



UNIVERSITAT DE
BARCELONA

**Síndrome de cáncer de mama y ovario hereditario:
Estudio *in vitro* de variantes *BRCA1* y *BRCA2* de
significado biológico desconocido y búsqueda de
nuevos genes responsables de este síndrome**

Francisco de Asís Quiles Vidal



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 3.0. Spain License.**



Facultat de Farmàcia

Departament de Bioquímica i Biologia Molecular

Programa de Doctorat de Biomedicina

**SÍNDROME DE CÁNCER DE MAMA Y OVARIO HEREDITARIO: ESTUDIO *IN VITRO*
DE VARIANTES *BRCA1* Y *BRCA2* DE SIGNIFICADO BIOLÓGICO DESCONOCIDO Y
BÚSQUEDA DE NUEVOS GENES RESPONSABLES DE ESTE SÍNDROME**

Francisco de Asís Quiles Vidal

2015

Memoria presentada por Francisco de Asís Quiles Vidal para optar al grado de
Doctor por la Universidad de Barcelona

Dra. Conxi Lázaro García
(Directora y tutora)

Francisco de Asís Quiles Vidal
(Doctorando)

A mis padres

A mi hermano

A mis allegados y amigos

AGRADECIMIENTOS

Hace cinco años, cuando estaba buscando que hacer con mi vida, una serie de eventos azarosos me permitieron conseguir una beca para realizar esta tesis doctoral. Gracias a ello, entré a formar parte de un gran equipo como el que hay el LRT2. De entre todas las cosas que podría destacar de este equipo me quedo con el gran compañerismo que hay en él. He aprendido muchas cosas tanto a nivel profesional como personal, a la vez que he conocido gente genial que más que compañeros de trabajo se han convertido en amigos.

En primer lugar tengo que agradecer por esta experiencia a **Conxi**, LA JEFA, por haberla hecho posible. Son muchas las cosas que tengo que agradecerle y me es difícil plasmarlo todo por escrito. Has sabido ser una jefa estricta cuando tocaba serlo, enseñándome a madurar como científico, y comprensiva y motivadora en los momentos más difíciles. Recuerdo aquel día en que, siendo yo aún un pipiolo, la cagué con una placa de secuencias y me fui a casa hundido... Juani (difícil de hablar de Conxi sin hablar de Juani) y tú me llamasteis en seguida para animarme... estos gestos te suben la moral a las nubes y hacen que te sientas respaldado! Valoro mucho que me hayas dejado hacer y hayas confiado en mí aún cuando no creías mucho en lo que estaba haciendo (tema ES-cells) y que me brindarás la oportunidad de ir a los USA a realizar un proyecto que me dejaste elegir a mí mismo. También agradezco muchísimo el esfuerzo final que has hecho para corregirme la tesis. En definitiva, muchas gracias por todo Conxi, sinceramente considero que esta experiencia no hubiera sido lo mismo sin una jefa como tú. Y, mi consejo para futuros becarios (si decides tener más después de aguantarnos a Joan y a mí...) es, hacer caso de los consejos de la JEFA! Ganareis tiempo y os ahorraréis en tropiezos!

El segundo pilar para mí en esta tesis has sido tú **Juani**. Mismas palabras de agradecimiento que he tenido para Conxi son aplicables para ti. Has sido mi guía en la poyata y en la toma de decisiones, has sabido descifrar mis pedidos disléxicos, me has defendido cuando todo el mundo pensaba que había hecho una "pacada" y me has parado los pies cuando me disponía a hacer alguna de ellas, hacer una tesis con alguien como tú a quién pedir ayuda lo hace todo mucho más fácil... las veces que te he levantado de la silla para preguntarte cosas y cada una de ellas te has levantado religiosamente a ayudarme. Me has escuchado cuando he estado "rallado" con algo, te has interesado en cada momento por como me iban las cosas. Además, considero que eres una persona genial y generosa y de la que he aprendido mucho a nivel personal, del trabajo en equipo y de tener conciencia de los demás. Estoy completamente seguro que el trabajo que he hecho en esta tesis y lo que he aprendido no hubiera sido posible sin tu ayuda. Conxi, no la dejes escapar por nada! Difícilmente encontraras una mano derecha en poyata como ella! ;-)

Gracias a todo el **Programa de Cáncer Hereditario**, como ya he dicho, formáis un equipo muy bueno, con un gran compañerismo. Se agradece el hecho de que puedas preguntar a cada uno de los jefes o técnicos sabiendo que te van a responder y ayudar de buena gana, algo así es digno de admirar en un equipo. Gracias a todos por ayudarme en mi trabajo, con vuestras correcciones en las publicaciones y por resolver cada una de las dudas que he tenido. Tengo que agradecer especialmente a **Mireia Menéndez** porque me has ayudado mucho con todo el tema del RNA y consejos a nivel de poyata y de escritura que han hecho que mi trabajo sea de mayor calidad. A **Nonia**, por resolver cada duda que he tenido, incluso las más absurdas... por divertirme con tus prisas, tus olvidos, tus razonamientos abstractos, tus super chistes. **Jesús** gracias a ti también por resolverme tantas dudas de una forma tan atenta y resolutiva, por tus consejos sobre la nieve y por tus datos precisos sobre cosas curiosas, jeje. **Eva Tornero** gracias por ayudarme a buscar cualquier variante, RNA, listas de datos, los programas *in silico*...y por ser tan divertida amenizando cada una de las comidas con tus historias. A! y gracias por adelantado por tu ayuda en el diseño de las diapos! jeje.

Gracias al resto de técnicos superiores **Mireia Morell, Dani y Xavi** y jefes **Laura Valle, Sara y Marta**, por estas ahí cuando he necesitado ayuda, por los consejos, las conversaciones en las comidas. Sois todos geniales!! Yo de mayor quiero ser como vos, jejeje. Por último, gracias a Gabi como cabeza del programa por brindarme la oportunidad de realizar esta tesis en este grupo.

Gracias al equipo de Diagnóstico, a Carol, **Raquel, Eva Montes, Ana, Susana, Anabel, Olga, Mónica y Gemma** por aguantar mis liadas! Cuantas veces habréis pensado en tirarme a la hoguera! jaja. En mi defensa tengo que decir que, como habéis podido observar desde que no ando por la poyata, no todas las “pacadas” de las que he sido culpado han sido culpa mía... pero bueno, entre asumir las culpas por 100 “pacadas” o 120 no importa! Así alguien sale impune gracias a mí y el karma me lo devolverá, jejeje. Fuera bromas, gracias por enseñarme muchas cosas a nivel de poyata, sobretodo orden, calidad de trabajo y respeto por el trabajo de los demás. Aunque todas me habéis enseñado algo en algún momento, Carol, con el RNA, Eva se hizo cargo de mí al principio con las PCR, Olga los micros...quiero darle especialmente las gracias a **Anabel** que estuvo ayudándome muchísimo en el proyecto de los exomas aún con un buen esquinco de tobillo y luego embarazada, estuvo ahí a tope hasta que conseguimos enviar las placas. Muchas gracias Anabel por ese esfuerzo, lo valoró mucho!

Gracias a **Natalia Casado** por todos los viajes a N₂ líquido con prisas y por el esfuerzo que hace en la sala de cultivos. También a Laura y Yolanda por todas las gestiones que han hecho y vuestro trabajo que facilita el nuestro.

Thank you to **Shyam Sharan** and his team (**Stacy, Kajal, Xia, Eswary, Susan, Lyn, Suzanne, Betty**) for providing me the opportunity to work in his team doing my own project. To show me another style of working and another style of live. Thank you too to all people I knew there that did this experience very funny, specially the Spanish team!

A **Xose Puente** y su grupo, especialmente a **Diana Puente**, por su colaboración y ayuda en el proyecto BRCAx, por responder rápidamente a cada duda que surgía y por el esfuerzo realizado.

Agradecer a todos mis compañeros becarios por haber hecho esta etapa tan divertida, por todo lo que hemos compartido dentro y fuera del laboratorio. Gracias a todos por aguantar mis despistes... A **Gardenia** por su orden, sus protocolos y sus sonrisas picaronas tras la pantalla del ordenador. **Elena** por estar siempre atenta cuando necesitaba ayuda en la poyata y hacerlo siempre con una sonrisa. **Fernando** por sus historias tan peculiares casi paranormales. A **Gau** por resolverme cualquier duda, por sus consejos y por ser la compañera ideal de Eva para amenizar las comidas! A **Estela y Maribel** por traer nuevo aires al laboratorio! A **Sami** le envió fuerzas porque se queda solo entre tanta mujer...que miedo. A **Pili** por pensar en mí cuando Conxi le dijo que había una beca disponible y por su simpatía sureña! Y por último los tres fantásticos! La **Vane**, la **Nuri** y el **Joanitu** porque hemos compartido mucho, muchas risas, muchas cenas, muchas liadas, viajes....Porque se atrevieron a venir a mi pueblo que parecía desierto! La **Vane** por su sonrisa constante y buen karma que te alegra solo de verla! El **Joanitu** por amenizar el laboratorio con tantos eventos, *oliverás* y groupone, por ser un gran compañero de máster y de trabajo y por repartirse conmigo los cabreros de la jefa... **Nuria**, mi compañera de ordenador, la fan de venti, la espontaneidad total...hemos compartido muchísimo Nuri no puedo escribirlo todo aquí, pero ya tu sabes!! Momentos buenos y no tan buenos, pero siempre intensos! Gracias por acogernos en tu pedazo casa de Menorca y por ser una gran amiga!

A la **gente del LRT1** por todas las cenas, Salamandras, barbacoas, partidillos de fútbol, días de voley...por hacer esta experiencia tan divertida! A **Miguel Ángel Pujana, Gorka y Paqui**, con los que he colaborado y me han ayudado en varios proyectos.

A **Adriana** por su ayuda en el proyecto de exomas y por ayudarme con la estadística en la que soy un poco pez.

A la **Chupipandi** por integrarme en su grupo cuando llegué al lab y darme tan buenos momentos, viajes, tapa-prats, cenas de navidad, vermutillos a las 4 de la tarde...y miles y cienas de cosas más. Sois muy grandes! Muy buena gente!

A **Natalia**, por estos más de tres años siendo compañeros de piso. Hemos tenido nuestros más y nuestros menos, y discrepamos en muchas cosas, pero lo que no hay duda es que es una muy buena amiga que sabes que estará ahí cuando la necesites.

A **Marta**, con la que compartí dos años de esta tesis y una magnífica experiencia en los USA. Fueron dos años geniales en los que pusiste cordura a mi vida desenfadada, jeje. Me ayudaste siempre y cuando lo necesité, me descubriste que era un 7 y me diste consejos que me han servido para afrontar futuros retos.

A todos los amigos tanto del mi pueblo como los que he ido conociendo en Barcelona, que han hecho los findes tan divertidos y los Lunes (y a veces hasta los Martes) tan duros!

A **Joan Giner** y **Sol Mazitelli** por ser mis primeros compañeros de piso en Barcelona y con los que he vivido muchos de los mejores momento de mi etapa aquí. Porque formábamos un equipo genial, la familia Serra, que tiempos aquellos.

A **Andrés**, mi *primo* y compañero de piso durante una etapa. Por enseñarme a ver el mundo con ojos más espirituales y por ayudarme a aprender a aporrear la guitarra y por introducirme con su grupo de amigos cuando llegué a BCN.

A **Gonzalo Ovejero** por, además de ser un buen amigo, diseñarme la portada de la tesis (si la hubiera hecho yo creo que me hubiera costado más que la tesis).

A **Maria Elizalde**, que me ha acompañado estos dos últimos años tan intensos, ese magnífico viaje a Menorca en Kayak, las vistas a Zuga, a Onda, roadtrips de Caravana, me has introducido en el buceo...han sido mil cosas. Pero sobretodo me quedo con tu forma de ser tan genial, honesta y de la que tanto he aprendido o he intentado... Siempre has estado cuando te he necesitado, muestra de ello es está tesis, que no hubiera podido depositar a tiempo si no hubieras estado a pie del cañón más de 12 horas al son de Cambio de sección impar...jejeje Casi nos volvemos locos, pero fue divertido y lo conseguimos! Mucchiismas gracias por tu ayuda en esa y en todas las otras ocasiones. Eres genial Maria!

Gracias a **mi familia** por ser tan cercana, tan divertida y por generar este ambiente que me ha ayudado a ver la vida con buenos ojos y optimismo.

Y, sobretodo, gracias a mis padres que sin ellos esto no hubiera sido posible. Me han inculcado los valores necesarios para llegar hasta aquí. Me han apoyado siempre en mis decisiones, por muy descabelladas que fueran y aunque no estuvieran de acuerdo con ellas, me han dejado tropezarme y siempre han estado ahí cuando lo he necesitado (que no han sido pocas veces dada mi tendencia a liarla). Han aguantado mis despistes, travesuras y demás. Por todo, por sentir que puedo caer pero siempre estaréis ahí. Y a mi hermanito **Pablo**! Que tiene una forma de ser genial, que admiro en muchos aspectos, con muchas cualidades que a mi me faltan y de las que tengo que aprender. Un abrazo Pablito! Que tu ya casi lo tiene también!

Espero no olvidarme a nadie, que dado que soy un despistado podría suceder, si es así espero que no me lo tengáis en cuenta. Muchas gracias a todos lo que me habéis ayudado en cada una de las cosas que me han ido surgiendo durante esta tesis.

ÍNDICE

ABREVIATURAS	I-II
INTRODUCCIÓN	1
1. Epidemiología del Cáncer de Mama.....	3
2. Histopatología de los tumores de mama. Diferencias entre tumores asociados a <i>BRCA1/2</i> y tumores esporádicos.....	4
3. Síndrome de Cáncer de Mama y Ovario Hereditario (SCMOH)	7
4. Genes <i>BRCA1</i> y <i>BRCA2</i> : Descripción, función y susceptibilidad	10
5. Clasificación clínica de variantes genéticas	16
6. Modelos de clasificación de variantes. Esquema de clasificación de variantes propuesto por la IARC.....	18
7. Mutaciones genéticas que alteran el correcto <i>splicing</i> del RNA.....	26
8. Ensayos funcionales en <i>BRCA1</i> y <i>BRCA2</i>	33
9. Asesoramiento genético en portadores de VSD	44
10. Predisposición al SCMOH más allá de mutaciones en <i>BRCA1</i> y <i>BRCA2</i>	45
ANTECEDENTES E HIPÓTESIS	51
OBJETIVOS	57
RESULTADOS	61
Capítulo 1. EVALUACIÓN DEL EFECTO DE VSD EN EL PROCESAMIENTO DEL mRNA DE LOS GENES <i>BRCA1/2</i>	65
Artículo 1: <i>Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing.</i>	67
Resultados sometidos a publicación: <i>Investigating the effect of 28 <i>BRCA1</i> and <i>BRCA2</i> mutations on their related transcribed mRNA.</i>	81
Capítulo 2. ESTUDIO FUNCIONAL DE VSD EN LAS PROTEÍNAS <i>BRCA1/2</i>	109
Artículo 2: <i>Functional and structural analysis of C-terminal <i>BRCA1</i> missense variants</i>	111
Resultados sometidos a publicación: <i>Identification of a founder <i>BRCA1</i> mutation in Moroccan population.</i>	119
Resultados no publicados. Análisis funcional de variantes de significado desconocido en <i>BRCA2</i> mediante un ensayo basado en células embrionarias de ratón.	135

Capítulo 3. ESTUDIO MUTACIONAL DE NUEVOS GENES IDENTIFICADOS EN LA VÍA DE ANEMIA DE FANCONI/BRCA EN FAMILIAS BRCAX	157
<i>Artículo 3: Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families</i>	159
<i>Artículo 4: Evaluation of rare variants in the new fanconi anemia gene ERCC4 (FANCO) as familial breast/ovarian cancer susceptibility alleles.</i>	167
Capítulo4. SECUENCIACIÓN DE EXOMAS EN FAMILIAS BRCAX	173
Resultados no publicados: Búsqueda de nuevos genes de predisposición al SCMOH mediante secuenciación de exomas de pacientes pertenecientes a familias BRCAX.....	175
RESUMEN DE RESULTADOS	193
DISCUSIÓN	195
Clasificación de Variantes de Significado Desconocido (VSD) en los genes BRCA1/2	197
Primera Aproximación: Evaluación del efecto de VSD identificadas en nuestro centro en el procesamiento del mRNA de los genes <i>BRCA1/2</i>	197
Segunda Aproximación: Estudio funcional de diferentes VSD en las proteínas <i>BRCA1/2</i>	201
Búsqueda de nuevos genes relacionados con el SCMOH	208
Primera Aproximación: <i>Análisis mutacional de nuevos genes identificados en la vía de Anemia de Fanconi/BRCA en familias BRCAX</i>	208
Segunda Aproximación: Búsqueda de nuevos genes de predisposición al SCMOH mediante secuenciación de exomas en familias BRCAX.	213
Sumario	217
CONCLUSION	221
BIBLIOGRAFÍA	225
ANEXOS	247
Artículo ANEXO: “Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes.”	249
INFORME DE PUBLICACIONES	265

ABREVIATURAS

A	Sitio de <i>Splice</i> acceptor
aa	Aminoácido
Align-GVGD	<i>Align-Grantham Variation_Grantham Deviation</i>
BAC	<i>Bacterial Artificial Chromosome</i>
BASC	<i>BRCA1-Associated Genome Surveillance Complex</i>
BIC	<i>Breast Cancer Information Core</i>
BRCAX	Familias de alto riesgo de sufrir cáncer de mama y/u ovario no portadoras de mutaciones en <i>BRCA1/2</i>
BRCT	<i>Breast Cancer Tandem repeat</i>
CCR	Cáncer Colorectal
CDI	Carcinoma Ductal Invasivo
CDIS	Carcinoma Ductal <i>In Situ</i>
cDNA	Ácido desoxirribonucleico complementario
CIMBA	<i>Consortium of Investigators of Modifiers of BRCA1/2</i>
CLI	Carcinoma Lobular Invasivo
CLIS	Carcinoma Lobular <i>In Situ</i>
CM	Cáncer de Mama
CMI	Cáncer de Mama Inflamatorio
CO	Cáncer de Ovario
CTP	Codón de Terminación Prematuro
D	Sitio de <i>splice</i> donador
DBD	Dominio de unión al DNA
DNA	Ácido desoxirribonucleico
ENIGMA	<i>Evidence-Based Network for the Interpretation of Germline Mutant Alleles</i>
ESE	<i>Exonic Splicing Enhancer</i>
ESP	<i>Exome Sequencing Project</i>
ESS	<i>Exonic Splicing Silencer</i>
fs	<i>frameshift</i>
G418	Gentamicina
GFP	<i>Green Fluorescent Protein</i>
GWAS	<i>Genome-Wide Association Study</i>
HAT	Hipoxantina-aminopterina-timidina
IARC	<i>International Agency for Research on Cancer</i>
ICL	<i>DNA Inter-strand Crosslink</i>
ICO	Instituto Catalán de Oncología
Indels	Inserciones y/o deleciones
IR	Radiación ionizante
ISE	<i>Intronic Splicing Enhancer</i>
ISS	<i>Intronic Splicing Silencer</i>
KDa	Kilo Daltons
LGR	Gra <i>Exonic Splicing Enhancer</i> ndes reordenamientos genómicos
LIF	<i>Leukemia Inhibitor Factor</i>
LOH	Pérdida de heterocigosidad
LOVD	<i>Leiden Open Variation Database</i>
MAF	<i>Minor Allele Frequency</i>
Mit C	Mitomomicina C
MLPA	Multiplex ligation-dependent probe amplification
MMS	Metil Metanosulfonato
mRNA	Ácido ribonucleico mensajero

MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NER	<i>Nucleotid excision repair</i>
NES	Señal de exportación nuclear
NGS	<i>Next Generation Sequencing</i> o Secuenciación masiva en paralelo
NLS	Señal de localización nuclear
NMD	<i>Non-sense Mediated Decay</i>
nt	Nucleótido
OH	Hidroxi
PCR	<i>Polymerase Chain Reaction</i>
PolyPhen2	<i>Polymorphism Phenotyping</i>
RE	Ratón Embrionarias
RH	Recombinación homóloga
RNA	Ácido ribonucleico
RT-PCR	<i>Retrotranscription Polymerase Chain Reaction</i>
SA	<i>Splicing</i> Alternativo
SCD	<i>Serine cluster domain</i>
SCMOH	Síndrome de Cáncer de Mama y Ovario Hereditario
SIFT	<i>Sorting Intolerant from Tolerant</i>
SMIPS	<i>Somatic Mutation Identification in Pooled Samples</i>
SNP	<i>Single Nucleotide Polymorphism</i>
sp	<i>Splice</i>
SS	<i>Splice Sites</i>
St	<i>Stop</i>
SWG	<i>Splicing Working Group</i>
UTR	<i>Untranslated region</i>
VSD	Variante de Significado Desconocido
wt	<i>wild-type</i>

INTRODUCCIÓN

1. Epidemiología del Cáncer de Mama

Según datos publicados en Globocan 2012^{*}, el cáncer es la segunda causa de defunción debido a enfermedades no transmisibles en países desarrollados y la primera en países en vías de desarrollo¹. De entre los diversos tipos de cáncer que se conocen, el cáncer de mama (CM) se presenta como el segundo cáncer más común en el mundo. En 2012, se diagnosticaron cerca de 1,7 millones de nuevos casos de CM a nivel mundial, lo que representa, aproximadamente, el 12% de todos los nuevos casos de cáncer diagnosticados en la población general y el 25% si consideramos sólo la población femenina (Globocan, 2012)¹. De este modo, se ha convertido en el tipo de cáncer más frecuente entre la población femenina, tanto en países desarrollados como en aquellos en vías de desarrollo¹⁻³. Además, la prevalencia[†] y la incidencia[‡] del CM están aumentando constantemente, principalmente debido al mayor envejecimiento de la población¹.

En los últimos años, se ha dado una mejoría sustancial en los programas de cribado y en los tratamientos asociados al CM⁴. Gracias a ello, se ha conseguido reducir considerablemente las tasa de mortalidad ligada a este tipo de cáncer a nivel global^{4, 5}. Aún así, sigue siendo la primera causa de muerte por cáncer en mujeres de las regiones menos desarrolladas (el 14,3% del total) y la segunda causa en mujeres de las regiones más desarrolladas (15,4% del total)¹.

En el caso concreto de España, según los datos de la Sociedad Española Oncológica Médica (SEOM, 2014), publicados en 2014 (<http://www.seom.org>), el CM es el de mayor incidencia entre la población femenina. Aún así, la incidencia del CM es menor que en otros países desarrollados, aunque similar al resto de países de Europa Mediterránea. En el 2012, se diagnostican unos 26.000 casos al año, lo que representa casi el 30% de todos los tumores diagnosticados en población femenina en nuestro país. La mayoría de estos casos se diagnostican entre los 35 y los 80 años, con un máximo entre los 45 y los 65 (SEOM, 2014). La tasa de mortalidad en España en el 2012 fue del 15,5% la más baja de todos los países Europeos¹. Si bien, es la primera causa de muerte

^{*} La Agencia Internacional para la Investigación del Cáncer (IARC), organismo especializado para el cáncer de la Organización Mundial de la Salud (OMS), cada cuatro años publica el informe GLOBOCAN (<http://globocan.iarc.fr>). El último fue publicado en 2012 con los datos oficiales más actualizados a nivel mundial sobre incidencia, prevalencia (1,3 y 5 años) y mortalidad para los 28 tipos de cáncer más frecuentes de 184 países.

[†] La prevalencia hace referencia al número de pacientes con cáncer en una determina población, independientemente del momento en que fueron diagnosticados. La prevalencia depende de la frecuencia de aparición de la enfermedad (incidencia) y de la letalidad de la misma (supervivencia). La prevalencia se suele estudiar a los 5 años, 3 años y 1 año. En este caso se utilizan datos de prevalencia a 5 años.

[‡] La incidencia es el número de casos nuevos de cáncer en un periodo determinado de tiempo, generalmente un año, en un número determinado de habitantes.

de por cáncer entre mujeres españolas. Además, a prevalencia a 5 años, estima que en el 2016 el CM constituirá el 40% de los casos diagnosticados en población femenina española (SEOM, 2014).

Factores de riesgo asociados al CM

El CM es una enfermedad compleja y asociada a diversos factores de riesgo⁶. Se puede diferenciar entre factores de riesgo modificables y no modificables. Por una parte, el cáncer es una enfermedad intrínsecamente ligada al envejecimiento⁷. Por ello, de entre los factores de riesgo no modificables, el envejecimiento es uno de los principales. De hecho, más del 70% de los casos son diagnosticados en mujeres de más de 50 años⁷. Por otra parte, el CM esté asociado a factores hormonales⁸. Por ello, variaciones en la expresión de estos puede influir en el riesgo a sufrir esta enfermedad. La mayor parte de factores de riesgo hormonales están relacionados con antecedentes reproductivos que modulan la expresión hormonal a lo largo de la vida. Entre ellos, la edad temprana de menarquía⁸, una menopausia tardía⁹ o una edad tardía del primer embarazo¹⁰ han sido asociadas a un incremento de riesgo. Los factores de riesgo modificables serían los que están ligados al estilo de vida. Entre ellos destacan, como factores asociados a un aumento de riesgo, el consumo excesivo de grasa¹¹, la obesidad¹² o el consumo de alcohol, entre otros. Existen otros factores que han demostrado tener un carácter protector, como realizar ejercicio de forma regular¹³.

No obstante, siempre que existan antecedentes familiares o personales de CM, éstos serán considerados como los factores de riesgo predominantes^{7,14}. En este sentido, datos epidemiológicos mostraron que en mujeres jóvenes diagnosticadas de CM el riesgo de desarrollar un segundo CM primario aumenta 4,5 veces respecto a la población sana y, además, el riesgo de desarrollar un cáncer de cualquier otro tipo se triplica¹³. Por otra parte, se ha demostrado que mujeres con antecedentes familiares de primer grado tienen el doble de riesgo de padecer la enfermedad¹⁵. Como se explicará a continuación, una proporción de estos casos familiares están ligados a una predisposición genética conocida^{16,17}.

2. Histopatología de los tumores de mama. Diferencias entre tumores asociados a *BRAC1/2* y tumores esporádicos.

Los tumores de mama se pueden clasificar atendiendo a diferentes características, como el tejido de origen, grado de diferenciación de las células tumorales o expresión de receptores hormonales o la proteína HER2.

El cáncer de mama puede comenzar en distintas áreas de la mama: los conductos, los lobulillos o, en algunos casos, el tejido intermedio. La mayoría de los tumores de

mama son carcinomas, un tipo de cáncer que comienza en las células epiteliales que revisten los órganos y tejidos internos. Los carcinomas pueden ser “in situ” (no invasivos), se mantienen en el tejido de origen, o invasivos, cuando se ha propagan hacia los tejidos mamarios que lo rodean.

Figura 1. Anatomía de la mama femenina

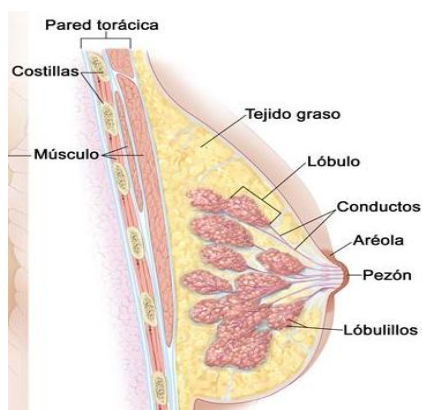


Figura 1. Esquema de los diferentes que forman la mama femenina. El CM puede originarse en cualquiera de ellos. (Figura extraída de la página web del *National Cancer Intitute* (USA) <http://www.cancer.gov/Common/PopUps/popDefinition.aspx?id=304766&version=Patient&language=Spanish>)

Los carcinomas ductal *in situ* (CDIS) o invasivo (CDI), son los tipos más comunes de CM no invasivo y invasivo, respectivamente. Se originan dentro de los conductos lácteos, que transportan la leche desde los lobulillos productores de leche al pezón. El CDI, afecta a alrededor del 80 % de las mujeres con CM invasivo¹⁸⁻²⁰. Se han descrito varios subtipos de carcinoma ductal invasivo poco frecuentes, entre ellos, el tubular que representa cerca del 1-2 % del total de tumores de mama, el medular, que representa entre el 3-5% o el carcinoma mucinoso que representa cerca del 2-3 %^{21,22}.

