</sup>, Eva Esteban-Cardenosa<sup>8</sup>, Anna Tenés<sup>1</sup>, Kira Yanowsky<sup>4,5</sup>, Alicia Barroso<sup>4,5</sup>, Gemma Montalban<sup>1</sup>, Ana Blanco<sup>9</sup>, Mònica Cornet<sup>10</sup>, Neus Gadea<sup>11</sup>, Mar Infante<sup>8</sup>, Trinidad Caldés<sup>3</sup>, Eduardo Díaz-Rubio<sup>12</sup>, Judith Balmaña<sup>13</sup>, Adriana Lasa<sup>10</sup>, Ana Vega<sup>9</sup>, Javier Benítez<sup>4,5</sup>, Miguel de la Hoya<sup>3</sup> and Orland Diez<sup>1,2,14</sup>

<sup>1</sup>Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>2</sup>Oncogenetics Group, Vall d'Hebron Research Institute (VHIR) and Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>3</sup>Laboratorio de Oncología Molecular, Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos, Madrid, Spain

<sup>4</sup>Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain

<sup>5</sup>Centre for Biomedical Network Research on Rare Diseases (CIBERER), Madrid, Spain

<sup>6</sup>Grupo de Medicina Xenómica-USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain

<sup>7</sup>Medical Oncology Department, Santa Creu i Sant Pau Hospital, Barcelona, Spain

<sup>8</sup>Instituto de Biología y Genética Molecular (UVa-CSIC), Genética del Cáncer, Valladolid, Spain

<sup>9</sup>Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica, CIBERER, IDIS, Santiago de Compostela, Spain

<sup>10</sup>Genetics Department, CIBERER U-705, Santa Creu i Sant Pau Hospital, Barcelona, Spain

<sup>11</sup>Medical Oncology Department, University Hospital of Vall d'Hebron, Barcelona, Spain

<sup>12</sup>Servicio de Oncología Médica, Hospital Clínico San Carlos, Madrid, Spain

<sup>13</sup>Medical Oncology Department, University Hospital of Vall d'Hebron, Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>14</sup>Oncogenetics Group, University Hospital of Vall d'Hebron, Barcelona, Spain

*RAD51D* mutations have been recently identified in breast (BC) and ovarian cancer (OC) families. Although an etiological role in OC appears to be present, the association of *RAD51D* mutations and BC risk is more unclear. We aimed to determine the prevalence of germline *RAD51D* mutations in Spanish BC/OC families negative for *BRCA1/BRCA2* mutations. We analyzed 842 index patients: 491 from BC/OC families, 171 BC families, 51 OC families and 129 patients without family history but with early-onset BC or OC or metachronous BC and OC. Mutation detection was performed with high-resolution melting, denaturing high-performance liquid chromatography or Sanger sequencing. Three mutations were found in four families with BC and OC cases (0.82%). Two were novel: c.1A>T (p.Met1?) and c.667+2\_667+23del, leading to the exon 7 skipping and one previously described: c.674C>T (p.Arg232\*). All were present in BC/OC families with only one OC. The c.667+2\_667+23del cosegregated in the family with one early-onset BC and two bilateral BC cases. We also identified the c.629C>T (p.Ala210Val) variant, which was predicted *in silico* to be potentially pathogenic. About 1% of the BC and OC Spanish families negative for *BRCA1/BRCA2* are carriers of *RAD51D* mutations. The presence of several BC mutation carriers, albeit in the context of familial OC, suggests an increased risk for BC, which should be taken into account in the follow-up and early detection measures. *RAD51D* testing should be considered in clinical setting for families with BC and OC, irrespective of the number of OC cases in the family.

**Key words:** *RAD51D*, genetic predisposition, familial breast and ovarian cancer

Additional Supporting Information may be found in the online version of this article.

Conflicts of interest: Nothing to report

**Grant sponsor:** Miguel Servet Program, Spanish Carlos III Health Institute; **Grant number:** CP10/00617; **Grant sponsor:** Fondo de Investigación Sanitaria (FIS) Research; **Grant number:** PI 12/00539; **Grant sponsor:** Intrasalud; **Grant number:** PI12/00070; **Grant sponsor:** Xunta de Galicia; **Grant number:** 10PXIB 9101297PR; **Grant sponsors:** Mutua Madrileña Foundation (FMMA); **Grant sponsor:** Acción Cooperativa y Colaborativa Intramural-CIBER2012, SAF2010-20493; **Grant sponsor:** Cancer Prevention Program of the Regional Government of Castilla y León

**DOI:** 10.1002/ijc.28540

**History:** Received 5 Aug 2013; Accepted 27 Sep 2013; Online 15 Oct 2013

**Correspondence to:** Dr. S. Gutiérrez-Enríquez, Oncogenetics Laboratory, Vall d'Hebron Institute of Oncology (VHIO), Planta 2 Modular Laboratoris Clinics (modular blau), Hospital of Vall d'Hebron, Barcelona, Pg Vall d'Hebron 119-129, Barcelona 08035, Spain, Tel.: +34-93-27460-00 ext 4826, Fax: +34-93-274-68-37, E-mail: sagutier@vhebron.net

Int. J. Cancer: 00, 00–00 (2013) © 2013 UICC

#### What's new?

*RAD51D* mutations have recently been identified in breast (BC) and ovarian (OC) cancer families. Although *RAD51D* mutations are associated with hereditary OC, such an association is less clear in BC. This study determined the prevalence of germline *RAD51D* mutations in Spanish BC/OC families testing negative for *BRCA1/BRCA2* mutations. *RAD51D* mutations were found in almost 1% (4/491) of BC/OC families. These families had only one OC case, with some carriers presenting BC, suggesting that *RAD51D* testing should be offered to all BC/OC families. The existence of an increased risk of BC should be considered when setting the follow-up and prevention measures.

Ovarian cancer (OC) represents the eighth most commonly diagnosed cancer among women in the world, and causes more deaths per year than any other cancer of the female reproductive system.<sup>1</sup> On a worldwide basis, 224,747 new cases (standardized incidence rate of 6.3 per 100,000 women) are diagnosed and 140,163 women die of OC annually (estimated in 2008).<sup>1</sup> Mortality is high because women typically present with late-stage disease when the overall 5-year relative survival rate is 45% (reviewed in Ref.2). The most important risk factor is having a family history of the disease (reviewed in Ref.3). Approximately 25% of newly diagnosed cases are due to a hereditary mutation in a single gene: 18% in *BRCA1* or *BRCA2*; 5% in other genes of the Fanconi-Anemia-BCRA DNA repair pathway, including *MRE11*, *BRIP1*, *PALB2*, *RAD51C*, *CHEK2*, *NBN*, *RAD50* and *BARD1*; 1% in genes of DNA mismatch repair pathways (*MLH1*, *MSH2*, *MSH6* and *PMS2*) and less than 1% in other genes such as *TP53*.<sup>3-5</sup>

The identification of *RAD51C* mutations in families with breast and ovarian cancer (BC/OC) prompted investigations on the role of another *RAD51* paralog, *RAD51D*, in cancer susceptibility.<sup>6</sup> The protein encoded by *RAD51D* is a member of the RAD51 protein family. The RAD51 protein is a key player in the homologous recombinational repair pathway and its activity appears to be strictly regulated by a number of cofactors including five RAD51 paralogs, namely RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3.<sup>7</sup> RAD51D protein forms a complex with RAD51B, RAD51C and XRCC2 (BCDX2 complex),<sup>8</sup> which is responsible for RAD51 recruitment or stabilization at DNA damage sites.<sup>9</sup> Additionally, it has been described that RAD51D plays a role in telomere protection against attrition and chromosome fusion.<sup>10</sup> Love-day *et al.*<sup>6</sup> identified inactivating mutations in *RAD51D* in eight of the 911 British cases with a familial history of BC/OC (0.88%) and in one of the 1,060 British controls (0.09%). The association was principally found with OC because the higher prevalence of mutations was present in families with more than one case of OC: four mutations in 235 families with two or more cases of OC (1.7%) and three mutations among the 59 families with three or more cases of OC (5.09%).<sup>6</sup> By contrast, no mutations were found in 737 unrelated individuals with only a history of familial BC. The authors estimated an OC relative risk of 6.30 and a BC relative risk of 1.32 for *RAD51D* mutation carriers.<sup>6</sup> Four further studies have explored the role of *RAD51D* mutations in BC/

OC families as well as in unselected OC cases.<sup>11-14</sup> These works show that *RAD51D* is an OC predisposition gene, but more studies in familial and sporadic OC series would be of value to further clarify the risks associated to OC and BC.<sup>3</sup>

The aim of our study was to determine the prevalence of germline *RAD51D* mutations in a cohort of Spanish BC and/or OC families previously found to be negative for *BRCA1* and *BRCA2* mutations.

#### Material and Methods

##### Patients

The study included a total of 842 unrelated index cases: 713 cases from families with a history of OC and/or BC cases and 129 cases with only a personal history of early-onset BC or OC or concomitant BC and OC. Six centers from Spain have contributed to this study: Hospital Universitari Vall d'Hebron, Barcelona (HUVH) ( $N = 281$ ); Hospital Clínico San Carlos (HCSC), Madrid ( $N = 168$ ); Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid ( $N = 165$ ); Fundación Pública Galega de Medicina Xenómica (FPGMX), Santiago de Compostela ( $N = 94$ ); Hospital de la Santa Creu i Sant Pau (HSCSP), Barcelona ( $N = 79$ ) and Instituto de Biología y Genética Molecular (IBGM), Valladolid ( $N = 55$ ). All index cases were previously screened for point mutations and large rearrangements in *BRCA1* and *BRCA2* genes and no disease-causing mutation was identified.

The studied families were as follows: (i) 491 families with one or more OC cases and one or more BC cases (77% of them had at least one case diagnosed before age 50); (ii) 171 site-specific female BC families with two or more cases (94% of families with at least one case diagnosed before age 50) and (iii) 51 families with two or more OC cases. All the probands analyzed, except 19 healthy cases, were diagnosed with BC (291) or OC (120) or BC + OC (55) or other cancers (6). The study also included patients without family history: 65 BC patients (95% of them diagnosed before age 50), 38 women with OC (82% of them diagnosed before age 50) and 26 with BC and OC. This study was approved by the corresponding institutional ethical committees and informed consent was obtained from all the participants.

##### Mutation analysis of *RAD51D*

Genomic DNA was extracted from patient peripheral blood samples by standard methods. The screening for germline variation of the coding region and exon-intron boundaries of

the *RAD51D* was performed using different techniques: direct sequencing (HUVH, FPGMX and HSCSP), high-resolution melting (HCSC and IBGM) and denaturing high-performance liquid chromatography (CNIO). All sequence variants were named according to Human Genome Variation Society (HGVS) nomenclature using the NCBI transcript ID NM\_002878.3 (or its equivalent sequence from Ensembl transcript ID ENST00000345365) as a reference sequence. The exons were numbered according to the NCBI Reference Sequence: NG\_031858.1.

#### *In silico* analysis of identified variants

All *RAD51D* genetic variants (coding and noncoding) were analyzed for their potential effect on splicing. For this analysis the variants are reported according to their location, relative to the 5' (donor) and 3' (acceptor) consensus sites described by Cartegni *et al.*,<sup>15</sup> that is 11 bases for the donor site (from the three last exonic to the eight first intronic bases) and 14 bases for the acceptor site (from the 12 last intronic to the first two exonic bases). Thus, the variants were classified in three categories: variants occurring at consensus donor or acceptor splice sites, exonic positions outside the consensus sites and intronic positions outside the consensus sites (Supporting Information Table S1). Splicing predictions were performed with Human Splicing Finder, SpliceSiteFinder-like, NNSplice, MaxEntScan and Genscan by using Alamut software v2.1 (Interactive biosoftware). For the nucleotide variants located in consensus sites, if the score estimated by MaxEntScan and SpliceSiteFinder-like was 15 and 5% lower than the wild-type score, respectively, the three remaining prediction algorithms were used (modified from Ref.16). Then, a consensus decrease across all algorithms was considered as indicative of disruption of normal splicing.<sup>6</sup> For variants occurring outside consensus splice sites, the exon skipping, the use of a cryptic site or of a *de novo* splice site, had to be predicted by at least the MaxEntScan and SpliceSiteFinder-like algorithms. We also evaluated the predicted effects of *RAD51D* missense variants on protein using PolyPhen, SIFT and Align GVDG by using Alamut v2.1 and Condel integrated score (<http://bg.upf.edu/condel/home>).<sup>17</sup>

#### cDNA analysis

To analyze splice-site mutations total RNA was purified from peripheral blood lymphocytes using the RNA blood mini kits (Qiagen, Hilden, Germany) or TRIzol (Invitrogen, Carlsbad, CA, US). At least two RNAs extracted from normal individuals were used as controls. To generate complementary DNA (cDNA), 1 µg RNA was retrotranscribed with either the SuperScriptII kit (Invitrogen, Carlsbad, CA, US) or the High-Capacity RNA-to-cDNA Master Mix (Applied Biosystems, Foster City, Ca, US). The cDNA was amplified with two sets of primers spanning exons 1–7 (1F-cctctctaggaaggggttaggg and 7R-cagttcctgaagaaccagtca) and exons 5–10 (5F-cagcgtaaaaccaggatga and 10R-ccaggtccaatgtctacca). The products were electrophoresed on agarose gels and the resolved cDNA

fragments were either directly sequenced or extracted from the gel using a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) before bidirectionally sequenced using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster city, Ca, US).

#### Results

We screened the *RAD51D* gene for germline sequence variations in a total of 842 Spanish index cases with either a family history of BC and/or OC or a personal history of BC or OC with early-onset or concurrent BC and OC. We identified 25 different *RAD51D* sequence variants (Table 1). Two variants overtly pathogenic and one likely pathogenic were detected in four families out of the 491 families with BC/OC cases. One of the clear mutations was a deletion affecting a canonical donor splice site leading to loss of exon 7 and a premature stop codon in the *RAD51D* transcript in one family (c.667+2\_667+23del, p.Val193Alafs\*4) (Fig. 1a). A nonsense variant in exon 8 was identified in two unrelated families (c.694C>T, p.Arg232\*). Moreover, a substitution likely pathogenic, c.1A>T (p.Met1?), affecting the translation initiation codon was also detected in one family with BC/OC cases. The *in silico* analysis predicted that this variant would cause the activation of a cryptic acceptor site at c.15 (Table 1). No evidence of a new aberrant transcript was obtained in the cDNA analysis of this variant and both wild-type and variant alleles were equally present in the sequence of the cDNA (Fig. 1b), suggesting that the c.1A>T does not induce a major alteration in the transcript expression. The cDNA analysis further revealed two alternative transcripts, one containing an out-of-frame skip of exon 3 and the other an in-frame skip of exons 3, 4 and 5 that were present both in the patient and all negative control samples (data not shown). The delta 3 isoform was also found in controls by Osher *et al.*<sup>11</sup> The delta 3, 4 and 5 isoform is described in NCBI database as an alternative transcript NM\_133629.2, also known as isoform 4 or TRAD-d3.

The families carrying the *RAD51D* pathogenic variants presented only one OC case (Fig. 2). The splicing variant c.667+2\_667+23del was present in the index case, a woman diagnosed of bilateral BC at the ages of 29 and 40, and in two sisters, one diagnosed with BC at the age of 49 and the other with bilateral BC at the age of 43 (Fig. 2). The nonsense variant c.694C>T (p.Arg232\*) was present in a woman diagnosed with OC at the age of 44 in one family and in a woman with bilateral OC at the age of 42 in the other family (Fig. 2). In one of these families three healthy sisters and one healthy brother also carried this nonsense variant (Fig. 2). Both the affected and healthy carriers (six subjects) also carried the missense change c.715C>T (p.Arg239Trp) in exon 8 (Table 1). The c.1A>T (p.Met1?) mutation was present in a woman with BC and OC diagnosed at the age of 51 and 64, respectively (Fig. 2).

In addition, we also identified seven *RAD51D* nonsynonymous variants (Table 1). The variants c.494G>A (p.Arg165Gln;

Cancer Genetics

Table 1. Germline RAD51D variants identified

Location exon/intron	Nucleotide change	Amino acid change	rs number	ALIGN		In silico analyses				MAF ESP EA <sup>1</sup>	Previously identified			
				GYGD	SIFT	POLYPHEN	Condel	Splice site	MAF (n)			MAF 1000 genomes		
Pathogenic mutations														
Exon 1	c.1A>T	p.Met1?								0.0006 (842)				
Intron 7	c.667+2_667+23del	p.Val193Alafs*4								0.0006 (842)				
Exon 8	c.694C>T	p.Arg232*								0.0012 (842)		Wickramanyake et al. <sup>10</sup> ; MAF 0.001		
Nonsynonymous														
Exon 1	c.26G>C	p.Cys9Ser	rs140825795	C0	Tolerated	Possibly damaging	Deleterious	No effect		0.0006 (842)		0.001	Loveday et al. <sup>5</sup> ; MAF < 1%; Wickramanyake et al. <sup>14</sup> ; MAF 0.001	
Exon 5	c.355T>C	p.Cys119Arg	rs201313861	C0	Tolerated	Benign	Neutral	No effect		0.0006 (842)		0	Osher et al. <sup>11</sup> ; MAF 0.003	
Exon 6	c.494 G>A	p.Arg165Gln	rs4796033	C0	Tolerated	Benign	Neutral	No effect		0.1087 (842)		0.143	Loveday et al. <sup>6</sup> ; MAF > 1%; Pelttari et al. <sup>12</sup> ; known polymorphism; Osher et al. <sup>11</sup> ; MAF > 1%; Wickramanyake et al. <sup>14</sup> ; MAF > 1%; Sellick et al. <sup>13</sup> ; MAF > 1%	
Exon 7	c.629C>T	p.Ala210Val		C65	Deleterious	Probably damaging	Deleterious	No effect		0.0006 (842)			0.0001	
Exon 8	c.695G>A	p.Arg232Gln	rs28363283	C0	Tolerated	Benign	Neutral	No effect		0.0029 (842)		0.0005	0.0001	Loveday et al. <sup>6</sup> ; MAF < 1%; Wickramanyake et al. <sup>14</sup> ; MAF 0.001
Exon 8	c.698A>G	p.Glu233Gly	rs28363284	C0	Tolerated	Possibly damaging	Neutral	No effect		0.0160 (842)		0.006	0.0168	Loveday et al. <sup>6</sup> ; MAF > 1%; Osher et al. <sup>11</sup> ;

Table 1. Germline *RAD51D* variants identified (Continued)

Location exon/intron	Nucleotide change	Amino acid change	rs number	In silico analyses							MAF ESP EA <sup>1</sup>	Previously identified	
				ALIGN GVDG	SIFT	Deleterious	Possibly damaging	Neutral	Splice site	MAF (n)			MAF 1000 genomes
Exon 8	c.715C>T <sup>2</sup>	p.Arg239Trp		C0				Neutral	No effect	0.0012 (842)	–	–	MAF > 1%; Wickramanyake <i>et al.</i> <sup>14</sup> ; MAF > 1%; Rodríguez-López <i>et al.</i> <sup>19</sup> ; MAF > 1%; Sellick <i>et al.</i> <sup>18</sup> ; MAF > 1%; Dowly <i>et al.</i> <sup>20</sup> ; MAF > 1%; Jara <i>et al.</i> <sup>21</sup> ; MAF > 1%
Synonymous													
Exon 3	c.234C>T	p.Ser78Ser	rs9901455	–	–	–	–		No effect	0.0736 (842)	0.191	0.082	Lovecay <i>et al.</i> <sup>5</sup> ; MAF > 1%; Osher <i>et al.</i> <sup>11</sup> ; MAF > 1%; Pelttari <i>et al.</i> <sup>12</sup> ; known polymorphism; Wickramanyake <i>et al.</i> <sup>14</sup> ; MAF > 1%; Rodríguez-López <i>et al.</i> <sup>19</sup> ; MAF > 1%; Sellick <i>et al.</i> <sup>18</sup> ; MAF > 1%
Exon 9	c.879G>A	p.Ala293Ala		–	–	–	–		No effect	0.0006 (842)	–	0.0001	Thompson <i>et al.</i> <sup>13</sup> ; MAF 0.001
Noncoding													
Exon 1 5'UTR	c.-60C>T			–	–	–	–		No effect	0.0006 (842)	–	–	
Exon 1 5'UTR	c.-111G>A			–	–	–	–		No effect	–	–	–	
Intron 1	c.82+128C>T		rs28363258	–	–	–	–		No effect	0.0088 (455)	0.005	–	Pelttari <i>et al.</i> <sup>12</sup> ; known polymorphism
Intron 1	c.83-4T>C			–	–	–	–		No effect	0.0006 (842)	–	–	
Intron 5	c.480+75T>G		rs8067688	–	–	–	–		c.480+74 de novo donor site	0.0029 (338)	0.018	–	

Cancer Genetics

Cancer Genetics

Table 1. Germline RAD51D variants identified (Continued)

Location exon/intron	Nucleotide change	Amino acid change	rs number	in silico analyses						MAF ESP EA <sup>1</sup>	MAF 1000 genomes	Previously identified
				ALIGN GVDG	SIFT	POLYPHEN	Condel	Splice site	MAF (n)			
Intron 5	c.480+158C>A		rs3816754	-	-	-	-	No effect	0.0479 (94)	0.089	-	
Intron 5	c.480+187T>C		rs936656	-	-	-	-	No effect	0.3564 (94)	0.478	-	
Intron 5	c.481-205A>G		rs8071313	-	-	-	-	No effect	0.0159 (94)	0.122	-	
Intron 5	c.481-144T>C		rs8067218	-	-	-	-	No effect	0.0053 (94)	0.018	-	
Intron 6	c.577-103G>C		rs188753384	-	-	-	-	No effect	0.0053 (281)	0.001	-	
Intron 8	c.739-103T>C		rs28363289	-	-	-	-	No effect	0.0202 (173)	0.062	-	
Intron 9	c.903+53C>T		rs45496096	-	-	-	-	c.903+65 cryptic acceptor enhanced	0.0169 (61.9)	0.006	Pelttari <i>et al.</i> <sup>12</sup> ; known polymorphism	
Intron 9	c.904-11T>A			-	-	-	-	Disruption of splicing acceptor site	0.0006 (842)	-	0	

<sup>1</sup>ESP EA: intron frequency allele reported by the NHLBI Exome Sequencing Project in European ancestry population.