El carcinoma lobular, *in situ* (CLIS) o invasivo (CLI), es el que se origina en las glándulas productoras de leche (lobulillos). El CLI representa el 10% de todos los CM detectados, representando el segundo tipo cáncer más común¹⁸⁻²⁰.

Otro tipos de carcinomas mamarios, poco frecuentes, son: el inflamatorio (CMI) que representa alrededor del 1 al 5 %, de los casos de CM, el papilar invasivo que engloba entre 1-2 , entre otros²³⁻²⁵.

El grado de un tumor es la descripción de éste basada en la morfología de las células que lo componen observadas bajo el microscopio. El grado tumoral es indicativo de la rapidez con la que éste se propagará. Los tumores se pueden dividir en tres grados (I, II y III)²⁶. En general, un tumor de grado I presenta células similares en morfología y comportamiento a las células sanas (crecimiento lento), se denominan tumores bien

diferenciados. Los tumores de grado 3 presentan células con morfología anómala y suelen ser de rápido crecimiento con mayor probabilidad de propagarse a otros tejidos, se denominan tumores indiferenciados. Los tumores de grado II (grado moderado) presenta una morfología y comportamiento intermedio y se denominan tumores escasamente diferenciados.

Los tumores de mama también se pueden clasificar en base a la presencia de receptores hormonales [estrógenos (ER), o progesterona (PR)] y la sobreexpresión de HER2^{27, 28}. Los tumores que contienen receptores de estrógeno, son referidos ER-positivos mientras que aquéllos que presentan receptores de progesterona se les denomina PR-positivos. Alrededor de dos de cada tres tumores de mama presentan al menos uno de estos receptores. Los tumores que no contienen receptores de estrógeno ni de progesterona, se dice que son negativos para receptores hormonales (ER-, PR-). Este tipo de tumores es más común en mujeres pre-menopáusicas. En cuanto a la expresión de la proteína de crecimiento HER2, los tumores de mama se pueden clasificar entre, HER2 positivo, los que sobreexpresan HER2 o contienen copias adicionales de este gen, o HER2 negativo, los que no presentan sobreexpresión de HER2. Aproximadamente, uno de cada cinco tumores de mama es HER2 positivos. Si las células tumorales no contienen receptores de estrógeno ni de progesterona ni sobreexpresan HER2, se les llama triple negativos. En caso que expresen los tres, se denominan triple positivos.

Diferencias entre tumores asociados a mutaciones germinales en BRCA1/2 y tumores esporádicos

Varios estudios comparativos de tumores asociados a pacientes con mutaciones germinales en los genes *BRCA1* y/o *BRCA2*, han demostrado que existen rasgos histopatológicos distintivos éstos y los tumores esporádicos.

Los tumores asociados a *BRCA1*, suelen presentar alto grado de diferenciación y una tasa de proliferación elevada (Grado III)²⁹. Además, suelen ser triple negativos^{30, 31} y suelen ser carcinomas ductales con una mayor proporción del subtipo medular³²⁻³⁶. Presentan sobreexpresión de TP53^{37, 38}. Además, estudios de expresión de cDNA en estos tumores han sugerido un fenotipo epitelial basal (patrones de expresión génica similar a las capas basales más profundas de los conductos y las glándulas del seno)³⁹ y se les asocia la expresión de citoqueratinas 5/6 (ref⁴⁰).

Por el contrario, los carcinomas asociados a *BRCA2* son más heterogéneos, por ello, en la mayoría de estudios no se han encontrado diferencias significativas entre los tumores asociados a *BRCA2* y los esporádicos^{33, 34, 41}. Como consecuencia, los marcadores inmunohistoquímicos asociados a los tumores asociados a *BRCA2* no suelen ser considerados como predictores en el diagnóstico de posibles portadores de mutaciones

germinales en *BRCA2*^{33, 34, 41}. No obstante, estos tumores presentan un rasgo característico no observado ni en los tumores asociados a *BRCA1* ni en los esporádicos, la sobreexpresión de la proteína nuclear RAD51 en el citoplasma^{42, 43}. Además, por lo general, muestran una tasa de proliferación baja⁴⁴ y un ligero aumento en la incidencia de carcinomas lobular o tubulolobular^{45, 46}. Respecto a los receptores hormonales suelen ser positivos para receptores de estrógenos en una tasa similar a la de los carcinomas esporádicos y no presentan sobreexpresión de HER2 (ref^{45, 47}).

3. Síndrome de Cáncer de Mama y Ovario Hereditario (SCMOH)

Aunque la mayor parte de los casos de CM son de origen esporádico, se estima que alrededor del 20 de los casos de CM presentan una historia familiar de la enfermedad, es decir, presentan agregación de varios casos de CM en la familia a través de varias generaciones (Fig. 2)⁴⁸. En un porcentaje de estos casos familiares, las causas genéticas han sido determinadas y se consideran, por tanto, hereditarios⁴⁸. Los casos de CM hereditarios son debidos a mutaciones germinales en genes que incrementan la susceptibilidad a padecer este tipo de cáncer¹⁶. Estos casos, se engloban dentro del denominado síndrome de Cáncer de Mama y Ovario Hereditario (SCMOH) y suponen cerca del 15% del total de casos de CM familiar (Fig. 2)⁴⁹. Las familias afectadas por este síndrome se caracterizan por la agregación de múltiples casos de cáncer de mama y/o de ovario en diferentes generaciones siguiendo un patrón de herencia autosómica dominante. Además, los tumores asociados a pacientes del SCMOH suelen aparecer a una edad más temprana que los esporádicos (edades pre-menopáusicas) y a menudo, se observa bilateralidad⁵⁰.

Las familias con riesgo de sufrir cáncer hereditario se pueden seleccionar en base a guías de consenso clínico desarrolladas por paneles de expertos, o en base a numerosos estudios que proporcionan datos empíricos sobre la probabilidad de tener una mutación genética en genes de predisposición^{7, 10, 51-54}. Teniendo en cuenta criterios clínicos, las familias pueden clasificarse en alto y moderado riesgo de sufrir cáncer de mama y ovario hereditario. A aquellas familias que cumplan criterios clínicos de alto riesgo se les ofrecerá la posibilidad de realizarse el test genético en genes de predisposición a la enfermedad.

En el caso de Cataluña, dichos criterios clínicos, establecidos en 2006 y recogidos en la Oncoguía del Consejo y Asesoramiento Genéticos en el Cáncer Hereditario de Cataluña (http://www.gencat.net/salut/depsan/pdf/og0601esgrr_din4.pdf) fueron los siguientes:

1. Tres o más individuos de primer grado afectados por cáncer de mama y/o ovario
2. Dos casos entre familiares de primer/segundo grado:
 - a. Dos casos de cáncer de ovario.
 - b. Un caso de cáncer de mama y otro de cáncer de ovario.
 - c. Un caso de cáncer de mama en un individuo de sexo masculino y otro caso de mama/ovario.
 - d. Dos casos de cáncer de mama en individuos menores de 50 años.
 - e. Un caso de cáncer de mama bilateral y otro de cáncer de mama (en una persona menor de 50 años).
3. Un caso de cáncer de mama en un paciente menor de 30 años.
4. Un caso de cáncer de mama y otro de ovario en el mismo paciente.
5. Un caso de cáncer de mama bilateral diagnosticado antes de los 40 años.

Actualmente estas guías están siendo modificadas y adaptadas a los conocimientos en este campo en los últimos diez años.

Bases Genéticas del SCMOH

Aunque han sido identificados varios genes de predisposición a este síndrome, sólo dos de ellos muestran una alta penetrancia[§] a la vez que alta prevalencia. Son los genes *BRCA1* (*Breast Cancer susceptibility gene 1*, MIM#113705) y *BRCA2* (*Breast Cancer susceptibility gene 2*, MIM#600185)^{16, 17, 55}. Mutaciones en la línea germinal en estos genes explican alrededor de un 15% de los casos del SCMOH (Fig. 2), aunque este porcentaje puede variar según la población estudiada⁵⁶. En algunas poblaciones aisladas, mutaciones en los genes *BRCA1/2* podrían explicar hasta un 50% de los casos⁵⁶.

§ Penetrancia genética se entiende como la proporción de individuos de una población que expresan el fenotipo patológico, entre todos los que presentan un genotipo portador de un alelo mutado. Se considera que mutaciones germinales en genes de alta penetrancia para una enfermedad confieren un aumento de riesgo 5 veces mayor que la población general a sufrir dicha enfermedad. Así pues, se estima que los genes de baja penetrancia lo aumentan menos de 1,5 veces y los de moderada penetrancia entre 1,5 y 5 veces (Jorde L CJ, Bamshad M, White R, editors. *Medical Genetics*. Mosby 2000; 2nd ed.)

Figura 2. Prevalencia de cáncer de mama familiar y hereditario y genes de predisposición al SCMOH

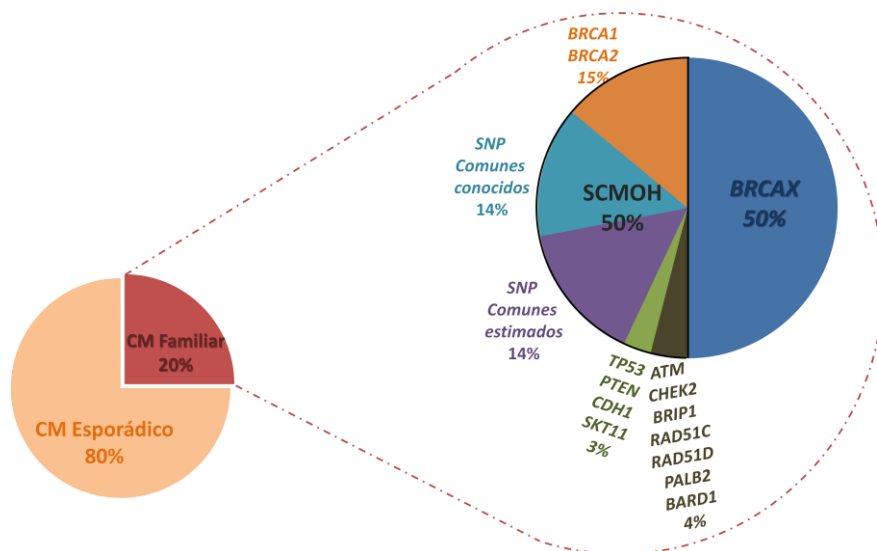


Figura 2. Alrededor del 80% del CM que afecta a la población general son de origen esporádico. El 20% restante presenta una agregación familiar. Hoy en día, las causas genéticas del CM familiar son conocidas para, aproximadamente, un 50% de los casos y se consideran, por tanto, hereditarios, debidos a mutaciones germinales. El 50% restante permanecen sin causa genética conocida hasta la actualidad y se conocen como casos BRCAX (esquema modificado de Couch et. al, 2014) (ref 57).

Al ser los genes *BRCA1/2* genes de alta penetrancia para esta enfermedad las mujeres portadoras de mutaciones en estos genes presentan un alto riesgo de sufrir cáncer de mama u ovario a lo largo de su vida^{16, 55}. En este sentido, los datos de riesgo medio acumulado con la edad, en portadores de mutaciones en genes *BRCA1/2*, varía según los estudios. Se estima que mujeres con mutaciones en *BRCA1* presentan un riesgo acumulado de hasta el 81% de sufrir CM y de hasta el 63% de sufrir cáncer de ovario a los 80 años. En cuanto a las portadoras de mutaciones en *BRCA2* se estima un riesgo a los 80 años mayor del 85% para el CM y del 25% para el CM y el CO, respectivamente^{58, 59}. En parte, las variaciones observadas entre poblaciones/pacientes se pueden explicar en función a la existencia de variantes comunes, identificadas a lo largo de todo el genoma, que actúan como modificadores de riesgo de CM y CO asociado a mutaciones en los genes *BRCA1/2*⁵⁸⁻⁶³. Estos modificadores de riesgo han sido básicamente identificados por estudios colaborativos en el sí del consorcio CIMBA (*Consortium of Investigators of Modifiers of BRCA1/2*), cuyo objetivo es la recolección de muestras biológicas y datos clínicos de portadores de mutaciones en estos genes en todo el mundo para poder realizar estudios a gran escala que permitan determinar con una alta probabilidad, la contribución de variantes genéticas en el fenotipo final observado en los portadores *BRCA*⁶⁴.

El resto de genes asociados a CM se consideran de baja prevalencia y, entre todos ellos, explicarían menos de un 10% de los casos asociados a este síndrome^{65, 66}. Algunos de estos genes adicionales son de alta penetrancia (*TP53*, *PTEN*, *STK11/LKB1* o *CDH1*) y están asociados a síndromes hereditarios concretos mientras que otros son de moderada o desconocida penetrancia (*CHEK2*, *PALB2*, *BRIP1*, *ATM*, *RAD51* entre otros)⁶⁷ (Fig. 2). Además, se considera que determinadas variantes frecuentes en población general podrían conferir un riesgo a CM, de modo que, la interacción entre varias de éstas podría explicar alrededor del 28% de las familias con SCMOH (*Apartado 10*) (Fig. 2)⁶⁷.

Aún así, más del 50% de los casos del SCMOH siguen sin causa genética conocida⁵⁷ y se denominan BRCAX (Fig. 2).

4. Genes *BRCA1* y *BRCA2*: Descripción, función y susceptibilidad

Como se ha dicho en el apartado anterior, *BRCA1* y *BRCA2* son los principales genes responsables del SCMOH. Fueron identificados a mediados de los años noventa mediante estudios de ligamiento en familias consideradas de alto riesgo que mostraban un patrón claro de hereditaria autosómica dominante^{16, 17}. El gen *BRCA1* se encuentra localizado en el brazo largo del cromosoma 17 (17q21), y el gen *BRCA2*, está localizado en el cromosoma 13 (13q12). Ambos son genes de gran tamaño. El gen *BRCA1* presenta una secuencia codificante de 5.592 nucleótidos (nt), repartidos en 24 exones (dos de ellos no codificantes), se extiende a lo largo de 81,2 Kb de DNA genómico y se traduce en una proteína de 1.863 aminoácidos (aa) (~220KDa). La secuencia codificante de *BRCA2*, por su parte, consta de 11.385 nt, distribuidos a lo largo de aproximadamente 70 Kb de DNA genómico y está compuesta de 27 exones, el primero de los cuales no se traduce. La proteína codificada por *BRCA2* consta de 3.418 aa (~380KDa). Cabe destacar que ambos genes poseen exones centrales muy largos: en *BRCA1* el exón 11 y en *BRCA2* los exones 10 y el 11 que codifican aproximadamente para el 60% de la proteína⁶⁸.

Función las proteínas BRCA1 y BRCA2. Vía de señalización FA/BRCA

Los genes *BRCA1/2* codifican para proteínas nucleares supresoras de tumores del tipo *caretaker*⁶⁹. Las proteínas *caretakers* se caracterizan por estar involucradas en el mantenimiento de la estabilidad genómica mediante el control del ciclo celular y la reparación del DNA.^{68, 70} Las proteínas *BRCA1/2* son proteínas multidominio y muchas de las funciones que realizan son dominio específicas o dependientes de sus interacciones con otras proteínas⁷¹.

El daño del DNA puede suceder de distintas formas como, por ejemplo, modificación de bases, roturas de doble cadena, entrecruzamiento inter-cadena de DNA (*DNA inter-strand crosslink*, (ICL)) entre otras^{72, 73}. Las células han evolucionado

generando distintos mecanismos de defensa para hacer frente a estos eventos de variada naturaleza⁷³. Uno de estos mecanismos es la recombinación homóloga (RH)⁷². La RH es un mecanismo de reparación de roturas de DNA de doble cadena, asociado a las fases S y G2 del ciclo celular, que utiliza dos cromátidas hermanas como plantilla para conseguir una reaparición precisa y libre de errores^{74, 75}. Defectos en la RH pueden conducir a aberraciones genómicas o reordenamientos cromosómicos, los cuales son eventos iniciadores de la tumorigénesis⁷⁶. Las proteínas BRCA1/2 son componentes esenciales de la vía de la RH, la cual es esencial para asegurar la integridad genómica (Fig. 3)^{74, 75}. BRCA1 tiene un papel central en la regulación de este proceso mediante la interacción con múltiples proteínas. Entre ellas, su interacción con BARD1⁷³, que potencia la actividad E3 ubiquitin ligasa de BRCA1, o su interacción con otros genes reparadores para formar un supercomplejo denominado BASC (*BRCA1-Associated Genome Surveillance Complex*) son claves para el reconocimiento y la reparación de daños de DNA de doble cadena⁷² (Fig. 3). El complejo BASC incluye las proteínas MSH2, MSH6, MLH1, ATM, BLM, Rad50-MRE11-NBS1 (complejo MRN) y RFC⁷². Por su parte, el papel principal de BRCA2 en la RH es el reclutamiento de RAD51 a la zonas de daño al DNA⁷⁷ (Fig. 3). RAD51 es una recombinasa involucrada en la invasión de la cadena de DNA sencilla durante la RH en humanos⁷⁸.

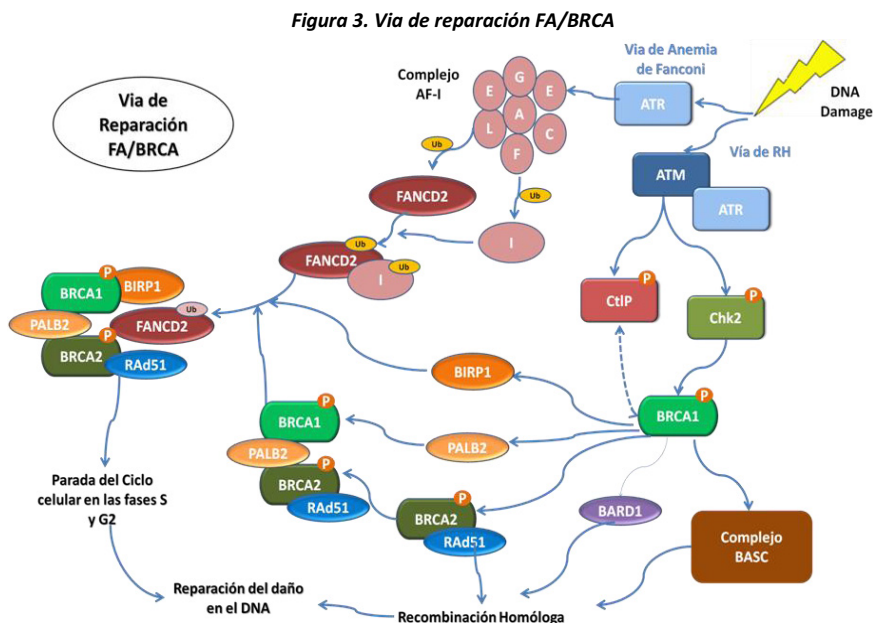


Figura 3. Esquema representando la vía de FA/BRCA. Las proteínas BRCA1/2 coexisten en un vía de reparación común: la vía de FA/BRCA. Esta vía se encarga de la reparación de los entrecruzamientos de cadena y se activa como respuesta a una horquilla de replicación bloqueada. Hoy en día se conocen 17 genes pertenecientes a esta vía. BRCA2 (FANCD1) tiene un papel directamente relacionado en la reparación del DNA, mediante el reclutamiento de RAD51 al lugar del daño al DNA. En cambio, BRCA1 (FANCS) tiene un papel regulador, mediante la interacción con otras proteínas de la vía, como PALB2, BRIP1 o BRCA2

Las proteínas BRCA1/2 coexisten en la vía de reparación denominada de Anemia de Fanconi/BRCA (FA) (Fig. 3). Está vía es la encargada de reparar los ICL⁷⁹. Defectos en

alguno de los 17 genes asociados a esta vía dan lugar a la Anemia de Fanconi⁸⁰⁻⁸². Esta enfermedad presenta un patrón de herencia autosómica recesiva y se caracteriza por una inestabilidad cromosómica que da lugar a múltiples anomalías congénitas, disfunción progresiva de la médula ósea y susceptibilidad a varios tipos de cáncer⁷⁹. La reparación de los ICL mediante esta vía requiere de la previa monoubiquitinización de FANCD y FANCI⁸⁰. Dicha ubiquitinización es mediada por un complejo con actividad ubiquitin ligasa formado por 8 de las proteínas de la vía de FA. Dicho complejo se forma como respuesta al daño en el DNA señalado por ATR⁸³. FANCD activo es capaz de reconocer los sitios de daño celular y unirse a la cromatina donde co-localiza con BRCA1, BRCA2, PALB2 y RAD51 los cuales son requeridos para la reparación mediante RH^{84, 85}. *BRCA2* (*FANCD1*) fue asociado a esta vía en el 2002⁸⁴. En cambio, aunque eran conocidas las conexiones entre *BRCA1* con los complejos D2-I y *BRCA2*-*PALB2*^{73, 80}, *BRCA1* no se había podido clasificar como un gen de FA hasta ahora. Esto es debido al hecho de que la mayor parte de las combinaciones bialélicas de mutaciones en *BRCA1* son letales⁸⁶. No obstante, un estudio reciente, ha reportado dos casos con mutaciones bialélicas en *BRCA1* mostrando un fenotipo de FA, lo cual ha permitido clasificar a *BRCA1* como un nuevo gen de esta vía, *FANCS*^{84, 85}.

Ahora bien, la función supresor tumoral de estos genes no se limita únicamente a su papel en la RH sino que también participan en otras funciones diferentes^{69, 87, 88}. Por ejemplo, la regulación del ciclo celular⁸⁹, la remodelación de la cromatina⁹⁰, la regulación de la duplicación de centrosomas u otros procesos celulares como la regulación de la transcripción génica de genes reparadores⁹¹.

Estructura de la proteína BRCA1

BRCA1 codifica para una fosfoproteína nuclear de 220 kDa. Presenta varios dominios funcionales, dos de ellos muy relevantes, el dominio RING y el dominio BRCT (Fig. 4)⁷¹.

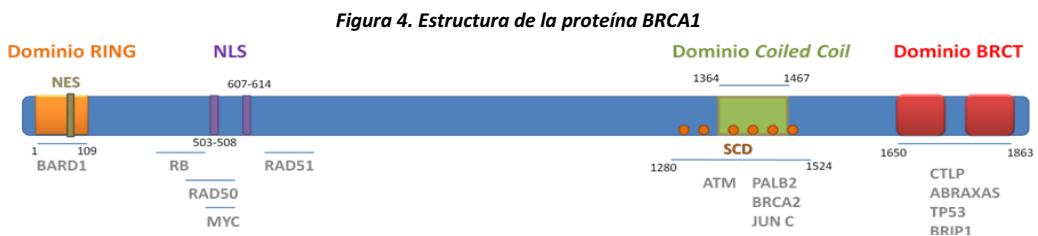


Figura 4. Diagrama mostrando la estructura de BRCA1. En este diagrama se muestran los principales dominios funcionales de la proteína BRCA1 y sus principales interacciones con otras proteínas.

El dominio RING se encuentra en el extremo N-terminal de la proteína (entre los aa 1-109). Es el responsable de la actividad E3 ubiquitin ligasa de BRCA1 y de la interacción con BARD1^{3392, 93}. En este dominio se encuentra una secuencia de 18 aa

hidrofóbicos (aa 81-99) que actúan como señal de exportación nuclear (NES) (Fig. 4)⁹⁴. Además, dicho dominio está involucrado en la regulación de la transcripción de muchos genes ligados a la reparación del DNA⁶⁹.

El dominio BRCT o dominio de activación transcripcional (Fig. 4). Es un dominio común entre las proteínas involucradas en la reparación del DNA^{95, 96}. Está localizado en la región C-terminal de BRCA1 (aa 1650-1863) en forma de repetición en tándem y participa en la regulación de la transcripción mediante la interacción con proteínas de la maquinaria transcripcional⁹⁷. Además, mediante este dominio, BRCA1 interacciona con una serie de proteínas fosforiladas, como BRIP1, para actuar en los puntos de control del ciclo celular de las fases S y G₂/M como consecuencia del daño en el DNA^{95, 98}.

La secuencia intermedia entre ambos dominios (aa 109-1649) da lugar a una región sin configuración tridimensional determinada. En esta región se encuentran dos señales de localización celular (NLS) (aa 503-508 y 607-614) y una región rica en residuos serina (SCD) (aa 1280-1524), las cuales son diana de fosforilación mediada por ATM/ATR como señalización de horquillas de replicación bloqueadas⁷⁴. También se han identificado en esta región una serie de motivos de unión a otras proteínas como RB, c-MYC, RAD50, RAD51 todas de vital importancia en la reparación del DNA durante el ciclo celular^{78, 99-101}. Además también podemos encontrar en esta zona intermedia el motivo *colied coil* (aa 1364-1467), mediante el cual BRCA1 interacciona con JUNC y PALB2¹⁰².

Estructura de la proteína BRCA2

BRCA2 codifica para una fosfoproteína nuclear de 384 kD¹⁰². La proteína BRCA2 presenta tres dominios funcionales principales: un dominio N-terminal, una serie de repeticiones BRC y un dominio de unión al DNA (DBD) (Fig. 5)¹⁰³.

Figura 5. Estructura de la proteína BRCA2

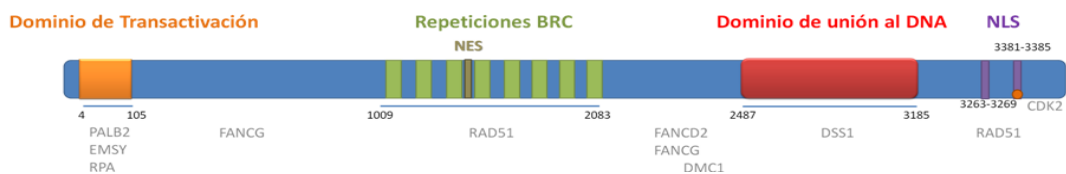


Figura 5. Diagrama mostrando la estructura de BRCA2. En este diagrama se muestran los principales dominios funcionales de la proteína BRCA2 y sus principales interacciones con otras proteínas.

El dominio N-terminal o dominio de transactivación (aa 4-105) presenta actividad acetiltransferasa de histonas. Gracias a su actividad acetiltransferasa BRCA2 está involucrada en la regulación de la transcripción del RNA¹⁰⁴. Además, este dominio está implicado en la interacción con PALB2, el cual interacciona físicamente con BRCA1¹⁰⁵. Esta

interacción es crítica para la función reparadora de estas tres proteínas¹⁰⁵. También mediante este dominio BRCA2 interactúa con EMSY y RPA, proteínas involucradas en la remodelación de la cromatina^{106, 107}.

El dominio BRC consta de 8 subdominios situados en tándem y codificados por el extenso exón 11 (aa 1.009-2.083)¹⁰². La principal función de este dominio es la interacción con RAD51, que como se ha comentado es esencial para la RH^{108, 109}. De hecho, la formación de los *foci* RAD51 como respuesta al daño en DNA es, en su mayor medida, dependiente de su interacción con BRCA2^{108, 109}.

El DBD situado en la región C-terminal, corresponde con una región altamente conservada en la evolución. Está formado por cinco motivos: un primer motivo rico en alfa-hélices, tres motivos de unión a oligosacáridos y oligonucleótidos (OB) y un motivo helicoidal que sobresale de la estructura, denominado Torre¹⁰³. Mediante este dominio BRCA2 se une a el DNA de doble cadena y de cadena sencilla durante la recombinación homóloga (RH)¹⁰³. Asimismo, la proteína DSS1 interactúa con BRCA2 a partir del DBD. Mediante esta interacción DSS1 estabiliza BRCA2 y controlar su localización nuclear¹¹⁰.

Otros aspectos importantes de la estructura de BRCA2 son su unión con FANCD2¹¹¹ (aa 2350-2545) mediante la región posterior al dominio BRC y el dominio PhePP (mediante el que se une a la proteína DMC1)¹¹². Además, en el extremo C-terminal contiene dos NLS, un sitio adicional de interacción con RAD51 y una diana de fosforilación regulada por CDK2¹¹³. Por otra parte, BRCA2 interactúa con FANCG a través de dos regiones situadas en la región central (aa 499-994 y 2.350-2.545).