<sup>2</sup>The carriers of this missense also presented the pathogenic mutation c.694C>T.

<sup>3</sup>Although no clearly stated in the article, the carrier of this missense also presented the pathogenic mutation c.694C>T.

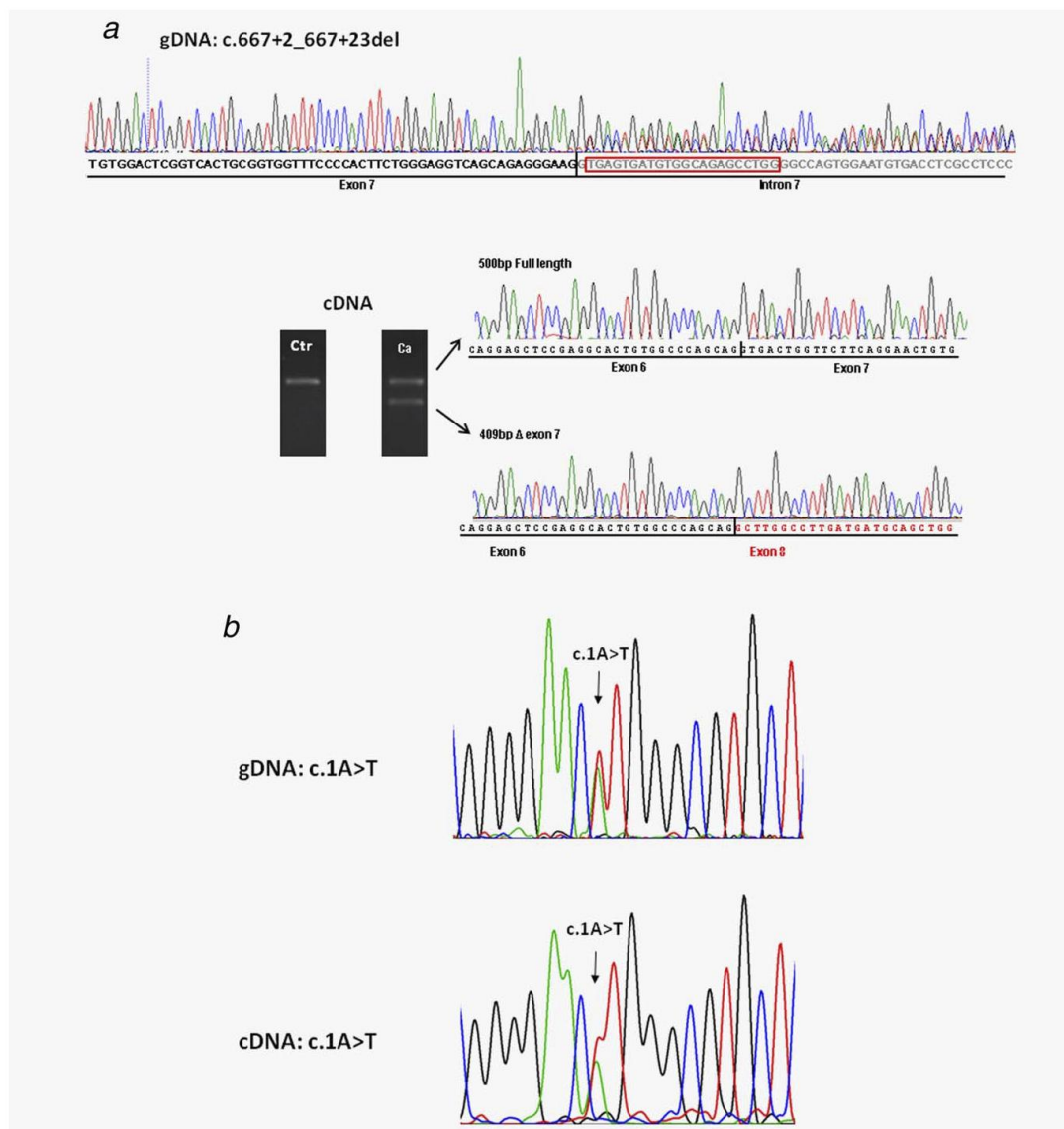
rs4796033) and c.698A>G (p.Glu233Gly; rs28363284) are reported in 1000 genomes and in NHLBI Exome Sequencing Project (ESP) databases as well as in previous reports<sup>18-21</sup> to have allele frequencies near to or greater than 1% (Table 1). Five missense variants, c.26G>C (p.Cys9Ser; rs140825795), c.355T>C (p.Cys119Arg; rs201313861), c.629C>T (p.Ala210Val), c.695G>A (p.Arg232Gln; rs28363283) and c.715C>T (p.Arg239Trp), have been reported with a minor allele frequency (MAF) of 0.0001 to 0.003 in either 1000 genomes and ESP databases as well as in previous studies<sup>11,14,22</sup> (Table 1). Both the c.26G>C and the c.715C>T variants (this last occurring in all the carriers of the deleterious c.694C>T variant) were predicted to be damaging by two of the four programs used to analyze the variant effect on the protein (Table 1). Interestingly, the nonsynonymous variant c.629C>T (p.Ala210Val) located in exon 7 of RAD51D was predicted to be deleterious by the four programs used, and is only reported in ESP project with an allele frequency of 0.0001 (Table 1). Similarly, it is also inferred to be potentially damaging based on the degree of conservation of the amino acid among the RAD51 paralogs (Fig. 3). This nonsynonymous variant was identified in a woman affected with OC at the age of 60 and a strong family history of BC/OC cases: one relative with BC and OC diagnosed both at the age of 40, one with bilateral OC diagnosed at the ages of 50 and 55 and one case of endometrial cancer at the age of 58.

We also detected two synonymous changes: c.234C>T (rs9901455) with a MAF >1% and c.879G>A with a MAF <1%. None of these variants were predicted to alter splicing (Table 1 and Supporting Information Table S1). In addition, the mutation screening also revealed 13 noncoding changes: two variants located in the 5'UTR of the gene and 11 intronic sequence alterations. The intronic variants c.480+75T>G, c.903+53C>T and c.904-11T>A were predicted to alter the splicing (Table 1 and Supporting Information Table S1). Unfortunately, no RNA samples were available to study their potential splicing effect.

Discussion

We identified germline RAD51D deleterious mutations in 0.82% of the families with at least one OC case and one BC case (4/491). Noteworthy, the four carrier families had seven BC cases. Five were diagnosed before the age 50, including two bilateral BC cases. Previous reports found similar prevalences of RAD51D pathogenic variants from 0.57 in Canadian and Belgian BC/OC families<sup>11</sup> to 0.88 in British BC/OC families.<sup>6</sup> Pelttari *et al.*<sup>12</sup> described the presence of a founder RAD51D mutation in 2.9% of the Finnish patients with a family history of both BC and OC. In contrast, Thompson *et al.*<sup>13</sup> detected no deleterious RAD51D changes in 303 Australian BC/OC families. It should be noted that in none of these studies, including ours, the screening techniques used do not allow to detect deletions/duplications affecting complete exons or the entire gene.

We did not identify any deleterious RAD51D variant in 171 unrelated individuals from pedigrees with BC cases and



Cancer Genetics

Figure 1. RNA analysis of two *RAD51D* pathogenic mutations. (a) Analysis of c.667+2\_667+23del; gDNA: electropherogram obtained from genomic DNA showing the 22-bp deletion in heterozygosity in intron 7; cDNA: agarose bands and sequences from cDNA amplification, showing the loss of exon 7 in the mutation carrier (Ca). Ctr: control cDNA. (b) Analysis of c.1A>T; sequences obtained from gDNA and cDNA showing the presence of the c.1A>T in heterozygosity. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

no OC cases, which is consistent with other studies that did not find any association between *RAD51D* mutations and BC risk.<sup>6,12-14</sup> Interestingly, no pathogenic variant was detected in our series of 51 pedigrees with only an OC family history. Our data agree with that obtained in previous reports in which a limited number of this type of families was analyzed: eight families<sup>12</sup> and 16 families.<sup>15</sup> Although more studies to

assess the implication of *RAD51D* in families with two or more OC cases are required, the results obtained in our larger cohort may suggest that other genes could confer a specific susceptibility to familial OC.

The alteration of the translation initiation codon (c.1A>T, p.Met1?) and the variant leading to a skipping of exon 7 (c.667+2\_667+23del, p. Val193Alafs\*4) had not been

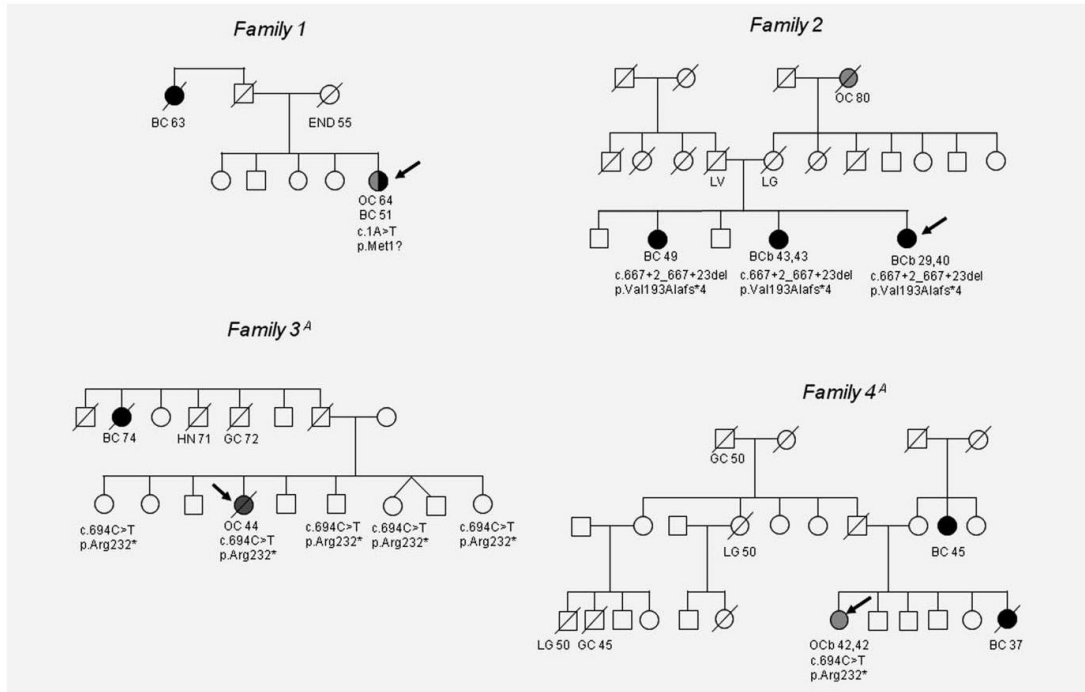


Figure 2. Pedigrees of four families with *RAD51D* pathogenic mutations. The arrow indicates the family proband. Individuals with ovarian cancer are shown as gray circles. Individuals with breast cancer are shown as black circles. Other cancers are shown as unfilled circle. BC: breast cancer; OC: ovarian cancer; BCb: bilateral breast cancer; OCb: bilateral ovarian cancer; HN: head and neck cancer; END: endometrial cancer; GC: gastric cancer; LG: lung cancer; LV: liver cancer. <sup>A</sup>The carriers of the nonsense mutation also presented the missense variant c.715C>T.

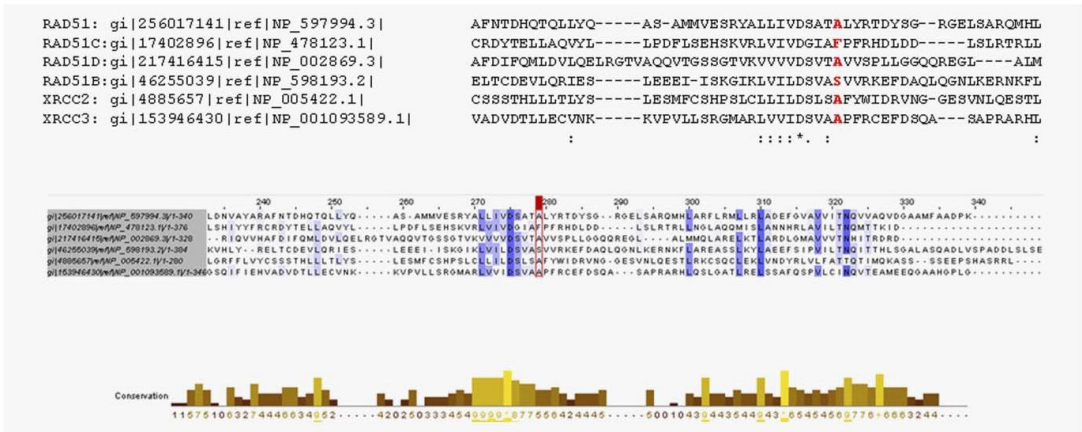


Figure 3. Conservation of site of the *RAD51D* missense variant c.629C>T (p.Ala210Val). RAD51 and its five paralogs were aligned using the Clustal Omega program (multiple sequence alignment program for proteins, <http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalo>). Fully or partially conserved residues are coloured. An \* (asterisk) indicates positions that have a single, fully conserved residue. A: (colon) indicates conservation between groups of strongly similar properties—scoring >0.5 in the Gonnet PAM 250 matrix. A. (period) indicates conservation between groups of weakly similar properties—scoring ≤0.5 in the Gonnet PAM 250 matrix. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

previously described. Given that other single base-pair substitutions located within the ATG translation initiation codon in different genes have been reported to perturb the initiation step of translation, and to reduce translational efficiency *in vivo* (reviewed in Ref.23), we consider the substitution c.1A>T (p.Met1?) pathogenic.

The nonsense variant in exon 9 (c.694C>T, p.Arg232\*) has also been identified in a group of American women with OC who had not been selected for age of onset or family history.<sup>14</sup> Interestingly, all carriers of this nonsense alteration in our series (two with OC and four healthy) and the affected carrier found by Wickramanyake *et al.*<sup>14</sup> also carried the nonsynonymous variant in exon 8 c.715C>T. Altogether, these data suggest that both variants are probably in *cis* and the nonsense nucleotide alteration could be a founder mutation.

The role of rare *RAD51D* missense variants in OC risk is not yet clear. Loveday *et al.*<sup>6</sup> did not find overall differences in the frequency of nontruncating *RAD51D* variants between cases and controls. Rare nonsynonymous variants have been further identified among families with BC only<sup>13</sup> and unselected women with OC,<sup>14</sup> but besides *in silico* analysis, were not further assessed for pathogenicity. We identified a nonsynonymous variant, c.629C>T (p.Ala210Val), which is potentially pathogenic based on SIFT, PolyPhen, Align GVG and Condel predictions, and the amino acid conservation among the *RAD51* paralogs (Fig. 3). The change Ala210Val is located in the C-terminal ATPase domain of *RAD51D* that is conserved among eukaryotic proteins (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238543>) as well as among the five human *RAD51* paralogs, and is closely related to the C-terminal domain of *RAD51*.<sup>24</sup> However, functional assays and cosegregation data with OC in the family are warranted to clarify the impact of this variant.

The studies published to date show that *RAD51D* mutations are more likely to be found in families with two or more OC cases and at least one BC case.<sup>6,11</sup> In contrast, in our study all *RAD51D* mutations were present in families reported with only one OC case and one or more BC cases (Table 2). Our findings suggest that *RAD51D* testing should be offered to affected women with a familiar history of BC/OC regardless of the number of OC cases in the family. Moreover, in our series the splicing variant c.667+2\_667+23del was present in three

women with BC (two of them had bilateral BC), and only one relative with OC. Loveday *et al.*,<sup>6</sup> Osher *et al.*<sup>11</sup> and Pelttari *et al.*<sup>12</sup> also described *RAD51D* deleterious variants in women affected with BC in BC/OC families. Although the frequency of *RAD51D* mutations in BC families with no OC could be anticipated to be very low,<sup>6</sup> these findings support a higher incidence of breast tumors associated to *RAD51D* mutations, albeit in the context of familial OC.

The apparent excess of BC carriers in these families could be attributed to an ascertainment bias selecting OC cases with a close family history of BC,<sup>6</sup> but also to the existence of other unknown variants that could modify the BC risk associated with *RAD51D* mutations in these families and warrants further investigations.<sup>11,12</sup> The early onset of BC in these families agrees with this latter hypothesis. Combining all data of *RAD51D*-associated BC cases published to date<sup>6,11,12</sup> and our data, a mean age of 44.77 years for BC onset was obtained (Supporting Information Table S2), which is lower than the mean age for sporadic breast tumors (an average age of onset in their 50s for Caucasian women referenced in Haffty *et al.*<sup>25</sup>) and similar to the mean age of BC in *BRCA1* or *BRCA2* mutation carriers reported in diverse populations: 43.6 years (Spanish), 41.6 (British) and 44.1 years (Canadian) for *BRCA1* and 42.8 (Spanish), 45.2 (British) and 47.3 years (Canadians) for *BRCA2*.<sup>26-28</sup> Alternatively, specific mutation risks may be attributed to the type and location of the variants.<sup>12</sup> Considering the mutations described in previously published articles and in this article, we did not observe any association between the BC or OC risk and location of the mutations along *RAD51D* (data not shown).

In conclusion, we identified three pathogenic variants, c.667+2\_667+23del, c.1A>T and c.694C>T in four out of the 491 probands (0.82%) from OC and BC families. Our results agree with the association of *RAD51D* mutations with the presence of OC cases in the family. Interestingly, all pathogenic variants were present in families reported with only one OC case, suggesting that *RAD51D* testing should be offered to any affected BC or OC subject irrespective of the number of OC cases in the family. In the c.667+2\_667+23del proband's family, both the proband and two relatives carried the mutated allele and were affected with BC (two of them bilaterals). Despite the modest proportion of *RAD51D* deleterious

Table 2. Prevalence of *RAD51D* pathogenic variants reported according to the number of ovarian cases in breast and ovarian cancer families

Reference	N° mutation carriers/n° BC + 10C families (%)	N° mutation carriers/n° BC + 20C families (%)	N° mutation carriers/n° BC + ≥30V families (%)	Total
Loveday <i>et al.</i> <sup>6</sup>	1/617 (0.16)	4/235 (1.7)	3/59 (5.1)	8/911 (0.9)
Osher <i>et al.</i> <sup>11</sup>	0/124 (0)	1/43 (2.32)	0/8 (0)	1/175 (0.5)
Pelttari <i>et al.</i> <sup>12</sup>	ns	ns	ns	3/105 (2.9)
Thompson <i>et al.</i> <sup>13</sup>	0/ns	0/ns	0/ns	0/303 (0)
Our study	4/282 (1.4)	0/189 (0)	0/20 (0)	4/491 (0.8)

Abbreviations: BC: breast cancer; OC: ovarian cancer; ns: not specified.

mutations identified, genetic testing of this gene should be considered into the clinical setting, at least for families with BC and OC. The presence of several BC mutation carriers, albeit in the context of familial OC, suggests an increased risk of this type of cancer, which should be considered when setting the follow-up and prevention measures, and future specific treatments. Additional studies in familial BC/OC series would be of value to provide more complete information on the actual risk of OC, as well as the potential risks of BC and any other cancers.