Susceptibilidad al Cáncer de Mama y Ovario asociada a los genes BRCA1/2

La tumorigénesis asociado a genes supresores, como los genes *BRCA1/2*, requiere de la inactivación bialélica de éstos^{68, 70}. De forma general, el proceso de inactivación de estos genes se puede explicar a partir del modelo de Knudson o *two hits*^{70, 114}. Este modelo propone que uno de los alelos se hereda inactivo debido a la presencia de una mutación en línea germinal y el segundo alelo se inactiva en una célula somática debido a un segundo evento al azar a lo largo de la vida. La herencia de un solo alelo activo confiere predisposición al cáncer, dado que, las posibilidades de que el gen quede inactivado aumentan en gran medida con respecto a la población normal¹¹⁵. El denominado segundo *hit* puede ser debido a errores en la replicación celular o a factores genotóxicos externos, en forma de cualquier tipo de mutación génica (mutaciones puntuales, inserciones-delecciones, *frameshift*, reordenamientos, etc.) o por silenciamiento del alelo activo por metilación del promotor^{114, 116}. La pérdida de función del segundo alelo es conocido como pérdida de heterocigosidad.

Aunque el modelo de Knudson es utilizado de forma general para explicar la susceptibilidad al cáncer asociada a genes supresores, existen evidencias que señalan que el mecanismo de inactivación de estos genes podría ser más complejo. Recientemente, se han publicado estudios que indican que podría estar ligado a la haploinsuficiencia derivada de la inactivación germinal del uno de sus alelos^{117, 118}. No obstante, no existen evidencias suficientes que lo confirmen.

Espectro mutacional de los genes BRCA1 y BRCA2

Los genes *BRCA1/2* muestran un patrón mutacional muy heterogéneo y complejo. Además, puede variar mucho en función de la población estudiada⁵³. Según datos obtenidos a partir de la principal base de datos para variantes detectadas en estos genes, denominada *Breast Cancer Information Core* (BIC)¹¹⁹, hay reportadas 1781 variantes en *BRCA1* y 2000 en *BRCA2* (Tabla 1). Las variantes reportadas en BIC son, por una parte, variantes que afectan a un solo nucleótido como: variantes intrónicas, variantes que afectan a un solo aminoácido (*missense*), que no afectan al aminoácido que codifican (*silent*) o variantes que generan un codón de terminación (mutaciones *STOP*). Y, por otra, pequeñas inserciones y deleciones que pueden alterar o no, el marco de lectura (*frameshift* e *in-frame*, respectivamente). Todas estas variantes se distribuyen a lo largo de las regiones codificantes y en las secuencias intrónicas flanqueantes. Además de éstas, existen otro tipo de variantes, no reportadas en BIC, que implican la reorganización de grandes segmentos de DNA y suelen afectar a uno o varios exones de estos genes. Son los denominados reordenamientos genómicos (LGR). Cabe destacar que, aunque, las variantes *frameshift* y los LGR se dan en ambos genes, son mucho más comunes en *BRCA1*, probablemente debido a la mayor frecuencia de secuencias repetitivas (Alu) en las regiones intrónicas de este gen^{120, 121}. De hecho, las mutaciones *frameshift* corresponden al 40% de las mutaciones patogénicas en *BRCA1*, y sólo al 20% en *BRCA2*, mientras que los LGR corresponden con el 14% en *BRCA1* y tan sólo con el 2,6% en *BRCA2*^{122, 123}.

Tabla 1. Espectro mutacional de los genes BRCA1 y BRCA2

Gen	Número de entradas en BIC	Número de variantes únicas	Número de variantes observadas solo una vez
<i>BRCA1</i>	15311	1781	964
<i>BRCA2</i>	14914	2000	1065

Datos obtenidos de BIC. Última modificación: 13-03-2015

El patrón mutacional que se da en estos genes variará mucho según la población estudiada, dando lugar a escenarios muy distinto en el análisis genético de los pacientes de alto riesgo. Por ejemplo, mientras que en poblaciones no endogámicas, la mayor parte de las mutaciones detectadas son específicas de una o de muy pocas familias, en ciertas poblaciones endogámicas, un número pequeño de mutaciones puede aparecer de

manera muy frecuente debido a un efecto fundador^{124, 125}. De hecho, en poblaciones no endogámicas, más del 50% de las variantes detectadas en estos genes han sido identificadas una única vez (Tabla 1). Como consecuencia, el test genético de pacientes de alto riesgo en este tipo de poblaciones, requerirá el análisis completo de ambos genes. Por el contrario, en las poblaciones con mutaciones fundadoras de alta frecuencia se realizará el cribado específico de dichas mutaciones antes que proceder al más costoso cribado mutacional completo de los genes *BRCA1/2*.

Los análisis genéticos rutinarios de *BRCA1/2* suelen centrarse en el cribado mutacional de las regiones codificantes y las regiones intrónicas flanqueadas más el análisis de alteraciones en el número de copias para detectar LGR. Por ello, la mayoría de estudios rutinarios no detectan cambios localizados en regiones promotoras, regiones UTR y regiones intrónicas alejadas de las regiones codificantes. Varios estudios han mostrado que este tipo de variantes pueden ser de relevancia clínica^{126, 127}. Por ejemplo, un estudio reciente ha detectado patrones de *splicing* (véase Apartado 7 para el *Splicing*) aberrantes en pacientes de alto riesgo en los que, a priori, no fueron detectadas mutaciones en los genes *BRCA1/2* mediante análisis genéticos rutinarios y que podrían estar asociados a mutaciones en regiones intrónicas profundas. No obstante, en la actualidad se están empezando implantar otras aproximaciones de cribado que incluyen regiones no codificantes¹²⁸.

5. Clasificación clínica de variantes genéticas

Las variantes genéticas se pueden clasificar clínicamente en varios tipos según su riesgo asociado:

Variantes patogénicas

Son las que producen la pérdida de función de la proteína y, por lo tanto, confieren un gran aumento en el riesgo de sufrir cáncer de mama y/u ovario a lo largo de la vida. Generalmente, corresponden con mutaciones que dan lugar a una proteína truncada no funcional debido a la creación de un codón de terminación prematuro¹²⁹. Tal evento puede ser debido a mutaciones de *STOP*, *frameshift*, a mutaciones que afectan el correcto *splicing* de estos genes o a LGR¹²⁹. Además, existen mutaciones *missense* que pueden generar la pérdida de función de la proteína. Para el gen *BRCA1* la gran mayoría de mutaciones *missense* patogénicas han sido identificadas en el dominio RING y el dominio BRCT, mientras que para el gen *BRCA2* en el dominio DBD^{16, 55, 130-132}. De hecho, un estudio multifactorial basado en alineamientos filogenéticos sugiere que la probabilidad de que una variante *missense* localizada fuera de estos dominios sea patogénica, es menor al 1%¹³³.

Variantes de significado desconocido (VSD)

Se definen como variaciones genéticas de las que no se dispone de suficiente información (genética, funcional y/o clínica) para determinar si afectan a la función de la proteína codificada y, por tanto, no se puede predecir su riesgo asociado. Generalmente, suelen ser variaciones del tipo *missense*, *silent*, *indels in-frame*, variantes intrónicas o LGR que afectan a zonas de desconocida importancia. La proporción de VSD detectadas en los genes *BRCA1/2* es muy variable y depende tanto de la población estudiada como de la experiencia y la el número de pacientes del laboratorio que las reporte, siendo mayor para población no-caucásica que para la caucásica^{53, 56, 134}. Por ejemplo, Myriad Genetics reporta entre un 5-6% de VSD como resultados los análisis genéticos de *BRCA1/2* para población de origen caucásico² en Estados Unidos. En cambio, para población Afro-Americana reporta cerca del 21%^{53, 56, 134}. Por otra parte, la mayoría de los laboratorios Europeos identifican alrededor de un 15% de VSD para población caucásica^{53, 134-137}. Como se explicará en el *Apartado 9*, la incertidumbre ligada al efecto de estas variantes supone un verdadero problema desde el punto de vista del consejo genético.

Variantes Neutras

Son aquellas variantes que no afectan la función de la proteína y por lo tanto, no se considera que aumenten el riesgo a sufrir cáncer. En el caso del SCMOH se consideran variantes neutras aquellos polimorfismos que aparecen en más del 1% de la población general o aquellas variantes que han mostrado *odds* a favor de su neutralidad en los modelos multifactoriales de clasificación de variantes. En el caso de los polimorfismos frecuentes (>1%), en general, no suelen tener un valor patogénico por si solos pero no se puede descartar que estén asociadas a pequeño aumento en la susceptibilidad al cáncer, como sucede con los modificadores de riesgo citados anteriormente^{138, 139}.

Variantes recurrentes en los genes BRCA1/2

La mayoría de las mutaciones detectadas en los genes *BRCA1/2* son específicas de un solo individuo o familia^{16, 17}. Sin embargo, se han descrito ciertas mutaciones que aparecen en un alto porcentaje de familias. Normalmente, estas mutaciones, denominadas recurrentes, están generalmente asociadas a determinadas poblaciones aisladas, aunque también se han identificado en poblaciones más abiertas¹⁴⁰.

Dentro de las mutaciones recurrentes podemos diferenciar entre las recurrentes fundadoras, si provienen de un único alelo común, o recurrentes no fundadoras si provienen de varios eventos mutacionales independientes. La comparación de haplotipos entre familias portadoras de una misma mutación recurrente nos permitirá discernir si tienen un posible origen común o no¹⁴¹. En el caso de las recurrentes no fundadoras, se

presume que su alta frecuencia está ligada a *hot spots* de mutación, es decir, a zonas inestables dentro del gen¹⁴¹. Por el contrario, la alta frecuencia en que aparecen las mutaciones fundadoras se explica mediante el denominado efecto fundador. Se entiende como efecto fundador el proceso por el cual una mutación rara se convierte en frecuente, dentro de una determinada población, debido al aislamiento y consecuente endogamia. Dichas mutaciones fundadoras, aunque en inicio pueden ser específicas de una población, con el tiempo, debido a efectos migratorios, se pueden encontrar, con representación significativa, en otras poblaciones¹⁴². La frecuencia con la que aparece una cierta mutación fundadora puede variar mucho dependiendo de la mutación y de la población estudiada^{125, 142}. Por ejemplo, si consideramos la población de judíos Azkenazi, las tres mutaciones fundadoras más características de esta población están presentes un 2,6% de esta población [*BRCA1* c.68_69delAG (1%), c.5266dupC (0,13%) y *BRCA2* c.5946delT (1,52)]^{143, 144}. Estas tres mutaciones explican más del 90% de los casos de SCMOH asociados a los genes *BRCA1/2* en esta población. Además, estas mutaciones se han identificado en diversos países con una representación significativa. Por ejemplo, si nos fijamos en la mutación *BRCA1* c.5266dupC se ha identificado en Rusia, Bielorrusia, Polonia, Letonia, República Checa, Grecia o Lituania con frecuencias que corresponden al 94%, 73%, 60%, 55%, 37-52%, 46% y 34%, respectivamente, del total de las mutaciones detectadas en *BRCA1* en estos países¹²⁵.

La identificación de este tipo de mutación facilita el proceso de diagnóstico genético en las poblaciones afectadas¹⁴². Como se ha comentado anteriormente, en los casos en los que la frecuencia de la mutación recurrente es considerablemente alta, el cribado específico de dichas mutaciones es más rápido y coste/efectivo en relación al cribado completo de ambos genes^{143, 144}. Además, la alta frecuencia con que aparecen, permite estimar, de una manera más precisa, su penetrancia y permite diseñar estudios para identificar factores modificadores de riesgo¹⁴².

6. Modelos de clasificación de variantes. Esquema de clasificación de variantes propuesto por la IARC.

Muchas de las variantes raras identificadas en los genes *BRCA1/2* claramente inactivan la función de la proteína por causar directa o indirectamente su truncamiento. En base a ello, estas variantes pueden ser clasificadas como patogénicas partir de su primera observación en el test genético^{16, 17, 55, 145}. Sin embargo, el análisis de estos genes suele identificar un alto porcentaje de VSD, lo que tiene cierta repercusión en el seguimiento de los pacientes afectados^{48, 146}. Por ello, elucidar su riesgo asociado con la mayor brevedad posible es de suma importancia para un correcto seguimiento clínico.

Existen varias líneas de evidencia que pueden aportar información sobre el riesgo asociado a estas variantes (*Siguiente subapartado*)¹¹⁵. Sin embargo, todas ellas presentan limitaciones intrínsecas o asociadas a la baja frecuencia poblacional que presentan la mayor parte de VSD en *BRCA1/2*¹⁴⁷. La mayor parte se han identificado únicamente en una o muy pocas familias. Debido a ello, la información genética y clínica disponible sobre ellas es muy escasa. Por ello, esclarecer la patogenicidad de estas variantes a partir de aproximaciones basadas en una única línea de evidencia no suele ser posible^{148, 149}. De hecho, la “*International Agency for Research on Cancer (IARC)*, en su “*2008 IARC Workshop*”, recomendó tomar ciertas precauciones a la hora de interpretar la información obtenida a partir de este tipo de evidencias¹⁵².

Con el fin de mejorar esta situación, se han diseñado una serie de modelos multifactoriales que integran varias de estas líneas de evidencia, lo que ha supuesto una mejoría sustancial en la interpretación del riesgo asociado a una VSD^{124, 147, 148, 150, 151}.

Aproximaciones utilizadas para la clasificación de variantes

Frecuencia Poblacional

Hace referencia a la frecuencia con que una variante es identificada en la población general. En la mayor parte de laboratorios se clasifica una variante como polimorfismo y, por tanto, se considera que tiene bajo o nulo impacto sobre el riesgo de sufrir cáncer, cuando es observada en $\geq 1\%$ en la población general. El hecho de que una variante pueda aparecer con diferentes frecuencias dependiendo de la población, implica que, un estudio caso-control adecuado para clasificar una variante como neutra debería estar realizado con individuos de una misma población con clínica conocida y de gran magnitud para obtener una estadística robusta. Sin embargo, acceder a muestras de este tipo no es posible en muchos casos, lo que supone la principal limitación de este modelo¹⁴⁷. Por ello, esta línea de evidencia ha sido excluida de los modelos multifactoriales. Recientemente, han sido publicadas bases de datos que contienen información de exomas completos de un gran número de casos de población control, aunque no indican la información clínica de los individuos, con lo cual no se puede saber con exactitud si tienen una historia de SCMOH o no, lo que limita en cierta medida su uso en este contexto^{152, 153}. En tal caso se suele tener en cuenta la frecuencia de dicho síndrome en dicha población general y se asume que si la frecuencia de la variante en dicha población es mayor, no puede ser causante del síndrome, considerándola neutra. En estos casos es esencial contar con un número de casos suficiente para considerar esta diferencia significativa al nivel de confianza exigido. A este respecto, ENIGMA considera no patogénicas las variantes descritas con una frecuencia alélica igual o superior al 1% en

grandes grupos de referencia sin alta consanguineidad, como los de *1000 Genomes* y *Exome Variant Server*.

Análisis de cosegregación

El análisis de cosegregación determina si una variante está asociada a la enfermedad en una familia determinada. Debido a la existencia de fenocopias asociadas al SCMOH, como consecuencia de la alta incidencia del CM y CO esporádico, y a la penetrancia incompleta de las mutaciones patogénicas en *BRCA1* y *BRCA2*, el análisis de cosegregación basado en una única familia es raramente concluyente^{147, 154}. Por ello, será mucho más determinante cuando esté basado en la combinación de varias familias que compartan la misma variante. En estos casos, la cosegregación se puede analizar utilizando modelos estadísticos específicos^{154, 155}. Para ello, se requiere asumir unos valores de penetrancia y frecuencia que infieren de las variantes patogénicas descritas. En ello radica la principal limitación de este modelo, ya que implica suponer que todas las variantes patogénicas tienen una misma penetrancia y que la VSD objeto de estudio se ajusta a este patrón de penetrancia elevada, lo cual no es siempre así (variantes hipomórficas con riesgo moderado)^{147, 154}. Además, para obtener valores significativos se requiere una gran cantidad de información familiar por lo que se necesita la participación activa de múltiples familiares siendo difícil en muchos casos. Aún así, la cosegregación es considerada como la herramienta ideal (*gold standard*) para clasificar una variante¹⁴⁷. De ahí que las variantes clasificadas sin más evidencias que las aportadas por la cosegregación pueden ser utilizadas para calibrar los modelos multifactoriales o los ensayos funcionales en términos de especificidad, sensibilidad y valor predictivo¹⁴⁸.

Ponderación de la historia familiar y personal

Esta aproximación se basa en la información obtenida a partir de la gran cantidad de afectos que han sido estudiados en los últimos años y de los cuales se ha podido obtener su historia familiar y personal (edad del diagnóstico, tipo de tumor, número de afectados de primer y segundo grado con sus respectivas edades de diagnóstico y tipos de tumor)^{124, 150}. Esta información permite establecer patrones de agregación para mutaciones patogénicas, los cuales suelen ser muy severos, y para polimorfismos benignos, que suelen ser inocuos. El patrón establecido para una determinada VSD, obtenido a partir de varias familias independientes portadoras de dicha VSD, se comparará con los patrones para mutaciones patogénicas y para polimorfismos benignos. De este modo, se podrá estimar la probabilidad de que la VSD sea patogénica o neutra¹⁵⁶. Al igual que el modelo de cosegregación, supone una penetrancia elevada para todas las variantes patogénicas, lo que no sería válido para mutaciones hipomórficas. Además, las familias seleccionadas para establecer los patrones de cosegregación deben haber sido

recopiladas utilizando los mismos criterios de selección que la variante testada, lo que requiere acceso a grandes bases de datos. Así pues, este método es de mayor valía en laboratorios donde el volumen de afectos testados es muy grande. Por ejemplo, *Myriad Genetics* reportó 290 VSD reclasificadas entre 2011 y 2010 en base a este método, lo cual corresponde al 56% de las variantes que reclasificaron en ese periodo¹³⁴.

Concurrencia “in trans” con mutaciones patogénicas

Debido a los fenotipos tan agresivos asociados a mutaciones patogénicas en homocigosis en los genes *BRCA1/2*^{157, 158} cuando una VSD es detectada en un mismo paciente en concurrencia con una variante patogénica pero en distintos alelos (en *trans*), la probabilidad de que dicha VSD sea neutra es muy elevada. En el caso de *BRCA1*, excepto muy pocas excepciones^{85, 159}, mutaciones patogénicas bialélicas son letales a nivel embrionario¹⁶⁰. En cuanto a *BRCA2*, mutaciones patogénicas en homocigosis resultan en Anemia de Fanconi (FA), con un fenotipo reconocible clínicamente¹⁵⁸. Recientemente, mutaciones bialélicas en *BRCA1* también han sido asociadas con el fenotipo FA⁸⁵. Por tanto, la ausencia del fenotipo de FA en un portador de una mutación patogénica y otra VSD nos indicará que dicha VSD tiene una alta probabilidad de ser neutra. Cabe destacar que se han identificado casos de FA atenuados asociados a *BRCA2*, con un fenotipo muy leve, lo que limita esta aproximación⁸⁴. Además, el riesgo de las variantes hipómorficas no podrá ser interpretado mediante este tipo de análisis^{161, 162}.

Para que el modelo de concurrencia sea informativo, se requiere determinar, en primera estancia, que las variantes están en *trans*¹⁴⁷. Cuando están en el mismo alelo (en *cis*) no se podrá determinar el efecto de la VSD, ya que el portador posee una copia con las dos alteraciones a la vez. La forma más directa para determinar si dos mutaciones están en *cis* o en *trans*, es genotiparlas en los progenitores del probando o algún otro familiar¹⁴⁷. Si no es posible, existen otros métodos como el clonaje del DNA del paciente.

Conservación evolutiva entre especies

La alta conservación filogenética de una determinada región génica sugiere que dicha región será de vital importancia para mantener un correcto funcionamiento de la proteína^{133, 163}. Por ello, cualquier cambio aminoacídico en una región altamente conservada puede tener un impacto drástico en la función o estructura de la proteína. Por tanto, una VSD localizada en una región de este tipo, que introduce un cambio no observado en ninguna otra especie, tiene mayor probabilidad de tener un efecto deletéreo sobre la proteína. Por otra parte, si la VSD si está localizada en una región poco conservada entre especies, se considera que la repercusión de esta variante en la función de la proteína será mucho menor, aún menor cuando el aminoácido resultante ha sido observado en otra especie¹⁶³.

Múltiples algoritmos computacionales basados en el alineamiento de proteínas de múltiples especies han sido diseñados para evaluar el significado evolutivo/funcional de un cambio de aminoácido. Sin embargo, debido a la alta tasa de falsos positivos y falsos negativos, sobre todo en las regiones no conservadas, así como a la disparidad de resultados entre los distintos programas, no pueden ser utilizados como única evidencia para clasificar variantes. Si bien, se consideran una buena herramienta para priorizar variantes para futuros estudios funcionales o como evidencia complementaria en análisis multifactoriales^{134, 147, 164, 165}.

El resultado de estos programas está sujeto tanto a la calidad de la secuencia de las especies alineadas como a la adecuada selección de las especies utilizadas en el alineamiento¹⁶⁶. De este modo, regiones pobremente anotadas se consideran insuficientes para este tipo de análisis¹³⁴. La clasificación de un mayor número de variantes *missense* mediante modelos genéticos o epidemiológicos, es necesario para poder validar y/o entrenar dichos programas de predicción¹³⁴.

Entre los modelos computacionales desarrollados a este efecto destacan *Align-GVGD*¹³³, *Sorting Intolerant from Tolerant (SIFT)*¹⁶⁷, *Mutation Taster*¹⁶⁸ y *Polymorphism Phenotyping (PolyPhen 2)*¹⁶⁹.

Por otra parte, publicaciones recientes indican que evaluar el impacto físico-químico que una *missense* puede tener en el correcto plegamiento de la proteína también puede ser un buen predictor para determinar el efecto de dicha variante en la función de la proteína^{170, 171}. Sin embargo, este método está limitado a los dominios funcionales que han podido ser cristalizados.

Estudios Funcionales

Por ser uno de los objetos principales de esta tesis, la aportación de los ensayos funcionales en la clasificación de variantes se explicará ampliamente en el *Apartado 8*. No obstante, a modo introductorio cabe decir que dichos ensayos evalúan el impacto que una mutación determinada puede tener en una o varias de las diferentes funciones conocidas de un gen. Debido a que BRCA1/2 son proteínas multidominio y multifunción, la interpretación del riesgo asociado a una de sus variantes, en base al resultado de un ensayo funcional, es muy complejo¹⁷². Por ello, aunque pueden aportar una información muy valiosa, aún no han sido incluidos en los modelos multifactoriales.

Análisis de splicing

El *splicing* y el análisis variantes de *splicing* es otro de los temas principales de la presente tesis, con lo que se tratarán extensamente en el *Apartado 7*. A modo

introdutorio, cabe decir que existen una serie de variantes que pueden afectar al proceso de *splicing*¹⁷³. Una variante que afecte al *splicing* puede, en muchos casos, ser clasificada directamente como patogénica o probablemente patogénica en el esquema de clasificación propuesto por la IARC (Tabla 2)^{129, 174-178}. De ahí, que el análisis experimental a nivel transcripcional de estas variantes es una de las técnicas más utilizadas para la clasificación de VSD en laboratorios de diagnóstico.

Perfil histopatológico y pérdida de heterocigosidad en tumor

Las características histopatológicas de los tumores asociados a una determinada VSD en los genes *BRCA1/2* pueden ser comparadas con los datos publicados para los tumores esporádicos y con los asociados a variantes patogénicas en estos genes. De este modo, se puede determinar la posible asociación de la VSD con el tumor. No obstante, los tumores asociados a mutaciones en *BRCA1* y *BRCA2*, no son uniformes, lo que supone una de las mayores limitaciones de este modelo³⁰. Aún así, esta aproximación puede ser utilizada como valor predictivo para mutaciones en *BRCA1*, dado que los tumores asociados a este gen muestran unas características histopatológicas concretas y, en cierta medida, diferentes de los esporádicos^{151, 179}. En cambio, para *BRCA2*, debido a la heterogeneidad que presentan sus tumores asociados, esta aproximación no tiene un gran valor predictivo¹⁴⁷.

La pérdida de heterocigosidad (LOH) es uno de los mecanismos comunes de inactivación del alelo *wild type* de los genes *BRCA1/2* en tumores asociado a dichos genes. Por ello, esta característica ha sido comúnmente utilizada para clasificar VSD^{151, 179-181}. Sin embargo, un estudio reciente sugiere que este modelo puede no ser suficientemente robusto para evaluar todas las variantes *missense de BRCA1/2*, dado que pueden mostrar perfiles patológicos diferentes a las mutaciones truncantes¹⁸². De hecho, actualmente, los estudios de LOH han sido excluidos de los modelos multifactoriales actuales.

Los valores obtenidos a favor de la patogenicidad de una variante en base a los modelos histopatológicos y de LOH, suelen ser calculados a partir de pequeños conjuntos de tumores, los cuales incluyen en gran parte, casos asociados a mutaciones truncantes, en los que la representación de otros tipos de variantes patogénicas es reducida (*missense, del/ins in frame*)¹¹⁵. La posibilidad de que existan características distintivas entre diferentes tipos de variantes patogénicas sugiere que estudios a gran escala de histopatología de tumor asociados a variantes no truncantes patogénicas conocidas (posición y riesgo asociado) son necesarios para elucidar bajo qué circunstancias la LOH y otras características tumorales, pueden ser utilizadas como una herramienta útil para clasificar este tipo de variantes¹⁴⁷. Además, en la práctica, no se suele disponer de

suficientes tumores asociados a una misma variante como para poder clasificarla en base a este modelo.

Modelos multifactoriales de clasificación de variantes

Los modelos multifactoriales son algoritmos cuantitativos que estiman la probabilidad que una variante tiene de ser patogénica en base a la combinación de varias líneas de evidencia independientes. Como se ha visto, algunas líneas de evidencia han sido excluidas de los modelos multifactoriales actuales debido a sus limitaciones^{147, 183}. De este modo, las líneas de evidencia integradas actualmente en dichos modelos son: la conservación evolutiva del aminoácido afectado junto la predicción del impacto físico-químico introducido por el cambio (obtenidas ambas a partir de *Align-GVGD*)¹³³, la cosegregación de la variante con la enfermedad, la concurrencia con otras variantes, la historia familiar y personal del paciente, la histopatología del tumor y el efecto de la variante en el *splicing* del RNA^{131, 147, 150}.

El uso de estos modelos ha permitido clasificar un gran número de VSD en *BRCA1/2* de forma robusta^{124, 133, 134, 147, 148, 150}. No obstante, la rareza asociada a las VSD, implica que en muchos casos ni siquiera estos métodos obtienen el poder estadístico suficiente para clasificarlas. Se espera que la inclusión, en un futuro cercano, de la información funcional en los citados modelos aporte una gran mejoría en este sentido^{130, 132}.

La limitación principal de estos modelos, como muchos de los modelos anteriores, es la ponderación de los valores de patogenicidad en base, principalmente, a las características de variantes patogénicas truncantes. Lo que implica que variantes *missense* hipomórficas o de moderada penetrancia serán difícilmente clasificadas en base a estos modelos¹⁴⁷.

En definitiva, la principal limitación de la mayoría de los modelos explicados, a parte de la rareza asociada a las VSD, son los sistemas de referencia de patogenicidad utilizados, que pueden no ser válidos para clasificar todo tipo de variantes. Para mejorar estos modelos, se requeriría una validación de dichos sistemas de referencia en series grandes de mutaciones que contuvieran una representación adecuada de variantes no truncantes correctamente anotadas. No obstante, no es una aproximación fácil y requiere del trabajo colaborativo a nivel mundial, con el propósito de compartir datos y recursos. El ejemplo más representativo de es el consorcio *Evidence-Based Network for the Interpretation of Germline Mutant Alleles* (ENIGMA)¹⁴⁹ que se fundó para la interpretación conjunta de variantes en los genes *BRCA1/2*. Se espera que trabajos interdisciplinarios diseñados por consorcios como ENIGMA ayuden a subsanar las limitaciones citadas anteriormente.

Esquema de clasificación de variantes propuesto por la IARC

En el contexto clínico, los modelos multifactoriales son considerados la aproximación más aceptable para evaluar el riesgo asociado a una VSD^{148, 149}. Por ello, la IARC mediante su “*Unclassified Genetic Variants Working Group*” ha propuesto un sistema numérico de clasificación de variantes basado en dichos modelos (Tabla 2). Dicho sistema consta de cinco clases y cada una de ellas está ligada a unas recomendaciones para el asesoramiento clínico¹⁸⁴.

Tabla 2. Esquema de clasificación de variantes propuesto por el grupo de trabajo *Unclassified Genetic Variants Working Group* de la IARC.