### Acknowledgements

The authors are grateful to the families for their cooperation and to the clinical personnel involved in aspects of recruitment and clinical data collection. SGE and AT are funded by

a Miguel Servet contract (CP10/00617) from the Spanish Carlos III Health Institute. The research in HUVH-VHIO (Barcelona) was supported by a Miguel Servet Project grant CP10/00617. HCSC contribution was supported by Fondo de Investigación Sanitaria (FIS) research grants PI 12/00539, an initiative of the Instituto de Salud Carlos III (Spain), partially supported by European Regional Development FEDER Funds. CNIO study was partially supported by Mutua Madrileña Foundation (FMMA), Intrasalud PII2/00070, Acción Cooperativa y Colaborativa Intramural-CIBER2012 and SAF2010-20493. In FPGMX, this study was supported by grants from the Xunta de Galicia (10PXIB 9101297PR) and FMM Foundation given to AV. The research in IBGM is funded by the Cancer Prevention Program of the Regional Government of Castilla y León.

### References

1. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893-917.
2. Permeth-Wey J, Sellers TA. Epidemiology of ovarian cancer. *Methods Mol Biol* 2009;472:413-37.
3. Weissman SM, Weiss SM, Newlin AC. Genetic testing by cancer site: ovary. *Cancer J* 2012;18:320-7.
4. Pennington KP, Swisher EM. Hereditary ovarian cancer: beyond the usual suspects. *Gynecol Oncol* 2012;124:347-53.
5. Walsh T, Casadei S, Lee MK, et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc Natl Acad Sci USA* 2011;108:18932-7.
6. Loveday C, Turnbull C, Ramsay E, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet* 2011;43:879-82.
7. Takata M, Sasaki MS, Tachiiri S, et al. Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol Cell Biol* 2001;21:2858-66.
8. Masson JY, Tarsounas MC, Stasiak AZ, et al. Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev* 2001;15:3296-307.
9. Chun J, Buechelmaier ES, Powell SN. Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway. *Mol Cell Biol* 2013;33:387-95.
10. Tarsounas M, Munoz P, Claas A, et al. Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* 2004;117:337-47.
11. Osher DJ, De LK, Michils G, et al. Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. *Br J Cancer* 2012;106:1460-3.
12. Peltari LM, Kiiski J, Nurminen R, et al. A Finnish founder mutation in RAD51D: analysis in breast, ovarian, prostate, and colorectal cancer. *J Med Genet* 2012;49:429-32.
13. Thompson ER, Rowley SM, Sawyer S, et al. Analysis of RAD51D in ovarian cancer patients and families with a history of ovarian or breast cancer. *PLoS One* 2013;8:e54772.
14. Wickramanyake A, Bernier G, Pennil C, et al. Loss of function germline mutations in RAD51D in women with ovarian carcinoma. *Gynecol Oncol* 2012;127:552-5.
15. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002;3:285-98.
16. Houdayer C, Caux-Moncoutier V, Krieger S, et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum Mutat* 2012;33:1228-38.
17. Gonzalez-Perez A, Lopez-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* 2011;88:440-9.
18. Sellick G, Fielding S, Qureshi M, et al. Germline mutations in RAD51, RAD51AP1, RAD51B, RAD51C, RAD51D, RAD52 and RAD54L do not contribute to familial chronic lymphocytic leukemia. *Leuk Lymphoma* 2008;49:130-3.
19. Rodriguez-Lopez R, Osorio A, Ribas G, et al. The variant E233G of the RAD51D gene could be a low-penetrance allele in high-risk breast cancer families without BRCA1/2 mutations. *Int J Cancer* 2004;110:845-9.
20. Dowty JG, Lose F, Jenkins MA, et al. The RAD51D E233G variant and breast cancer risk: population-based and clinic-based family studies of Australian women. *Breast Cancer Res Treat* 2008;112:35-9.
21. Jara L, Dubois K, Gaete D, et al. Variants in DNA double-strand break repair genes and risk of familial breast cancer in a South American population. *Breast Cancer Res Treat* 2010;122:813-22.
22. Loveday C, Turnbull C, Ruark E, et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet* 2012;44:475-6.
23. Wolf A, Caliebe A, Thomas NS, et al. Single base-pair substitutions at the translation initiation sites of human genes as a cause of inherited disease. *Hum Mutat* 2011;32:1137-43.
24. Kim YM, Choi BS. Structural and functional characterization of the N-terminal domain of human Rad51D. *Int J Biochem Cell Biol* 2011;43:416-22.
25. Haffty BG, Choi DH, Goyal S, et al. Breast cancer in young women (YBC): prevalence of BRCA1/2 mutations and risk of secondary malignancies across diverse racial groups. *Ann Oncol* 2009;20:1653-9.
26. Diez O, Osorio A, Duran M, et al. Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: a high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat* 2003;22:301-12.
27. Mavaddat N, Pharoah PD, Blows F, et al. Familial relative risks for breast cancer by pathological subtype: a population-based cohort study. *Breast Cancer Res* 2010;12:R10.
28. Panchal S, Bordeleau L, Poll A, et al. Does family history predict the age at onset of new breast cancers in BRCA1 and BRCA2 mutation-positive families? *Clin Genet* 2010;77:273-9.

Supplementary Table 1. Detailed analysis of splicing alteration predictions

Nucleotide change	Location, SS, distance	Variant Categories according their location relative to the 5' and 3' consensus site	Splicing predictions at nearest natural junction (5' or 3' score modification)	Proximal Cryptic/ De novo (% wild type)
<b>Pathogenic Mutations</b>				
c.1 A>T	Exon 1, D, 82	Exonic occurring outside consensus positions		c.15 cryptic acceptor enhanced: SSF-L: — ⇒ 74.67; MaxEnt: 1.63 ⇒ 2.97 (+82.8%); HSF: = 80.32
c.667+2_667+23del	Intron 8, D, 2	Donor consensus splice site (GT canonical splice site)	Donor: SSF-L: 95.33 ⇒ —; MaxEnt: 10.47 ⇒ —; NNSPLICE: 0.99 ⇒ —; GeneSplicer: 8.22 ⇒ —; HSF: 97.89 ⇒ —	
c.694 C>T	Exon 9, A, 27	Exonic occurring outside consensus positions	Acceptor: GeneSplicer: 17.67 ⇒ 17.27 (-2.3%). Donor: GeneSplicer: 4.51 ⇒ 4.12 (-8.7%)	c.706 cryptic acceptor enhanced: HSF: 75%
<b>Non-synonymous</b>				
c.26 G>C	Exon 1, D, 57	Exonic occurring outside consensus positions	Donor: GeneSplicer: 8.12 ⇒ 7.58 (-6.6%).	

<b>Nucleotide change</b>	<b>Location, SS, distance</b>	<b>Variant Categories according their location relative to the 5' and 3' consensus site</b>	<b>Splicing predictions at nearest natural jonction (5' or 3' score modification)</b>	<b>Proximal Cryptic/ De novo (% wild type)</b>
c.355 T>C	Exon 6, A, 10	Exonic occurring outside consensus positions	Acceptor: NNSPLICE: 0.58 ⇒ 0.40 (-30.6%).	c.355 cryptic donor enhanced: MaxEnt: 15%; c.346-1 cryptic donor enhanced: GeneSplicer: 53%
c.494 G>A	Exon 7, A, 14	Exonic occurring outside consensus positions	Acceptor: NNSPLICE: 0.97 ⇒ 0.97 (-0.1%). GeneSplicer: 10.12 ⇒ 9.92 (-2.0%).	c.496 de novo site: HSF: — ⇒ 82.48. c.498 cryptic acceptor enhanced: HSF: 90%. c.532 cryptic acceptor enhanced: GeneSplicer: 19%
c.629 C>T	Exon 8, D, 39	Exonic occurring outside consensus positions	Acceptor: GeneSplicer: 6.01 ⇒ 5.72 (-4.8%). Donor: GeneSplicer: 8.22 ⇒ 7.49 (-8.8%)	c.577-1 cryptic donor enhanced: GeneSplicer: 41%
c.695 G>A	Exon 9, A, 28	Exonic occurring outside consensus positions	Acceptor: GeneSplicer: 17.67 ⇒ 16.98 (-3.9%). Donor: GeneSplicer: 4.51 ⇒ 3.75 (-16.9%)	c.698: cryptic acceptor enhanced: HSF: 77%. c.700 cryptic acceptor enhanced: HSF: 78%.

<b>Nucleotide change</b>	<b>Location, SS, distance</b>	<b>Variant Categories according their location relative to the 5' and 3' consensus site</b>	<b>Splicing predictions at nearest natural junction (5' or 3' score modification)</b>	<b>Proximal Cryptic/ De novo (% wild type)</b>
c.698 A>G	Exon 9, A, 31	Exonic occurring outside consensus positions	Acceptor: GeneSplicer: 17.67 $\Rightarrow$ 17.39 (-1.6%). Donor: GeneSplicer: 4.51 $\Rightarrow$ 4.01 (-11.1%)	c.698 cryptic acceptor enhanced: HSF: 80%. c.706 cryptic acceptor enhanced: HSF: 74%
c.715 C>T*	Exon 9, D, 24	Exonic occurring outside consensus positions	Acceptor: GeneSplicer: 17.67 $\Rightarrow$ 17.30 (-2.1%). Donor GeneSplicer: 4.51 $\Rightarrow$ 4.11 (-8.9%)	c.734 cryptic acceptor enhanced: MaxEnt: 14%
<b>Synonymous</b>				
c.234 C>T	Exon 3, D, 30	Exonic occurring outside consensus positions	Donor: GeneSplicer: 3.87 $\Rightarrow$ 3.59 (-7.1%)	
c.879 G>A	Exon 10, D, 25	Exonic occurring outside consensus positions	Donor: GeneSplicer: 10.59 $\Rightarrow$ 10.61 (+0.2%)	c.903+1 cryptic acceptor enhanced: GeneSplicer: 56%. c.876 de novo donor site: SSF-L: 87%. c.880 cryptic donor enhanced: SSF-L: 76%
<b>Non-coding</b>				

<b>Nucleotide change</b>	<b>Location, SS, distance</b>	<b>Variant Categories according their location relative to the 5' and 3' consensus site</b>	<b>Splicing predictions at nearest natural jonction (5' or 3' score modification)</b>	<b>Proximal Cryptic/ De novo (% wild type)</b>
c.-60 T>C	Exon 1 5' UTR, D, 142	Exonic occurring outside consensus positions		c.-116 cryptic acceptor enhanced: GeneSplicer: 12.37 ⇒ 12.39 (+0.2%). c.-92 cryptic acceptor enhanced: GeneSplicer: 7.57 ⇒ 7.59 (+0.3%)
c.-111 G>A	Exon 1 5' UTR, A, 146	Exonic occurring outside consensus positions		c.-92 cryptic acceptor enhanced: GeneSplicer: 7.57 ⇒ 8.04 (+6.2%)
c.82+128 C>T	Intron 1, D, 128	Intronic occurring outside consensus positions		c.82+134 cryptic acceptor enhanced: SSF-L: — ⇒ 70.78
c.83-4 T>C	Intron 1, A, 4	Acceptor consensus slice site	Acceptor SSF-L: = 93.07; MaxEnt: 8.52 ⇒ 8.08 (-5.2%); NNSPLICE: 0.95 ⇒ 0.91 (-4.9%); GeneSplicer: 6.77 ⇒ 6.51 (-3.8%); HSF: 90.83 ⇒ 91.49 (+0.7%)	c.83-2 de novo donor site: HSF: 64%
c.480+75 T>G	Intron 6, D, 75	Intronic occurring outside consensus positions	Donor: GeneSplicer: 8.05 ⇒ 7.82 (-2.8%)	c.480+74 de novo donor site: SSF-L: 77%; MaxEnt: 25%; HSF: 80%
c.480+158 C>A	Intron 6, D, 158	Intronic occurring outside consensus positions		
c.480+187 T>C	Intron 6, D, 187	Intronic occurring outside consensus positions		

<b>Nucleotide change</b>	<b>Location, SS, distance</b>	<b>Variant Categories according their location relative to the 5' and 3' consensus site</b>	<b>Splicing predictions at nearest natural jonction (5' or 3' score modification)</b>	<b>Proximal Cryptic/ De novo (% wild type)</b>
c.481-205 A>G	Intron 6, A, 205	Intronic occurring outside consensus positions		
c.481-144 T>C	Intron 6, A, 144	Intronic occurring outside consensus positions		c.481-83 cryptic donor enhanced: GeneSplicer: 31%
c.577-103 G>C	Intron 7, A, 103	Intronic occurring outside consensus positions		c.577-82 cryptic donor enhanced NNSPLICE: 84%
c.739-103 T>C	Intron 9, A, 103	Intronic occurring outside consensus positions		c.739-90 cryptic acceptor enhanced MaxEnt: 70%
c.903+53 C>T	Intron 10, D, 53	Intronic occurring outside consensus positions	Donor: GeneSplicer: 10.59 $\Rightarrow$ 10.42 (-1.6%)	c.903+65 cryptic acceptor enhanced SSF-L: 89%; MaxEnt: 84%; GeneSplicer: 24%; HSF: 92%
c.904-11 T>A	Intron 10, A, 11	Acceptor consensus slice site	Acceptor: SSF-L: 88.46 $\Rightarrow$ 84.65 (-4.3%); MaxEnt: 8.16 $\Rightarrow$ 5.24 (-35.8%); NNSPLICE: 0.95 $\Rightarrow$ 0.77 (-18.8%); GeneSplicer: 10.95 $\Rightarrow$ 7.33 (-33.0%); HSF: 93.02 $\Rightarrow$ 89.11 (-4.2%)	

**Supplementary Table 2. Age onset of BC among RAD51D carriers**

Reference	Mutation	BC cases with RAD51D deleterious mutations	BC onset age (years)
Gutiérrez-Enriquez et al., 2013	c.1 A>T, p.Met1Leu	BC and OC	51
	c.667+2_667+23del, p.Gly223AlafsX4	Bilateral BC	29
	c.667+2_667+23del, p.Gly223AlafsX4	Bilateral BC	43
	c.667+2_667+23del, p.Gly223AlafsX4	BC	49
Loveday et al., 2011	c.363delA	Bilateral BC	34
	c.556C>T; p.Arg186X	BC	39
	c.556C>T; p.Arg186X	BC	58
	c.556C>T; p.Arg186X	BC	53
	c.480+1G>A	BC	51
	c.556C>T; p.Arg186X	BC	35
	c.757C>T; p.Arg253X	BC	47
Osher et al., 2012	c.556C>T; p.Arg186X	BC	34
	c.556C>T; p.Arg186X	DCIS	47
		<b>mean</b>	<b>43.8</b>
		min	29
		max	58
Pelttari et al., 2012	c.576+1G>A	5 BC patients	
		<b>mean</b>	<b>45.7</b>

## Anexo II

**Tabla 9:** Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios

Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Series de familias BRCA1													
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
					Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
2	c.-31G>T		NA	UTR	1	0,0004			1	0,002								
2	c.60G>A	p. =	rs200430973	Silent	1	0,0004								1	0,002			
2	c.85C>T	p.Arg29Cys	rs144832924	Missense	1	0,0004										1	0,001	
2	c.86G>A	p.Arg29His	rs149117119	Missense	6	0,0023			1	0,002						5	0,007	
2	c.90C>T	p. =	rs118089506	Silent	42	0,0160			20	0,041			7	0,009	15	0,014		
2	c.112C>T	p.Leu38Phe	rs141167501	Missense	2	0,0008							1	0,001			1	0,001
2	c.145G>A	p.Glu49Lys	rs139757085	Missense	1	0,0004												
2	c.237T>C	p. =	rs375620914	Silent	2	0,0008							2	0,002				
2	c.244A>G	p.Asn82Asp	rs202237877	Missense	1	0,0004								1	0,002			
2	c.247G>A	p.Gly83Ser	rs200319106	Missense	1	0,0004								1	0,002			
2	c.248G>C	p.Gly83Ala	NA	Missense	1	0,0004	1	0,011										
2	c.299C>A	p.Thr100Asn	rs201163062	Missense	1	0,0004								1	0,002			
2	c.339T>C	p. =	rs144326379	Silent	1	0,0004	1	0,011										
2	c.376A>G	p.Thr126Ala	NA	Missense	1	0,0004			1	0,002								
2	c.421G>T	p.Gly141Trp	rs137976282	Missense	9	0,0034	1	0,011	1	0,002			2	0,002	2	0,004	3	0,004
2	c.452C>T	p.Pro151Leu	rs200191906	Missense	1	0,0004								1	0,002			
2	c.465A>G	p. =	rs138512851	Silent	2	0,0008			1	0,002			1	0,001				
2	c.497A>C	p.Gln166Pro	NA	Missense	3	0,0011											3	0,004
3	c.536-80C>T		rs137869443	Intronic	1	0,0004							1	0,001				

Tabla 9: Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios

					Series de familias BRCA1													
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
3	c.553G>A	p.Asp185Asn	rs201769293	Missense	1	0,0004									1	0,002		
3	c.555C>T	p. =	rs74640850	Silent	170	0,0648	9	0,095	42	0,086	10	0,19	67	0,091	42	0,08		
3	c.590T>C	p.Val197Ala	rs147826749	Missense	12	0,0046	1	0,011	3	0,006	1	0,02	2	0,002	3	0,006	2	0,003
3	c.610C>T	p.Arg204Cys	rs79842542	Missense	224	0,0853	10	0,106	42	0,086	10	0,19	65	0,089	42	0,08	55	0,075
3	c.678C>T	p. =	rs28516461	Silent	69 (1)	0,0267	3	0,032	26	0,053	1	0,02	12	0,016	27 (1)	0,052		
3	c.707C>T	p.Ala236Val	rs200803578	Missense	1	0,0004									1	0,002		
3	c.710G>A	p.Arg237Gln	rs138615800	Missense	17	0,0065	1	0,011			1	0,02	9	0,012	5	0,009	1	0,001
3	c.734C>T	p.Pro245Leu	rs199929086	Missense	2	0,0008									2	0,004		
3	c.742G>A	p.Glu248Lys	rs148547201	Missense	1	0,0004									1	0,002		
3	c.753G>A	p. =	rs8061528	Silent	466 (30)	0,1890	26 (2)	0,276	52 (1)	0,109	14 (2)	0,307	202	0,277	172 (25)	0,374		
3	c.760+10C>G		NA	Intronic	7	0,0027			7	0,014								
4	c.761-32T>G		rs118098382	Intronic	43	0,0164	2	0,021					29	0,039	12	0,022		
4	c.832C>T	p.Arg278Trp	rs141597706	Missense	1	0,0004			1	0,002								
4	c.833G>A	p.Arg278Gln	rs201192909	Missense	1	0,0004									1	0,002		
4	c.851A>G	p.His284Arg	NA	Missense	1	0,0004							1	0,001				
4	c.950+29C>T		NA	Intronic	1	0,0004							1	0,001				
5	c.999C>T	p. =	rs7198338	Silent	2	0,0008							1	0,001	1	0,002		
5	c.1065G>A	p. =	rs200183071	Silent	1	0,0004									1	0,002		
5	c.1077G>T	p.Lys359Asn	rs149470704	Missense	1	0,0004											1	0,001
5	c.1114C>T	p.Arg372Trp	NA	Missense	1	0,0004			1	0,002								
5	c.1132C>A	p.Pro378Arg	NA	Missense	1	0,0004							1	0,001				
5	c.1152A>G	p. =	rs112511042	Silent	171 (1)	0,0655	11	0,117	44	0,095	10	0,19	65	0,089	41 (1)	0,078		
5	c.1153C>A	p.Pro385Thr	rs115694169	Missense	18	0,0069	1	0,011	3	0,006			2	0,002	6	0,011	6	0,008
5	c.1156A>G	p.Met386Val	rs113490934	Missense	171 (1)	0,0655	11	0,117	44	0,095	10	0,19	65	0,089	41 (1)	0,078	30	0,041