Clase	Probabilidad de ser patogénica	Recomendaciones de cribado en familiares para el uso clínico	Recomendaciones para el asesoramiento genético	Recomendaciones de cribado en familiares para investigación
5 Patogénica	>0,99	Valor predictivo-Testar familiares	Protocolos de seguimiento establecidos para pacientes de alto riesgo	No indicado
4 Probablemente Patogénica	0,95-0,99	Valor predictivo-Testar familiares	Protocolos de seguimiento establecidos para pacientes de alto riesgo	Recomendable para clasificar la variante
3 Variante de significado desconocido	0,05-0,945	No usar la variante como valor predictivo de riesgo-No testar familiares*	Basado en historia familiar y otros factores de riesgo	Recomendable para clasificar la variante
2 Probablemente no patogénica o de baja significancia clínica	0,001-0,049	No usar la variante como valor predictivo de riesgo-No testar familiares*	Tratar como si no se hubiera encontrado la variante	Recomendable para clasificar la variante
1 No patogénica o de baja significancia clínica	<0,001	No usar la variante como valor predictivo de riesgo-No testar familiares*	Tratar como si no se hubiera encontrado la variante	No indicado

*Recomendar al paciente otras opciones de cribado como, por ejemplo, grandes reordenamientos. (Tabla traducida de Spurdle et. al, 2012)

El sistema propuesto por la IARC aporta una serie de avances en comparación a sistemas anteriores de clasificación, como por ejemplo el sistema de 3 clases (Patogénica, VSD y Neutra/Bajo Impacto Clínico) que se utilizaba hasta ese momento¹⁸⁵. El sistema propuesto por la IARC introduce dos nuevas clases. Estas nuevas clases, Clase 4 o probablemente patogénica y Clase 2 o probablemente neutra, permiten discernir entre las variantes que realmente tienen un significado desconocido (VSD o Clase 3) y aquellas de las cuales hay evidencias significativas pero no irrefutables sobre su patogenicidad o

neutralidad, permitiendo así un asesoramiento clínico más apropiado a los portadores de estas variantes. Además, sugiere cuándo y cómo realizar el test predictivo en familiares del probando así como el asesoramiento clínico de los pacientes con estas variantes.

Cabe destacar que, la IARC ha publicado una base de datos dónde se recopila toda la información reportada en la literatura para más de 1000 variantes en los genes *BRCA1/2*, incluyendo la patogenicidad asignada a cada variante en cada publicación (<http://chromium.liacs.nl/LOVD2/cancer/home.php>)¹⁴⁷. Además ha creado una base de datos que recoge las VSD que han sido reclasificadas mediante modelos multifactoriales, indicando su nueva clasificación (<http://brca.iarc.fr/LOVD/home.php>)¹⁸⁶.

7. Mutaciones genéticas que alteran el correcto *splicing* del RNA

Conceptos básicos sobre el splicing

El *splicing* consiste en un mecanismo co-transcripcional mediante el cual, a partir de un transcrito de RNA primario, resultante de la transcripción de toda la secuencia de un gen, se obtiene un RNA maduro (mRNA). Para ello, se eliminan de las regiones no codificantes (intrones) del transcrito primario y el empalman las regiones codificantes (exones)^{187, 188}. El mRNA es el que servirá como plantilla para la síntesis de proteínas¹⁸⁹. Aunque el *splicing* es un proceso común a cualquier tipo celular, es mucho más común en células eucariotas¹⁹⁰.

En el *splicing* intervienen elemento en *cis* y en *trans*. Los elementos en *cis* son aquellos localizados en la secuencia del transcrito primario que va a ser procesado. Corresponden a secuencias reguladoras, secuencias de reconocimiento de puntos de corte y los puntos de corte en sí mismos (Fig. 8). Los elementos en *trans*, la maquinaria de *splicing* propiamente dicha, consisten en una serie de complejos proteicos y factores de transcripción que se encargan del proceso de corte y empalme del RNA primario. Los elementos en *trans* reconocen a los elementos en *cis* para realizar su función^{191, 192}.

Los puntos de corte o sitios de *splice* (SS, *splice sites*)¹⁹² son secuencias dinucleótidas localizadas en los extremos 5' y 3' de cada intrón, delimitando la frontera entre exones e intrones. Son secuencias universalmente conservadas y se denominan SS donador y aceptor, respectivamente. Generalmente, el sitio donador (D) consiste en un dinucleótido GU y el sitio aceptor (A) con un dinucleótido AG (Fig. 8)¹⁷³. Las secuencias colindantes a los SS, aunque presentan una gran degeneración, suelen ser del tipo **GURUGU** en el sitio D y (Y)n**CAG** en el sitio A, donde (Y)n representa una secuencia polipirimidínica y R cualquier purina¹⁹². Además, en la región 3' del intrón, existe otra secuencia conservada de reconocimiento, el denominado punto de ramificación (Fig. 8)¹⁹³. El punto de ramificación, se encuentra en cualquier lugar entre los nucleótidos 20 y

50 de la zona anterior a la A del sitio aceptor y se caracteriza por poseer una adenina universalmente conservada dentro de una secuencia del tipo YNYYRAY, donde Y indica una pirimidina, N cualquier nucleótido, R cualquier purina, y A la adenina. Cabe mencionar, que han sido descritos lugares SS distintos a los citados, aunque solo representan el 1% de los casos. Corresponden a la secuencia AU para la región 5' y AC para la región 3' (ref. ¹⁹²).

Figura 8. Esquema mostrando los elementos en cis de reconocimiento de splicing

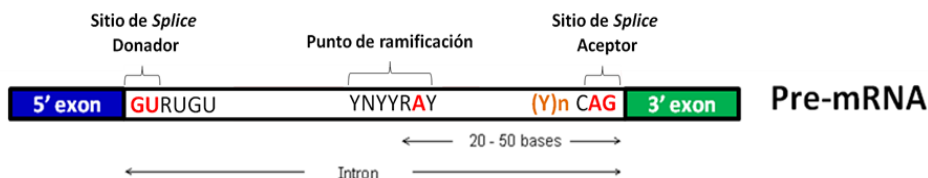


Figura 8. En este esquema se representan los diferentes sitios en *cis* involucrados en el splicing. El sitio donador de *splicing* (GU), el sitio aceptor de *splicing* (AG) y la adenina del punto de ramificación están marcados en rojo. En amarillo, la secuencia polipirimidínica colindante al sitio aceptor. Y, pirimidina; N, cualquier nucleótido; R, cualquier purina.

En cuanto a los elementos en *trans*, el más relevantes es el complejo proteico denominado *spliceosome*^{194, 195}. Consiste en un complejo dinámico formado por 5 subunidades que van cambiando durante el proceso de maduración del RNA (Fig. 9). Las 5 subunidades, denominadas U1, U2, U4, U5 y U6, están compuestas por ribonucleoproteínas nucleares pequeñas o snRNP. Estas proteínas se caracterizan por contener una secuencia de RNA que utilizan para unirse a los intrones durante el proceso de *splicing* por complementariedad de bases¹⁹⁶. Otros elementos en *trans* implicados en la regulación del *splicing* son los diferentes factores de transcripción encargados de regular el *splicing*. Entre ellos, las proteínas SR (*serine-arginine rich*) y las ribonucleoproteínas heterogéneas nucleares o hnRNPs (*heterogeneous nuclear RNPs*).

Mecanismo general de splicing

El mecanismo de *splicing* consta de dos etapas caracterizadas por dos reacciones de trans-esterificación catalizadas por el *spliceosome*¹⁹⁶⁻¹⁹⁹. Inicialmente, tiene lugar la reclusión de las distintas subunidades del *spliceosome* en la región intrónica, seguida de una serie de cambios conformacionales de éste que darán lugar al *spliceosome* en su forma catalítica activa. Para ello, en primer lugar la subunidad U1 reconoce el sitio D por complementariedad de bases con la secuencia de RNA (Fig. 9). Al mismo tiempo, la proteína SF1 (*Splicing Factor 1*) se une al punto de ramificación y el factor auxiliar heterodimérico U2AF65/35 se une al sitio A (AG) y a la secuencia de polipirimidinas colindante, mediante su subunidad menor y subunidad mayor, respectivamente. Al complejo formado se le denomina estadio temprano del *spliceosome* o complejo E. A

continuación, la subunidad U2, también por complementariedad de bases, se une a la región de ramificación, desplazando a SF1. El complejo así formado se denomina complejo A. Posteriormente, son recluidas las subunidades U4/U6 y U5 formando el complejo B. Finalmente, se liberan las subunidades U1 y U4 produciéndose la activación catalítica del *spliceosome* (complejo C) y dando lugar a la primera reacción de transesterificación. En esta reacción, la adenina conservada de la región de ramificación, mediante su grupo 2' OH, ataca al grupo fosfato de la guanina inicial del sitio D. Como consecuencia, el exón situado en la región 5' queda liberado y el intrón queda formando un lazo, dando lugar al denominado complejo intermedio C1. Inmediatamente tiene lugar un nuevo reordenamiento conformacional del *spliceosome* que dará lugar a la segunda reacción de trans-esferificación²⁰⁰. En esta segunda reacción, el grupo 3' OH libre del exón 5' ataca el grupo fosfato del primer nucleótido del exón adyacente (3'), generándose una unión covalente entre ambos exones y la liberación del intrón junto con las subunidades U2, U5 y U6 del *spliceosome*. Las secuencias consenso minoritarias 5'AU, 3'AC dan lugar a un mecanismo de *splicing* similar al descrito, mediante el conocido como *spliceosome menor*²⁰¹.

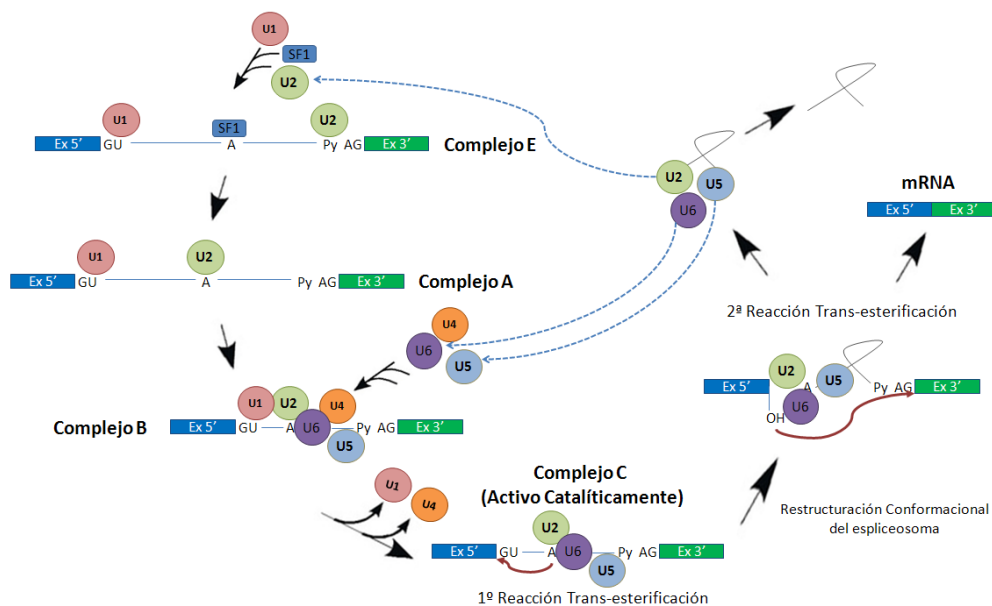
Figura 9. Esquema representando el mecanismo de *splicing* de pre-mRNA

Figura 9. Esquema representando las diferentes etapas del proceso de *splicing*. El *splicing* sucede en dos etapas principales, caracterizadas por dos reacciones de trans-esterificación consecutivas catalizadas por el *spliceosome*. Durante el *splicing* las diferentes subunidades del *spliceosome* (U1, U2, U4, U5, U6) son reclutadas gradualmente. Antes de que tenga lugar la reacción de trans-esterificación, una serie de cambios conformacionales de estas subunidades darán lugar al *spliceosome* en su forma catalíticamente activa. Mediante estas reacciones, se cortan y eliminan las regiones intrónicas y se empalman las exónicas, dando lugar a un RNA maduro a partir de un transcrito primario que contiene toda la secuencia complementaria al DNA de un gen. (Figura adaptada de la página web del grupo de investigación de Hennig Urlaub, <http://www.mpibpc.mpg.de/289605/07-spliceosome>).

Regulación del *splicing*. *Splicing alternativo*.

El *splicing* alternativo (SA) es el mecanismo mediante el cual se producen varios mRNA a partir de un mismo transcrito primario. Es debido a la gran degeneración de los elementos de reconocimiento de *splicing* en células eucariotas, que posibilita la inclusión diferencial de exones durante el procesamiento del pre-mRNA^{187, 202-204}. Como consecuencia del SA, para muchos genes, el número de transcritos alternativos generados es muy amplio. La expresión de los diferentes transcritos posibles estará determinada por factores como el tipo de tejido, estímulos intra/extra celulares o las distintas etapas del desarrollo^{203, 204}. Concretamente, para los genes *BRCA1/2*, estudios recientes han identificado 73 y 24 transcritos alternativos en sangre, número que no varía de forma significativa en tejido mamario (ref^{205, 206} y datos no publicados de ENIGMA).

No obstante, el *splicing* consta de un sistema de regulación interno que asegura que una serie de transcritos alternativos se produzcan de forma predominante. Dicho sistema consta de una serie de secuencias reguladoras en *cis*, que interactúan con factores de transcripción para promover (activadoras) o reprimir (silenciadoras) la

inclusión de un determinado exón en el mRNA^{194, 195}. Dichos elementos se pueden encontrar tanto en regiones intrónicas como exónicas y se denominan *Exonic/Intronic Splicing Enhancers* (ESE o ISE) o *Exonic/Intronic Splicing Silencers* (ESS o ISS). Entre los factores de transcripción que interactúan con ellos destacan las proteínas SR, que suelen actuar como activadoras y las proteínas hnRNP, que suelen actuar como represoras²⁰⁷⁻²¹⁰.

Variantes que provocan defectos en el sistema de regulación del *splicing* han sido asociadas con varios tipos de enfermedades, entre ellas diferentes tipos de enfermedades hereditarias²¹¹⁻²¹³.

Alteraciones de *splicing* asociadas al SCMOH

Hay estudios que indican que el alrededor del 60% de las mutaciones asociadas a enfermedades hereditarias causan alteraciones del *splicing*²¹⁴. Además, la lista de mutaciones puntuales vinculadas a defectos de *splicing* en este tipo de enfermedades es creciente. De hecho, se han reportado más de 80 estudios relacionados con variantes que afectan el *splicing* de *BRCA1/2*²¹⁵. Debido a ello, el evaluar el impacto biológico de variantes puntuales en el *splicing* de genes de predisposición, se ha convertido en una práctica habitual en el diagnóstico clínico^{174, 195, 211, 214, 216}.

Las mutaciones de *splicing* pueden afectar a los componentes de la maquinaria *splicing* (mutaciones en *trans*) o, a la secuencia del gen que va a ser procesado (mutaciones en *cis*)²⁸⁴. En el caso del SCMOH las mutaciones de *splicing* son variantes en *cis* en los genes *BRCA1/2*.

Las variantes de *splicing* en *cis* pueden ser de diferente naturaleza. Dependiendo de su ubicación, algunas son consideradas directamente como probablemente patogénicas por afectar de forma clara al *splicing*²¹⁵, son las variantes que afectan a los dinucleótidos invariables de los sitios donador y aceptor de *splicing*. Sin embargo, se han reportado otro tipo de variantes cuyo efecto en el *splicing* es difícil de predecir a nivel genómico^{214, 217, 218}. Serían variantes que generan un SS *de novo*, que activan SS crípticos, o que alteran secuencias reguladoras (activadoras, silenciadoras, punto de ramificación o la secuencia polypyrimidínica). Para determinar el efecto de estas variantes en el *splicing* se requiere realizar ensayos funcionales a nivel de RNA¹²⁹.

Priorización de variantes para realizar estudios de RNA

Idealmente, en el contexto clínico, se debería determinar el efecto en *splicing* de todas las VSD identificadas. Sin embargo, la alta frecuencia con que estas variantes aparecen en el test genético de *BRCA1/2* hace que no sea logísticamente posible en todos

los laboratorios¹⁷⁶. En este sentido, el uso de los programas de predicción *in silico*, que permiten determinar la probabilidad de que una variante afecte el *splicing*¹²⁹, han mostrado ser una buena herramienta para priorizar que variantes tienen más posibilidades de alterar el procesamiento correcto del mRNA y deberían ser estudiadas a nivel funcional¹⁷⁶. Existen dos tipos de programas de predicción *in silico*, los que permiten analizar el efecto de una variante en el reconocimiento de los SS y los que permiten analizarlos en las secuencias reguladoras ESE/ISE y ESS/ISS²¹⁹⁻¹²⁹. Los programas de reconocimiento de SS, permiten evaluar los SS canónicos, SS crípticos o predecir la formación de un SS *de novo*. Estos programas han mostrado ser muy eficientes para predecir alteraciones en los SS canónicos, pero presentan ciertas carencias a la hora de reconocer SS crípticos o *de novo*¹²⁹. Los programas de reconocimiento de secuencias reguladoras, reconocen múltiples secuencias que potencialmente puede interaccionar con los factores de *splicing* pero no determinan cuales de ellas serán utilizadas¹²⁹. Además, el hecho de que estas secuencias estén muy degeneradas y muchas de ellas pueden ser no funcionales sugiere que los datos *in silico* obtenidos a partir de este tipo de programas debe ser tratados con mucha precaución¹⁹⁴.

Numerosos estudios han comparado los programas de predicción entre ellos, observándose discrepancias en muchos casos¹⁰⁴. Por ello, es preferible combinar varios de ellos cuando se quiere seleccionar variantes para el realizar ensayos de RNA. Algunos estudios han comparado con los resultados de los ensayos de RNA^{129, 175}. También en estos estudios se han encontrado varias discrepancias, especialmente en las variantes que afectan a las secuencias reguladoras. Esto indica la necesidad de una mayor validación de estos programas, especialmente cuando se espera que pronto sean integrados en los modelos multifactoriales^{129, 175, 176, 215, 220}.

Análisis funcional de variantes que pueden afectar el *splicing*

Los estudios *in vitro* a nivel de RNA, suponen una manera fácil y robusta para determinar el efecto de una variante sobre el *splicing*¹⁷⁶. Además, como se ha comentado anteriormente, la variante puede ser clasificada directamente en función de este resultado, lo que hace estos estudios aún más valiosos a nivel clínico^{149, 184}.

Los ensayos de RNA más utilizados son los basados en la extracción de RNA a partir de sangre del paciente, posterior amplificación del cDNA mediante PCR de retro-transcripción (RT-PCR) y secuenciación del producto de PCR para identificar la presencia de posibles transcritos aberrantes^{129, 176, 221}. No obstante, existen otras metodologías de detección como: la PCR cuantitativa, electroforesis en geles de agarosa, electroforesis capilar o ensayo de desequilibrio alélico²²²⁻²²⁴. Estas aproximaciones se pueden utilizar por si solas o en combinación.

Es importante recalcar que una gran proporción de los transcrito alternativos que generen un CTP son naturalmente eliminados por la célula para reducir la posibilidad de errores en la expresión génica mediante un mecanismo denominado *Nonsense-Mediated Decay* (NMD)²²⁵. Por ello, para la detección de este tipo de transcritos en los ensayos de RNA, es necesario utilizar inhibidores de este sistema, como por ejemplo, la puomicina o la ciclohexamida^{129, 176, 221}.

Como ya se ha dicho anteriormente, los patrones de expresión de los transcritos puede ser tejido específicos, por ello, idealmente, la extracción de RNA debería ser realizada a partir del tejido afectado¹⁷⁴. No obstante, el hecho de que no existan diferencias significativas entre los patrones de expresión sangre/mama para los genes *BRCA1/2*, indica que el uso de sangre para los estudios de RNA en estos genes supone una buena aproximación (ref^{205, 206} y datos no publicados de ENIGMA). Esto conlleva una gran ventaja, ya que en la mayoría de los casos no se dispone de tejido mamario del paciente. En otras situaciones, no se dispone de ningún tipo de tejido del paciente. En estos casos, se puede diseñar un mini-gen que englobe la región afectada (exon afectado y exones/intrones colindantes), transfectarlo en un modelo celular y extraer el RNA expresado a partir del mini-gen^{176, 226}. En este sentido, el ensayo funcional basado en células embrionarias de ratón aporta grandes ventajas, dado que permite, el análisis del gen en toda su amplitud²²⁷.

Dificultades en la interpretación de los resultados de los ensayos de RNA de BRCA1/2

Una de las principales limitaciones en la interpretación de los análisis de RNA de *BRCA1/2* es el gran número transcritos alternativos que se generan de forma constitutiva en estos genes^{175, 205, 215, 224}. El uso de un amplio número de controles sin mutación en estos ensayos, es esencial para poder discriminar entre transcritos alternativos constitutivos y los causados por la variante objeto de estudio²²⁴. Aún así, existen variantes cuyo efecto sólo varía la proporción de transcritos alternativos constitutivos o genera transcritos aberrantes de forma parcial²¹⁵. En estos casos, en términos de relevancia clínica, es muy difícil interpretar el efecto de la variante, dado que no está bien definido como pueden afectar los cambios en los niveles de expresión de los diferentes transcritos en la función de la proteína. Para determinar el efecto de estas variantes, sería necesario realizar ensayos funcionales adicionales que evalúen su efecto en la función de la proteína²²⁸. Además, es necesario diseñar estudios multicéntricos a gran escala que permitan establecer la correlación entre las proporciones de transcritos aberrantes y la clínica de los pacientes²²⁴.

Por otra parte, el uso de diferentes metodologías de detección puede llevar a diferentes resultados, lo que dificulta aún más la interpretación de estos análisis²²⁴. Con el

fin de detectar qué factores son críticos en esta variabilidad y estandarizar los protocolos utilizados, el consorcio ENIGMA, con su grupo *Splicing Working Group*, realizó un estudio comparativo entre los ensayos de RNA realizados en diferentes laboratorios. De este estudio se concluyó que, el diseño de la PCR (localización de primer, tiempo de elongación) y la metodología de detección del producto de PCR junto con un conocimiento previo de los transcritos alternativos constitutivos, son los factores más relevantes a la hora de realizar un ensayo de RNA²²⁴.

8. Ensayos funcionales en *BRCA1* y *BRCA2*

El uso de los ensayos funcionales tiene como principal propósito conseguir información adicional sobre una VSD para poder clasificarla clínicamente²²⁹. En un sentido estricto, debido a las limitaciones asociadas a su interpretación, la información funcional únicamente puede utilizarse para evaluar el efecto de una variante en una determinada función de la proteína y no para determinar su patogenicidad o fenotipo clínico^{51, 230}. Para esto último hará falta completar la información funcional con información proveniente de otras líneas de evidencia. De hecho, la información funcional ha demostrado ser de un gran valor añadido para interpretar la patogenicidad de una variante cuando se utiliza como información suplementaria a evidencias aportadas por otras fuentes^{131, 170, 231, 232}. Aún así, en un contexto en el que la baja frecuencia poblacional de muchas VSD hace que su clasificación mediante modelos epidemiológico, genéticos o multifactoriales sea prácticamente imposible, la información funcional es de gran relevancia¹⁴⁷. Ya que, es de suponer que, tanto los pacientes como los clínicos pueden preferir tomar decisiones clínicas en base a una información no inequívoca sobre el riesgo asociado a una variante, estimada en base a su efecto funcional, antes que no utilizar ninguna información^{124, 150, 172, 184, 233}. Por ello, la IARC establece que la información cualitativa proporcionada por los ensayos funcionales puede ser categorizada según el sistema de clasificación que ellos mismos proponen, lo que permite hacer uso de las recomendaciones clínicas pertinentes a cada clase¹⁸⁴. Si bien, consideran necesario que dicha información esté adecuadamente contrastada mediante grupos de especialistas con un amplio conocimiento de las funciones del gen y de las vías en las que está implicado. Aunque la decisión final de utilizar esta información con fines clínicos debe ser abordada conjuntamente por el panel de especialistas de laboratorio, facultativos clínicos y por los propios pacientes. No obstante, la decisión final dependerá de si los pacientes se sienten cómodos utilizando esta información, aún sabiendo que puede ser incorrecta¹⁷². Por ello, es mandatorio que los pacientes entiendan completamente las limitaciones del ensayo utilizado, los riesgos asociados a una interpretación errónea del resultado y del significado de cada una de las clases en el esquema de clasificación de variantes¹⁷². Además, los clínicos y asesores genéticos necesitarán un entrenamiento específico para interpretar estos resultados funcionales y saber cómo entregar la información al paciente. En este sentido, con el fin

de facilitar el proceso de análisis e interpretación de los datos funcionales, recientemente se ha publicado en BIC una herramienta bioinformática interactiva que proporciona toda la información funcional disponible sobre cada variante identificada en *BRCA1* (<https://research.nhgri.nih.gov/projects/bic/circos/index.shtml>)²³⁰.

De todos modos, la información funcional adquiriría mucha más relevancia clínica si fuera incluida en los modelos estadísticos multifactoriales de clasificación de variantes¹³². Con ello se conseguiría una gran mejoría de estos modelos y se aceleraría, en gran medida, la clasificación de las VSD. Sin embargo, debido a la complejidad asociada a la interpretación de los ensayos funcionales diseñados para las proteínas *BRCA1/2*, la información funcional que éstos aportan aún no ha podido ser integrada en los citados modelos.

Problemática de la inclusión de los ensayos funcionales en los modelos multifactoriales

La difícil interpretación de los resultados funcionales deriva, principalmente, de que *BRCA1/2* son genes que codifican para proteínas con múltiples funciones y no se conoce con exactitud cuáles de ellas contribuyen a su función supresora de tumores. Así, una variante no se podrá considerar como patogénica únicamente en base al resultado de un ensayo funcional, si no se está completamente seguro que la pérdida de la función testada inactivará su función supresora de tumores^{131, 132, 172}. Cada variante debe ser analizada con un ensayo, o ensayos, diseñados para el dominio en el que esté localizada, de otra forma, se podría caer en el error de clasificarla como neutra, cuando se está analizando una función en la que el aminoácido afectado no está involucrado¹⁷².

Por otra parte, los ensayos funcionales pueden identificar variantes hipomórficas, las cuales pueden tener efectos intermedios o moderados sobre el riesgo a sufrir cáncer^{131, 161, 234}. Estas variantes no responden plenamente a un fenotipo deficiente y pueden mostrar resultados discordantes en diferentes ensayos funcionales, lo que dificulta la interpretación de su impacto en la función supresora. Como se ha comentado, este tipo de variantes no tienen cabida en los modelos multifactorial actuales, ya que éstos parten de la hipótesis que una variantes es patogénica con penetrancia elevada o neutra y no consideran un efecto intermedio^{124, 150}.

El primer paso para que la información proporcionada por un ensayo funcional pueda ser introducida en los modelos multifactoriales es determinar el valor predictivo de dicho ensayo. Para ello, es esencial que sean correctamente validados, haciendo uso de un amplio panel adecuado de variantes patogénicas y neutras conocidas, para establecer su sensibilidad y especificidad¹⁷². Las variantes utilizadas en la serie de validación deben ser variantes que hayan sido clasificadas previamente mediante modelos basados en evidencias genéticas directas no funcionales (por ejemplo, historia familiar, cosegregación

o concurrencia). Desafortunadamente, muchos de los ensayos funcionales actuales aún no han podido ser adecuadamente validados^{130, 233}. Si bien, los grandes esfuerzos realizados en este campo, han dado sus primeros frutos. Recientemente, se han publicado los primeros modelos computacionales que permiten convertir los resultados de determinados ensayos funcional en valores de probabilidad, que pueden ser integrados en los modelos multifactoriales^{131, 132, 231, 232}.

Consideraciones técnicas en la realización de ensayos funcionales

Además de las limitaciones intrínsecas en la interpretación de la información funcional, los ensayos funcionales presentan una serie de limitaciones técnicas que pueden llevar a una interpretación errónea de los resultados. Para evitar caer en errores de interpretación hay que tener en cuenta una serie de consideraciones cuando se realice un ensayo funcional^{130, 233}:

- En el caso de detección de variantes hipomórficas, cuando la especificidad, sensibilidad y valor predictivo del ensayo no hayan sido adecuadamente validados y los límites de patogenicidad/neutralidad no hayan sido establecidos, las variantes cuyo efecto sea intermedio en comparación al de los controles patogénicos y neutros se consideran de efecto desconocido y el resultado no tendrá valor predictivo.
- Se deben usar controles adecuados en cada experimento. Al menos uno variante neutra y una mutación patogénica conocidas. Estos controles deben estar localizados en el dominio interrogado por el ensayo funcional.
- Se deben incluir triplicados de cada muestra en cada experimento independiente y cada experimento independiente debe ser repetido al menos dos veces.
- Siempre que los ensayos se realizan utilizando métodos de transfección transitoria, la eficiencia de transfección debe ser medida y normalizada para cada plásmido transfectado. Además, las transfección transitoria da lugar a niveles de expresión de proteína muy variables, en comparación con transfecciones estables. Por ello, para asegurar una correcta interpretación, es esencial examinar si las proteínas se expresan en un nivel similar al fisiológico.
- Al realizar experimentos de transfección transitoria, en el mismo experimento se debe utilizar una única preparación de plásmido. Esto permitirá evitar la variabilidad causada por diferencias en la calidad del plásmido.

- El uso de proteínas parciales puede conducir a errores de interpretación de la relevancia clínica de ciertas variantes. Por lo tanto, es claramente preferible el uso de ensayos que consideren la proteína completa, a pesar de los problemas asociados con la detección de proteínas de gran tamaño como las BRCA1/2. En este sentido, los grandes avances que se están dando en el campo de la ingeniería genética permiten, actualmente, introducir la variante en el gen endógeno, lo que conllevará una gran mejora en la interpretación de los resultados^{235, 236}.

Ensayos funcionales comunes para BRCA1 y BRCA2

Ensayo de localización subcelular

Las proteínas BRCA1/2 desarrollan su función supresora en el núcleo^{237, 238}. Por tanto, su deslocalización en otras regiones de la célula, debido al efecto de una variante, puede derivar en una alteración de su función. Para examinar si estas proteínas están deslocalizadas a causa de una variante, se transfecta transitoriamente un plásmido *GFP-BRCA1/2* en células MCF7 (para BRCA1) y V-C8 o HEK293T (para BRCA2). Los vectores *GFP* (Proteína verde Fluorescente) emiten fluorescencia, lo que permite cuantificar los *focis* BRCA1/2 nucleares formados en respuesta a un daño en el DNA mediante microscopio de fluorescencia. Para ello, las células transgénicas (*BRCA2* variante y *BRCA2 wild type*) se tratan con radiación IR y se compara el número de *foci* BRCA1/2 producidos por las variantes con respecto a lo producidos por *BRCA2 wild -type (wt)*²³⁹⁻²⁴².

La ponderación de este ensayo resulta complicada debido al uso de plásmidos que incluyen una gran variabilidad en la expresión de la proteína, lo que puede afectar a su deslocalización¹³⁰.

Ensayo funcional basado en el uso de células madre de ratón embrionarias (Células-RE)

El ensayo funcional basado en células-RE se basa en la observación de que *BRCA1/2* funcional es esencial para la viabilidad de estas células cuando son deficientes en *Brca2*^{243, 244}. Para realizar este ensayo se utilizan células-RE genéticamente modificadas. Estas células contienen un alelo no funcional de *Brca1/2* y otro que puede ser deletado mediante recombinación inducida por Cre. Para realizar este ensayo, se se transfecta un BAC que contiene *BRCA1/2* humano *wt* o portando las mutaciones objeto de estudio en dichas células, de modo que, la capacidad que tenga cada variante para rescatarlas cuando se induce la delección del alelo condicional de *Brca1/2*, servirá para evaluar el efecto de la variante en la función de *BRCA1/2*. Las variantes que no sean capaces de rescatar dichas células serán consideradas patológicas y las que sean capaces,

se analizaran mediante una serie de ensayos biológicos relacionados con las funciones de los genes *BRCA1/2* (sensibilidad a radiación IR, sensibilidad a drogas que dañan el DNA, recombinación homóloga o estabilidad cromosómica). Las variantes que son deficientes para alguno de estos ensayos se consideran hipomórficas^{228, 234, 245}. Este sistema presenta dos ventajas principales respecto a otros ensayos funcionales. Por una parte, las células embrionarias no son propensas a la inestabilidad genómica que caracteriza a la mayor parte de líneas celulares cancerígenas utilizadas en otros ensayos. Por otra, utilizan un BAC que contiene todo el gen *BRCA1/2*. Esto asegura niveles de expresión de *BRCA1/2* similares a los fisiológicos, evitando la sobreexpresión ligada al uso de plásmido. Por otra parte, permite testar variantes en el contexto de expresión de la proteína completa, a diferencia de muchos ensayos de compensación que se basan en la expresión del cDNA de dominios funcionales puntuales, lo que puede llevar a conclusiones erróneas. Además, el uso de BAC permite testar zonas reguladoras o regiones intrónicas haciendo de este sistema una buena aproximación para analizar variantes que pueden afectar al *splicing*²²⁷.

Debido a las ventajas que este ensayo aporta, fué el elegido en la presente tesis para testar una serie de VSD posiblemente asociadas al SCMOH. La metodología seguida para realizar este ensayo se describe detalladamente en el *Capítulo 2* de resultados: “Análisis funcional de variantes de significado desconocido en *BRCA2* mediante un ensayo basado en células embrionarias de ratón”.

Ensayo de recombinación homóloga dirigida (HDR)

Para testar el efecto de una variante puede tener en la capacidad de *BRCA1/2* de realizar la RH, se generó un constructo de recombinación reportero *DR-GFP* es transfectado de forma estable en células deficientes en *BRCA1/2*^{75, 242, 246-249}. El *DR-GFP* se compone de dos repeticiones *GFP* diferencialmente mutadas (*SceGFP* y *iGFP*) y separados por un gen de resistencia a la puomicina (3,7Kb). En una de las repeticiones *GFP*, *Sce-GFP*, se generó un sitio de corte para la encima de restricción *I-Sce I*, de modo que, cuando *I-Sce I* es expresada en estas células se producirá la rotura de doble cadena en *Sce-GFP*. Además, el sitio de corte *I-Sce I* introducido consta de dos codones STOP *in-frame*, por tanto, *Sce-GFP* es inactivo. Por otra parte, *iGFP* está truncado en 5' y 3', de modo que es también inactivo, de manera que las células no expresan *GFP*. Cuando *Sce-GFP* es reparado mediante RH utilizando el *iGFP* como plantilla, éste será funcional y las células expresarán *GFP*. La cuantificación de células *GFP* positivas se mide mediante clitómetro. Este ensayo ha sido correctamente validado mostrando un 100% de especificidad y sensibilidad utilizando 18 mutaciones no patogénicas y 13 patogénicas conocidas. A raíz de ahí, se pudo desarrollar un método estadístico de clasificación de

VSD basado en el resultado de este ensayo¹³¹. Estos valores de probabilidad podrán ser incluidos en los modelos multifactoriales¹³².

Ensayo de recombinación homóloga espontánea

El ensayo descrito anteriormente fué diseñado para testar variantes en la RH inducida. No obstante, existe otro ensayo similar para medir el efecto de variantes en estos genes en la RH espontánea^{250, 251}.

Ensayo de recombinación homóloga espontánea utiliza el modelo celular HeLaG1, que contiene dos alelos de resistencia a higromicina (Hygr) integrados de forma estable e inactivada mediante mutaciones en lugares diferentes. Un evento de RH espontáneo intracromosomal puede dar lugar a un alelo activo de Hygr. La frecuencia de este tipo de eventos espontáneos se puede medir como el número de células viables después de la selección con Hygr respecto al número de células sembradas. Para medir el efecto de una variante en la RH espontánea, se transfecta en estas células un plásmido que contiene el cDNA de *BRCA1/2* (wt o variante), se siembran las células y se seleccionan con Hygr durante dos semanas. Las colonias resistentes a Hygr obtenidas se tiñen, se cuantifican y se comparan wt Vs variante.

Ensayo de recombinación en levadura

Este ensayo evalúa el efecto de la expresión de las proteínas BRCA1/2, portando la VSD, sobre la RH en la cepa diploide de levadura *Saccharomyces cerevisiae*^{252, 253}. Este sistema mide, por una parte, la RH intracromosomal mediante dos alelos del gen *HIS* que comparten 400 pb de homología. Uno de ellos contiene una deleción en la región 3'-terminal y el otro en la 5'-terminal. Y, por otra parte, la RH intercromosomal utilizando los dos alelos *ade2-40* y *ade2-101*, que están localizados en dos cromosomas homólogos. La expresión de *BRCA1 wt* no tiene ningún efecto en la RH, en cambio, la expresión de variantes patogénicas induce un incremento significativo en ambas, la RH intracromosomal e intercromosomal. Con *BRCA2* sucede lo contrario, la expresión de *BRCA2 wt* incrementa ambas, la RH intracromosomal y intercromosomal, sin afectar al número de RAD51 *foci*, lo que sugiere que *BRCA2* funcional interfiere en la reparación de DNA en esta levadura. Por tanto, las variantes en *BRCA2* que incrementen la RH se consideran no patogénicas y las que no estimulan la RH podrían estar asociadas con un incremento en el riesgo a sufrir cáncer. Para *BRCA1* sucedería al revés.

Amplificación del centrosoma

Este ensayo se basa en la observación de que la reducción en la expresión o la inactivación de los genes *BRCA1/2* conduce a la amplificación del centrosoma^{242, 254, 255}.

Para evaluar los defectos funcionales en BRCA1/2 debidos a una variante, este ensayo mide la capacidad de una variante de restablecer el número de centrosomas cuando es transfectada en células en las que BRCA1/2 ha sido silenciado con un RNA de interferencia. Para ello, en estas células un plásmido que contiene el cDNA de BRCA1/2 (*wt* o mutante) fusionado con *GFP*. Los centriolos y centrosomas se visualizan y enumeran mediante inmunofluorescencia 72h después de la transfección. Las variantes patogénicas no serán capaces de restablecer el número de centrosomas en comparación al *wt*, mientras que las neutras si lo harán.

Ensayo de resistencia a agentes genotóxicos

La hipersensibilidad a agentes que dañan el DNA como agentes *crosslinkers* o metilantes es una de las características mostradas por muchas de las líneas celulares deficientes en BRCA1/2, lo que sugiere que esta propiedad está ligada a la pérdida de funcionalidad de estos genes^{244, 256-259}. Además, dicha hipersensibilidad puede ser reconstituida cuando se transfecta BRCA1/2 funcional en dichas células²⁶⁰. En base a esta observación, la reconstitución de líneas celulares deficientes en estos genes, complementadas con *BRCA1/2* (*wt* o variante) y posteriormente tratadas con agentes genotóxicos puede ser utilizado para determinar el efecto funcional de una variante^{234, 261}. La viabilidad celular después del tratamiento se puede medir mediante ensayos de supervivencia clonogénica²⁴². Los resultados se comparan con células complementadas con *BRCA1/2 wt* y *BRCA1/2* portando una *missense* patogénica. Algunos de los agentes genotóxicos comúnmente utilizados son Mitomicina C, Cisplatino, metil metanoxulfonato, radiación ionizante o radiación UV^{227, 242, 259, 262}, entre otros.

Aproximaciones agnósticas

Los estudios basados en los perfiles de expresión de un gen se presentan como una aproximación prometedora para detectar variantes patogénicas^{263, 264}. Estos estudios, basados en microarrays de expresión, se apoyan en la premisa de que la expresión de ciertos genes asociadas a *BRCA1/2* y/o a la modulación de las vías de señalización en las que éstos están involucrados, estará alterada cuando *BRCA1/2* estén mutados en heterocigosis. El uso de microarrays tiene la ventaja que permite analizar cambios de expresión en varias vías de una manera agnóstica. No obstante, estos estudios requieren de una correcta validación.

Evaluación de variantes localizadas en las regiones reguladoras de BRCA1/2

Mutaciones en las regiones reguladoras de BRCA1/2 incapacitan su transcripción y se han identificado en pacientes de cáncer de mama^{265, 266}. En un contexto en el que los avances en las tecnologías de secuenciación permiten la secuenciación del loci *BRCA1/2*

enteros de forma coste-efectiva, es de prever un aumento considerable en el número de variantes identificadas en las regiones no codificantes¹²⁸. Por tanto, el desarrollo de ensayos funcionales que nos permitan evaluar estas variantes de forma sistemática, es necesario. Dado que muchas de estas variantes, a priori, serán inocuas, es aconsejable priorizar, mediante métodos bioinformáticos, cuáles de ellas analizar mediante experimentos funcionales. Para ello, la conservación evolutiva es aceptada como un buen indicador para seleccionarlas.

El efecto de variantes localizadas en las regiones UTR, o en regiones activadoras o represoras de la expresión, se puede medir mediante la capacidad que tiene esa variante en la modulación de la transcripción de un gen reportero (ej. *Luciferasa*) transfectado de forma transitoria en células mamíferas y fusionado con dichas regiones promotoras²⁶⁷. Este método requiere del uso de un adecuado tipo celular humano como huésped dado que se pueden dar grandes variaciones asociadas a efectos específicos de la tipo celular. Además, permanece incierto el nivel de expresión del transcrito que debe considerarse como límite a partir del cual se produce un defecto funcional. Para determinar la expresión límite es necesario testar un gran número de variantes previamente clasificadas situadas en estas regiones.

Ensayos funcionales descritos para BRCA1

Ensayo de actividad E3 ubiquitín ligasa y de interacción con BARD1

BRCA1 interacciona con BARD1 mediante el dominio RING situado en la región N-terminal. El complejo formado con esta proteína es necesario para la actividad ubiquitín ligasa de BRCA1²⁶⁸⁻²⁷⁰. En este ensayo, se evalúa, paralelamente, el impacto de VSD localizadas en la región N-terminal de BRCA1 en la capacidad del heterodímero BRCA1: BARD1 de autoubiquitinizarse, con el impacto que tienen dichas variantes en la interacción BRCA1/BARD1, utilizando para esto último un ensayo de doble híbrido²⁷¹. Para el ensayo de ubiquitín ligasa, la región BRCA1N-terminal (2–101) y la región BARD1 comprendida entre los aminoácidos 27-146, fueron amplificadas a partir de cDNA humano y fusionadas cada una a una poli-histona. Estos fragmentos son clonados en el vector *pET15b* y expresados en *Escherichia coli* en presencia de ubiquitina. Posteriormente, el heterodímero BRCA1/BARD1 es clonado para medir la capacidad de este compuesto de auto-ubiquitinizarse.

Ensayo de la actividad transcripcional (TA)

Debido a que este ensayo ha sido uno de los utilizados en esta tesis, se explicará más detalladamente que el resto.

BRCA1 interviene en la regulación transcripcional mediante su dominio BRCT²⁷². Así pues, el ensayo de la actividad transcripcional está basado en la capacidad de la región C-terminal de BRCA1 (aa 1396-1863) de transactivar un gen reportero, cuando éste se encuentra fusionado a un dominio de unión al DNA (DBD)^{273, 274}. Este ensayo permite testar variantes localizadas en la región C-terminal de BRCA1, la cual incluye el dominio BRCT en tándem (aa 1646-1736 y aa 1760-1855), y ha demostrado ser una medida directa de la integridad estructural de BRCA1^{231, 275}.

Figura 6. Fragmento de la región C-terminal que puede ser analizavz en el ensayo de trasnactivación

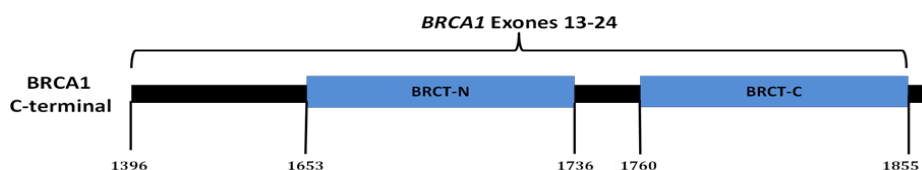
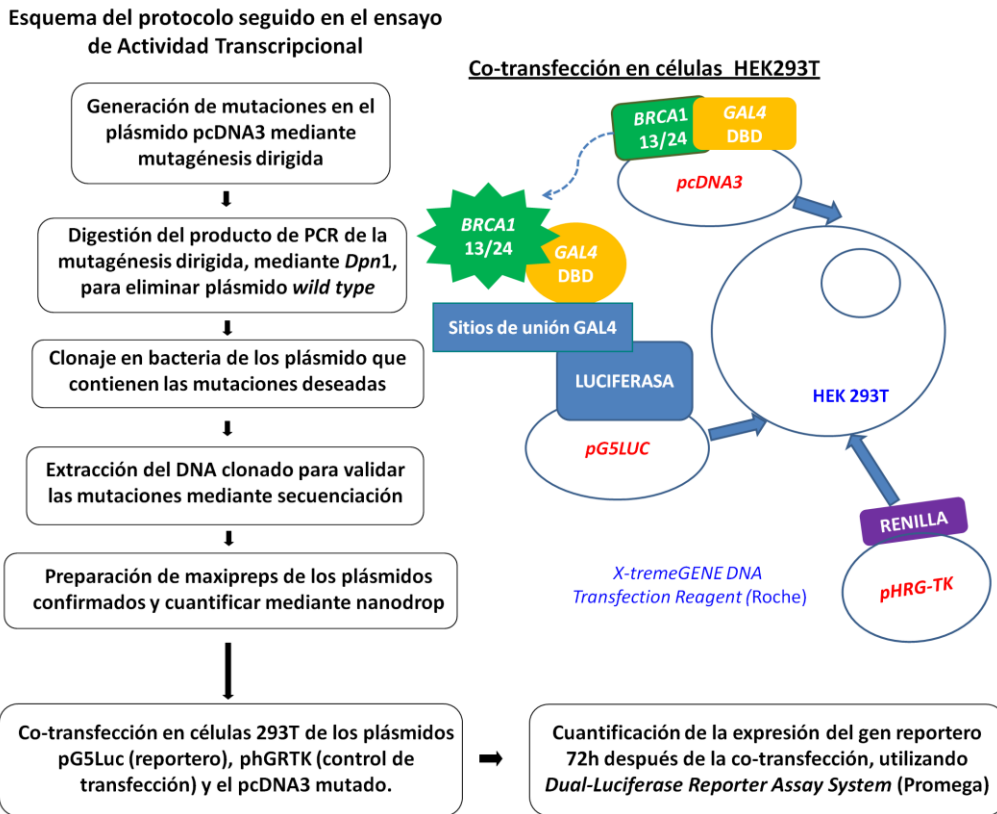


Figura 6. En esta figura se representa la región C-terminal de la proteína BRCA1, la cual puede ser testada por el ensayo de activación transcripcional. Esta región comprendida entre los aminoácidos 1396-1864, que esta codificada por exones 13-24 de BRCA1, incluye el dominio en tándem BRCT, que cosnta de dos repeticiones, BRCT-N (aa 1653-1736) y BRCT-C (aa 1760-1855).

Un esquema detallado del procedimiento seguido en este ensayo se muestra en la Figura 7. El ensayo de actividad transcripcional consta de dos etapas. En primer lugar, se genera, mediante mutación dirigida, la mutación deseada en un plásmido que contiene la región terminal de BRCA1 (aa 1386-1863) fusionada a 4 sitios de unión al DNA GAL4 (pcDNA3). Una vez generada la mutación, se clona en bacteria, se validan las colonias que contienen la variante por secuenciación y se prepara una “maxiprep” de DNA a partir de las colonias positivas. En la segunda etapa, el pcDNA3 con la mutación se co-transfecta en células mamíferas HEK293T junto con un plásmido reportero, que contiene el gen luminiscente de *Luciferasa* (pG5LUC) y un plásmido de *Renilla Luciferasa* que está bajo el control de un promotor constucional (pHRG-TK). El plásmido *Renilla* se utiliza como control de la eficiencia de transfección y de viabilidad celular (Fig. 7). La expresión del gen luciferasa está regulada por 5 DBD GAL4 situados en su región promotora, por tanto, la expresión del gen luciferasa servirá como medida indirecta de la capacidad de BRCA1 para activar la transcripción de GAL4. De este modo, BRCA1 funcional activará la transcripción de GAL4 que se unirá a los sitios de unión del promotor *Luciferasa* induciendo su expresión. En cambio, si la variante afecta la función de BRCA1, la expresión de luciferasa será menor o nula. En cada experimento se utiliza un control wt y dos mutaciones patogénicas que conducen a la pérdida de expresión de *Luciferasa*.

Figura 7. Esquema del protocolo seguido en el ensayo de transcripcional de BRCA1



La especificidad y sensibilidad de este ensayo ha sido ampliamente validado mediante un set de más de 70 variantes, incluyendo 13 variantes patogénicas y 7 neutras previamente clasificadas por modelos que no utilizaban información funcional. Gracias a ello, se ha podido desarrollar un modelo computacional que integra los resultados de este ensayo con los de la historia familiar, lo que supone el primer paso en la inclusión de los resultados funcionales en los modelos multifactoriales¹³².

Ensayo de unión a proteínas fosforiladas

El dominio BRCT de BRCA1 interacciona con proteínas fosforiladas que contiene la secuencia pSer-x-x-Phe^{98, 276}. En este ensayo, el dominio BRCT portando la VSD estudiada es expresado *in vitro* en bacteria y precipitado utilizando esferas de agarosa unidas a un péptido con la secuencia pSer-x-x-Phe^{229, 277}. Esto permite testar la capacidad de la VSD para unirse a al péptido pSer-x-x-Phe en comparación con el wt. Además, la especificad de dicha unión se puede medir utilizando esferas de agarosa unidas a un péptido no fosforilado como control. Para ello, se compara la unión de la VSD con el péptido fosforilado con la unión al péptido no fosforilado²²⁹.

Sensibilidad a la proteasa

Varios estudios han demostrado que el dominio BRCT presenta mayor susceptibilidad a la digestión mediante proteasas cuando contiene una mutación deletérea, en comparación con el *wt*²⁷⁸⁻²⁸⁰. Este ensayo, nos permite identificar variantes que provocan defectos en el plegamiento de BRCA1. Está limitado a variantes localizadas en la zona interior de la proteína, dado que las localizadas en la superficie no afectarán al plegamiento. Por ello, la especificidad de este ensayo es menor que el ensayo de la TA y el de unión a proteínas fosforiladas²²⁹.

El ensayo de sensibilidad a la proteasa y el ensayo de unión a proteínas fosforiladas han sido extensivamente utilizados para analizar VSD en BRCA1. Por tanto, al igual que el ensayo de actividad transcripcional, la sensibilidad y especificidad de estos ensayos ha sido robustamente determinada, mostrando una fuerte correlación entre los resultados de estos tres ensayos y la predisposición al cáncer. Así pues, estos dos ensayos también podrán ser incorporados en los modelos multifactoriales en un futuro cercano^{132, 229}.

Ensayo de fenotipo de colonias pequeñas (SCP)

El SCP deriva de la observación que la expresión de BRCA1 funcional en levadura inhibe su crecimiento dando lugar a colonias pequeñas^{281, 282}. El fenotipo puede ser testado cualitativamente observando el tamaño de las colonias o cuantitativamente contando el número de células por colonia. Aunque las bases bioquímicas del SCP no están aún claramente definidas, se ha observado una fuerte correlación entre la supresión del crecimiento y la expresión de BRCA1 *wt* o portando variantes neutras. Igualmente, se ha observado que mutaciones truncantes y *missense* patogénicas no afectan el crecimiento²⁸³. Parece ser, que la integridad del dominio BRCT es crítica para el SCP. De hecho, Este ensayo, este ensayo está actualmente limitado a variantes localizadas en el dominio BRCT, para las cuales ha mostrado una alta sensibilidad y especificidad y los resultados obtenidos correlacionan con los obtenidos en el ensayo de TA^{281, 282}. No obstante, actualmente está siendo testado para variantes en el dominio de RING.

Ensayo de fenotipo de localización en levadura (YLP)

Mediante el YLP se puede determinar la capacidad que tiene una mutación patogénica en BRCA1 de inducir la deslocalización de dicha proteína al citoplasma. Para determinar este efecto, se transfecta BRCA1 (*wt* o con la variante) fusionado con la proteína fluorescente mCherry. Esta proteína fusionada, con BRCA1 *wt*, se acumula en el núcleo en, aproximadamente, un tercio de las células y se reduce cuando BRCA1 contiene

una variante patogénica. Este efecto puede ser determinado cuantitativamente mediante microscopia de fluorescencia. Los resultados entre YLP y SCP correlacionan fuertemente sugiriendo que los defectos de crecimiento asociados a BRCA1 en levadura podrían ser debidos a su acumulación en el núcleo^{281, 282, 284}. El YLP si ha sido validado para testar variantes localizadas en el dominio de RING.

Ensayos funcionales descritos para BRCA2

Ensayo de interacción entre proteínas

Como se ha descrito en apartados anteriores, algunas de las funciones más importantes de BRCA2 dependen de la interacción con otras proteínas. Por tanto, medir diferencias en la afinidad de interacción, debidas a la presencia de una variante en uno de los dominios de unión de BRCA2, puede ayudar a determinar el efecto funcional de dicha variante. Las diferencias de afinidad se pueden analizar directamente mediante varios métodos bioquímicos como *pull downs*, co-inmunoprecipitaciones o ensayos de doble híbrido. Las interacciones de BRCA2 más estudiadas, hasta ahora, han sido con PALB2^{105, 227} y APRIN²⁸⁵. Se ha demostrado que varias variantes que afectan la interacción PALB2-BRCA2 están asociadas a un mayor riesgo de sufrir cáncer de mama^{105, 228}. Esta asociación aún no ha sido establecida para variantes que afectan otras interacciones de BRCA2.

Modelo syngeneic basado en líneas celulares cancerígenas humanas con BRCA noqueado (SyVal model)

Este ensayo está basado en células cancerígenas del epitelio colorectal humano (DLD1) modificadas genéticamente. En las células DLD1 alelo de BRCA2 es noqueado mediante la delección del exón11. La introducción de variantes en el alelo funcional endógeno de BRCA2 servirá para analizar su efecto funcional²⁸⁶. En los clones portadores de la variante se evalúa el efecto que dichas variantes tienen en la formación del RAD51 foci y en la sensibilidad a la Mitomicina C.

9. Asesoramiento genético en portadores de VSD

Como se ha explicado, debido a su naturaleza, el riesgo clínico asociado a una VSD es desconocido por su propia definición. Esto implica que la VSD no podrá ser utilizada como valor predictivo en el contexto clínico y, por tanto, la detección de una VSD en el test genético se deberá usar como un resultado indeterminado del estudio realizado. Por ello, los portadores seguirán los protocolos de seguimiento clínico acordes a este estado, que estarán principalmente basados en la historia personal y familiar de la enfermedad^{115, 146, 287}.

Sin embargo, en manos de facultativos clínicos sin formación específica en genética, las VSD puede generar cierta subjetividad, condicionando decisiones clínicas futuras^{136, 288}. Con ello, se puede incurrir en un alto riesgo de perjudicar al paciente, ya que las decisiones que éstos tomaran respecto a cirugías profilácticas preventivas u otras intervenciones dependerán, en gran medida, de la percepción de riesgo que reciban a partir de las recomendaciones de sus facultativos clínicos^{115, 136, 288}. Además, ha sido reportado que, alrededor del 30% de los portadores de una VSD no comprenden correctamente el impacto que ésta tiene sobre el riesgo a desarrollar cáncer. Y, generalmente, de manera errónea, la asocian a un incremento de riesgo, lo que provoca en ellos un gran impacto psicológico y puede tener repercusiones en sus decisiones futuras²⁸⁷⁻²⁹¹.

Este escenario refleja la gran dificultad ligada al seguimiento clínico de los portadores de VSD, y por tanto, la importancia que tiene el clasificarlas correctamente en el menor tiempo posible¹⁴⁶. Cuando la VSD sea reclasificada como patogénica, tanto el paciente como sus familiares obtendrán un gran beneficio, ya que los portadores podrán beneficiarse de las estrategias de prevención establecidas para su nuevo estado de riesgo y la variante se podrá utilizar como valor predictivo de riesgo^{115, 146, 287}. Con ello, los no portadores quedarán liberados del estrés asociados a su situación de incertidumbre en cuanto a su riesgo. Evidentemente cuando una VSD sea reclasificada como benigna, el beneficio para los pacientes y familiares será menos significativo, dado que seguirán siendo considerados pacientes de alto riesgo sin causa genética conocida y su seguimiento clínico no variará. Si bien, puede reducir la subjetividad que una VSD puede generar en el facultativo clínico^{136, 288}.