Series de familias BRCAX																		
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
5	c.1163+10C>T		rs80116508	Intronic	162 (1)	0,0621	1	0,011	44	0,095	10	0,19	66	0,090	41 (1)	0,078		
5	c.1163+11T>C		rs76350200	Intronic	1	0,0004			1	0,002								
5	c.1163+194C>T		NA	Intronic	1	0,0004							1	0,001				
6	c.1164-16T>C		NA	Intronic	1	0,0004	1	0,011										
6	c.1164-40C>A		NA	Intronic	1	0,0004	1	0,011										
6	c.1164-66T>A		rs200497436	Intronic	3	0,0011	1	0,011							2	0,004		
6	c.1164-75C>G		rs59622164	Intronic	118	0,0450	1	0,011			10	0,19	65	0,089	42	0,078		
6	c.1243G>A	p.Glu415Lys	rs146021821	Missense	4	0,0015											4	0,005
6	c.1271C>T	p.Ala424Val	NA	Missense	1	0,0004							1	0,001				
6	c.1366+11T>C		rs76350200	Intronic	137	0,0522	12	0,128			10	0,19	67	0,091	48	0,091		
7	C.1367-44G>T		rs185962726	Intronic	2	0,0008							2	0,002				
7	c.1371T>G	p.Asn457Lys	rs74319927	Missense	183	0,0697	10	0,106			10	0,19	64	0,087	38	0,072	61	0,083
7	c.1372A>G	p.Lys458Glu	rs149126845	Missense	3	0,0011							1	0,001			2	0,003
7	c.1419C>T	p. =	NA	Silent	1	0,0004	1	0,011										
7	c.1513G>A	p.Ala505Thr	NA	Missense	1	0,0004							1	0,001				
7	c.1517G>A	p.Ser506Asn	NA	Missense	1	0,0004							1	0,001				
7	c.1637A>G	p.Tyr546Cys	rs150547487	Missense	1	0,0004			1	0,002								
7	c.1641G>A	p. =	rs200497436	Silent	2	0,0008							1	0,001	1	0,002		
8	c.1702G>A	p.Val568Met	rs371825444	Missense	1	0,0004							1	0,001				
8	c.1736T>C	p.Leu579Pro	NA	Missense	1	0,0004							1	0,001				
8	c.1755C>T	p. =	rs114016359	Silent	8	0,0030							2	0,002	6	0,011		
8	c.1755C>A	p. =	rs114016359	Silent	1	0,0004									1	0,002		
8	c.1803G>A	p. =	rs144892556	Silent	8	0,0030							2	0,002	6	0,011		
8	c.1832C>A	p.Ala611Asp	rs201641321	Missense	1	0,0004									1	0,002		

Tabla 9: Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios

					Series de familias BRCA													
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
8	c.1845C>T	p. =	NA	Silent	1	0,0004							1	0,001				
8	c.1846G>A	p.Val616Met	rs199863797	Missense	2	0,0008	1	0,011							1	0,002		
8	c.1896G>C	p. =	rs200859735	Silent	1	0,0004									1	0,002		
8	c.1898G>A	p.Gly633Asp	rs1056085	Missense	3	0,0011									1	0,002	2	0,003
8	c.1911G>A	p. =	rs200013924	Silent	1	0,0004									1	0,002		
8	c.1924+5G>A		NA	Intronic	1	0,0004							1	0,001				
8	c.1924+59C>T		NA	Intronic	1	0,0004							1	0,001				
8	c.1924+127G>A		NA	Intronic	1	0,0004							1	0,001				
9	c.1925-122C>T		NA	Intronic	1	0,0004							1	0,001				
9	c.1925-95G>A		rs180740748	Intronic	14	0,0053							14	0,019				
9	c.1925-30G>A		rs149916101	Intronic	2	0,0008							2	0,002				
9	c.1925-106T>G		NA	Intronic	1	0,0004							1	0,001				
9	c.1925G>T	p.Gly642Val	NA	Missense	1	0,0004											1	0,001
9	c.1991C>T	p.Pro664Leu	rs369805557	Missense	1	0,0004			1	0,002								
9	c.2006G>A	p.Arg669Asp	rs200807331	Missense	1	0,0004									1	0,002		
9	c.2012T>C	p.Leu671Ser	rs77985244	Missense	233	0,0888	11	0,117	46	0,009	10	0,19	65	0,089	44	0,084	57	0,077
9	c.2013+2T>A		NA	Intronic	1	0,0004							1	0,001				
9	c.2013+21G>T		rs139099529	Intronic	1	0,0004			1	0,002								
9	c.2013+23G>A		rs112226642	Intronic	164	0,0625	11	0,117	45	0,009			65	0,089	43	0,081		
9	c.2013+130C>G		NA	Intronic	1	0,0004							1	0,001				
9	c.2013+137G>C		rs80186343	Intronic	76	0,0290	11	0,117					65					
10	c.2014-28G>A		rs200455476	Intronic	1	0,0004							1	0,001				
10	c.2098G>A	p.Gly700Arg	NA	Missense	1	0,0004					1	0,02						
10	c.2160+50C>T		rs75762935	Intronic	128	0,0488	10	0,106			10	0,19	66	0,090	42	0,079		

Series de familias BRCA1																		
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
11	c.2161-136G>A		rs139226142	Intronic	1	0,0004							1	0,001				
11	c.2235C>T	p. =	rs75184268	Silent	3	0,0011			1	0,002					2	0,004		
11	c.2305G>C	p.Glu769Gln	rs150712805	Missense	5	0,0019			3	0,006							2	0,003
12	c.2328-35del13		NA	Intronic	2	0,0008							2	0,002				
12	c.2328-9G>C		rs73505419	Intronic	1	0,0004							1	0,001				
12	c.2346C>T	p. =	NA	Silent	1	0,0004	1	0,011										
12	c.2359G>A	p.Glu787Lys	rs140600202	Missense	16	0,0061			2	0,004			3	0,004	8	0,015	3	0,004
12	c.2437G>A	p.Glu813Lys	NA	Missense	1	0,0004											1	0,001
12	c.2469G>C	p.Trp823Cys	NA	Missense	1	0,0004	1	0,011										
12	c.2469G>A	p.Trp823*	NA	Non Sense	1	0,0004											1	0,001
12	c.2597A>C	p.Gln866Pro	rs201539019	Missense	1	0,0004									1	0,002		
12	c.2609C>T	p.Ala870Val	rs149584080	Missense	4	0,0015							4	0,005				
12	c.2628C>T	p. =	rs201279467	Silent	1	0,0004							1	0,001				
12	c.2654C>T	p.Pro885Leu	rs147492092	Missense	1	0,0004			1	0,002								
12	c.2667A>G	p. =	rs146790631	Silent	2	0,0008			1	0,002			1	0,001				
12	c.2681T>G	p.Val894Gly	rs145137472	Missense	2	0,0008							2	0,002				
12	c.2786C>T	p.Pro929Leu	rs117707719	Missense	1	0,0004							1	0,001				
12	c.2824G>C	p.Glu942Gln	rs114014006	Missense	16	0,0061			8	0,016			2	0,002			6	0,008
12	c.2854G>A	p.Ala952Thr	rs59939128	Missense	6	0,0023			1	0,002					2	0,004	3	0,004
12	c.2854_55GC>AT	p.Ala952Met	rs78637028	Missense	218	0,0830	8	0,085	47	0,009	10	0,19	55	0,075	42	0,079	56	0,076
12	c.2924C>T	p.Pro975Leu	rs114472821	Missense	23	0,0088			8	0,016			1	0,001	7	0,013	7	0,009
12	c.2949C>T	p. =	rs139287784	Silent	1	0,0004							1	0,001				
12	c.2975G>A	p.Gly992Glu	rs139287784	Missense	3	0,0011			1	0,002					1	0,002	1	0,001
12	c.3019C>A	p.Gln1007Lys	rs138798067	Missense	4	0,0015							1	0,001			3	0,004

Tabla 9: Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios

					Series de familias BRCA1													
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
12	c.3034A>G	p.Arg1012Gly	NA	Missense	1	0,0004											1	0,001
12	c.3062G>A	p.Arg1021His	rs200842643	Missense	1	0,0004									1	0,002		
12	c.3101T>C	p.Arg1034His	rs150453226	Missense	1	0,0004					1	0,02						
12	c.3109T>C	p. =	rs58735123	Silent	2	0,0008							1	0,001	1	0,002		
12	c.3118C>T	p.Pro1040Ser	rs201359490	Missense	1	0,0004											1	0,001
12	c.3162G>A	p. =	rs76488917	Silent	59 (1)	0,0229	1	0,011	23	0,047	1	0,02	9	0,012	25 (1)	0,049		
12	c.3178C>T	p.Arg1060Trp	rs144273492	Missense	10	0,0038							2	0,002			8	0,011
12	c.3189C>T	p. =	rs200742809	Silent	12 (3)	0,0057									12 (3)	0,028		
12	c.3308G>A	p.Arg1103His	rs201539019	Missense	2	0,0008			1	0,002					1	0,002		
12	c.3315C>T	p. =	rs183029626	Silent	1	0,0004							1	0,001				
12	c.3316G>A	p.Val1106Met	rs200057338	Missense	1	0,0004									1	0,002		
12	c.3337G>C	p.Gly1113Arg	NA	Missense	1	0,0004											1	0,001
12	c.3365C>T	p.Pro1122Leu	rs714181	Missense	345 (2)	0,1322	12	0,128	80 (1)	0,166	10	0,19	76	0,104	63 (1)	0,121	104	0,141
12	c.3368C>A	p.Ser1123Tyr	rs144647122	Missense	2	0,0008							1	0,001			1	0,001
12	c.3420A>G	p. =	rs201446870	Silent	1	0,0004									1	0,002		
12	c.3583_3585del ATT	p.Ile1195del	rs199897550	In Frame Deletion	1	0,0004									1	0,002		
12	c.3662C>T	p.Ala1221Val	rs3827530	Missense	226	0,0861	10	0,106	47	0,009	10	0,19	68	0,093	35	0,066	56	0,076
12	c.3676C>T	p.Arg1226Trp	rs142008398	Missense	2	0,0008											2	0,003
12	c.3783G>A	p. =	rs77699867	Silent	8	0,0030			1	0,002			1	0,001	6	0,011		
12	c.3812C>T	p.Ser1271Phe	rs3810813	Missense	241 (1)	0,0922	4	0,042	46	0,009	4 (1)	0,096	54	0,074	52	0,098	81	0,110
12	c.3849C>G	p. =	rs200928592	Silent	1	0,0004									1	0,002		
12	c.3857C>T	p.Ala1286Val	rs149011965	Missense	6	0,0023							6	0,082				
12	c.3860T>G	p.Val1287Gly	NA	Missense	1	0,0004							1	0,001				

Series de familias BRCAX																		
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
12	c.3868C>A	p.His1290Asn	rs112596894	Missense	1	0,0004			1	0,002								
12	c.3872C>T	p.Thr1291Met	NA	Missense	1	0,0004	1	0,011										
12	c.3913G>A	p.Ala1305Thr	NA	Missense	1	0,0004											1	0,001
12	c.3939C>A	p. =	rs200057338	Silent	1	0,0004			1	0,002								
12	c.3963G>A	p. =	rs116781836	Silent	4	0,0015			1	0,002			2	0,002	1	0,002		
12	c.4024A>G	p.Ser1342Gly	rs140051968	Missense	3	0,0011							3	0,004				
12	c.4057C>A	p.His1353Asn	rs142205392	Missense	1	0,0004											1	0,001
12	c.4068G>A	p. =	rs115491049	Silent	1	0,0004									1	0,002		
12	c.4155G>A	p.Arg1372Gln	rs79174372	Missense	2	0,0008			2	0,004								
12	c.4241C>T	p.Pro1414Leu	rs140892471	Missense	1	0,0004			1	0,002								
12	c.4261A>T	p.Ile1421Phe	rs141567438	Missense	5	0,0019	1	0,011					2	0,002			2	0,003
12	c.4272C>T	p. =	NA	Silent	1	0,0004							1	0,001				
12	c.4333C>T	p.Arg1445Trp	NA	Missense	1	0,0004			1	0,002								
12	c.4338C>T	p.Thr1446Ser	rs77718962	Missense	1	0,0004									1	0,002		
12	c.4344C>T	p. =	rs368388100	Silent	1	0,0004			1	0,002								
12	c.4403G>A	p.Arg1468His	NA	Missense	1	0,0004											1	0,001
12	c.4409C>T	p.Pro1470Leu	rs72778139	Missense	4	0,0015	1	0,011	1	0,002							2	0,003
12	c.4410G>A	p. =	NA	Silent	1	0,0004							1	0,001				
12	c.4427C>G	p.Thr1476Ser	rs372321470	Missense	1	0,0004							1	0,001				
12	c.4494G>A	p. =	rs146532299	Silent	1	0,0004			1	0,002								
12	c.4500T>C	p. =	rs3810812	Silent	1025 (178)	0,4583	63 (21)	0,67	239 (114)	0,726	29 (12)	0,788	448	0,614	246 (145)	0,743		
12	c.4523C>T	p.Ser1508Leu	rs112694849	Missense	4	0,0015											4	0,005
12	c.4530G>T	p. =	rs139254595	Silent	3	0,0011	1	0,011	1	0,002			1	0,001				
12	c.4549G>T	p.Glu1517*	NA	Non Sense	1	0,0004			1	0,002								

Tabla 9: Variantes encontradas en el gen SLX4 en casos HBOC en seis estudios

Series de familias BRCA1																		
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
12	c.4563T>C	p. =	rs200225783	Silent	1	0,0004									1	0,002		
12	c.4580C>T	p.Pro1527Leu	rs149362820	Missense	1	0,0004									1	0,002		
12	c.4581G>A	p. =	rs78635099	Silent	14	0,0053			4	0,008			1	0,001	9	0,017		
12	c.4597G>T	p.Ala1533Ser	rs200377191	Missense	1	0,0004									1	0,002		
12	c.4597GC>TA	p.Ala1533Thr	NA	Missense	1	0,0004											1	0,001
12	c.4600G>A	p.Gly1534Ser	rs78770603	Missense	5	0,0019			1	0,002					1	0,002	3	0,004
13	c.4637-110T>C		rs74005837	Intronic	1	0,0004							1	0,001				
13	c.4637-125C>T		NA	Intronic	1	0,0004	1	0,011										
13	C.4637-210A>		NA	Intronic	1	0,0004							1	0,001				
13	c.4637-227C>T		rs75693937	Intronic	74	0,0282	9	0,096					65	0,089				
13	C.4637-234G>A		NA	Intronic	1	0,0004							1	0,001				
13	c.4637-267C>T		NA	Intronic	1	0,0004							1	0,001				
13	c.4648C>T	p.Arg1550Trp	rs77021998	Missense	10	0,0038			3	0,006			3	0,004	1	0,002	3	0,004
13	c.4649G>A	p.Arg1550Gln	rs369852048	Missense	1	0,0004											1	0,001
13	c.4739+10C>T		NA	Intronic	1	0,0004	1	0,011										
13	c.4739+24G>T		rs12933120	Intronic	403 (14)	0,1749	22 (2)	0,234	93	0,191	12	0,23	133	0,182	143 (12)	0,294		
13	c.4739+31G>A		NA	Intronic	2	0,0008							2	0,002				
13	c.4739+79G>A		rs74762428	Intronic	20	0,0076							20	0,027				
13	c.4739+90C>G		NA	Intronic	2	0,0008							2	0,002				
14	c.4797C>T	p. =	NA	Silent	1	0,0004							1	0,001				
14	c.4865A>G	p.Gln1622Arg	rs201504997	Missense	1	0,0004									1	0,002		
14	c.4887C>T	p. =	NA	Silent	1	0,0004							1	0,001				
14	c.5029C>T	p.Pro1677Ser	rs7196345	Missense	2	0,0008											2	0,003
14	c.5040G>C	p.Arg1680Ser	rs199592185	Missense	1	0,0004									1	0,002		

Series de familias BRCA1																		
					Total N=2625		España-BCN N=94		España-MAD N=486		Alemania N=52		Holanda N=729		Italia N=526		USA N=738	
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.	Portadores (Homo)	Frec.
14	c.5072A>G	p.Asn1691Ser	NA	Missense	1	0,0004	1	0,011										
14	c.5081C>T	p.Ala1694Val	NA	Missense	1	0,0004							1	0,001				
14	c.5146T>A	p.Ser1716Thr	rs75182789	Missense	3	0,0011									1	0,002	2	0,003
14	c.5150A>C	p.Gln1717Pro	NA	Missense	1	0,0004											1	0,001
15	c.5154-28C>T		rs201912097	Intronic	1	0,0004									1	0,002		
15	c.5155T>A	p.Ser1719Tyr	rs200074692	Missense	1	0,0004									1	0,002		
15	c.5183T>G	p.Phe1728Cys	rs200991968	Missense	1	0,0004									1	0,002		
15	c.5189C>T	p.Ser1730Phe	NA	Missense	1	0,0004											1	0,001
15	c.5389C>T	p. =	NA	Silent	1	0,0004	1	0,011										
15	c.5440C>T	p.Arg1814Cys	NA	Missense	2	0,0008							1	0,001			1	0,001
15	c.5501A>G	p.Asn1834Ser	rs111738042	Missense	24	0,0091	2	0,022			1	0,02	7	0,009	4	0,007	10	0,014
15	c.*8A>G		rs3751839	Intronic	126	0,0480	9	0,097			10	0,19	57	0,078	50	0,095		
15	c.*102G>A		NA	Intronic	2	0,0008	1	0,011					1	0,001				
15	c.*113C>T		rs76661336	Intronic	64	0,0244	8	0,086					56	0,076				



**Tabla 10:** Variantes encontradas en el gen *RAD51D* en casos *HBOC* y esporádicos en siete estudios

Series de familias BRCAx y casos esporádicos																				
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Total N=4289		Finlandia N=95		Australia N=1305		Reino Unido N=911*		EE.UU. N= 814		Canada-Bélgica N=175		Francia N=142		España N= 842	
					Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec
5'UTR	c.-111G>A		-	UTR	1	0,0002													1	0,001
5'UTR	c.-60C>T		-	UTR	1	0,0002													1	0,001
5'UTR	c.-39C>T		-	UTR	1	0,0002			1	0,0008										
1	c.1A>T	p.Met1?	-	pérdida codón de inicio	1	0,0002													1	0,001
1	c.26G>C	p.Cys9Ser	rs140825795	no sinónima	5	0,0012					2	0,002	1	0,001			1	0,007	1	0,001
1	c.47T>C	p.Met16Thr	-	no sinónima	1	0,0002			1	0,0008										
1	c.82+128C>T		rs28363258	intrónica	15	0,0036	***												15	0,018
2	c.83-26A>G		rs202048781	intrónica	1	0,0002								1	0,006					
2	c.83-20T>C		rs182793287	intrónica	1	0,0002			1	0,0008										
2	c.83-4T>C		-	intrónica	1	0,0002													1	0,001
2	c.117A>T	p.=	-	sinónima	1	0,0002			1	0,0008										
2	c.131_144+24del38	p.Val28Glyfs*21	-	cambio de fase (splicing)	1	0,0002							1	0,001						
2	c.137C>G	p.Ser46Cys	-	no sinónima	2	0,0005					1	0,001	1	0,001						
3	c.185C>T	p.Ser62Leu	rs374357106	no sinónima	1	0,0002					1	0,001								
3	c.211C>T	p.Leu71Phe	-	no sinónima	1	0,0002											1	0,007		
3	c.234C>T	p.=	rs9901455	sinónima	354	0,0844	***				124	0,136	70	0,086	36	0,206			124	0,147
4	c.264-6C>T		-	intrónica	1	0,0002			1	0,0008										
4	c.264-50_264-49del		-	intrónica	1	0,0002											1	0,007		
4	c.270_271dup	p.Lys91Ilefs*13	-	cambio de fase	1	0,0002					1	0,001								
4	c.286G>T	p.Gly96Cys	-	no sinónima	1	0,0002			1	0,0008										
4	c.292G>A	p.Asp98Asn	-	no sinónima	1	0,0002											1	0,007		

Tabla 10. Variantes encontradas en el gen RAD51D en casos HBOC y esporádicos en siete estudios

Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Series de familias BRCAx y casos esporádicos															
					Total N=4289		Finlandia N=95		Australia N=1305		Reino Unido N=911*		EE.UU. N= 814		Canadá-Bélgica N=175		Francia N=142		España N= 842	
					Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec
5	c.345G>C	p.Gln115His	-	no sinónima ( <i>splicing</i> )	1	0,0002					1	0,001								
5	c.355T>C	p.Cys119Arg	rs201313861	no sinónima	2	0,0005								1	0,006			1	0,001	
5	c.363delA	p.Ala122Glnfs*14	-	cambio de fase	1	0,0002					1	0,001								
5	c.394G>A	p.Val132Ile	rs201141245	no sinónima	1	0,0002			1	0,0008										
5	c.480+1G>A		-	intrónica ( <i>splicing</i> )	1	0,0002					1	0,001								
5	c.480+75T>G		rs8067688	intrónica	5	0,0012												5	0,015	
5	c.480+158C>A		rs3816754	intrónica	9	0,0021												9	0,027	
5	c.480+187T>C		rs936656	intrónica	67	0,0156												67	0,713	
6	c.481-205A>G		rs8071313	intrónica	3	0,0007												3	0,032	
6	c.481-144T>C		rs8067218	intrónica	1	0,0002												1	0,011	
6	c.494G>A	p.Arg165Gln	rs4796033	no sinónima	589	0,1404	***				264	0,290	101	0,124	41	0,234		183	0,217	
6	c.556C>T	p.Arg186*	-	codon de parada	4	0,0009			1	0,0008	2	0,002		1	0,006					
6	c.576+1G>A		-	intrónica ( <i>splicing</i> )	2	0,0005	2	0,001												
7	c.577-103G>C		rs188753384	intrónica	3	0,0007												3	0,004	
7	c.580delA	p.Thr194Leufs*6	-	cambio de fase	1	0,0002							1	0,001						
7	c.620C>T	p.Ser207Leu	-	no sinónima	3	0,0007					1	0,001	1	0,001			1	0,007		
7	c.629C>T	p.Ala210Val	-	no sinónima	1	0,0002												1	0,001	
7	c.630G>A	p. =	-	sinónima	1	0,0002					1	0,001								
7	c.667+2_667+23del	p.Val193Alafs*4	-	intrónica ( <i>splicing</i> )	1	0,0002												1	0,001	
8	c.694C>T	p.Arg232*	-	codon de parada	3	0,0007							1	0,001				2	0,002	
8	c.695G>A	p.Arg232Gln	rs28363283	no sinónima	7	0,0016					1	0,001	1	0,001				5	0,006	
8	c.698A>G	p.Glu233Gly	rs28363284	no sinónima	87	0,0203					43	0,047	10	0,012	7	0,040		27	0,032	
8	c.715C>T	p.Arg239Trp	-	no sinónima	3	0,0007							1	0,001				2	0,002	

Series de familias BRCA y casos esporádicos																					
Exón	Cambio de nucleótido	Cambio de aminoácido	dbSNP	Tipo de cambio	Total N=4289		Finlandia N=95		Australia N=1305		Reino Unido N=911*		EE.UU. N= 814		Canada-Bélgica N=175		Francia N=142		España N= 842		
					Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	Portadores (Homo)	Frec	
8	c.738+54T>C		-	intrónica	1	0,0002												1	0,007		
9	c.739-103T>C		rs28363289	intrónica	7	0,0016														7	0,040
9	c.739-10T>C		rs199998187	intrónica	1	0,0002					1	0,001									
9	c.748delC	p.His250Thrfs*2	-	cambio de fase	**	-															
9	c.753A>G	p.Ile251Met	-	no sinónima	1	0,0002					1	0,001									
9	c.757C>T	p.Arg253*	-	codon de parada	1	0,0002					1	0,001									
9	c.793G>A	p.Gly265Arg	rs140285068	no sinónima	2	0,0005			1	0,0008				1	0,006						
9	c.796C>T	p.Arg266Cys	-	no sinónima	1	0,0002			1	0,0008											
9	c.803G>A	p.Trp268*	-	codon de parada	2	0,0005			1	0,0008	1	0,001									
9	c.872G>A	p.Arg291His	rs150134822	no sinónima	1	0,0002					1	0,001									
9	c.879G>A	p.=	-	sinónima	2	0,0005			1	0,0008										1	0,001
9	c.911G>A	p.Gly304Asp	rs200615280	no sinónima	1	0,0002					1	0,001									
9	c.903+53C>T		rs45496096	intrónica	21	0,0050	***											1	0,007	20	0,032
10	c.904-11T>A		-	intrónica	1	0,0002														1	0,001
10	c.904-3C>T		rs45478491	intrónica	1	0,0002								1	0,006						
10	c.932T>A	p.Ile311Asn	rs145309168	no sinónima	1	0,0002							1	0,001							

\* Este estudio incluye una serie adicional de 737 casos de cáncer de mama familiar cuyo resultado se describe como ausencia de variantes deletéreas sin detalles adicionales y por ello no ha sido incluido en esta tabla.

\*\* Esta variante se ha descrito únicamente en controles, se ha incluido en la tabla por tener un efecto deletéreo. Aquellas variantes de significado clínico incierto encontradas únicamente en controles no han sido incluidas en la tabla.

\*\*\*Estas variantes se han descrito como presentes en la población sin especificarse su frecuencia.



## Referencias

### Sin autor

Genetics/familial high risk assesment:Breast and Ovarian Clinical Practice. Version 1.2012. In Clinical Practice Guidelines in Oncology.

([http://www.nccn.org/professionals/physician\\_gls/pdf/genetics\\_screening.pdf](http://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf), National Comprehensive Cancer Network).

(1985). Ataxia-telangiectasia: genetics, neuropathology, and immunology of a degenerative disease of childhood. Proceedings of a conference. Solvang, California, January 16-20, 1984. Kroc Foundation series 19, 1-407.

(2000). Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. British journal of cancer 83, 1301-1308.

(2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57-74.

### A

Acedo, A., Sanz, D.J., Duran, M., Infante, M., Perez-Cabornero, L., Miner, C., and Velasco, E.A. (2012). Comprehensive splicing functional analysis of DNA variants of the BRCA2 gene by hybrid minigenes. Breast cancer research : BCR 14, R87.

Akbari, M.R., Tonin, P., Foulkes, W.D., Ghadirian, P., Tischkowitz, M., and Narod, S.A. (2010). RAD51C germline mutations in breast and ovarian cancer patients. Breast cancer research : BCR 12, 404.

Alter, B.P., Giri, N., Savage, S.A., Peters, J.A., Loud, J.T., Leathwood, L., Carr, A.G., Greene, M.H., and Rosenberg, P.S. (2010). Malignancies and survival patterns in the National Cancer Institute inherited bone marrow failure syndromes cohort study. British journal of haematology 150, 179-188.

Ameziane, N., van den Ouweland, A.M., Adank, M.A., Vijzelaar, R.N., Errami, A., Dorsman, J.C., Joenje, H., Meijers-Heijboer, H., and Waisfisz, Q. (2009). Lack of large genomic deletions in BRIP1, PALB2, and FANCD2 genes in BRCA1/2 negative familial breast cancer. *Breast cancer research and treatment* 118, 651-653.

Anczukow, O., Buisson, M., Leone, M., Coutanson, C., Lasset, C., Calender, A., Sinilnikova, O.M., and Mazoyer, S. (2012). BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy. *Clinical cancer research : an official journal of the American Association for Cancer Research* 18, 4903-4909.

Andersen, S.L., Bergstralh, D.T., Kohl, K.P., LaRocque, J.R., Moore, C.B., and Sekelsky, J. (2009). Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. *Molecular cell* 35, 128-135.

Antoniou, A., Pharoah, P.D., Narod, S., Risch, H.A., Eyfjord, J.E., Hopper, J.L., Loman, N., Olsson, H., Johannsson, O., Borg, A., *et al.* (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *American journal of human genetics* 72, 1117-1130.

Antoniou, A.C., Gayther, S.A., Stratton, J.F., Ponder, B.A., and Easton, D.F. (2000). Risk models for familial ovarian and breast cancer. *Genetic epidemiology* 18, 173-190.

Antoniou, A.C., Kuchenbaecker, K.B., Soucy, P., Beesley, J., Chen, X., McGuffog, L., Lee, A., Barrowdale, D., Healey, S., Sinilnikova, O.M., *et al.* (2012). Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers. *Breast cancer research : BCR* 14, R33.

Antoniou, A.C., Pharoah, P.D., McMullan, G., Day, N.E., Stratton, M.R., Peto, J., Ponder, B.J., and Easton, D.F. (2002). A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *British journal of cancer* 86, 76-83.

Arason, A., Gunnarsson, H., Johannesdottir, G., Jonasson, K., Bendahl, P.O., Gillanders, E.M., Agnarsson, B.A., Jonsson, G., Pylkas, K., Mustonen, A., *et al.* (2010). Genome-wide search for breast cancer linkage in large Icelandic non-BRCA1/2 families. *Breast cancer research : BCR* 12, R50.

Armes, J.E., Trute, L., White, D., Southey, M.C., Hammet, F., Tesoriero, A., Hutchins, A.M., Dite, G.S., McCredie, M.R., Giles, G.G., *et al.* (1999). Distinct molecular pathogeneses of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study. *Cancer research* 59, 2011-2017.

## **B**

Bakker, J.L., van Mil, S.E., Crossan, G., Sabbaghian, N., De Leeneer, K., Poppe, B., Adank, M., Gille, H., Verheul, H., Meijers-Heijboer, H., *et al.* (2013). Analysis of the Novel Fanconi Anemia Gene SLX4/FANCP in Familial Breast Cancer Cases. *Human mutation* 34, 70-73.

Barroso, E., Milne, R.L., Fernandez, L.P., Zamora, P., Arias, J.I., Benitez, J., and Ribas, G. (2006). FANCD2 associated with sporadic breast cancer risk. *Carcinogenesis* 27, 1930-1937.

Bartkova, J., Tommiska, J., Oplustilova, L., Aaltonen, K., Tamminen, A., Heikkinen, T., Mistrik, M., Aittomaki, K., Blomqvist, C., Heikkila, P., *et al.* (2008). Aberrations of the MRE11-RAD50-NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. *Molecular oncology* 2, 296-316.

Beristain, E., Guerra, I., Vidaurrazaga, N., Burgos-Bretones, J., and Tejada, M.I. (2010). LOH analysis should not be used as a tool to assess whether UVs of BRCA1/2 are pathogenic or not. *Familial cancer* 9, 289-290.

Bernholtz, S., Laitman, Y., Kaufman, B., Shimon-Paluch, S., and Friedman, E. (2012). Phenocopy breast cancer rates in Israeli BRCA1 BRCA2 mutation carrier families: is the risk increased in non-carriers? *Breast cancer research and treatment* 132, 669-673.

Berwick, M., Satagopan, J.M., Ben-Porat, L., Carlson, A., Mah, K., Henry, R., Diotti, R., Milton, K., Pujara, K., Landers, T., *et al.* (2007). Genetic heterogeneity among Fanconi anemia heterozygotes and risk of cancer. *Cancer research* 67, 9591-9596.

Blanco, A., de la Hoya, M., Balmana, J., Ramon y Cajal, T., Teule, A., Miramar, M.D., Esteban, E., Infante, M., Benitez, J., Torres, A., *et al.* (2012). Detection of a large rearrangement in PALB2 in Spanish breast cancer families with male breast cancer. *Breast cancer research and treatment* 132, 307-315.

Blay, P., Santamaria, I., Pitiot, A.S., Luque, M., Alvarado, M.G., Lastra, A., Fernandez, Y., Paredes, A., Freije, J.M., and Balbin, M. (2013). Mutational analysis of BRCA1 and BRCA2 in hereditary breast and ovarian cancer families from Asturias (Northern Spain). *BMC cancer* 13, 243.

Bochar, D.A., Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W.S., Wang, W., Kashanchi, F., and Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell* 102, 257-265.

Bogdanova, N., Feshchenko, S., Schurmann, P., Waltes, R., Wieland, B., Hillemanns, P., Rogov, Y.I., Dammann, O., Bremer, M., Karstens, J.H., *et al.* (2008). Nijmegen Breakage Syndrome mutations and risk of breast cancer. *International journal of cancer. Journal international du cancer* 122, 802-806.

Bonatti, F., Pepe, C., Tancredi, M., Lombardi, G., Aretini, P., Sensi, E., Falaschi, E., Cipollini, G., Bevilacqua, G., and Caligo, M.A. (2006). RNA-based analysis of BRCA1 and BRCA2 gene alterations. *Cancer genetics and cytogenetics* 170, 93-101.

Bork, P., Hofmann, K., Bucher, P., Neuwald, A.F., Altschul, S.F., and Koonin, E.V. (1997). A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 11, 68-76.

Boyd, N.F., Stone, J., Vogt, K.N., Connelly, B.S., Martin, L.J., and Minkin, S. (2003). Dietary fat and breast cancer risk revisited: a meta-analysis of the published literature. *British journal of cancer* 89, 1672-1685.

Bray, F., McCarron, P., and Parkin, D.M. (2004). The changing global patterns of female breast cancer incidence and mortality. *Breast cancer research : BCR* 6, 229-239.

Brewster, A., and Helzlsouer, K. (2001). Breast cancer epidemiology, prevention, and early detection. *Current opinion in oncology* 13, 420-425.

Broeks, A., Urbanus, J.H., Floore, A.N., Dahler, E.C., Klijn, J.G., Rutgers, E.J., Devilee, P., Russell, N.S., van Leeuwen, F.E., and van 't Veer, L.J. (2000). ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. *American journal of human genetics* 66, 494-500.

Brooks-Wilson, A.R., Kaurah, P., Suriano, G., Leach, S., Senz, J., Grehan, N., Butterfield, Y.S., Jeyes, J., Schinas, J., Bacani, J., *et al.* (2004). Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. *Journal of medical genetics* 41, 508-517.

Brown, M.A., Xu, C.F., Nicolai, H., Griffiths, B., Chambers, J.A., Black, D., and Solomon, E. (1996). The 5' end of the BRCA1 gene lies within a duplicated region of human chromosome 17q21. *Oncogene* 12, 2507-2513.

Buslov, K.G., Iyevleva, A.G., Chekmariova, E.V., Suspitsin, E.N., Togo, A.V., Kuligina, E., Sokolenko, A.P., Matsko, D.E., Turkevich, E.A., Lazareva, Y.R., *et al.* (2005). NBS1 657del5 mutation may contribute only to a limited fraction of breast cancer cases in Russia. *International journal of cancer. Journal international du cancer* 114, 585-589.

**C**

Calvo, V., and Beato, M. (2011). BRCA1 counteracts progesterone action by ubiquitination leading to progesterone receptor degradation and epigenetic silencing of target promoters. *Cancer research* 71, 3422-3431.

Callebaut, I., and Mornon, J.P. (1997). From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. *FEBS letters* 400, 25-30.

Campos, B., Diez, O., Alvarez, C., Palma, L., Domenech, M., Balmana, J., Sanz, J., Ramirez, A., Alonso, C., Carvallo, P., *et al.* (2004). [Haplotype of the BRCA2 6857delAA mutation in 4 families with breast/ovarian cancer]. *Medicina clinica* 123, 543-545.

Campos, B., Diez, O., Odefrey, F., Domenech, M., Moncoutier, V., Martinez-Ferrandis, J.I., Osorio, A., Balmana, J., Barroso, A., Armengod, M.E., *et al.* (2003). Haplotype analysis of the BRCA2 9254delATCAT recurrent mutation in breast/ovarian cancer families from Spain. *Human mutation* 21, 452.

Cancer, C.G.o.H.F.i.B. (1997). Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. . *Lancet* 350, 1047-1059.

Cantor, S.B., Bell, D.W., Ganesan, S., Kass, E.M., Drapkin, R., Grossman, S., Wahrer, D.C., Sgroi, D.C., Lane, W.S., Haber, D.A., *et al.* (2001). BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* 105, 149-160.

Caputo, S., Benboudjema, L., Sinilnikova, O., Rouleau, E., Beroud, C., and Lidereau, R. (2012). Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. *Nucleic acids research* 40, D992-1002.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature reviews. Genetics* 3, 285-298.

Casadei, S., Norquist, B.M., Walsh, T., Stray, S., Mandell, J.B., Lee, M.K., Stamatoyannopoulos, J.A., and King, M.C. (2011). Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer. *Cancer research* 71, 2222-2229.

Catucci, I., Colombo, M., Verderio, P., Bernard, L., Ficarazzi, F., Mariette, F., Barile, M., Peissel, B., Cattaneo, E., Manoukian, S., *et al.* (2012). Sequencing analysis of SLX4/FANCP gene in Italian familial breast cancer cases. *PloS one* 7, e31038.

Clague, J., Wilhoite, G., Adamson, A., Bailis, A., Weitzel, J.N., and Neuhausen, S.L. (2011). RAD51C germline mutations in breast and ovarian cancer cases from high-risk families. *PloS one* 6, e25632.

Clark, S.L., Rodriguez, A.M., Snyder, R.R., Hankins, G.D., and Boehning, D. (2012). Structure-Function Of The Tumor Suppressor BRCA1. *Computational and structural biotechnology journal* 1.

Claus, E.B., Risch, N.J., and Thompson, W.D. (1990). Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* 131, 961-972.

Claus, E.B., Schildkraut, J.M., Thompson, W.D., and Risch, N.J. (1996). The genetic attributable risk of breast and ovarian cancer. *Cancer* 77, 2318-2324.

Collaborative Group on Hormonal Factors in Breast, C. (2001). Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358, 1389-1399.

Collaborative Group on Hormonal Factors in Breast, C. (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet* 360, 187-195.

Couch, F.J., Rasmussen, L.J., Hofstra, R., Monteiro, A.N., Greenblatt, M.S., and de Wind, N. (2008). Assessment of functional effects of unclassified genetic variants. *Human mutation* 29, 1314-1326.

Coulet, F., Fajac, A., Colas, C., Eyries, M., Dion-Miniere, A., Rouzier, R., Uzan, S., Lefranc, J.P., Carbonnel, M., Cornelis, F., *et al.* (2012). Germline RAD51C mutations in ovarian cancer susceptibility. *Clinical genetics*.

Chlebowski, R.T., Hendrix, S.L., Langer, R.D., Stefanick, M.L., Gass, M., Lane, D., Rodabough, R.J., Gilligan, M.A., Cyr, M.G., Thomson, C.A., *et al.* (2003). Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. *JAMA : the journal of the American Medical Association* 289, 3243-3253.

Chun, J., Buechelmaier, E.S., and Powell, S.N. (2013). Rad51 Paralog Complexes BCDX2 and CX3 Act at Different Stages in the BRCA1-BRCA2-Dependent Homologous Recombination Pathway. *Molecular and cellular biology* 33, 387-395.

## D

D'Andrea, A.D. (2010). Susceptibility pathways in Fanconi's anemia and breast cancer. *The New England journal of medicine* 362, 1909-1919.

de Garibay, G.R., Diaz, A., Gavina, B., Romero, A., Garre, P., Vega, A., Blanco, A., Tosar, A., Diez, O., Perez-Segura, P., *et al.* (2012). Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families. *European journal of human genetics : EJHG*.

de Juan Jimenez, I., Esteban Cardenosa, E., Palanca Suela, S., Barragan Gonzalez, E., Aznar Carretero, I., Munarriz Gandia, B., Santaballa Bertran, A., Torregrosa Maicas, M.D., Guillen Ponce, C., Sanchez Heras, A.B., *et al.* (2012). Low prevalence of BRCA1 and BRCA2 mutations in the sporadic breast cancer of Spanish population. *Familial cancer* 11, 49-56.

de Juan Jimenez, I., Garcia Casado, Z., Palanca Suela, S., Esteban Cardenosa, E., Lopez Guerrero, J.A., Segura Huerta, A., Chirivella Gonzalez, I., Sanchez Heras, A.B., Juan Fita, M.J., Tena Garcia, I., *et al.* (2013). Novel and recurrent BRCA1/BRCA2 mutations in early onset and familial breast and ovarian cancer detected in the Program of Genetic Counseling in Cancer of Valencian Community (eastern Spain). Relationship of family phenotypes with mutation prevalence. *Familial cancer*.

de la Hoya, M. (2004). Documento de consenso sobre cáncer de mama hereditario. In *Documentos de Consenso en Cáncer Hereditario* (Madrid, Sociedad Española de Oncología Médica), pp. 13-15.

de la Hoya, M., Gutierrez-Enriquez, S., Velasco, E., Osorio, A., Sanchez de Abajo, A., Vega, A., Salazar, R., Esteban, E., Llord, G., Gonzalez-Sarmiento, R., *et al.* (2006). Genomic rearrangements at the BRCA1 locus in Spanish families with breast/ovarian cancer. *Clinical chemistry* 52, 1480-1485.

De Leeneer, K., Van Bockstal, M., De Brouwer, S., Swietek, N., Schietecatte, P., Sabbaghian, N., Van den Ende, J., Willocx, S., Storm, K., Blaumeiser, B., *et al.* (2012). Evaluation of RAD51C as cancer susceptibility gene in a large breast-ovarian cancer patient population referred for genetic testing. *Breast cancer research and treatment* 133, 393-398.

Del Valle, J., Campos, O., Velasco, A., Darder, E., Menendez, M., Feliubadalo, L., Tornero, E., Blanco, I., Izquierdo, A., Brunet, J., *et al.* (2011). Identification of a new complex rearrangement affecting exon 20 of BRCA1. *Breast cancer research and treatment* 130, 341-344.

del Valle, J., Feliubadalo, L., Nadal, M., Teule, A., Miro, R., Cuesta, R., Tornero, E., Menendez, M., Darder, E., Brunet, J., *et al.* (2010). Identification and comprehensive characterization of large genomic rearrangements in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 122, 733-743.