Hoy en día, la media temporal de reclasificación de VSD es demasiado grande en comparación al tiempo en que las pacientes portadoras tardan en tomar una decisión al respecto de las cirugías profilácticas^{128,134, 136, 146}, lo que implica la necesidad de mejorar los modelos de clasificación de variantes actuales¹⁴⁷.

10. Predisposición al SCMOH más allá de mutaciones en *BRCA1* y *BRCA2*.

Otros genes de predisposición al SCMOH

Con el paso de los años, después del descubrimiento de los genes *BRCA1/2*, quedó en evidencia que la genética de susceptibilidad al CM iba mucho más allá de estos dos genes. De hecho, *BRCA1/2* sólo son capaces de explicar alrededor de un 15-20% de los casos de CM familiar⁵⁷. Por ello, la búsqueda de nuevos genes y modelos de predisposición ha sido constante desde entonces.

Existen una serie de síndromes de predisposición al cáncer, que entre otras características clínicas, muestran una mayor prevalencia de CM entre sus afectos⁶⁷. Estos síndromes se caracterizan por mutaciones germinales en distintos genes de predisposición. Entre ellos, *TP53*, que predispone al síndrome de Li-Fraumeni; *PTEN*, al síndrome de Cowden; *LKB1-STK11*, causante del síndrome de Peutz-Jeghers o *CDH1* que predispone al síndrome de cáncer gástrico difuso²⁹²⁻²⁹⁴. Todos estos síndromes confieren un elevado riesgo al CM, aunque presentan frecuencias poblacionales muy bajas y contribuyen, entre todos ellos a un 3% de los casos de SCMOH⁶⁷ (Fig. 2).

Estudios dirigidos hacia genes interactores de los genes *BRCA1/2* y genes que actúan en vías de reparación comunes a éstos en familias de alto riesgo de CM, identificaron una serie de nuevos genes de predisposición considerados de moderada penetrancia⁶⁷. Sin embargo, entre todos estos genes tan sólo se pueden explicar un 3% adicional de los casos de SCMOH⁵⁷ (Fig. 2). Por una parte, se identificaron varios genes pertenecientes a la vía de la Anemia de Fanconi. Entre ellos: *PALB2 (FANCN)*, *BRIP1 (FANCI)*, *RAD51C (FANCD1)*, *FANCD2* e interactores de éstos como *XRCC2* y *RAD51D*²⁹⁵⁻³⁰⁰. Entre ellos, se ha descrito que genes como *BRIP1* y *RAD51C* confieren un alto riesgo a sufrir cáncer de ovario. Por otra parte, se han identificado genes involucrados en la vía de reparación de RH, común con *BRCA1*. Entre ellos *BARD1*, *ATM*, *CHEK2* y los genes del complejo *MRN (MRE11-RAD50-NBS1)*³⁰¹⁻³⁰⁶. Destacar que, *ATM* fue identificado por predisponer al síndrome de ataxia-telangiectasia, una enfermedad autosómica recesiva, cuyo fenotipo sugería que, la población femenina afectada por este síndrome tenía un riesgo moderado de sufrir CM. Estudios epidemiológicos y moleculares posteriores demostraron que mutaciones en heterocigosis en este gen confieren un riesgo moderado a sufrir CM³⁰¹.

Además, mediante una serie de estudios de asociación en el genoma completo (GWAS), que comparan la frecuencia con que aparecen ciertos polimorfismos comunes (SNP) entre poblaciones de alto riesgo de CM y población general, se han identificado alrededor de 90 SNP que aparecen con frecuencias significativamente mayor en poblaciones afectas por el CM; lo que sugiere que están asociados a un pequeño aumento en el riesgo de esta enfermedad^{307, 308}. Mediante estos SNP de baja penetrancia se puede explicar alrededor de un 14% de los casos de CM familiar (Fig. 2). Además, se espera que mediante futuros estudios GWAS en tamaños muestrales más grandes se podrán identificar nuevos SNP comunes asociados al CM que permitirán explicar un 14% adicional de los casos de SCMOH^{307, 309}. La susceptibilidad al CM asociada a estos alelos de baja penetrancia parece estar ligada a un modelo poligénico de predisposición^{124, 310}. Según este modelo, los genes de alta, moderada y baja penetrancia interactuarían entre ellos y/o con factores asociados al ambiente o el estilo de vida¹²⁴. Un modelo de interacción entre estos factores asume que cada uno actúa de forma independiente y el

efecto común es multiplicativo¹²⁴. Sin embargo, aunque los estudios de co-ocurrencia realizados hasta ahora, con SNP de baja penetrancia, parece indicar que éstos interactúan de forma multiplicativa entre ellos, una situación mucho más compleja se deriva de su interacción con otros factores de riesgo³¹¹⁻³¹⁴.

Paneles genéticos

Durante mucho tiempo, los genes *BRCA1/2* han sido los únicos genes considerados en el diagnóstico genético rutinario de pacientes con sospecha de sufrir el SCMOH. La baja frecuencia poblacional con que aparecen mutaciones en el resto de genes de alta o moderada penetrancia identificados para esta enfermedad ha hecho que no hayan sido incluidos en los test genéticos hasta ahora, a excepción de los genes de alta penetrancia asociados a los síndromes citados anteriormente (*TP53*, *PTEN*, *SKT11* y *CDH1*) cuando la paciente cumplía criterios clínicos para dichos síndromes³⁰¹. Sin embargo, la aparición de las nuevas técnicas de secuenciación masiva y el abaratamiento de los costes de éstas, ha propiciado el desarrollado de paneles genéticos que permiten el análisis simultáneo de éstos, dando cabida a los genes de predisposición de menor prevalencia en el test genético rutinario de una forma coste/efectiva. Entre todos los paneles desarrollados para el cáncer hereditarios se pueden analizar hasta un centenar de genes de predisposición al cáncer, siendo aproximadamente unos veinte los directamente asociados a la predisposición al SCMOH³¹⁵. El primero de ellos diseñado para el SCMOH, denominado BROCA, fue desarrollado por Walsh et. al en el año 2010. Dicho panel incluía 21 genes de predisposición a cáncer de mama incluyendo *BRCA1/2*. La validación de este panel en 20 mujeres diagnosticadas de CM o CO portadoras de una mutación conocida identificó todas las mutaciones conocidas y ningún falso positivo, demostrando a su vez ser una técnica coste efectiva en comparación a los test rutinarios¹²⁸. La aplicación posterior de este panel en 360 paciente con CO, identificó 63 mutaciones en *BRCA1/2* lo que corresponde con el 18% de los pacientes testados y 23 mutaciones en otros 10 genes, que corresponde con el 6,1% de los pacientes³¹⁶. En los años siguientes varios estudios basado en paneles multigénicos han identificado mutaciones en genes de predisposición distintos a *BRCA1/2* en frecuencias que varían entre el 3% y el 14% de los pacientes analizados que varía en relación al número de genes incluidos en el panel y el tipo y número de pacientes seleccionados para el estudio³¹⁷⁻³²². Esta situación ha propiciado que varios laboratorios privados como *Myriad Genetics*, *Invitae* o *Ambry Genetics* entre otras hayan desarrollado diferentes paneles para el uso comercial (un listado de los paneles comerciales desarrollados hasta ahora se puede encontrar en la revisión de Easton et. al 2014) (ref³¹⁵).

Aunque es evidente que la implantación de los paneles genéticos conlleva una mejora en el diagnóstico genético, ya que la posibilidad de testar simultáneamente varios

permite aumentar el porcentaje de pacientes con mutación identificada y por tanto, una mayor estratificación del riesgo pudiendo ayudar a seleccionar pacientes para determinados tratamientos (por ejemplo, iPARP para portadores de mutaciones patogénicas en *BRCA*)³²³⁻³²⁵, la utilidad clínica actual de estos paneles esta aún por determinar. Este hecho es debido a que los avances en el poder de interpretación de los resultados obtenidos en los paneles no ha seguido el mismo ritmo de desarrollo que las tecnologías de secuenciación. Por ejemplo, se desconoce la penetrancia real de mutaciones (principalmente de las *missense*) en algunos de los genes de predisposición incluidos en dichos paneles. Por otra parte, cuando se utilizan paneles de cáncer hereditario generales, existe la posibilidad de encontrar mutaciones no esperadas en genes de predisposición a enfermedades diferentes a las que iba dirigido el test (*incidental findings*). Además, el número de VSD aumentará en gran medida³²⁶. Por todo ello, cuando se utilicen paneles en el diagnóstico genético, se debe explicar correctamente al paciente que test se le va a realizar y cuáles son los posibles resultados y limitaciones del estudio realizado³²⁷. Para esto, será necesario establecer unas guías clínicas de actuación consensuadas que indiquen cuándo se deben utilizar dichos paneles, cómo manejar la información obtenida y qué información proporcionar al paciente^{57, 300}.

Una mayor conocimiento de las bases genética de susceptibilidad al SCMOH, que nos permita una correcta interpretación del riesgo en base a los modelos poligénicos, es necesario para mejorar el asesoramiento genético asociado a los paneles³²⁸. Análisis genéticos en series grandes de pacientes y la detección de nuevos genes de predisposición ayudarán a mejorar este entendimiento²⁷⁶.

Consideraciones en la búsqueda de nuevos genes de predisposición al SCMOH

Desde el descubrimientos de los genes *BRCA1/2*, han sido varios los intentos de identificar nuevos genes de predisposición al SCMOH con las mismas características que éstos. Sin embargo, los estudios realizados hasta ahora, parecen indicar que, aunque existen más genes de alta penetrancia como *BRCA1/2*, éstos solo explicarán un porcentaje muy pequeño de familias de alto riesgo^{329, 330}. Estos estudios, principalmente, han consistido en estudios de ligamiento diseñados en base a la presima de un patrón de herencia autosómica dominante en familias no asociadas a mutaciones en *BRCA1/2* (conocidas como BRCAX)^{329, 331, 332}.

La aparición de las nuevas técnicas de NGS se presentó como una herramienta prometedora para la identificación de nuevos genes. Aunque, hasta ahora, no ha tenido el éxito esperado en lo referente a la detección de genes de alta penetrancia^{152, 333}, algunos estudios basados en la secuenciación de exomas de varios pacientes pertenecientes a familias de alto riesgo con un patrón de herencia autosómica

dominante, han identificado algunos genes de predisposición al SCMOH aunque el riesgo asociado a mutaciones en estos genes es aún desconocido (*BLM*, *FANCC*, *REQL*, *XRCC2*,)^{299, 334-336}.

La gran heterogeneidad que presenta el grupo de familias BRCA, tanto a nivel genético (modelos de herencia poligénicos, diferentes *loci* asociados con la enfermedad, posibilidad de herencia autosómica recesiva, presencia de fenocopias) como a nivel histopatológico, es una de las principales causas del poco éxito en la búsqueda de nuevos genes^{333, 337}. Los estudios realizados hasta ahora, la mayoría basados en patrones de herencia autosómica dominante, no consideran esta heterogeneidad en su diseño. Se espera, que el estudio de subgrupos histopatológicos concretos dentro de los casos BRCA o el diseño de estudios que se ajusten a diferentes patrones de herencia podrá ayudar en la identificación de nuevos genes de alta penetrancia asociados al SCMOH³³⁸⁻³⁴¹. Por otra parte, la existencia de nuevos genes de moderada penetrancia es segura, sin embargo, es difícil predecir cuantos serán y que proporción de casos de SCMOH serán atribuibles a ellos. Hasta ahora, todos los genes de moderada penetrancia identificados han sido genes involucrados en vías de reparación de DNA, principalmente en las que también participaban los genes *BRCA1/2*⁶⁷. Esto indica que, la re-secuenciación de genes reparadores en grupos de familias de alto riesgo, *BRCA1/2* negativas, es la aproximación más óptima para detectar nuevos genes de predisposición moderada²⁹⁹.

ANTECEDENTES E HIPÓTESIS

Antecedentes que han motivado la presente tesis doctoral

El Cáncer de Mama (CM) es el segundo tipo de cáncer más común a nivel mundial. Se considera que alrededor de un 25% del CM se presenta en un contexto de agregación familiar. Si consideramos que en el año 2012, se diagnosticaron más de 1.700.000 de casos de CM a nivel mundial (Globocan 2012), el CM familiar afectaría a más de 400.000 mujeres en todo el mundo.

El CM familiar está causado por mutaciones genéticas en genes de susceptibilidad al cáncer que predisponen al SCMOH. En la actualidad, sólo se han identificado dos genes de alto riesgo asociados al SCMOH que presentan alta prevalencia, *BRCA1* y *BRCA2*. Mutaciones germinales en estos genes explican menos del 20% de los casos de CM familiar. Sin embargo, el estudio mutacional de *BRCA1/2* identifica un alto porcentaje de variantes cuyo significado biológico es desconocido (VSD), no pudiendo ser utilizadas para el diagnóstico genético.

El porcentaje de VSD detectadas en *BRCA1/2* varía entre 2-15% dependiendo de la población analizada y la casuística del laboratorio que las reporta. La gran mayoría de VSD son específicas de una o muy pocas familias lo que dificulta su clasificación mediante modelos epidemiológicos, genéticos o multifactoriales. En este contexto, analizar estas VSD mediante ensayos que evalúen su efecto en la función del gen puede aportar nuevas evidencias que ayuden a su clasificación.

Los esfuerzos realizados los últimos años, así como una mejor comprensión de las bases genéticas y moleculares del SCMOH ha llevado a la identificación de otros genes de predisposición a esta enfermedad. La mayoría de ellos son considerados de penetrancia moderada/baja (*PALB2*, *CHEK2*, *BRIP1*, *ATM*...) y están implicados en las vías de reparación genómico en las que están involucrados los genes *BRCA1/2*. Además, se han descrito otros síndromes hereditarios poco frecuentes, asociados a cáncer de mama pero con otras características clínicas asociadas como son los síndromes de Li-Fraumeni (causado por mutaciones en *TP53*), el síndrome de Cowden (causado por mutaciones en *PTEN*), el síndrome de Peutz-Jeghers (causado por mutaciones en *STK11*) o el síndrome de cáncer gástrico difuso (causado por mutaciones en *CDH1*). No obstante, se estima que más de un 50% de familias de alto riesgo siguen sin causa genética conocida (son los casos denominados casos BRCAX). Los estudios realizados hasta ahora indican una naturaleza muy heterogénea de estos casos, con existencia de diferentes patrones de herencia. Elucidar las causas genéticas del SCMOH es necesario para mejorar el diagnóstico y el consejo genético así como las opciones terapéuticas de los pacientes afectados por esta enfermedad.

Hipótesis de trabajo

El estudio funcional de VSD en los genes *BRCA1/2* puede aportar nuevas evidencias que nos ayuden a elucidar el impacto clínico de dichas variantes. Como consecuencia, algunas de ellas se podrán clasificar como variantes patogénicas o neutras, beneficiando el manejo clínico de los pacientes portadores de las mismas.

La secuenciación de genes candidatos involucrados en vías de reparación comunes a los genes *BRCA1/2* o de genes que interaccionan con éstos puede identificar nuevos genes de susceptibilidad al SCMOH.

La secuenciación de exomas de pacientes pertenecientes a familias BRCAX, seleccionadas en base a un patrón de herencia autosómica dominante, puede identificar nuevos genes de susceptibilidad de alta y/o moderada penetrancia a esta enfermedad.

En definitiva, la clasificación de la relevancia clínica de VSD en los genes *BRCA1/2* y la identificación de nuevos genes de susceptibilidad al SCMOH contribuirá a una mejor comprensión de las bases moleculares de esta enfermedad. Con ello, se espera mejorar el diagnóstico y el consejo genético de los pacientes con SCMOH.

OBJETIVOS

Objetivo

El objetivo global de esta tesis doctoral es mejorar el diagnóstico genético y el consejo genético de los pacientes con síndrome de cáncer de mama y ovario hereditario (SCMOH) haciendo uso de diversas herramientas de análisis genéticos, moleculares y celulares.

Objetivos específicos

Con el fin de abordar el objetivo principal propuesto en esta tesis, se proponen dos grandes subobjetivos, tal y como se detalla a continuación:

1. Análisis *in vitro* de VSD identificadas en nuestra población en los genes *BRCA1* y *BRCA2*.
 - 1.1 Evaluar el efecto de VSD identificadas en nuestro centro en el procesamiento del mRNA de los genes *BRCA1/2*.
 - 1.2 Estudio del posible efecto funcional de diferentes VSD en las proteínas *BRCA1/2*.
2. Búsqueda de nuevos genes implicados en la génesis del SCMOH.
 - 2.1. Secuenciación, mediante tecnología Sanger, de genes candidatos de la vía de Anemia de Fanconi/*BRCA*, así como de genes candidatos por ser interactores de las proteínas *BRCA1/2*.
 - 2.2. Secuenciación de exomas de diversos pacientes pertenecientes a familias *BRCA*X distintas, con elevada carga familiar y que muestren un patrón de herencia autosómico dominante.

RESULTADOS

Los resultados de esta tesis se dividen en cuatro apartados, de acuerdo con los subobjetivos propuestos. Los resultados ya publicados o enviados a publicación se presentan adjuntando la publicación correspondiente. Los resultados aún no publicados se presentan en formato de artículo científico, redactado en castellano. Los resultados se presentan en el siguiente orden:

1. Evaluación del efecto de VSD en el procesamiento del mRNA de los genes *BRCA1/2*.

Artículo 1

"Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing."

Whiley et. al, Clinical Chemistry, 2014

Resultados sometidos a publicación

"Investigating the effect of 28 BRCA1 and BRCA2 variants on their related transcribed mRNA."

Quiles et. al, Breast Cancer Research and Treatment

(In press)

2. Estudio funcional de VSD en las proteínas *BRCA1/2*.

Artículo 2

"Functional and structural analysis of C-terminal BRCA1 missense variants."

Quiles and Fernández-Rodríguez et. al, Plos One, 2013

Resultados sometidos a publicación

"Identification of a founder BRCA1 mutation in the Moroccan population"

Quiles et. al, Clinical Genetics (en revisión Diciembre 2015)

Resultados no publicados

"Análisis funcional de variantes de significado desconocido en BRCA2 mediante un ensayo basado en células embrionarias de ratón."

3. Estudio mutacional de nuevos genes identificados en la vía de Anemia de Fanconi/BRCA en familias BRCA

Artículo 3

“Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families.”

Fernández-Rodríguez* and Quiles* et. al, BMC Cancer, 2012

Artículo 4

“Evaluation of rare variants in the new fanconi anemia gene ERCC4 (FANCO) as familial breast/ovarian cancer susceptibility alleles.”

Osorio A et. al, Human Mutation, 2013

4. Secuenciación de exomas en familias BRCA.

Resultados no publicados

“Búsqueda de nuevos genes de predisposición al SCMOH mediante secuenciación de exomas de pacientes pertenecientes a familias BRCA.”

Capítulo 1

EVALUACIÓN DEL EFECTO DE VSD EN EL PROCESAMIENTO DEL mRNA DE LOS GENES *BRCA1/2*

ARTÍCULO 1

Artículo 1: Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing.

Whiley PJ¹, de la Hoya M, Thomassen M, Becker A, Brandão R, Pedersen IS, Montagna M, Menéndez M, **Quiles F**, Gutiérrez-Enríquez S, De Leeneer K, Tenés A, Montalban G, Tserpelis D, Yoshimatsu T, Tirapo C, Raponi M, Caldes T, Blanco A, Santamariña M, Guidugli L, de Garibay GR, Wong M, Tancredi M, Fachal L, Ding YC, Kruse T, Lattimore V, Kwong A, Chan TL, Colombo M, De Vecchi G, Caligo M, Baralle D, Lázaro C, Couch F, Radice P, Southey MC, Neuhausen S, Houdayer C, Fackenthal J, Hansen TV, Vega A, Diez O, Blok R, Claes K, Wappenschmidt B, Walker L, Spurdle AB, Brown MA; ENIGMA consortium.

Clinical Chemistry. 2014 Feb; 60(2):341-52.

Factor de Impacto (2014): 7,911

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

Phillip J. Whiley,^{1,2} Miguel de la Hoya,³ Mads Thomassen,⁴ Alexandra Becker,^{5,6} Rita Brandão,⁷ Inge Sokilde Pedersen,⁸ Marco Montagna,⁹ Mireia Menéndez,¹⁰ Francisco Quiles,¹⁰ Sara Gutiérrez-Enríquez,¹¹ Kim De Leener,¹² Anna Tenés,¹¹ Gemma Montalban,¹¹ Demis Tserpelis,⁷ Toshio Yoshimatsu,¹³ Carole Tirapo,¹⁴ Michela Raponi,¹⁵ Trinidad Caldes,³ Ana Blanco,¹⁶ Marta Santamariña,¹⁷ Lucia Guidugli,¹⁸ Gorka Ruiz de Garibay,³ Ming Wong,¹⁹ Mariella Tancredi,²⁰ Laura Fachal,¹⁶ Yuan Chun Ding,²¹ Torben Kruse,⁴ Vanessa Lattimore,²² Ava Kwong,²³ Tsun Leung Chan,²³ Mara Colombo,²⁴ Giovanni De Vecchi,²⁴ Maria Caligo,¹⁹ Diana Baralle,¹⁵ Conxi Lázaro,¹⁰ Fergus Couch,¹⁷ Paolo Radice,²⁴ Melissa C. Southey,¹⁸ Susan Neuhausen,²¹ Claude Houdayer,¹⁴ Jim Fackenthal,¹³ Thomas Van Overeem Hansen,²⁵ Ana Vega,¹⁶ Orland Diez,¹¹ Rien Blok,⁷ Kathleen Claes,¹² Barbara Wappenschmidt,^{5,6} Logan Walker,²² Amanda B. Spurdle,¹ and Melissa A. Brown²
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 re-

sults 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.

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

de Compostela, Spain; ¹⁷ Grupo de Medicina Xenómica—USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain; ¹⁸ Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, NY; ¹⁹ Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia; ²⁰ Istituto di Anatomia Patologica, Università di Pisa, Pisa, Italy; ²¹ Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA; ²² Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; ²³ Center for Genomic Medicine, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ²⁴ Department of Pathology University of Otago, Christchurch, New Zealand; ¹² Center for Medical Genetics, Ghent University Hospital, De Pintelaan, Ghent, Belgium; ²³ Division of Molecular Pathology, Department of Pathology, Hong Kong Sanatorium & Hospital, Hong Kong, China.

* Address correspondence to this author at: The University of Queensland, St. Lucia, Qld, 4072, Australia. E-mail melissa.brown@uq.edu.au.

Disclaimer: 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/clinchem.2013.210658

For example, because of the position of primers and PCR extension times, several isoforms associated with *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 Δ 19,20 and *BRCA1* c.135-1G>T Δ 5q and Δ 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*)²⁶ 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 ma-

majority of these used reverse transcriptase PCR (RT-PCR)²⁷ 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.cmg.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 sub-

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

²⁷ 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.

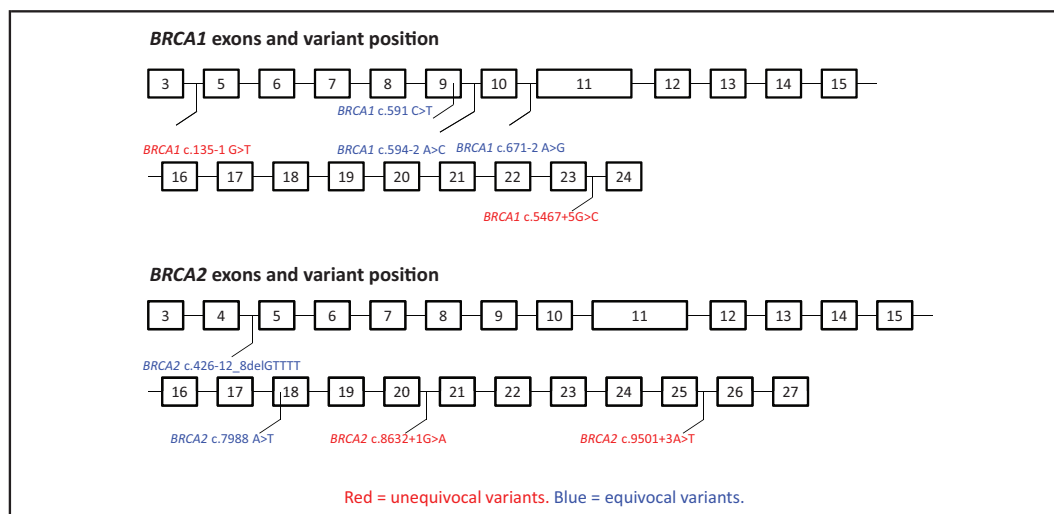


Fig. 1. *BRCA1* and *BRCA2* exons showing the positions of the variants studied.

BRCA1 c.135-1 G>T, *BRCA1* c.5467+5 G>C and *BRCA2* c.9501 + 3A>T and c.8632 + 1G>A were considered to produce unequivocal splicing aberrations. *BRCA1* c.591 C>T, c.594-2 A>C and c.671-2 A>G and *BRCA2* c.426-12_8delGTTTT and c.7988 A>T were considered to produce equivocal splicing aberrations.

mitted to ENIGMA from more than 43 sites in 19 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 cap-

illary electrophoresis (CE) in phase 2. Seven sites that participated in phase 1 repeated the assays under the controlled conditions of phase 2. An additional 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 $\mu\text{g}/\text{mL}$, 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 $\Delta 5\text{q}$ 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 $\Delta 9/10/11\text{q}$ 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 unequivocal 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).

Table 1. Phase 1 results submitted by 17 sites.^a

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

Continued on page 346

Table 1. Phase 1 results submitted by 17 sites.^a (Continued from page 345)

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

^a +, Detection of the transcript; -, no detection.
^b 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.
^c NA, not assessed.
^d Transcripts detected in the variant carrier only. Other transcripts were identified in at least 1 control.

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 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.

Table 2. Summary of results from phase 2 of the study, in which PCR primers and conditions were controlled.^{a,b}

Variant	Transcripts	Phase 2 results (10 sites) ^c									
		1*	2 [^]	3*	4*	8 [^]	9 [^]	16*	17*	18*	19*
Unequivocal											
<i>BRCA2</i> c.9501+3A>T	Full length	+	+	+	+	+	+	+	+	+	+
	Δ25 ^d	+	+	+	+	+	+	+	+	+	+
	ins i23 ^d	-	+	-	-	-	-	-	-	-	-
<i>BRCA2</i> c.8632+1G>A	Full length	+	+	+	+	+	+	NA	+	+	+
	Δ19 ^d	-	-	-	-	+	-	-	-	-	-
	Δ20 ^d	+	+	+	+	+	+	+	+	+	+
	Δ19&20 ^d	+	+	-	-	+	+	-	+	-	-
	ins i20 ^d	-	+	-	+	-	-	-	-	-	-
<i>BRCA1</i> c.135-1G>T	Extra peaks ^d	-	+	-	-	-	-	-	-	-	-
	Full length	+	+	+	+	+	+	NA	+	+	+
	Δ5q ^d	+	+	-	-	+	-	-	+	-	-
	Δ5 ^d	+	+	+	+	+	+	+	+	+	+
	Δ3 ^d	-	+	-	-	+	-	-	-	-	-
	Δ3,5 ^d	-	+	-	-	+	-	-	-	-	-
	Δ5,6 ^d	-	+	-	-	-	-	-	-	-	-
<i>BRCA1</i> c.5467+5G>C	ins i3 ^d	-	+	-	-	+	-	-	-	-	-
	ins i3+Δ5 ^d	-	+	-	-	-	-	-	-	-	-
	Full length	+	+	+	+	+	+	+	+	+	+
	Δ21 ^d	-	-	-	-	+	-	-	-	-	-
	Δ21,23 ^d	-	-	-	-	+	-	-	-	-	-
	Δ22,23 ^d	-	+	-	-	+	-	-	+	-	-
	Δ23 ^d	+	+	+	+	+	+	+	+	+	+
Δ22 ^d	-	+	-	+	+	-	-	-	-	-	
Equivocal											
<i>BRCA2</i> c.7988A>T	Full length	+	+	+	+	+	+	+	+	+	+
	Δ18	+	+	+	+	+	+	+	+	+	+
	Δ17&18	+	+	+	+	+	+	+	+	+	+
<i>BRCA2</i> c.426-12_8delGTTTT*	FL	+	+	+	+	+	+	NA	+	+	+
	Δ6q,7	+	+	-	+	+	-	-	+	-	-
	ins 18bp ^d	-	-	-	-	-	+	-	-	-	-
	Δ5 ^d	+	+	+	+	+	+	-	+	+	+
	Δ5,6,7 ^d	-	+	-	-	+	-	-	+	-	-
<i>BRCA1</i> c.591C>T	Δ6,7 ^d	-	-	-	-	-	+	-	-	-	+
	Full length	+	+	+	+	+	+	+	+	+	NA
	Δ9	-	+	+	+	+	-	+	+	-	-
	Δ9,10	+	+	+	+	+	+	+	+	+	+
	Δ9,10,11q	+	+	+	-	+	+	+	+	-	-
	Δ9,10,11	-	+	-	+	+	-	-	-	-	-
	Δ11q	+	+	+	-	+	+	+	+	-	-
	Δ9,11q ^d	+	+	-	-	+	-	+	-	-	-
	Δ10,11q ^d	-	-	-	-	+	-	-	-	-	-
	Δ10	-	+	-	-	+	-	-	-	-	-
	Δ11q+ins13 ^d	-	+	-	-	+	-	-	-	-	-
ins i21 intron 9 ^d	-	-	-	-	+	-	-	-	-	-	
Extra peaks ^d	-	+	-	-	-	-	-	-	-	-	

Continued on page 348

Table 2. Summary of results from phase 2 of the study, in which PCR primers and conditions were controlled.^{a,b} (Continued from page 347)

Variant	Transcripts	Phase 2 results (10 sites) ^c											
		1*	2 [^]	3*	4*	8 [^]	9 [^]	16*	17*	18*	19*		
<i>BRCA1</i> c.594-2A>C	Full length	+	+	+	+	+	+	+	+	+	NA		
	Δ9	-	+	-	-	+	+	-	-	-			
	Δ9,10	+	+	+	+	+	+	+	+	+			
	Δ9,10,11q	+	+	+	-	+	+	+	+	-			
	Δ9,10,11	+	+	-	-	-	+	-	-	-			
	Δ10 ^d	+	+	+	+	+	+	+	+	+			
	Δ11q	+	+	+	+	+	+	+	+	-			
	ins i21 ^d	-	+	-	-	+	-	-	-	-			
	Δ9,11q ^d	-	+	-	-	+	+	+	-	-			
	Δ10,11q ^d	-	+	+	-	+	+	+	+	-			
Δ11q+ins13 ^d	-	+	-	-	+	-	-	-	-				
		Phase 2 <i>BRCA1</i> c.671-2A>G results (12 sites) ^c											
		1*	2 [^]	3*	4*	8 [^]	9 [^]	16*	17*	18*	20 [^]	21*	22 [^]
<i>BRCA1</i> c.671-2A>G	Full length	+	+	+	+	+	+	+	+	+	+	+	+
	Δ9	-	+	+	-	+	-	-	-	-	+	-	-
	Δ9,10	+	+	+	+	+	+	+	+	+	+	+	-
	Δ9,10,11	+	+	+	+	+	+	+	+	-	+	+	-
	Δ9,10,11q	+	+	-	-	+	+	-	-	+	-	-	-
	Δ10 ^d	-	+	-	-	+	-	-	-	-	-	-	-
	Δ11	+	+	+	+	-	+	+	+	+	+	-	+
	Δ11q	-	+	-	+	+	+	-	-	+	+	-	+
	Δ9,10,11,12 ^d	-	-	-	-	-	-	+	-	-	-	-	-
	Δ9,11 ^d	-	+	-	-	-	-	-	-	-	-	-	-
	Δ10,11 ^d	+	+	+	+	+	+	+	+	+	-	-	+
	Δ10,11q ^d	-	-	-	-	+	-	-	-	-	-	-	-
	Δ11&12 ^d	-	+	-	-	-	-	-	-	-	+	-	-
	Δ~3.2kb exon 11 ^d	+	-	-	-	-	-	-	-	-	-	-	+
	Δ9,11q ^d	-	+	-	-	+	-	-	-	-	-	-	-
	ins i21 ^d	-	-	-	-	+	-	+	-	-	-	-	-
Δ11q+ins13 ^d	-	+	-	-	+	-	-	-	-	-	-	-	

^a +, Detection of the transcript; -, no detection.
^b 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.
^c *, 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.
^d Transcripts detected in the variant carrier only. Other transcripts were identified in at least 1 control.

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. 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).

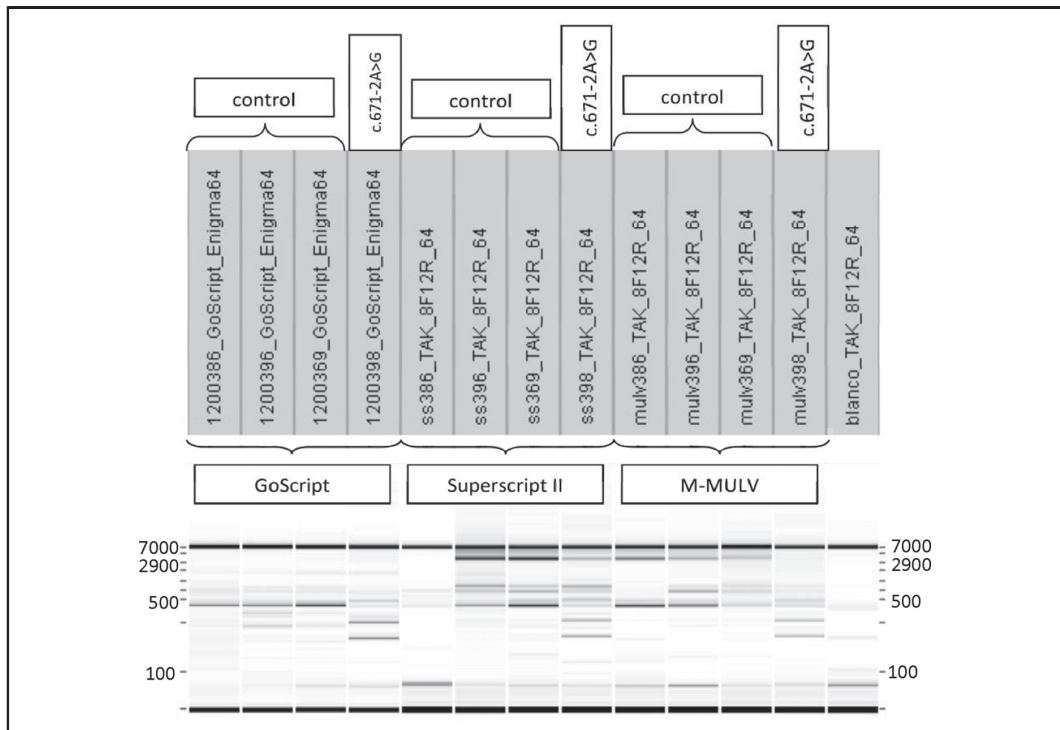


Fig. 2. Comparison of cDNA synthesis enzymes for the detection of different isoforms arising from the variant *BRCA1*: c.671-2A>G variant.

RT-PCR results (obtained by capillary electrophoresis on the Labchip, Caliper) obtained by using the same RNA, *Taq* polymerase (Takara), and PCR program. cDNA was synthesized with 3 different kits [GoScript (Promega), M-MuLV (New England BioLabs), SuperscriptII (Invitrogen)].

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 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 equivocal variant *BRCA2* c.8632+1G>A detected the ins i21bp intron 9 in phase 1 but not in phase 2, despite using the same primers and PCR conditions.

Some sites sequenced the products of PCR reactions experienced 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,

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.

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 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 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, The Beckman Research Institute of City of Hope.

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 #628353. 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: ISCIIIRETIC: 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. Gutiérrez-Enríquez, 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.
14. Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornero L, Esteban-Cardenosa 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.
 15. 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, Machackova 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. Machackova 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, Castellague 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, Revillion 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, Sinilnikova 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.

Resultados sometidos a publicación

Investigating the effect of 28 BRCA1 and BRCA2 mutations on their related transcribed mRNA.

Francisco Quiles¹, Mireia Menéndez¹, Eva Tornero¹, Jesús del Valle¹, Àlex Teulé¹, Sarai Palanca², Angel Izquierdo¹, Carolina Gómez¹, Olga Campos¹, Raúl Santamaria³, Joan Brunet¹, Gabriel Capellá¹, Lúdia Feliubadaló¹ and Conxi Lázaro¹.

Breast Cancer Research and Treatment

Factor de Impacto (2015): 3,940

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Full title:

Investigating the effect of 28 BRCA1 and BRCA2 mutations on their related transcribed mRNA

Authors:

Francisco Quiles¹, Mireia Menéndez¹, Eva Tornero¹, Jesús del Valle¹, Àlex Teulé¹, Sarai Palanca², Angel Izquierdo¹, Carolina Gómez¹, Olga Campos¹, Raúl Santamaria³, Joan Brunet¹, Gabriel Capellá¹, Lúdia Feliubadaló¹ and Conxi Lázaro

Authors' affiliations:

¹Hereditary Cancer Program, Catalan Institute of Oncology (ICO-IDIBELL, ICO-IdIBGi, ICO-IGTP), L'Hospitalet de Llobregat, Barcelona, Spain.

²Molecular Biology Laboratory, Service of Clinical Analysis, University and Polytechnic Hospital La Fe, Valencia (Spain).

³Genética Molecular, Laboratorio Echevarne, Barcelona (Spain)

Corresponding author:

Conxi Lázaro Ph.D., Hereditary Cancer Program, Molecular Diagnostics Unit, Catalan Institute of Oncology (ICO-IDIBELL), Hospital Duran i Reynals, Gran Via 199, L'Hospitalet de Llobregat, Barcelona 08908, Spain. Tel.: +34932607342; Fax: +34932607466; E-mail: clazaro@iconcologia.net

Abstract

Purpose: Germline inactivating mutations in the *BRCA1* and *BRCA2* genes are responsible for Hereditary Breast and Ovarian Cancer Syndrome (HBOCS). Genetic testing of these genes identifies a significant proportion of variants of uncertain significance (VUS). Elucidation of the clinical impact of these variants is an important challenge in genetic diagnostics and counseling. In this study, we assess the RNA effect of 28 *BRCA1* and *BRCA2* VUS identified in our set of HBOCS families with the aim of gaining insight into their clinical relevance.

Materials and Methods: mRNA was extracted from VUS carriers and controls lymphocytes cultured for 5-6 days and treated with puromycin. RNA was reverse transcribed to perform transcriptional analysis for the study of splicing aberrations. *In silico* prediction tools were used to select those variants most likely to affect the RNA splicing process.

Results: Six out of the 28 variants analyzed showed an aberrant splicing pattern and could therefore be classified as probably pathogenic mutations.

Conclusion: Reclassification of VUS improves the genetic counseling and clinical surveillance of carriers of these mutations and highlights the importance of RNA studies in routine diagnostic laboratories.

Keywords: *BRCA1*; *BRCA2*; VUS; splicing; RNA

Introduction

1 Germline mutations in the *BRCA1* and *BRCA2* genes significantly increase the lifetime risk of
2 breast and ovarian cancer, among other neoplasias [1-4]. Inactivating germline mutations in
3 these genes account for around 20-25% of familial cases [5]. However, 2-15% of patients tested
4 for *BRCA* mutations are carriers of variants of unknown significance (VUS); the exact figure
5 changes constantly and depends on the expertise and experience of the intervening laboratory as
6 well as on the population studied [6,7]. The clinical uncertainty generated by VUS greatly
7 hinders the clinical management of carriers. In addition, the clinical ambiguity can be difficult
8 for carriers to interpret and may impact on their psychosocial outcomes [8-10]. Thus, the
9 reclassification of VUS is paramount for a better clinical management of carriers. The efforts
10 made in recent years to gather more information about VUS have allowed the use of
11 multifactorial models and led to the classifications of some variants as deleterious mutations or
12 benign polymorphisms [11-13]. As consequence, the number of VUS in these genes is falling
13 and large private laboratories with databases of thousands of mutations achieved a drastic
14 decline in their VUS rate. For instance, Myriad laboratories claimed a drop in VUS from 12.8%
15 in 2002 to 2.7% in 2012 [14,15]. Unfortunately, in many cases there is still not enough data for
16 these variants to be classified or the existing data is not available because not necessarily all
17 laboratories share their information. In such cases, functional assays that evaluate specific
18 impact of VUS in *BRCA1* or *BRCA2* proteins may help to classify the variants [16,17]. In this
19 context, several studies have also demonstrated that analyzing the effect of VUS on splicing
20 may provide robust evidence to assist classification [18-22]. In this area, the splicing working
21 group of the ENIGMA (Evidence-Based Network for the Interpretation of Germline Mutant
22 Alleles) consortium put considerable effort into standardizing protocols between different
23 laboratories in order to obtain comparable and reliable results for the effect of VUS on the
24 correct splicing of *BRCA1* and *BRCA2* [23].
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

The aim of the present study is to evaluate the putative splicing effect of 28
VUS in *BRCA1* and *BRCA2* genes identified in Spanish Hereditary Breast and Ovarian Cancer
Syndrome (HBOCS) families studied at our center.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Materials and methods

Patients and variants selection

Since its creation in 1998, the Hereditary Cancer Program at the Catalan Institute of Oncology (*Institut Catala d'Oncologia*, ICO) has identified a set of 2397 high-risk families with suspected HBOCS. Following the Catalan Consensus OncoGuidelines (http://www.gencat.net/salut/depsan/pdf/og0601esgrr_din4.pdf) on genetic testing for this syndrome, patients were tested for mutations in the *BRCA1* and *BRCA2* genes after receiving appropriate genetic counseling and providing written informed consent (unpublished data). As result of this genetic testing, the Hereditary Cancer Program compiled a cohort of 279 VUS in *BRCA1* and *BRCA2*. A total of 28 variants were selected for this study (Table 1). These variants were selected according to availability of lymphocytes and because, to our knowledge, they had never been tested at the RNA level, except *BRCA1* c.5123C>A variant, which effect in RNA splicing remains unclear in the literature[20,24,25].

Several *in silico* tools were used to select those variants most likely to affect the RNA splicing process. Disruption/creation of splice sites was evaluated with Alamut Visual 2.4 (Interactive Biosoftware), a new splicing module, integrating a number of prediction methods: Splice Site Prediction by Neural Network (NNSPLICE) (http://www.fruitfly.org/seq_tools/splice.html), SpliceSiteFinder-like (<http://www.genet.sickkids.on.ca/~ali/splicesitefinder.html>), MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentseq_scoreseq.html), GeneSplicer (http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml), Human Splicing Finder (<http://www.umd.be/HSF/>). When the variant was located in the coding region, two programs integrated in Alamut Visual 2.4 were also used to determine whether an alteration of putative exonic splicing enhancer (ESE) elements could be observed: ESEfinder (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) and RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>).

Cell culture, RNA isolation and RT-PCR

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Frozen human lymphocytes from carriers and controls were cultured for 5-6 days and treated with puromycin for 4-6 h before RNA extraction in order to avoid Nonsense-Mediated Decay (NMD). Total RNA was isolated using either RNAeasy Kit (Qiagen) or Trizol reagent according to the manufacturer's instructions[19]. A total of 1.0 µg of total RNA was reverse transcribed using random primers and SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. To better evaluate splicing aberrations, transcripts from carriers were compared with transcripts from control lymphocyte cultures both by agarose gel analysis and by Sanger sequencing. PCR primers were designed following ENIGMA recommendations (i.e. forward and reverse primers when possible must be at least one whole exon 5' or 3' of the variant, respectively) (Online Resource 3).

Mutation nomenclature

Human Genome Variation Society (HGVS) approved guidelines (<http://www.hgvs.org/mutnomen/>) were used for *BRCA1* and *BRCA2* nomenclature. For mutation nomenclature, nucleotide 1 is the A of the ATG-translation initiation codon.

Results

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Mutational screening of *BRCA1* and *BRCA2* genes in Spanish HBOCS families constantly identifies new VUS. Abnormal splicing alterations have been detected in a large proportion of VUS, even those that are not located at the canonical splice sites [26]. In this context, 28 VUS from a cohort of mutations were selected for this study, 16 of them located in *BRCA1* and 12 in *BRCA2*; 10 were intronic variants, 4 were copy number alterations (also known as large genomic rearrangements, LGR) involving just one exon and detected by MLPA analysis and the remaining 14 were missense variants (Table 1). Thirteen of these variants had previously been reported in the BIC (<http://research.nhgri.nih.gov/bic/>) database or in the Leiden Open Variation Database (LOVD) (<http://brca.iarc.fr/LOVD/home.php>); the remaining 15 variants were reported for the first time in this study. To our knowledge, there are no previously published data on the transcriptional effect of any of these 28 variants except for the *BRCA1* c.5123C>A variant, whose effect on RNA splicing has yet to be clearly elucidated in the literature [20,24,25].

To experimentally evaluate the impact of the selected VUS on splicing, RNA was extracted from the lymphocytes of variant carriers and from a control group. The transcriptional profiles of carriers were compared, as minimum, to those of 6 independent control samples (not all showed in the Figures) by agarose gel electrophoresis. Sanger sequencing were performed for the carrier and two controls in all cases. Six out of the 28 variants led to an aberrant transcription profile: three LGR and one intronic variant in *BRCA1*, and one LGR and one intronic variant in *BRCA2*. In addition, the *BRCA1* c.5123C>A variant was found to have different splicing effects depending on the biological source used to isolate RNA. No evidence of aberrant splicing was detected when RNA was extracted from cultured lymphocytes treated with puromycin. However, partial in-frame skipping of exon 18 (p.Ala1693_Trp1718del) was detected when RNA was extracted directly from fresh lymphocytes (Online Resource 1). This variant has been previously reported as a recurrent mutation in the Spanish population[27] and

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

has been detected in 22 independent families at our center. The remaining 21 variants had no apparent effect on mRNA splicing as observed under our protocol (Table 1).

The results obtained by *in silico* prediction tools were in agreement with the observed experimental results, except for GeneSplicer predictions, which showed a variation of the scores that our RNA analysis did not confirm in approximately half of the mutations that were predicted not to alter the consensus splice sites by the other programs (Supp. Table 1). Consequently, we have not taken into account the results obtained by GeneSplicer. In agreement with the other *in silico* prediction programs, two intronic variants (*BRCA1* c.4987-20_4987-11del and *BRCA2* c.681+5G>C) were predicted to reduce recognition of the consensus acceptor and donor splice sites, respectively (Fig. 1 c) and our RNA assay showed the complete skipping of *BRCA1* exon 17 and *BRCA2* exon 8, respectively (Fig. 1 a, b). Eight intronic mutations were predicted not to significantly alter the consensus splice sites. Only slight differences were observed in *BRCA1* c.81-11T>C, *BRCA2* c.-39-12_-39-10delTCT and *BRCA2* c.632-3C>T by at least three prediction programs, but no alterations at the RNA level were observed in any of these eight cases. None of the missense variants led to a clear alteration in donor/acceptor splice sites but they showed some alteration in exonic enhancer elements (Online Resource 2). In all of these cases, the RNA profile was not altered. The *in silico* tools used here were not designed to study the effect of copy number alterations; moreover, the breakpoints of these rearrangements are unknown when using MLPA as the method of detection. Consequently, RNA analysis was performed for LGRs affecting a single exon in order to determine the putative effect of these LGRs at the protein level, to classify their clinical relevance and its pathogenicity.

The effect in the RNA of the six variants showing aberrant splicing pattern evidenced that all of them led to a truncated protein. The intronic variants *BRCA1* c.4987-20_4987-11del and *BRCA2* c.681+5G>C located in introns 16 and 8, respectively, caused complete skipping of exon 17 (r.4987_5074del) and exon 8 (r.632_681del) generating, in both cases, an out-of-frame deletion that would create a premature stop codon, p.Asp1692Metfs*10 and p.Val211Glufs*10 respectively (Fig. 1 a, b). *BRCA1* c.(441+1_442-1)_(547+1_548-1)del and *BRCA1*

1 c.(4357+1_4358-1)_(4484+1_4485-1)del, encompassing exons 8 and 14, respectively, led to the
2 complete deletion of these exons at the RNA level (r.442_547del and r.4358_4484del,
3 respectively) (Fig. 1 d, e). In both cases, aberrant splicing generated an out-of-frame deletion,
4 which was predicted to create a premature stop codon, p.Gly148Aspfs*51 and
5
6 p.Ala1453Glyfs*10, respectively. *BRCA1* c.(5193+1_5194-1)_(5277+1_5278-1)del, a LGR
7 affecting exon 20, led to an aberrant transcript in which exons 18, 19 and 20 were skipped
8 (r.5075_5278del) (Fig. 1 f). This aberrant transcript would produce a premature stop codon,
9 p.Asp1691Aspfs*70. *BRCA2* c.(8632+1_8633-1)_(8754+1_8755-1)dup, encompassing exon 21,
10 led to an exon 21 duplication in tandem (r.8633_8754)dup (Fig. 1 g), which would generate a
11 premature stop codon, p.Gly2919Alafs*13. Pedigrees showing the clinical characteristics of
12 patients with these VUS and co-segregation analyses performed are presented in Figure 2.
13
14
15
16
17
18
19
20
21
22
23

24 The study reported here is a qualitative study designed to determine the presence of
25 abnormal transcripts generated from VUS. However, as PCR reactions were set up to end before
26 reaching the plateau, a semiquantitative analysis based on agarose gel band quantification
27 showed an equal proportion of aberrant versus wild-type transcripts in all cases except for
28 *BRCA1* c.(441+1_442-1)_(547+1_548-1)del, where the ratio between aberrant and wild-type
29 transcripts was 0.34 (SD=0.11) (data not shown).
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Discussion

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

The clinical impact of some DNA variants (i.e. missense, small in-frame deletions/insertions, silent or intronic variants) is difficult to predict, hence most of these variants are classified as VUS. Determining the pathogenicity of VUS is of great importance, as classification would allow for better-personalized clinical management of carriers. Multifactorial algorithms have been developed for several hereditary cancer syndrome genes[13,11]. Functional assays can provide new evidence when the available clinical data are insufficient for classification. Among these assays, the mRNA splicing analysis is a simple and robust method to identify possible deleterious effects. It is currently possible to classify variants as pathogenic or probably pathogenic in the 5-tier scheme proposed for the IARC according to RNA results[22], with important implications in the clinical setting. In our study, we used an RNA analysis approach to examine the mRNA effect of 28 *BRCA1* and *BRCA2* VUS identified in HBOC patients at our center. First, we used *in silico* prediction programs to select those variants most likely to affect the RNA splicing process, which produced largely consistent results, with only slight differences observed, except for the GeneSplicer program. The use of *in silico* prediction tools filters out those variants with a low probability of being deleterious, reducing the search to fewer candidate variants for further experimental validation. Then, we performed a qualitative analysis of the selected variants, following ENIGMA recommendations for the most critical steps (primer design, use of agents to avoid NMD, RT and PCR conditions). Aberrant splicing was identified in 6 variants that had never been studied before. Two intronic variants, *BRCA1* c.4987-20_4987-11del and *BRCA2* c.681+5G>C were correctly predicted by *in silico* programs to reduce the efficiency of the nearest consensus splice sites. These predictions were confirmed by the RNA analysis, which showed complete skipping of exon 17 of *BRCA1* and exon 8 of *BRCA2*, respectively, producing in both cases frameshift transcripts that created a premature stop codon and would result in a truncated protein (Table 1). Three variants, *BRCA1* c.(441+1_442-1)_(547+1_548-1)del, *BRCA1* c.(4357+1_4358-1)_(4484+1_4485-1)del, and *BRCA2* c.(8632+1_8633-1)_(8754+1_8755-1)dup, were LGRs which caused complete skipping of the unique exon affected by the deletion. Variant *BRCA1* c.(5193+1_5194-1)_(5277+1_5278-

1)del was an LGR that caused the loss of exon 20 at the genomic level but the skipping of exons 18, 19 and 20 at the RNA level. Defining the real effect on splicing of single exon deletions is paramount to assess their clinical relevance, especially when dealing with small in frame exons. In all cases analyzed here, single exon LGRs produce abnormal splicing and putative truncated proteins. All together, our study showed that the six variants altering the correct splicing produced aberrant frameshift transcripts that created a premature stop codon and would result in truncated proteins. For this reason we can classify them as probably pathogenic variants (Table 1).

Although our study was not designed to perform an accurate quantitative analysis of transcripts it can be considered a semi-quantitative approach, and in all cases but one the agarose electrophoresis image showed equal expression of aberrant transcripts in comparison with the wild-type transcript. Variant *BRCA1* c.(441+1_442-1)_(547+1_548-1)del showed less expression of the aberrant transcript compared to the wild-type transcript. This difference cannot be attributed to the production of normal transcripts from the mutated allele, as it has been described for other mutations²⁸, due to the fact that the mutation in this case is a genomic deletion. The presence of other transcripts not captured by our design or differences in transcription level between the deleted and wild-type allele could be responsible for such differences but elucidation of the exact mechanism is out of the scope of the present study.

It should be noted that the *BRCA1* c.5123C>A variant had been previously characterized at the RNA level in earlier studies[20,24,25]. However, the literature reveals a degree of ambiguity concerning the effect of c.5123C>A on RNA splicing. Lovelock et al.[24] reported no aberrant splicing for this variant when RNA was isolated from lymphocytes cultured and treated with cycloheximide to inhibit NMD (Online Resource 1). By contrast, the same authors observed skipping of exon 18 (r.5075_5152del) when RNA was isolated directly from peripheral-blood leukocytes, as reported by Millevoi et al.[25]. The study of this mutation in our patients produced the same results as reported previously; no aberrant skipping event was observed when RNA was extracted from cultured lymphocytes treated with puromycin, but

1 partial skipping of exon 18 (r.5075_5152del) was observed when we used freshly extracted
2 RNA from lymphocytes. As already described for other genes[28], the proportion of abnormal
3 transcripts from a given mutation can vary depending on the type of cells used for the RNA
4 analysis and the different cell culture conditions applied. This again highlights the importance
5 of standardizing the protocols used in different laboratories when performing RNA analyses, as
6 suggested by the ENIGMA consortium [23].
7
8

9
10
11 The 21 remaining VUS did not show a pattern of aberrant splicing under our
12 experimental design, so they would be unlikely to alter the splicing of the affected region.
13
14

15
16 To conclude, analysis of the mRNA effect of 28 VUS in the *BRCA1* and *BRCA2* genes
17 showed an aberrant pattern of splicing for six variants, which could therefore be classified as
18 probably pathogenic mutations according to the ENIGMA classification rules. This
19 classification allows for better clinical management of both carriers and non-carriers[8,9],
20 facilitates the genetic counseling process for these patients and their families and can help
21 patients to take future therapeutic and reproductive decisions (e.g. risk-reducing surgeries in
22 carriers, etc.). Consequently, we recommend standardized RNA analysis of VUS as a routine
23 step in genetic testing laboratories.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Acknowledgments

We thank all patients who contributed to the study. The authors would like to particularly acknowledge the support of the Asociación Española Contra el Cáncer (AECC). Contract grant sponsor: Instituto de Salud Carlos III, organismo adscrito al Ministerio de Economía y Competitividad and “Fondo Europeo de Desarrollo Regional (FEDER), una manera de hacer Europa”, Catalan Health Institute and Autonomous Government of Catalonia. Contract grant numbers: PI10/01422, PI13/00285, PIE13/00022, RD12/0036/0008, RD12/0036/0014, RD12/0036/0031, 2009SGR290 and 2014SGR338.