Diez, O., Osorio, A., Duran, M., Martinez-Ferrandis, J.I., de la Hoya, M., Salazar, R., Vega, A., Campos, B., Rodriguez-Lopez, R., Velasco, E., *et al.* (2003). Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: a high proportion of mutations unique to Spain and evidence of founder effects. *Human mutation* 22, 301-312.

Diez, O., Osorio, A., Robledo, M., Barroso, A., Domenech, M., Cortes, J., Albertos, J., Sanz, J., Brunet, J., SanRoman, J.M., *et al.* (1999). Prevalence of BRCA1 and BRCA2 Jewish mutations in Spanish breast cancer patients. *British journal of cancer* 79, 1302-1303.

Domchek, S.M., Tang, J.B., Stopfer, J., Lilli, D.R., Hamel, N., Tischkowitz, M., Monteiro, A.N., Messick, T.E., Powers, J., Yonker, A., *et al.* (2012). Biallelic Deleterious BRCA1 Mutations in a Woman with Early-Onset Ovarian Cancer. *Cancer discovery*.

Dosil, V., Tosar, A., Canadas, C., Perez-Segura, P., Diaz-Rubio, E., Caldes, T., and de la Hoya, M. (2010). Alternative splicing and molecular characterization of splice site variants: BRCA1 c.591C>T as a case study. *Clinical chemistry* 56, 53-61.

## E

Easton, D., Ford, D., and Peto, J. (1993). Inherited susceptibility to breast cancer. *Cancer surveys* 18, 95-113.

Easton, D.F., Steele, L., Fields, P., Ormiston, W., Averill, D., Daly, P.A., McManus, R., Neuhausen, S.L., Ford, D., Wooster, R., *et al.* (1997). Cancer risks in two large breast cancer families linked to BRCA2 on chromosome 13q12-13. *American journal of human genetics* 61, 120-128.

Eccles, D.M., Evans, D.G., and Mackay, J. (2000). Guidelines for a genetic risk based approach to advising women with a family history of breast cancer. UK Cancer Family Study Group (UKCFSG). *Journal of medical genetics* 37, 203-209.

Ellis, N.A., and Offit, K. (2012). Heterozygous mutations in DNA repair genes and hereditary breast cancer: a question of power. *PLoS genetics* 8, e1003008.

Emi, M., Matsushima, M., Katagiri, T., Yoshimoto, M., Kasumi, F., Yokota, T., Nakata, T., Miki, Y., and Nakamura, Y. (1998). Multiplex mutation screening of the BRCA1 gene in 1000 Japanese breast cancers. *Japanese journal of cancer research : Gann* 89, 12-16.

Esashi, F., Christ, N., Gannon, J., Liu, Y., Hunt, T., Jasin, M., and West, S.C. (2005). CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* 434, 598-604.

Evers, B., and Jonkers, J. (2006). Mouse models of BRCA1 and BRCA2 deficiency: past lessons, current understanding and future prospects. *Oncogene* 25, 5885-5897.

## F

Fackenthal, J.D., Lee, Y., and Olopade, O.I. (2012a). Hidden dangers: a cryptic exon disrupts BRCA2 mRNA. *Clinical cancer research : an official journal of the American Association for Cancer Research* 18, 4865-4867.

Fackenthal, J.D., and Olopade, O.I. (2007). Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nature reviews. Cancer* 7, 937-948.

Fackenthal, J.D., Zhang, J., Zhang, B., Zheng, Y., Hagos, F., Burrill, D.R., Niu, Q., Huo, D., Sveen, W.E., Ogundiran, T., *et al.* (2012b). High prevalence of BRCA1 and BRCA2 mutations in unselected Nigerian breast cancer patients. *International journal of cancer. Journal international du cancer* 131, 1114-1123.

Fan, S., Wang, J., Yuan, R., Ma, Y., Meng, Q., Erdos, M.R., Pestell, R.G., Yuan, F., Auborn, K.J., Goldberg, I.D., *et al.* (1999). BRCA1 inhibition of estrogen receptor signaling in transfected cells. *Science (New York, N.Y.)* 284, 1354-1356.

Fekairi, S., Scaglione, S., Chahwan, C., Taylor, E.R., Tissier, A., Coulon, S., Dong, M.Q., Ruse, C., Yates, J.R., 3rd, Russell, P., *et al.* (2009). Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* 138, 78-89.

Feng, B.J., Tavtigian, S.V., Southey, M.C., and Goldgar, D.E. (2011). Design considerations for massively parallel sequencing studies of complex human disease. *PloS one* 6, e23221.

Ferlay, J., Parkin, D.M., and Steliarova-Foucher, E. (2010). Estimates of cancer incidence and mortality in Europe in 2008. *European journal of cancer* 46, 765-781.

Fernandez-Rodriguez, J., Quiles, F., Blanco, I., Teule, A., Feliubadalo, L., Valle, J.D., Salinas, M., Izquierdo, A., Darder, E., Schindler, D., *et al.* (2012). Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. *BMC cancer* 12, 84.

Ford, D., Easton, D.F., Bishop, D.T., Narod, S.A., and Goldgar, D.E. (1994). Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet* 343, 692-695.

Foulkes, W.D., Metcalfe, K., Sun, P., Hanna, W.M., Lynch, H.T., Ghadirian, P., Tung, N., Olopade, O.I., Weber, B.L., McLennan, J., *et al.* (2004). Estrogen receptor status in BRCA1- and BRCA2-related breast cancer: the influence of age, grade, and histological type. *Clinical cancer research : an official journal of the American Association for Cancer Research* 10, 2029-2034.

## G

Gaildrat, P., Krieger, S., Di Giacomo, D., Abdat, J., Revillion, F., Caputo, S., Vaur, D., Jamard, E., Bohers, E., Ledemeney, D., *et al.* (2012). Multiple sequence variants of BRCA2 exon 7 alter splicing regulation. *Journal of medical genetics* 49, 609-617.

Garcia-Closas, M., Brinton, L.A., Lissowska, J., Chatterjee, N., Peplonska, B., Anderson, W.F., Szeszenia-Dabrowska, N., Bardin-Mikolajczak, A., Zatonski, W., Blair, A., *et al.* (2006). Established breast cancer risk factors by clinically important tumour characteristics. *British journal of cancer* 95, 123-129.

Garcia-Patino, E., Gomendio, B., Provencio, M., Silva, J.M., Garcia, J.M., Espana, P., and Bonilla, F. (1998). Germ-line BRCA1 mutations in women with sporadic breast cancer: clinical

correlations. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 16, 115-120.

Garcia, M.J., Fernandez, V., Osorio, A., Barroso, A., Fernandez, F., Urioste, M., and Benitez, J. (2009a). Mutational analysis of FANCL, FANCM and the recently identified FANCI suggests that among the 13 known Fanconi Anemia genes, only FANCD1/BRCA2 plays a major role in high-risk breast cancer predisposition. *Carcinogenesis* 30, 1898-1902.

Garcia, M.J., Fernandez, V., Osorio, A., Barroso, A., Llorca, G., Lazaro, C., Blanco, I., Caldes, T., de la Hoya, M., Ramon, Y.C.T., *et al.* (2009b). Analysis of FANCB and FANCN/PALB2 fanconi anemia genes in BRCA1/2-negative Spanish breast cancer families. *Breast cancer research and treatment* 113, 545-551.

Gayther, S.A., Goringe, K.L., Ramus, S.J., Huntsman, D., Roviello, F., Grehan, N., Machado, J.C., Pinto, E., Seruca, R., Halling, K., *et al.* (1998). Identification of germ-line E-cadherin mutations in gastric cancer families of European origin. *Cancer research* 58, 4086-4089.

Gayther, S.A., Mangion, J., Russell, P., Seal, S., Barfoot, R., Ponder, B.A., Stratton, M.R., and Easton, D. (1997). Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. *Nature genetics* 15, 103-105.

Ghimenti, C., Sensi, E., Presciuttini, S., Brunetti, I.M., Conte, P., Bevilacqua, G., and Caligo, M.A. (2002). Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations. *Genes, chromosomes & cancer* 33, 235-242.

Giardiello, F.M., Brensinger, J.D., Tersmette, A.C., Goodman, S.N., Petersen, G.M., Booker, S.V., Cruz-Correa, M., and Offerhaus, J.A. (2000). Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology* 119, 1447-1453.

Glover-Collins, K., and Thompson, M.E. (2008). Nuclear export of BRCA1 occurs during early S phase and is calcium-dependent. *Cellular signalling* 20, 958-968.

Go, R.C., King, M.C., Bailey-Wilson, J., Elston, R.C., and Lynch, H.T. (1983). Genetic epidemiology of breast cancer and associated cancers in high-risk families. I. Segregation analysis. *Journal of the National Cancer Institute* *71*, 455-461.

Goldgar, D.E., Easton, D.F., Byrnes, G.B., Spurdle, A.B., Iversen, E.S., and Greenblatt, M.S. (2008). Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Human mutation* *29*, 1265-1272.

Golmard, L., Caux-Moncoutier, V., Davy, G., Al Ageeli, E., Poirot, B., Tirapo, C., Michaux, D., Barbaroux, C., CD, D.E., Nicolas, A., *et al.* (2013). Germline mutation in the RAD51B gene confers predisposition to breast cancer. *BMC cancer* *13*, 484.

Gomez Garcia, E.B., Oosterwijk, J.C., Timmermans, M., van Asperen, C.J., Hogervorst, F.B., Hoogerbrugge, N., Oldenburg, R., Verhoef, S., Dommering, C.J., Ausems, M.G., *et al.* (2009). A method to assess the clinical significance of unclassified variants in the BRCA1 and BRCA2 genes based on cancer family history. *Breast cancer research : BCR* *11*, R8.

Gonzalez-Neira, A., Rosa-Rosa, J.M., Osorio, A., Gonzalez, E., Southey, M., Sinilnikova, O., Lynch, H., Oldenburg, R.A., van Asperen, C.J., Hoogerbrugge, N., *et al.* (2007). Genomewide high-density SNP linkage analysis of non-BRCA1/2 breast cancer families identifies various candidate regions and has greater power than microsatellite studies. *BMC genomics* *8*, 299.

Gorski, B., Debniak, T., Masojc, B., Mierzejewski, M., Medrek, K., Cybulski, C., Jakubowska, A., Kurzawski, G., Chosia, M., Scott, R., *et al.* (2003). Germline 657del5 mutation in the NBS1 gene in breast cancer patients. *International journal of cancer. Journal international du cancer* *106*, 379-381.

Gracia-Aznarez, F.J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., Duran, M., Osorio, A., Moreno, L., Gonzalez-Neira, A., *et al.* (2013). Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. *PLoS one* *8*, e55681.

---

Green, A.M., and Kupfer, G.M. (2009). Fanconi anemia. *Hematology/oncology clinics of North America* 23, 193-214.

Gronwald, J., Raczynski, A., Tarhoni, M., Blachowski, M., Huzarski, T., Byrski, T., Toloczko-Grabarek, A., Debniak, T., Cybulski, C., Huzarska, J., *et al.* (2006). Population screening for cancer family syndromes in the west pomeranian region of poland. *Hereditary cancer in clinical practice* 4, 56.

Gudas, J.M., Li, T., Nguyen, H., Jensen, D., Rauscher, F.J., 3rd, and Cowan, K.H. (1996). Cell cycle regulation of BRCA1 messenger RNA in human breast epithelial cells. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* 7, 717-723.

Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., Taite, H., Scoular, R., Miller, A., and Reeve, A.E. (1998). E-cadherin germline mutations in familial gastric cancer. *Nature* 392, 402-405.

## H

Haffty, B.G., Choi, D.H., Goyal, S., Silber, A., Ranieri, K., Matloff, E., Lee, M.H., Nissenblatt, M., Toppmeyer, D., and Moran, M.S. (2009). Breast cancer in young women (YBC): prevalence of BRCA1/2 mutations and risk of secondary malignancies across diverse racial groups. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 20, 1653-1659.

Haitjema, A., Brandt, B.W., Ameziane, N., May, P., Heringa, J., de Winter, J.P., Joenje, H., and Dorsman, J.C. (2013). A protein prioritization approach tailored for the FA/BRCA pathway. *PLoS one* 8, e62017.

Hall, J.M., Lee, M.K., Newman, B., Morrow, J.E., Anderson, L.A., Huey, B., and King, M.C. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. *Science (New York, N.Y.)* 250, 1684-1689.

Hashizume, R., Fukuda, M., Maeda, I., Nishikawa, H., Oyake, D., Yabuki, Y., Ogata, H., and Ohta, T. (2001). The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *The Journal of biological chemistry* 276, 14537-14540.

Hearle, N., Schumacher, V., Menko, F.H., Olschwang, S., Boardman, L.A., Gille, J.J., Keller, J.J., Westerman, A.M., Scott, R.J., Lim, W., *et al.* (2006). Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. *Clinical cancer research : an official journal of the American Association for Cancer Research* 12, 3209-3215.

Heikkinen, K., Rapakko, K., Karppinen, S.M., Erkkö, H., Knuutila, S., Lundan, T., Mannermaa, A., Borresen-Dale, A.L., Borg, A., Barkardottir, R.B., *et al.* (2006). RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. *Carcinogenesis* 27, 1593-1599.

Heine, G.F., and Parvin, J.D. (2007). BRCA1 control of steroid receptor ubiquitination. *Science's STKE : signal transduction knowledge environment* 2007, pe34.

Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A., Bignell, G., Warren, W., Aminoff, M., Hoglund, P., *et al.* (1998). A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 391, 184-187.

Henderson, B.R. (2005). Regulation of BRCA1, BRCA2 and BARD1 intracellular trafficking. *BioEssays : news and reviews in molecular, cellular and developmental biology* 27, 884-893.

Hernan, I., Borrás, E., de Sousa Dias, M., Gamundi, M.J., Mane, B., Llorca, G., Agundez, J.A., Blanca, M., and Carballo, M. (2012). Detection of genomic variations in BRCA1 and BRCA2 genes by long-range PCR and next-generation sequencing. *The Journal of molecular diagnostics : JMD* 14, 286-293.

Hery, C., Ferlay, J., Boniol, M., and Autier, P. (2008). Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 19, 1187-1194.

Hilbers, F.S., Meijers, C.M., Laros, J.F., van Galen, M., Hoogerbrugge, N., Vasen, H.F., Nederlof, P.M., Wijnen, J.T., van Asperen, C.J., and Devilee, P. (2013). Exome sequencing of germline DNA from non-BRCA1/2 familial breast cancer cases selected on the basis of aCGH tumor profiling. *PLoS one* 8, e55734.

Hilbers, F.S., Wijnen, J.T., Hoogerbrugge, N., Oosterwijk, J.C., Collee, M.J., Peterlongo, P., Radice, P., Manoukian, S., Feroce, I., Capra, F., *et al.* (2012). Rare variants in XRCC2 as breast cancer susceptibility alleles. *Journal of medical genetics* 49, 618-620.

Hodson, C., and Walden, H. (2012). Towards a molecular understanding of the fanconi anemia core complex. *Anemia* 2012, 926787.

Hofstra, R.M., Spurdle, A.B., Eccles, D., Foulkes, W.D., de Wind, N., Hoogerbrugge, N., and Hogervorst, F.B. (2008). Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance. *Human mutation* 29, 1292-1303.

Hollestelle, A., Wasielewski, M., Martens, J.W., and Schutte, M. (2010). Discovering moderate-risk breast cancer susceptibility genes. *Current opinion in genetics & development* 20, 268-276.

Hopper, J.L., Southey, M.C., Dite, G.S., Jolley, D.J., Giles, G.G., McCredie, M.R., Easton, D.F., and Venter, D.J. (1999). Population-based estimate of the average age-specific cumulative risk of breast cancer for a defined set of protein-truncating mutations in BRCA1 and BRCA2. Australian Breast Cancer Family Study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 8, 741-747.

Houdayer, C., Caux-Moncoutier, V., Krieger, S., Barrois, M., Bonnet, F., Bourdon, V., Bronner, M., Buisson, M., Coulet, F., Gaildrat, P., *et al.* (2012). Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Human mutation* 33, 1228-1238.

Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G., *et al.* (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. *Science (New York, N.Y.)* 297, 606-609.

Huen, M.S., Sy, S.M., and Chen, J. (2010). BRCA1 and its toolbox for the maintenance of genome integrity. *Nature reviews. Molecular cell biology* 11, 138-148.

Hughes-Davies, L., Huntsman, D., Ruas, M., Fuks, F., Bye, J., Chin, S.F., Milner, J., Brown, L.A., Hsu, F., Gilks, B., *et al.* (2003). EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* 115, 523-535.

Hussain, S., Wilson, J.B., Medhurst, A.L., Hejna, J., Witt, E., Ananth, S., Davies, A., Masson, J.Y., Moses, R., West, S.C., *et al.* (2004). Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Human molecular genetics* 13, 1241-1248.

## I

Infante, M., Duran, M., Lasa, A., Acedo, A., de la Hoya, M., Esteban-Cardenosa, E., Sanz, D.J., Perez-Cabornero, L., Lastra, E., Miner, C., *et al.* (2010). Two founder BRCA2 mutations predispose to breast cancer in young women. *Breast cancer research and treatment* 122, 567-571.

Ishitobi, M., Miyoshi, Y., Hasegawa, S., Egawa, C., Tamaki, Y., Monden, M., and Noguchi, S. (2003). Mutational analysis of BARD1 in familial breast cancer patients in Japan. *Cancer letters* 200, 1-7.

## J

Jacobsen, O., and Fraser, R. (1946). *Heredity in Breast Cancer: A Genetic and Clinical Study of Two Hundred Probands* (Nyt Nordisk Forlag).

Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA: a cancer journal for clinicians* 61, 69-90.

---

Johannsson, O.T., Idvall, I., Anderson, C., Borg, A., Barkardottir, R.B., Egilsson, V., and Olsson, H. (1997). Tumour biological features of BRCA1-induced breast and ovarian cancer. *European journal of cancer* 33, 362-371.

Johnson, J., Healey, S., Khanna, K.K., kConFab, and Chenevix-Trench, G. (2011). Mutation analysis of RAD51L1 (RAD51B/REC2) in multiple-case, non-BRCA1/2 breast cancer families. *Breast cancer research and treatment* 129, 255-263.

Judkins, T., Rosenthal, E., Arnell, C., Burbidge, L.A., Geary, W., Barrus, T., Schoenberger, J., Trost, J., Wenstrup, R.J., and Roa, B.B. (2012). Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer* 118, 5210-5216.

## K

Kainu, T., Juo, S.H., Desper, R., Schaffer, A.A., Gillanders, E., Rozenblum, E., Freas-Lutz, D., Weaver, D., Stephan, D., Bailey-Wilson, J., *et al.* (2000). Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proceedings of the National Academy of Sciences of the United States of America* 97, 9603-9608.

Karp, S.E., Tonin, P.N., Begin, L.R., Martinez, J.J., Zhang, J.C., Pollak, M.N., and Foulkes, W.D. (1997). Influence of BRCA1 mutations on nuclear grade and estrogen receptor status of breast carcinoma in Ashkenazi Jewish women. *Cancer* 80, 435-441.

Kato, M., Yano, K., Matsuo, F., Saito, H., Katagiri, T., Kurumizaka, H., Yoshimoto, M., Kasumi, F., Akiyama, F., Sakamoto, G., *et al.* (2000). Identification of Rad51 alteration in patients with bilateral breast cancer. *Journal of human genetics* 45, 133-137.

Kee, Y., and D'Andrea, A.D. (2012). Molecular pathogenesis and clinical management of Fanconi anemia. *The Journal of clinical investigation* 122, 3799-3806.

Key, T.J., Verkasalo, P.K., and Banks, E. (2001). Epidemiology of breast cancer. *The lancet oncology* 2, 133-140.

Kim, H., Huang, J., and Chen, J. (2007). CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response. *Nature structural & molecular biology* *14*, 710-715.

Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the SLX4 gene in Fanconi anemia. *Nature genetics* *43*, 142-146.

Kleiman, F.E., Wu-Baer, F., Fonseca, D., Kaneko, S., Baer, R., and Manley, J.L. (2005). BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. *Genes & development* *19*, 1227-1237.

Korde, L.A., Mueller, C.M., Loud, J.T., Struewing, J.P., Nichols, K., Greene, M.H., and Mai, P.L. (2011). No evidence of excess breast cancer risk among mutation-negative women from BRCA mutation-positive families. *Breast cancer research and treatment* *125*, 169-173.

Korenberg, J.R., and Rykowski, M.C. (1988). Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. *Cell* *53*, 391-400.

Kushnir, A., Laitman, Y., Shimon, S.P., Berger, R., and Friedman, E. (2012). Germline mutations in RAD51C in Jewish high cancer risk families. *Breast cancer research and treatment* *136*, 869-874.

## L

Lakhani, S.R., Reis-Filho, J.S., Fulford, L., Penault-Llorca, F., van der Vijver, M., Parry, S., Bishop, T., Benitez, J., Rivas, C., Bignon, Y.J., *et al.* (2005). Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clinical cancer research : an official journal of the American Association for Cancer Research* *11*, 5175-5180.

Lakhani, S.R., Van De Vijver, M.J., Jacquemier, J., Anderson, T.J., Osin, P.P., McGuffog, L., and Easton, D.F. (2002). The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* *20*, 2310-2318.

Landwehr, R., Bogdanova, N.V., Antonenkova, N., Meyer, A., Bremer, M., Park-Simon, T.W., Hillemanns, P., Karstens, J.H., Schindler, D., and Dork, T. (2011). Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer. *Breast cancer research and treatment* 130, 1021-1028.

Levy-Lahad, E., and Friedman, E. (2007). Cancer risks among BRCA1 and BRCA2 mutation carriers. *British journal of cancer* 96, 11-15.

Lewis, A.G., Flanagan, J., Marsh, A., Pupo, G.M., Mann, G., Spurdle, A.B., Lindeman, G.J., Visvader, J.E., Brown, M.A., Chenevix-Trench, G., *et al.* (2005). Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. *Breast cancer research : BCR* 7, R1005-1016.

Li, F.P., and Fraumeni, J.F., Jr. (1969). Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. *Journal of the National Cancer Institute* 43, 1365-1373.

Lim, W., Olschwang, S., Keller, J.J., Westerman, A.M., Menko, F.H., Boardman, L.A., Scott, R.J., Trimbath, J., Giardiello, F.M., Gruber, S.B., *et al.* (2004). Relative frequency and morphology of cancers in STK11 mutation carriers. *Gastroenterology* 126, 1788-1794.

Lindor, N.M., Guidugli, L., Wang, X., Vallee, M.P., Monteiro, A.N., Tavtigian, S., Goldgar, D.E., and Couch, F.J. (2012). A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Human mutation* 33, 8-21.

Lorick, K.L., Jensen, J.P., Fang, S., Ong, A.M., Hatakeyama, S., and Weissman, A.M. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proceedings of the National Academy of Sciences of the United States of America* 96, 11364-11369.

Lose, F., Lovelock, P., Chenevix-Trench, G., Mann, G.J., Pupo, G.M., Spurdle, A.B., and Kathleen Cuninghame Foundation Consortium for Research into Familial Breast, C. (2006). Variation in the RAD51 gene and familial breast cancer. *Breast cancer research : BCR* 8, R26.

Loveday, C., Turnbull, C., Ramsay, E., Hughes, D., Ruark, E., Frankum, J.R., Bowden, G., Kalmyrzaev, B., Warren-Perry, M., Snape, K., *et al.* (2011). Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nature genetics* 43, 879-882.

Loveday, C., Turnbull, C., Ruark, E., Xicola, R.M., Ramsay, E., Hughes, D., Warren-Perry, M., Snape, K., Breast Cancer Susceptibility, C., Eccles, D., *et al.* (2012). Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nature genetics* 44, 475-476; author reply 476.

Lu, W., Wang, X., Lin, H., Lindor, N.M., and Couch, F.J. (2012). Mutation screening of RAD51C in high-risk breast and ovarian cancer families. *Familial cancer* 11, 381-385.

Lubinski, J., Phelan, C.M., Ghadirian, P., Lynch, H.T., Garber, J., Weber, B., Tung, N., Horsman, D., Isaacs, C., Monteiro, A.N., *et al.* (2004). Cancer variation associated with the position of the mutation in the BRCA2 gene. *Familial cancer* 3, 1-10.

Lynch, H., Wen, H., Kim, Y.C., Snyder, C., Kinarsky, Y., Chen, P.X., Xiao, F., Goldgar, D., Cowan, K.H., and Wang, S.M. (2013). Can unknown predisposition in familial breast cancer be family-specific? *The breast journal* 19, 520-528.

Lynch, H.T. (1967). *Hereditary Factors in Carcinoma* (Springer).

Lynch, H.T., and Krush, A.J. (1971). Carcinoma of the breast and ovary in three families. *Surgery, gynecology & obstetrics* 133, 644-648.

Lynch, H.T., Krush, A.J., and Guirgis, H. (1973). Genetic factors in families with combined gastrointestinal and breast cancer. *The American journal of gastroenterology* 59, 31-40.

Lynch, H.T., Krush, A.J., Lemon, H.M., Kaplan, A.R., Condit, P.T., and Bottomley, R.H. (1972). Tumor variation in families with breast cancer. *JAMA : the journal of the American Medical Association* 222, 1631-1635.

---

Lynch, H.T., and Lynch, J.F. (2002). Hereditary cancer: family history, diagnosis, molecular genetics, ecogenetics, and management strategies. *Biochimie* 84, 3-17.

## M

Ma, Y., Fan, S., Hu, C., Meng, Q., Fuqua, S.A., Pestell, R.G., Tomita, Y.A., and Rosen, E.M. (2010). BRCA1 regulates acetylation and ubiquitination of estrogen receptor-alpha. *Molecular endocrinology* 24, 76-90.

Mack, T.M., Hamilton, A.S., Press, M.F., Diep, A., and Rappaport, E.B. (2002). Heritable breast cancer in twins. *British journal of cancer* 87, 294-300.

Macklin, M.T. (1959). Comparison of the number of breast-cancer deaths observed in relatives of breast-cancer patients, and the number expected on the basis of mortality rates. *Journal of the National Cancer Institute* 22, 927-951.

Machackova, E., Foretova, L., Lukesova, M., Vasickova, P., Navratilova, M., Coene, I., Pavlu, H., Kosinova, V., Kuklova, J., and Claes, K. (2008). Spectrum and characterisation of BRCA1 and BRCA2 deleterious mutations in high-risk Czech patients with breast and/or ovarian cancer. *BMC cancer* 8, 140.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Jr., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z., Tainsky, M.A., *et al.* (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science (New York, N.Y.)* 250, 1233-1238.

Malvezzi, M., Bertuccio, P., Levi, F., La Vecchia, C., and Negri, E. (2012). European cancer mortality predictions for the year 2012. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 23, 1044-1052.

Martin, S.T., Matsubayashi, H., Rogers, C.D., Philips, J., Couch, F.J., Brune, K., Yeo, C.J., Kern, S.E., Hruban, R.H., and Goggins, M. (2005). Increased prevalence of the BRCA2 polymorphic

stop codon K3326X among individuals with familial pancreatic cancer. *Oncogene* 24, 3652-3656.

Martinez-Ferrandis, J.I., Vega, A., Chirivella, I., Marin-Garcia, P., Insa, A., Lluch, A., Carracedo, A., Chaves, F.J., Garcia-Conde, J., Cervantes, A., *et al.* (2003). Mutational analysis of BRCA1 and BRCA2 in Mediterranean Spanish women with early-onset breast cancer: identification of three novel pathogenic mutations. *Human mutation* 22, 417-418.

Mavaddat, N., Pharoah, P.D., Blows, F., Driver, K.E., Provenzano, E., Thompson, D., Macinnis, R.J., Shah, M., Easton, D.F., and Antoniou, A.C. (2010). Familial relative risks for breast cancer by pathological subtype: a population-based cohort study. *Breast cancer research : BCR* 12, R10.

Maxwell, C.A., Benitez, J., Gomez-Baldo, L., Osorio, A., Bonifaci, N., Fernandez-Ramires, R., Costes, S.V., Guino, E., Chen, H., Evans, G.J., *et al.* (2011). Interplay between BRCA1 and RHAMM regulates epithelial apicobasal polarization and may influence risk of breast cancer. *PLoS biology* 9, e1001199.

Mazoyer, S., Dunning, A.M., Serova, O., Dearden, J., Puget, N., Healey, C.S., Gayther, S.A., Mangion, J., Stratton, M.R., Lynch, H.T., *et al.* (1996). A polymorphic stop codon in BRCA2. *Nature genetics* 14, 253-254.

Menendez, M., Castellsague, J., Mirete, M., Pros, E., Feliubadalo, L., Osorio, A., Calaf, M., Tornero, E., del Valle, J., Fernandez-Rodriguez, J., *et al.* (2012). Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 132, 979-992.

Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R.L., Schmidt, M.K., Chang-Claude, J., Bojesen, S.E., Bolla, M.K., *et al.* (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nature genetics* 45, 353-361.

Miki, Y., Katagiri, T., Kasumi, F., Yoshimoto, T., and Nakamura, Y. (1996). Mutation analysis in the BRCA2 gene in primary breast cancers. *Nature genetics* 13, 245-247.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W., *et al.* (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (New York, N.Y.)* 266, 66-71.

Milne, R.L., Osorio, A., Cajal, T.R., Vega, A., Llorca, G., de la Hoya, M., Diez, O., Alonso, M.C., Lazaro, C., Blanco, I., *et al.* (2008). The average cumulative risks of breast and ovarian cancer for carriers of mutations in BRCA1 and BRCA2 attending genetic counseling units in Spain. *Clinical cancer research : an official journal of the American Association for Cancer Research* 14, 2861-2869.

Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M., and Kouzarides, T. (1997). Transcriptional activation functions in BRCA2. *Nature* 386, 772-773.

Miramar, M.D., Calvo, M.T., Rodriguez, A., Anton, A., Lorente, F., Barrio, E., Herrero, A., Burriel, J., and Garcia de Jalon, A. (2008). Genetic analysis of BRCA1 and BRCA2 in breast/ovarian cancer families from Aragon (Spain): two novel truncating mutations and a large genomic deletion in BRCA1. *Breast cancer research and treatment* 112, 353-358.

Misra, S., Sharma, S., Agarwal, A., Khedkar, S.V., Tripathi, M.K., Mittal, M.K., and Chaudhuri, G. (2010). Cell cycle-dependent regulation of the bi-directional overlapping promoter of human BRCA2/ZAR2 genes in breast cancer cells. *Molecular cancer* 9, 50.

Mohammadi, L., Vreeswijk, M.P., Oldenburg, R., van den Ouweland, A., Oosterwijk, J.C., van der Hout, A.H., Hoogerbrugge, N., Ligtenberg, M., Ausems, M.G., van der Luijt, R.B., *et al.* (2009). A simple method for co-segregation analysis to evaluate the pathogenicity of unclassified variants; BRCA1 and BRCA2 as an example. *BMC cancer* 9, 211.

Mucaki, E.J., Ainsworth, P., and Rogan, P.K. (2011). Comprehensive prediction of mRNA splicing effects of BRCA1 and BRCA2 variants. *Human mutation* 32, 735-742.

Munoz, I.M., Hain, K., Declais, A.C., Gardiner, M., Toh, G.W., Sanchez-Pulido, L., Heuckmann, J.M., Toth, R., Macartney, T., Eppink, B., *et al.* (2009). Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Molecular cell* 35, 116-127.

## N

Nelen, M.R., Padberg, G.W., Peeters, E.A., Lin, A.Y., van den Helm, B., Frants, R.R., Coulon, V., Goldstein, A.M., van Reen, M.M., Easton, D.F., *et al.* (1996). Localization of the gene for Cowden disease to chromosome 10q22-23. *Nature genetics* 13, 114-116.

Newman, B., Austin, M.A., Lee, M., and King, M.C. (1988). Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proceedings of the National Academy of Sciences of the United States of America* 85, 3044-3048.

## O

Oldenburg, R.A., Kroeze-Jansema, K.H., Houwing-Duistermaat, J.J., Bayley, J.P., Dambrot, C., van Asperen, C.J., van den Ouweland, A.M., Bakker, B., van Beers, E.H., Nederlof, P.M., *et al.* (2008). Genome-wide linkage scan in Dutch hereditary non-BRCA1/2 breast cancer families identifies 9q21-22 as a putative breast cancer susceptibility locus. *Genes, chromosomes & cancer* 47, 947-956.

Olsen, J.H., Hahnemann, J.M., Borresen-Dale, A.L., Tretli, S., Kleinerman, R., Sankila, R., Hammarstrom, L., Robsahm, T.E., Kaariainen, H., Bregard, A., *et al.* (2005). Breast and other cancers in 1445 blood relatives of 75 Nordic patients with ataxia telangiectasia. *British journal of cancer* 93, 260-265.

Osher, D.J., De Leeneer, K., Michils, G., Hamel, N., Tomiak, E., Poppe, B., Leunen, K., Legius, E., Shuen, A., Smith, E., *et al.* (2012). Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. *British journal of cancer* 106, 1460-1463.

Osorio, A., Endt, D., Fernandez, F., Eirich, K., de la Hoya, M., Schmutzler, R., Caldes, T., Meindl, A., Schindler, D., and Benitez, J. (2012). Predominance of pathogenic missense variants in the RAD51C gene occurring in breast and ovarian cancer families. *Human molecular genetics* 21, 2889-2898.

---

**P**

Palacios, J., Honrado, E., Osorio, A., Cazorla, A., Sarrio, D., Barroso, A., Rodriguez, S., Cigudosa, J.C., Diez, O., Alonso, C., *et al.* (2005). Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast cancer research and treatment* *90*, 5-14.

Palanca, S., de Juan, I., Perez-Simo, G., Barragan, E., Chirivella, I., Martinez, E., Fuster, O., and Bolufer, P. (2013). The deletion of exons 3-5 of BRCA1 is the first founder rearrangement identified in breast and/or ovarian cancer Spanish families. *Familial cancer* *12*, 119-123.

Palanca Suela, S., Esteban Cardenosa, E., Barragan Gonzalez, E., Oltra Soler, S., de Juan Jimenez, I., Chirivella Gonzalez, I., Segura Huerta, A., Guillen Ponce, C., Martinez de Duenas, E., and Bolufer Gilabert, P. (2008). Identification of a novel BRCA1 large genomic rearrangement in a Spanish breast/ovarian cancer family. *Breast cancer research and treatment* *112*, 63-67.

Panchal, S., Bordeleau, L., Poll, A., Llacuachaqui, M., Shachar, O., Ainsworth, P., Armel, S., Eisen, A., Sun, P., and Narod, S.A. (2010). Does family history predict the age at onset of new breast cancers in BRCA1 and BRCA2 mutation-positive families? *Clinical genetics* *77*, 273-279.

Pang, Z., Yao, L., Zhang, J., Ouyang, T., Li, J., Wang, T., Fan, Z., Fan, T., Lin, B., and Xie, Y. (2011). RAD51C germline mutations in Chinese women with familial breast cancer. *Breast cancer research and treatment* *129*, 1019-1020.

Park, D.J., Lesueur, F., Nguyen-Dumont, T., Pertesi, M., Odefrey, F., Hammet, F., Neuhausen, S.L., John, E.M., Andrulis, I.L., Terry, M.B., *et al.* (2012). Rare mutations in XRCC2 increase the risk of breast cancer. *American journal of human genetics* *90*, 734-739.

Peixoto, A., Santos, C., Rocha, P., Pinheiro, M., Principe, S., Pereira, D., Rodrigues, H., Castro, F., Abreu, J., Gusmao, L., *et al.* (2009). The c.156\_157insAlu BRCA2 rearrangement accounts

for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. *Breast cancer research and treatment* 114, 31-38.

Pelttari, L.M., Heikkinen, T., Thompson, D., Kallioniemi, A., Schleutker, J., Holli, K., Blomqvist, C., Aittomaki, K., Butzow, R., and Nevanlinna, H. (2011). RAD51C is a susceptibility gene for ovarian cancer. *Human molecular genetics* 20, 3278-3288.

Pelttari, L.M., Kiiski, J., Nurminen, R., Kallioniemi, A., Schleutker, J., Gylfe, A., Aaltonen, L.A., Leminen, A., Heikkila, P., Blomqvist, C., *et al.* (2012). A Finnish founder mutation in RAD51D: analysis in breast, ovarian, prostate, and colorectal cancer. *Journal of medical genetics* 49, 429-432.

Pellegrini, L., Yu, D.S., Lo, T., Anand, S., Lee, M., Blundell, T.L., and Venkitaraman, A.R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 420, 287-293.

Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, D.F., Evans, C., Deacon, J., and Stratton, M.R. (1999). Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *Journal of the National Cancer Institute* 91, 943-949.

Peto, J., and Mack, T.M. (2000). High constant incidence in twins and other relatives of women with breast cancer. *Nature genetics* 26, 411-414.

Pharoah, P.D., Day, N.E., Duffy, S., Easton, D.F., and Ponder, B.A. (1997). Family history and the risk of breast cancer: a systematic review and meta-analysis. *International journal of cancer. Journal international du cancer* 71, 800-809.

Pharoah, P.D., Guilford, P., Caldas, C., and International Gastric Cancer Linkage, C. (2001). Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology* 121, 1348-1353.

Phelan, C.M., Kwan, E., Jack, E., Li, S., Morgan, C., Aube, J., Hanna, D., and Narod, S.A. (2002). A low frequency of non-founder BRCA1 mutations in Ashkenazi Jewish breast-ovarian cancer families. *Human mutation* 20, 352-357.

Phillips, K.A. (2000). Immunophenotypic and pathologic differences between BRCA1 and BRCA2 hereditary breast cancers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 18, 107S-112S.

Plon, S.E., Eccles, D.M., Easton, D., Foulkes, W.D., Genuardi, M., Greenblatt, M.S., Hogervorst, F.B., Hoogerbrugge, N., Spurdle, A.B., and Tavtigian, S.V. (2008). Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human mutation* 29, 1282-1291.

Prokofyeva, D., Bogdanova, N., Dubrowinskaja, N., Bermisheva, M., Takhirova, Z., Antonenkova, N., Turmanov, N., Datsyuk, I., Gantsev, S., Christiansen, H., *et al.* (2013). Nonsense mutation p.Q548X in BLM, the gene mutated in Bloom's syndrome, is associated with breast cancer in Slavic populations. *Breast cancer research and treatment* 137, 533-539.

Pros, E., Gomez, C., Martin, T., Fabregas, P., Serra, E., and Lazaro, C. (2008). Nature and mRNA effect of 282 different NF1 point mutations: focus on splicing alterations. *Human mutation* 29, E173-193.

## R

Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T., *et al.* (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature genetics* 39, 165-167.

Rahman, N., Teare, M.D., Seal, S., Renard, H., Mangion, J., Cour, C., Thompson, D., Shugart, Y., Eccles, D., Devilee, P., *et al.* (2000). Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene* 19, 4170-4173.

Rajendra, E., and Venkitaraman, A.R. (2010). Two modules in the BRC repeats of BRCA2 mediate structural and functional interactions with the RAD51 recombinase. *Nucleic acids research* 38, 82-96.

Rapakko, K., Heikkinen, K., Karppinen, S.M., and Winqvist, R. (2006). Screening for RAD51 and BRCA2 BRC repeat mutations in breast and ovarian cancer families. *Cancer letters* 236, 142-147.

Reeves, G.K., Pirie, K., Beral, V., Green, J., Spencer, E., Bull, D., and Million Women Study, C. (2007). Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ* 335, 1134.

Reid, S., Schindler, D., Hanenberg, H., Barker, K., Hanks, S., Kalb, R., Neveling, K., Kelly, P., Seal, S., Freund, M., *et al.* (2007). Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nature genetics* 39, 162-164.

Renwick, A., Thompson, D., Seal, S., Kelly, P., Chagtai, T., Ahmed, M., North, B., Jayatilake, H., Barfoot, R., Spanova, K., *et al.* (2006). ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature genetics* 38, 873-875.

Roa, B.B., Boyd, A.A., Volcik, K., and Richards, C.S. (1996). Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nature genetics* 14, 185-187.

Robson, M.E., Storm, C.D., Weitzel, J., Wollins, D.S., Offit, K., and American Society of Clinical, O. (2010). American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 28, 893-901.

Rodriguez-Lopez, R., Osorio, A., Ribas, G., Pollan, M., Sanchez-Pulido, L., de la Hoya, M., Ruibal, A., Zamora, P., Arias, J.I., Salazar, R., *et al.* (2004). The variant E233G of the RAD51D gene could be a low-penetrance allele in high-risk breast cancer families without BRCA1/2 mutations. *International journal of cancer. Journal international du cancer* 110, 845-849.

Rodriguez-Lopez, R., Osorio, A., Sanchez-Pulido, L., De La Hoya, M., Barroso, A., Caldes, T., and Benitez, J. (2003). No mutations in the XRCC2 gene in BRCA1/2-negative high-risk breast cancer families. *International journal of cancer. Journal international du cancer* 103, 136-137.

Rodriguez, J.A., and Henderson, B.R. (2000). Identification of a functional nuclear export sequence in BRCA1. *The Journal of biological chemistry* 275, 38589-38596.

Romano, M., Buratti, E., and Baralle, D. (2013). Role of Pseudoexons and Pseudointrons in Human Cancer. *International journal of cell biology* 2013, 810572.

Romero, A., Perez-Segura, P., Tosar, A., Garcia-Saenz, J.A., Diaz-Rubio, E., Caldes, T., and de la Hoya, M. (2011). A HRM-based screening method detects RAD51C germ-line deleterious mutations in Spanish breast and ovarian cancer families. *Breast cancer research and treatment* 129, 939-946.

Rosa-Rosa, J.M., Pita, G., Urioste, M., Llort, G., Brunet, J., Lazaro, C., Blanco, I., Ramon y Cajal, T., Diez, O., de la Hoya, M., *et al.* (2009). Genome-wide linkage scan reveals three putative breast-cancer-susceptibility loci. *American journal of human genetics* 84, 115-122.

Rosenberg, P.S., Alter, B.P., and Ebell, W. (2008). Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica* 93, 511-517.

Rosenberg, P.S., Greene, M.H., and Alter, B.P. (2003). Cancer incidence in persons with Fanconi anemia. *Blood* 101, 822-826.

Roy, R., Chun, J., and Powell, S.N. (2012). BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nature reviews. Cancer* 12, 68-78.

**S**

Santarosa, M., and Maestro, R. (2012). BRACKing news on triple-negative/basal-like breast cancers: how BRCA1 deficiency may result in the development of a selective tumor subtype. *Cancer metastasis reviews* 31, 131-142.

Sanz, D.J., Acedo, A., Infante, M., Duran, M., Perez-Cabornero, L., Esteban-Cardenosa, E., Lastra, E., Pagani, F., Miner, C., and Velasco, E.A. (2010). A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* 16, 1957-1967.

Sasieni, P.D., Shelton, J., Ormiston-Smith, N., Thomson, C.S., and Silcocks, P.B. (2011). What is the lifetime risk of developing cancer?: the effect of adjusting for multiple primaries. *British journal of cancer* 105, 460-465.

Satagopan, J.M., Offit, K., Foulkes, W., Robson, M.E., Wacholder, S., Eng, C.M., Karp, S.E., and Begg, C.B. (2001). The lifetime risks of breast cancer in Ashkenazi Jewish carriers of BRCA1 and BRCA2 mutations. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 10, 467-473.

Sauer, M.K., and Andrulis, I.L. (2005). Identification and characterization of missense alterations in the BRCA1 associated RING domain (BARD1) gene in breast and ovarian cancer. *Journal of medical genetics* 42, 633-638.

Saunus, J.M., Edwards, S.L., French, J.D., Smart, C.E., and Brown, M.A. (2007). Regulation of BRCA1 messenger RNA stability in human epithelial cell lines and during cell cycle progression. *FEBS letters* 581, 3435-3442.

Scheuner, M.T., McNeel, T.S., and Freedman, A.N. (2010). Population prevalence of familial cancer and common hereditary cancer syndromes. *The 2005 California Health Interview*

Survey. *Genetics in medicine : official journal of the American College of Medical Genetics* 12, 726-735.

Schuster, B., Knies, K., Stoepker, C., Velleuer, E., Friedl, R., Gottwald-Muhlhauser, B., de Winter, J.P., and Schindler, D. (2013). Whole Exome Sequencing Reveals Uncommon Mutations in the Recently Identified Fanconi Anemia Gene SLX4/FANCP. *Human mutation* 34, 93-96.

Seal, S., Barfoot, R., Jayatilake, H., Smith, P., Renwick, A., Bascombe, L., McGuffog, L., Evans, D.G., Eccles, D., Easton, D.F., *et al.* (2003). Evaluation of Fanconi Anemia genes in familial breast cancer predisposition. *Cancer research* 63, 8596-8599.

Seal, S., Thompson, D., Renwick, A., Elliott, A., Kelly, P., Barfoot, R., Chagtai, T., Jayatilake, H., Ahmed, M., Spanova, K., *et al.* (2006). Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nature genetics* 38, 1239-1241.

Seitz, S., Rohde, K., Bender, E., Nothnagel, A., Kolble, K., Schlag, P.M., and Scherneck, S. (1997). Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. *Oncogene* 14, 741-743.

Shah, S., Kim, Y., Ostrovnaya, I., Murali, R., Schrader, K.A., Lach, F.P., Sarrel, K., Rau-Murthy, R., Hansen, N., Zhang, L., *et al.* (2013). Assessment of Mutations in Hereditary Breast Cancers. *PloS one* 8, e66961.

Shamseldin, H.E., Elfaki, M., and Alkuraya, F.S. (2012). Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. *Journal of medical genetics* 49, 184-186.

Shin, S., and Verma, I.M. (2003). BRCA2 cooperates with histone acetyltransferases in androgen receptor-mediated transcription. *Proceedings of the National Academy of Sciences of the United States of America* 100, 7201-7206.

Silvestri, V., Rizzolo, P., Falchetti, M., Zanna, I., Masala, G., Palli, D., and Ottini, L. (2011). Mutation screening of RAD51C in male breast cancer patients. *Breast cancer research : BCR* 13, 404.

Sinilnikova, O.M., Mazoyer, S., Bonnardel, C., Lynch, H.T., Narod, S.A., and Lenoir, G.M. (2006). BRCA1 and BRCA2 mutations in breast and ovarian cancer syndrome: reflection on the Creighton University historical series of high risk families. *Familial cancer* 5, 15-20.

Sluiter, M.D., and van Rensburg, E.J. (2011). Large genomic rearrangements of the BRCA1 and BRCA2 genes: review of the literature and report of a novel BRCA1 mutation. *Breast cancer research and treatment* 125, 325-349.

Smith, P., McGuffog, L., Easton, D.F., Mann, G.J., Pupo, G.M., Newman, B., Chenevix-Trench, G., Szabo, C., Southey, M., Renard, H., *et al.* (2006). A genome wide linkage search for breast cancer susceptibility genes. *Genes, chromosomes & cancer* 45, 646-655.

Snape, K., Ruark, E., Tarpey, P., Renwick, A., Turnbull, C., Seal, S., Murray, A., Hanks, S., Douglas, J., Stratton, M.R., *et al.* (2012). Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast cancer research and treatment* 134, 429-433.

Sokolenko, A.P., Iyevleva, A.G., Preobrazhenskaya, E.V., Mitiushkina, N.V., Alysheva, S.N., Suspitsin, E.N., Kuligina, E., Gorodnova, T.V., Pfeifer, W., Togo, A.V., *et al.* (2012). High prevalence and breast cancer predisposing role of the BLM c.1642 C>T (Q548X) mutation in Russia. *International journal of cancer. Journal international du cancer* 130, 2867-2873.

Solyom, S., Winqvist, R., Nikkila, J., Rapakko, K., Hirvikoski, P., Kokkonen, H., and Pylkas, K. (2011). Screening for large genomic rearrangements in the FANCA gene reveals extensive deletion in a Finnish breast cancer family. *Cancer letters* 302, 113-118.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., *et al.* (2003). Repeated observation of breast tumor subtypes in

independent gene expression data sets. *Proceedings of the National Academy of Sciences of the United States of America* *100*, 8418-8423.

Spurdle, A.B., Couch, F.J., Hogervorst, F.B., Radice, P., and Sinilnikova, O.M. (2008a). Prediction and assessment of splicing alterations: implications for clinical testing. *Human mutation* *29*, 1304-1313.

Spurdle, A.B., Healey, S., Devereau, A., Hogervorst, F.B., Monteiro, A.N., Nathanson, K.L., Radice, P., Stoppa-Lyonnet, D., Tavtigian, S., Wappenschmidt, B., *et al.* (2012). ENIGMA--evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Human mutation* *33*, 2-7.

Spurdle, A.B., Lakhani, S.R., Healey, S., Parry, S., Da Silva, L.M., Brinkworth, R., Hopper, J.L., Brown, M.A., Babikyan, D., Chenevix-Trench, G., *et al.* (2008b). Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis--a report from the kConFab Investigators. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* *26*, 1657-1663.

Steffen, J., Nowakowska, D., Niwinska, A., Czapczak, D., Kluska, A., Piatkowska, M., Wisniewska, A., and Paszko, Z. (2006). Germline mutations 657del5 of the NBS1 gene contribute significantly to the incidence of breast cancer in Central Poland. *International journal of cancer. Journal international du cancer* *119*, 472-475.

Stephens, F.E., Gardner, E.J., and Woolf, C.M. (1958). A recheck of kindred 107, which has shown a high frequency of breast cancer. *Cancer* *11*, 967-972.

Stoepker, C., Hain, K., Schuster, B., Hilhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W., *et al.* (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nature genetics* *43*, 138-141.

Stormorken, A.T., Hoff, G., Norstein, J., Bowitz-Lothe, I.M., Hanslien, E., Grindedal, E., and Moller, P. (2006). Estimated prevalence of hereditary cancers and the need for surveillance in a Norwegian county, Telemark. *Scandinavian journal of gastroenterology* 41, 71-79.

Struewing, J.P., Hartge, P., Wacholder, S., Baker, S.M., Berlin, M., McAdams, M., Timmerman, M.M., Brody, L.C., and Tucker, M.A. (1997). The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *The New England journal of medicine* 336, 1401-1408.

Stuckey, A. (2011). Breast cancer: epidemiology and risk factors. *Clinical obstetrics and gynecology* 54, 96-102.

Suen, T.C., Tang, M.S., and Goss, P.E. (2005). Model of transcriptional regulation of the BRCA1-NBR2 bi-directional transcriptional unit. *Biochimica et biophysica acta* 1728, 126-134.

Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* 138, 63-77.

Swift, M., Reitnauer, P.J., Morrell, D., and Chase, C.L. (1987). Breast and other cancers in families with ataxia-telangiectasia. *The New England journal of medicine* 316, 1289-1294.

Szabo, C., Masiello, A., Ryan, J.F., and Brody, L.C. (2000). The breast cancer information core: database design, structure, and scope. *Human mutation* 16, 123-131.

## T

Tamary, H., Nishri, D., Yacobovich, J., Zilber, R., Dgany, O., Krasnov, T., Aviner, S., Stepensky, P., Ravel-Vilk, S., Bitan, M., *et al.* (2010). Frequency and natural history of inherited bone marrow failure syndromes: the Israeli Inherited Bone Marrow Failure Registry. *Haematologica* 95, 1300-1307.

Tarabeux, J., Zeitouni, B., Moncoutier, V., Tenreiro, H., Abidallah, K., Lair, S., Legoix-Ne, P., Leroy, Q., Rouleau, E., Golmard, L., *et al.* (2013). Streamlined ion torrent PGM-based diagnostics: BRCA1 and BRCA2 genes as a model. *European journal of human genetics : EJHG*.

Tavtigian, S.V., Byrnes, G.B., Goldgar, D.E., and Thomas, A. (2008a). Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. *Human mutation* 29, 1342-1354.

Tavtigian, S.V., Greenblatt, M.S., Lesueur, F., and Byrnes, G.B. (2008b). In silico analysis of missense substitutions using sequence-alignment based methods. *Human mutation* 29, 1327-1336.

Tesoriero, A.A., Wong, E.M., Jenkins, M.A., Hopper, J.L., Brown, M.A., Chenevix-Trench, G., Spurdle, A.B., and Southey, M.C. (2005). Molecular characterization and cancer risk associated with BRCA1 and BRCA2 splice site variants identified in multiple-case breast cancer families. *Human mutation* 26, 495.

Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., and De Greve, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Human mutation* 26, 284.

Thompson, D., and Easton, D. (2001). Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *American journal of human genetics* 68, 410-419.

Thompson, D., Szabo, C.I., Mangion, J., Oldenburg, R.A., Odefrey, F., Seal, S., Barfoot, R., Kroeze-Jansema, K., Teare, D., Rahman, N., *et al.* (2002). Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proceedings of the National Academy of Sciences of the United States of America* 99, 827-831.

Thompson, E.R., Boyle, S.E., Johnson, J., Ryland, G.L., Sawyer, S., Choong, D.Y., kConFab, Chenevix-Trench, G., Trainer, A.H., Lindeman, G.J., *et al.* (2012a). Analysis of RAD51C

germline mutations in high-risk breast and ovarian cancer families and ovarian cancer patients. *Human mutation* 33, 95-99.

Thompson, E.R., Doyle, M.A., Ryland, G.L., Rowley, S.M., Choong, D.Y., Tothill, R.W., Thorne, H., kConFab, Barnes, D.R., Li, J., *et al.* (2012b). Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS genetics* 8, e1002894.

Thompson, E.R., Rowley, S.M., Sawyer, S., kConFab, Eccles, D.M., Trainer, A.H., Mitchell, G., James, P.A., and Campbell, I.G. (2013). Analysis of RAD51D in ovarian cancer patients and families with a history of ovarian or breast cancer. *PloS one* 8, e54772.

Thorlacius, S., Struewing, J.P., Hartge, P., Olafsdottir, G.H., Sigvaldason, H., Tryggvadottir, L., Wacholder, S., Tulinius, H., and Eyfjord, J.E. (1998). Population-based study of risk of breast cancer in carriers of BRCA2 mutation. *Lancet* 352, 1337-1339.

Thorslund, T., Esashi, F., and West, S.C. (2007). Interactions between human BRCA2 protein and the meiosis-specific recombinase DMC1. *The EMBO journal* 26, 2915-2922.

Tommiska, J., Seal, S., Renwick, A., Barfoot, R., Baskcomb, L., Jayatilake, H., Bartkova, J., Tallila, J., Kaare, M., Tamminen, A., *et al.* (2006). Evaluation of RAD50 in familial breast cancer predisposition. *International journal of cancer. Journal international du cancer* 118, 2911-2916.

## V

Vallee, M.P., Francy, T.C., Judkins, M.K., Babikyan, D., Lesueur, F., Gammon, A., Goldgar, D.E., Couch, F.J., and Tavtigian, S.V. (2012). Classification of missense substitutions in the BRCA genes: a database dedicated to Ex-UVs. *Human mutation* 33, 22-28.

Van Der Looij, M., Szabo, C., Besznyak, I., Liszka, G., Csokay, B., Pulay, T., Toth, J., Devilee, P., King, M.C., and Olah, E. (2000). Prevalence of founder BRCA1 and BRCA2 mutations among

breast and ovarian cancer patients in Hungary. *International journal of cancer. Journal international du cancer* 86, 737-740.

Vanags, A., Strumfa, I., Gardovskis, A., Borosenko, V., Abolins, A., Teibe, U., Trofimovics, G., Miklasevics, E., and Gardovskis, J. (2010). Population screening for hereditary and familial cancer syndromes in Valka district of Latvia. *Hereditary cancer in clinical practice* 8, 8.

Vaz, F., Hanenberg, H., Schuster, B., Barker, K., Wiek, C., Erven, V., Neveling, K., Endt, D., Kesterton, I., Autore, F., *et al.* (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. *Nature genetics* 42, 406-409.

Vega, A., Campos, B., Bressac-De-Paillerets, B., Bond, P.M., Janin, N., Douglas, F.S., Domenech, M., Baena, M., Pericay, C., Alonso, C., *et al.* (2001). The R71G BRCA1 is a founder Spanish mutation and leads to aberrant splicing of the transcript. *Human mutation* 17, 520-521.

Vreeswijk, M.P., Kraan, J.N., van der Klift, H.M., Vink, G.R., Cornelisse, C.J., Wijnen, J.T., Bakker, E., van Asperen, C.J., and Devilee, P. (2009). Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. *Human mutation* 30, 107-114.

Vuorela, M., Pylkas, K., Hartikainen, J.M., Sundfeldt, K., Lindblom, A., von Wachenfeldt Wappling, A., Haanpaa, M., Puistola, U., Rosengren, A., Anttila, M., *et al.* (2011). Further evidence for the contribution of the RAD51C gene in hereditary breast and ovarian cancer susceptibility. *Breast cancer research and treatment* 130, 1003-1010.

## **W**

Walker, L.C., Whiley, P.J., Houdayer, C., Hansen, T.V., Vega, A., Santamarina, M., Blanco, A., Fachal, L., Southey, M.C., Lafferty, A., *et al.* (2013). Evaluation of a 5-Tier Scheme Proposed for Classification of Sequence Variants Using Bioinformatic and Splicing Assay Data: Inter-Reviewer Variability and Promotion of Minimum Reporting Guidelines. *Human mutation*.

Walsh, T., Casadei, S., Lee, M.K., Pennil, C.C., Nord, A.S., Thornton, A.M., Roeb, W., Agnew, K.J., Stray, S.M., Wickramanayake, A., *et al.* (2011). Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 18032-18037.

Walsh, T., Lee, M.K., Casadei, S., Thornton, A.M., Stray, S.M., Pennil, C., Nord, A.S., Mandell, J.B., Swisher, E.M., and King, M.C. (2010). Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. *Proceedings of the National Academy of Sciences of the United States of America* *107*, 12629-12633.

Wang, J., Zhang, J., Li, K., Zhao, W., and Cui, Q. (2012). SpliceDisease database: linking RNA splicing and disease. *Nucleic acids research* *40*, D1055-1059.

Wang, S.C., Lin, S.H., Su, L.K., and Hung, M.C. (1997). Changes in BRCA2 expression during progression of the cell cycle. *Biochemical and biophysical research communications* *234*, 247-251.

Wappenschmidt, B., Becker, A.A., Hauke, J., Weber, U., Engert, S., Kohler, J., Kast, K., Arnold, N., Rhiem, K., Hahnen, E., *et al.* (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*, e50800.

Warner, E., Foulkes, W., Goodwin, P., Meschino, W., Blondal, J., Paterson, C., Ozelik, H., Goss, P., Allingham-Hawkins, D., Hamel, N., *et al.* (1999). Prevalence and penetrance of BRCA1 and BRCA2 gene mutations in unselected Ashkenazi Jewish women with breast cancer. *Journal of the National Cancer Institute* *91*, 1241-1247.

Welsh, P.L., and King, M.C. (2001). BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Human molecular genetics* *10*, 705-713.

WHO (2011). Cause-specific mortality: regional estimates for 2008.

WHO (2012). Estadísticas sanitarias mundiales 2012.

Wickramanyake, A., Bernier, G., Pennil, C., Casadei, S., Agnew, K.J., Stray, S.M., Mandell, J., Garcia, R.L., Walsh, T., King, M.C., *et al.* (2012). Loss of function germline mutations in RAD51D in women with ovarian carcinoma. *Gynecologic oncology* 127, 552-555.

Williams, W.R., and Anderson, D.E. (1984). Genetic epidemiology of breast cancer: segregation analysis of 200 Danish pedigrees. *Genetic epidemiology* 1, 7-20.

Wong, J.M., Ionescu, D., and Ingles, C.J. (2003). Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. *Oncogene* 22, 28-33.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., and Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378, 789-792.

Wooster, R., Neuhausen, S.L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., *et al.* (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science (New York, N.Y.)* 265, 2088-2090.

## X

Xia, B., Sheng, Q., Nakanishi, K., Ohashi, A., Wu, J., Christ, N., Liu, X., Jasin, M., Couch, F.J., and Livingston, D.M. (2006). Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Molecular cell* 22, 719-729.

## Y

Yamane, K., Katayama, E., and Tsuruo, T. (2000). The BRCT regions of tumor suppressor BRCA1 and of XRCC1 show DNA end binding activity with a multimerizing feature. *Biochemical and biophysical research communications* 279, 678-684.

Yang, H., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H., and Pavletich, N.P. (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science (New York, N.Y.)* 297, 1837-1848.

Yarden, R.I., and Brody, L.C. (1999). BRCA1 interacts with components of the histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America* 96, 4983-4988.

You, Z., and Bailis, J.M. (2010). DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. *Trends in cell biology* 20, 402-409.

Yu, X., and Chen, J. (2004). DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Molecular and cellular biology* 24, 9478-9486.

Yu, X., Chini, C.C., He, M., Mer, G., and Chen, J. (2003). The BRCT domain is a phospho-protein binding domain. *Science (New York, N.Y.)* 302, 639-642.

## Z

Zheng, Y., Zhang, J., Hope, K., Niu, Q., Huo, D., and Olopade, O.I. (2010). Screening RAD51C nucleotide alterations in patients with a family history of breast and ovarian cancer. *Breast cancer research and treatment* 124, 857-861.

Zhu, Q., Pao, G.M., Huynh, A.M., Suh, H., Tonnu, N., Nederlof, P.M., Gage, F.H., and Verma, I.M. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* 477, 179-184.