Conflict of interest statement

The authors of this paper certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

References

1. Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, et al. (1994) BRCA1 mutations in primary breast and ovarian carcinomas. *Science (New York, NY)* 266 (5182):120-122
2. 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 (New York, NY)* 266 (5182):66-71
3. 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 (6559):789-792
4. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, Ponder BJ, Easton DF (2002) A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *British journal of cancer* 86 (1):76-83
5. Kurian AW (2010) BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Current opinion in obstetrics & gynecology* 22 (1):72-78
6. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpfer KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC (2002) Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 20 (6):1480-1490
7. Murray ML, Cerrato F, Bennett RL, Jarvik GP (2011) Follow-up of carriers of BRCA1 and BRCA2 variants of unknown significance: variant reclassification and surgical decisions. *Genet Med* 13 (12):998-1005
8. Garcia C, Lyon L, Littell RD, Powell CB (2014) Comparison of risk management strategies between women testing positive for a BRCA variant of unknown significance and women with known BRCA deleterious mutations. *Genet Med* 16 (12):896-902
9. Richter S, Haroun I, Graham TC, Eisen A, Kiss A, Warner E (2013) Variants of unknown significance in BRCA testing: impact on risk perception, worry, prevention and counseling. *Ann Oncol* 24 Suppl 8:viii69-viii74
10. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV (2008) Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human mutation* 29 (11):1282-1291
11. 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 (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 (1):2-7
12. Vallee MP, Francy TC, Judkins MK, Babikyan D, Lesueur F, Gammon A, Goldgar DE, Couch FJ, Tavtigian SV (2012) Classification of missense substitutions in the BRCA genes: a database dedicated to Ex-UVs. *Human mutation* 33 (1):22-28
13. Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS (2008) Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Human mutation* 29 (11):1265-1272
14. Eggington JM, Bowles KR, Moyes K, Manley S, Esterling L, Sizemore S, Rosenthal E, Theisen A, Saam J, Arnell C, Pruss D, Bennett J, Burbidge LA, Roa B, Wenstrup RJ (2014) A comprehensive laboratory-based program for classification of variants of uncertain significance in hereditary cancer genes. *Clinical genetics* 86 (3):229-237
15. Pruss D, Morris B, Hughes E, Eggington JM, Esterling L, Robinson BS, van Kan A, Fernandes PH, Roa BB, Gutin A, Wenstrup RJ, Bowles KR (2014) Development and validation

of a new algorithm for the reclassification of genetic variants identified in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 147 (1):119-132

16. Millot GA, Carvalho MA, Caputo SM, Vreeswijk MP, Brown MA, Webb M, Rouleau E, Neuhausen SL, Hansen T, Galli A, Brandao RD, Blok MJ, Velkova A, Couch FJ, Monteiro AN (2012) A guide for functional analysis of BRCA1 variants of uncertain significance. *Human mutation* 33 (11):1526-1537

17. Guidugli L, Carreira A, Caputo SM, Ehlen A, Galli A, Monteiro AN, Neuhausen SL, Hansen TV, Couch FJ, Vreeswijk MP (2014) Functional assays for analysis of variants of uncertain significance in BRCA2. *Human mutation* 35 (2):151-164

18. Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Leone M, Mazoyer S, Muller D, Remenieras A, Revillion F, Rouleau E, Sokolowska J, Vert JP, Lidereau R, Soubrier F, Sobol H, Sevenet N, Bressac-de Paillerets B, Hardouin A, Tosi M, Sinilnikova OM, Stoppa-Lyonnet D (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 (8):1228-1238

19. Menendez M, Castellsague J, Mirete M, Pros E, Feliubadalo L, Osorio A, Calaf M, Tornero E, del Valle J, Fernandez-Rodriguez J, Quiles F, Salinas M, Velasco A, Teule A, Brunet J, Blanco I, Capella G, Lazaro C (2012) Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 132 (3):979-992

20. 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 (6):1957-1967

21. Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, 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 (1):107-114

22. 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, Radice P, Spurdle AB (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* 34 (10):1424-1431

23. Whiley PJ, de la Hoya M, Thomassen M, Becker A, Brandao R, Pedersen IS, Montagna M, Menendez M, Quiles F, Gutierrez-Enriquez S, De Leener K, Tenes A, Montalban G, Tserpelis D, Yoshimatsu T, Tirapo C, Raponi M, Caldes T, Blanco A, Santamarina M, Guidugli L, de Garibay GR, Wong M, Tancredi M, Fachal L, Ding YC, Kruse T, Lattimore V, Kwong A, Chan TL, Colombo M, De Vecchi G, Caligo M, Baralle D, Lazaro C, Couch F, Radice P, Southey MC, Neuhausen S, Houdayer C, Fackenthal J, Hansen TV, Vega A, Diez O, Blok R, Claes K, Wappenschmidt B, Walker L, Spurdle AB, Brown MA (2014) Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing. *Clinical chemistry* 60 (2):341-352

24. Lovelock PK, Healey S, Au W, Sum EY, Tesoriero A, Wong EM, Hinson S, Brinkworth R, Bekessy A, Diez O, Izatt L, Solomon E, Jenkins M, Renard H, Hopper J, Waring P, Tavtigian SV, Goldgar D, Lindeman GJ, Visvader JE, Couch FJ, Henderson BR, Southey M, Chenevix-Trench G, Spurdle AB, Brown MA (2006) Genetic, functional, and histopathological evaluation of two C-terminal BRCA1 missense variants. *Journal of medical genetics* 43 (1):74-83

25. Millevoi S, Bernat S, Telly D, Fouque F, Gladieff L, Favre G, Vagner S, Toulas C (2010) The c.5242C>A BRCA1 missense variant induces exon skipping by increasing splicing repressors binding. *Breast cancer research and treatment* 120 (2):391-399

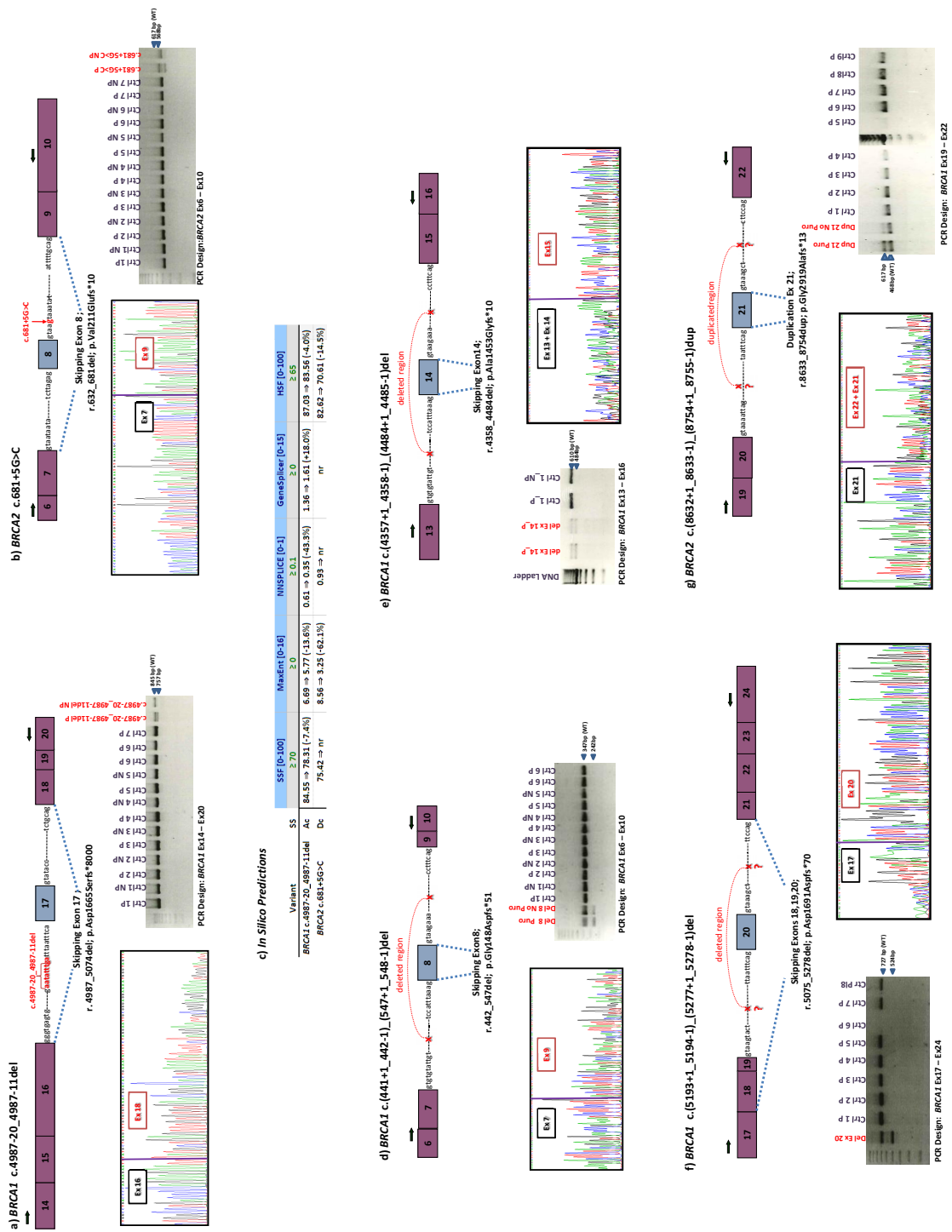
26. Ars E, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Human molecular genetics* 9 (2):237-247

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
27. Diez O, Osorio A, Duran M, Martinez-Ferrandis JI, de la Hoya M, Salazar R, Vega A, Campos B, Rodriguez-Lopez R, Velasco E, Chaves J, Diaz-Rubio E, Jesus Cruz J, Torres M, Esteban E, Cervantes A, Alonso C, San Roman JM, Gonzalez-Sarmiento R, Miner C, Carracedo A, Eugenia Armengod M, Caldes T, Benitez 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. *Human mutation* 22 (4):301-312
28. Pros E, Larriba S, Lopez E, Ravella A, Gili ML, Kruyer H, Valls J, Serra E, Lazaro C (2006) NF1 mutation rather than individual genetic variability is the main determinant of the NF1-transcriptional profile of mutations affecting splicing. *Human mutation* 27 (11):1104-1114

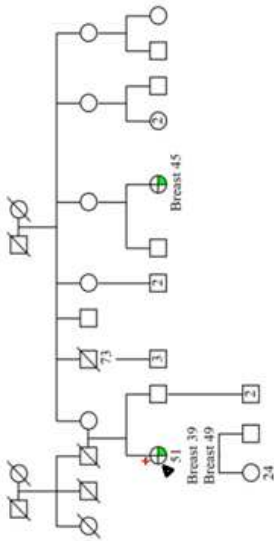
Gene	Variant (HGVS)	Protein prediction if splicing is not altered	BIC	LOVD	RNA Analysis	Putative Protein Effect after RNA analysis
BRC41	c.81-11T>C	p.(=)	Not reported	Not reported	r.=	p.=
	c.181T>A	p.(Cys61Ser)	Not reported	VUS	r.181u>a	p.Cys61Ser
	c.(441+1_442-1)_(547+1_548-1)del	p.?	Not reported	Not reported	r.442_547del	p.Gly148Aspfs*51
	c.(4357+1_4358-1)_(4484+1_4485-1)del	p.?	Not reported	Not reported	r.4358_4484del	p.Ala1453Glyfs*10
	c.4675+11A>G	p.(=)	Not reported	Not reported	r.=	p.=
	c.4882A>G	p.(Met1628Val)	6, VUS	Neutral	r.4882a>g	p.Met1628Val
	c.4956G>A	p.(Met1652Ile)	60, VUS	No consensus	r.4956g>a	p.Met1652Ile
	c.4987-20_4987-11del	p.(=)	Not reported	Not reported	r.4987_5074del	p.Val1665Serfs*8
	c.5075-49delC	p.(=)	Not reported	Not reported	r.=	p.=
	c.5100A>G	p.(=)	Not reported	Not reported	r.5100a>g	p.=
	c.5117G>A	p.(Gly1706Glu)	7, VUS	Deleterious	r.=	p.Gly1706Glu
	c.5123C>A	p.(Ala1708Glu)	45, PAT	Deleterious	r.[5123c>a, 5075_5152del]	p.[Ala1708Glu; Ala1693Trp1718delinsGly]
	c.5154G>A	p.(Trp1718*)	3, PAT	Deleterious	r.5154g>a	p.Trp1718*
	c.5175A>G	p.(=)	Not reported	Not reported	r.5175a>g	p.=
	c.(5193+1_5194-1)_(5277+1_5278-1)del	p.?	Not reported	Not reported	r.5075_5278del	p.Asp1691Aspfs*70
	c.5467+9C>A	p.(=)	Not reported	Not reported	r.=	p.=
	BRC42	c.-39-12_-39-10delTCT	p.(=)	7, VUS	Not reported	r.=
c.599C>T		p.(Thr200Ile)	Not reported	Not reported	r.599c>u	p.Thr200Ile
c.632-3C>T		p.(=)	Not reported	Not reported	r.=	p.=
c.681+5G>C		p.(=)	Not reported	Not reported	r.632_681del	p.Val211Glufs*10
c.682-30A>C		p.(=)	Not reported	Not reported	r.=	p.=

c.1550A>G	p.(Asn517Ser)	2, VUS	Neutral	r.1550a>g	p.Asn517Ser
c.1792A>G	p.(Thr598Ala)	62, VUS	Neutral	r.1792a>g	p.Thr598Ala
c.(8632+1_8633-1) (8754+1_8755-1)dup	p.?	Not reported	Not reported	r.8633_8754dup	p.Gly2919Lysfs*13
c.8754+135G>A	p.(=)	1, VUS	Not reported	r.=	p.=
c.8905G>A	p.(Val2969Met)	62, VUS	Neutral	r.8905g>a	p.Val2969Met
c.9270C>T	p.(=)	Not reported	Not reported	r.9270c>u	p.=
c.9292T>C	p.(Tyr3098His)	39, VUS	Neutral	r.9292u>c	Tyr3098His

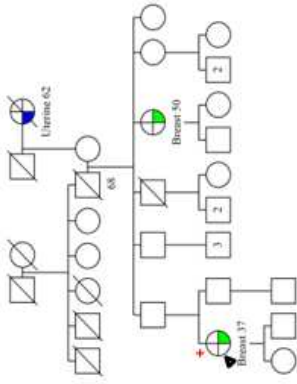
Figure



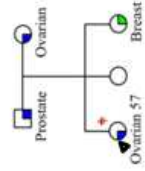
a) BRCA1 c.4987-20_4987-11del



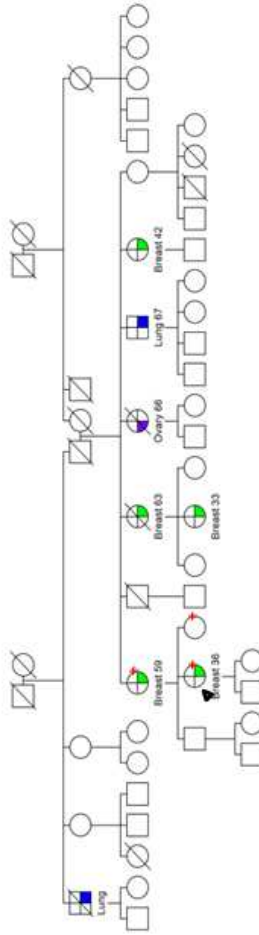
b) BRCA2 c.681+5G>C



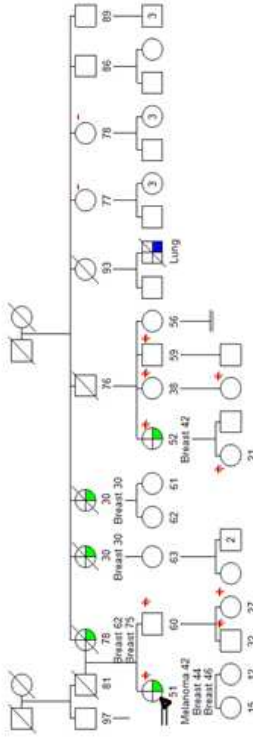
c) BRCA1 c.(441+1_442-1)_(547+1_548-1)del



d) BRCA1 c.(4357+1_4358-1)_(4484+1_4485-1)del



e) BRCA1 c.(5193+1_5194-1)_(5277+1_5278-1)del



f) BRCA2 c.(8632+1_8633-1)_(8754+1_8755-1)dup

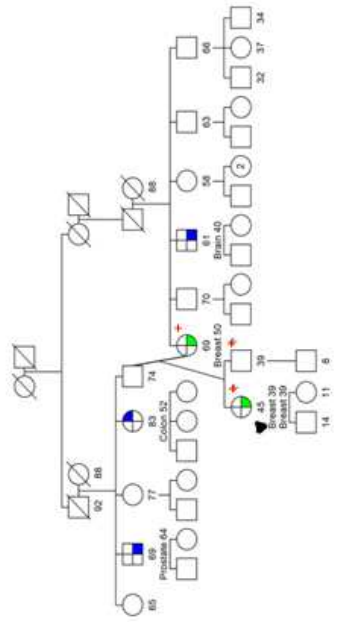


Figure Legends

Fig 1. Schematic representation of the different aberrant transcripts identified. **a)** Schematic representation of the aberrant splicing caused by the *BRCA1* c.4987-20_4987-11del variant. cDNA amplification of the *BRCA1* c.4987-20_4987-11del showed a double band in the agarose gel, one corresponding to the wild-type transcript (wt) (845 bp) and the other one corresponding to the transcript lacking exon 17 (757 bp). Sequencing of aberrant transcript cDNA showed complete skipping of exon 17. **b)** cDNA amplification and sequencing of *BRCA2* in the patient harboring c.681+5G>C showed the complete deletion of exon 8. Agarose gel for the cDNA amplification from the *BRCA2* c.681+5G>C carrier showed a double band, one corresponding to the wt allele (617 bp) and the other one to a transcript lacking exon 8 (568 bp). **c)** Scores of *in silico* predictions tools (SSF, MaxEnt, NNSplice, GeneSplicer, HSF) for the *BRCA1* c.4987-20_4987-11del and the *BRCA2* c.681+5G>C variants. Both variants are predicted to reduce the recognition of the affected splice site. The threshold used in each program is indicated in green under its name. nr, not recognized. **d)** Schematic representation of the aberrant splicing caused by the LGR *BRCA1* c.(441+1_442-1)_(547+1_548-1)del variant. cDNA amplification showed a double band, one corresponding to the wt transcript (347 bp) and the other to a transcript lacking exon 8 (242 bp), as shown in the electropherogram on the left. **e)** RT-PCR analysis of the LGR *BRCA1* c.(4357+1_4358-1)_(4484+1_4485-1)del transcript showed a double band, one corresponding to the wt transcript and the other one belonging to the deletion of exon 14 (610 bp). Reverse strand sequencing of the RT-PCR product showed an overlapped sequence after exon 14. **f)** Schematic representation of the aberrant splicing caused by the LGR *BRCA1* c.(5193+1_5194-1)_(5277+1_5278-1)del. cDNA amplification showed a double band, one corresponding to the wt transcript (727 bp) and the other to the transcript lacking exons 18, 19 and 20 (524 bp), as shown in the electropherogram on the right. **g)** cDNA amplification of the patient carrying the LGR *BRCA2* c.(8632+1_8633-1)_(8754+1_8755-1)dup led to a double band, one corresponding to the wt transcript (617 bp) and the other to the transcript carrying an exon 21 tandem duplication. Sequencing of the double band showed an overlapped sequence

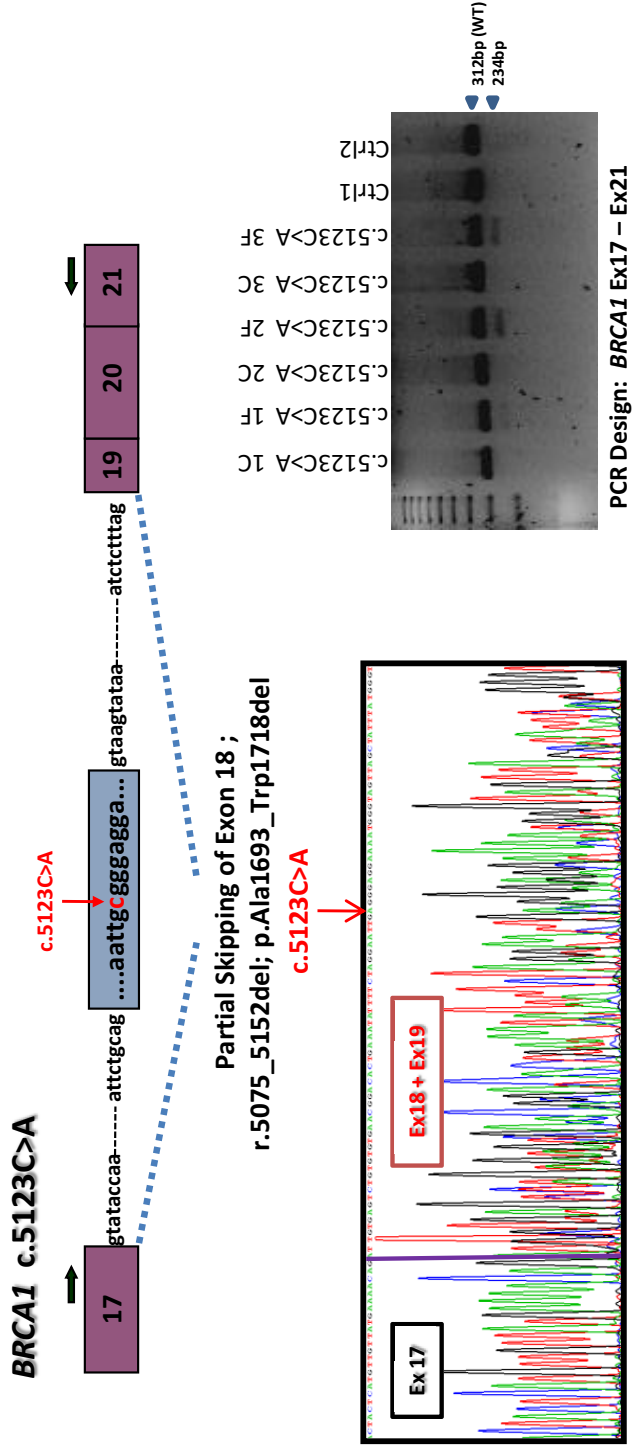
1 from the beginning of exon 22 corresponding to the sequences of exon 22 and the duplicated
2 exon 21.

3
4 **Fig 2.** Pedigrees from families of patients carrying the six VUS that showed an aberrant effect
5 in the RNA splicing. Black arrowhead was used to point the probands and red plus symbols to
6 mark carriers. Either, current age of individuals or age of death (when available) were located
7 under the individual symbol. Type of cancer with the age of onset at right side is written below
8 age.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Investigating the effect of 28 BRCA1 and BRCA2 mutations on their related transcribed mRNA. Breast Cancer Research and Treatment Journal. Francisco Quiles, Mireia Menéndez, Eva Tornero, Jesús del Valle, Àlex Teulé, Sarai Palanca, Angel Izquierdo, Carolina Gómez, Olga Campos, Raúl Santamaría, Joan Brunet, Gabriel Capellá, Lúdia Feliubadaló and Conxi Lázaro.

Corresponding Author: Conxi Lázaro Ph.D., Hereditary Cancer Program, Molecular Diagnostics Unit, Catalan Institute of Oncology (ICO-IDIBELL), Hospital Duran i Reynals, Gran Via 199, L'Hospitalet de Llobregat, Barcelona 08908, Spain. Tel.: +34932607342; Fax: +34932607466; E-mail: clazaro@iconcologia.net

Online resource 1. Schematic representation of the aberrant transcript identified for variant BRCA1 c.5123C>A



Online resource 1. Schematic representation of the aberrant splicing caused by the BRCA2 c.5123C>A variant. RNA from three BRCA2 c.5123C>A variant carriers (1, 2, 3) belonging to independent families was extracted from both cultured and frozen lymphocytes. Agarose gel showed a single band when RNA was extracted from cultured lymphocytes (C) corresponding to the wt transcript. In contrast, a double band was detected when RNA was extracted from frozen lymphocytes (F), one corresponding to the wt transcript (312 bp) and the other, in less proportion, corresponding to the transcript lacking exon 18 (234 bp), as shown in the electropherogram on the left.

Investigating the effect of 28 BRCA1 and BRCA2 mutations on their related transcribed mRNA. Breast Cancer Research and Treatment Journal (Francisco Quiles, Mireia Miréndez, Eva Tornero, Jesús del Valle, Alex Trufé, Sara Palanca, Angel Liquezudo, Carolina Gómez Olga Campos, Rial Santamaría, Joan Brunet, Gabriel Capells, Lidia Feliubadala and Conxi Lázaro. Corresponding Author: Conxi Lázaro Ph.D., Hereditary Cancer Program, Molecular Diagnostics Unit, Catalan Institute of Oncology (ICO-IDIBELL), Hospital Duran Reynals, Gram Via 199, L'Hospitalet del Llobregat, Barcelona, 08908, Spain. Tel.: +34932607342; Fax: +34932607466; E-mail: clazar@iconologia.net

Online Resource 2. Table showing the *in silico* predictions for variants analyzed in this study

Gene	Variant (HGVS)	Effect in the splicing	<i>In silico</i> prediction tools integrated in Alamut software					ESFinder	RESCLEE					
			SSF (0-100)**	MaxEnt (0-16)	NNSplice (0-1)	GeneSplicer (0-15)	HSF (0-100)							
BRCA1	c.8111T>C	No effect on splicing	D	70.09	>70***	7.05	6.26 (11.2%)	0.52	0.48 (7.0%)	77.26	75.42 (2.2%)	NA	No changes	
	c.181T>A	No effect on splicing	A	= 62.92	= 8.19	= 0.97	= 0.97	= 0.97	= 0.97	= 77.83	= 77.83	NA	NA	
	c.441+1_442-1_647+1_546-1del	r.442_547del	D	= 77.37	= 7.84	= 0.92	= 0.92	= 0.92	= 0.92	= 77.08	= 77.08	NA	NA	
	c.4357+1_4358-1_4484+1_4485-1del	r.4358_4484del	...	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
	c.6075+11A>G	No effect on splicing	D	= 78.15	= 6.84	= 0.44	= 0.44	= 0.44	= 0.44	= 83.37	= 83.37	NA	NA	
	c.4882A>G	No effect on splicing	D	nr	nr	nr	nr	nr	nr	nr	nr	NA	3 ESE were disrupted	
	c.4956G>A	No effect on splicing	D	= 70.38	= 5.91	= 0.66	= 0.66	= 0.66	= 0.66	= 81.24	= 81.24	NA	NA	
	c.4987_50_4987+1del	r.4987_5074del	D	= 84.55	= 78.31 (7.4%)	= 6.69	= 5.77 (13.6%)	= 0.61	= 0.35 (-43.3%)	= 1.36	= 1.61 (+18.0%)	= 87.03	= 83.96 (-4.0%)	NA
	c.5075-494A>C	No effect on splicing	A	= 92.62	= 8.96	= 0.89	= 0.89	= 0.89	= 0.89	= 91.29	= 91.29	NA	NA	
	c.5100A>G	No effect on splicing	A	= 92.62	= 8.96	= 0.89	= 0.89	= 0.89	= 0.89	= 91.29	= 91.29	NA	NA	
	c.5117G>A	No effect on splicing	D	= 92.62	= 8.96	= 0.89	= 0.89	= 0.89	= 0.89	= 91.29	= 91.29	NA	NA	
	c.5123C>A	r.5075_5152del (partial)	A	= 92.62	= 8.96	= 0.89	= 0.89	= 0.89	= 0.89	= 91.29	= 91.29	NA	NA	
	c.5154G>A	No effect on splicing	A	= 81.95	= 8.64 (11.5%)	= 0.96	= 0.95 (-1.0%)	= 0.96	= 0.95 (-1.0%)	= 83.79	= 83.62 (0.2%)	NA	NA	
	c.5175A>G	No effect on splicing	D	= 96.95	= 11.08	= 1.00	= 1.00	= 1.00	= 1.00	= 98.24	= 98.24	NA	NA	
	c.5193+1_5194-1_5277+1_5278-1del	r.5075_5278del	...	NA	NA	NA	NA	NA	NA	NA	NA	NA	1 new ESE	
	c.5467A>C	No effect on splicing	D	= 96.16	= 9.33	= 0.99	= 0.99	= 0.99	= 0.99	= 92.36	= 92.36	NA	NA	NA
BRCA2	c.39_12_39_10del1CT	No effect on splicing	New D (c.5467A>C)	71.85	82.68 (+15.0%)	5.39	6.64 (+23.2%)	0.28	0.56 (+96.6%)	77.07	85.87 (+11.4%)	NA	NA	
	c.599C>T	No effect on splicing	D	= 91.50	= 85.69 (-6.4%)	= 9.79	= 8.97 (-8.3%)	= 0.92	= 0.77 (-16.6%)	= 6.62	= 4.33 (-34.5%)	= 88.77	= 88.77	NA
	c.682A>C	No effect on splicing	D	= 78.15	= 6.84	= 0.44	= 0.44	= 0.44	= 0.44	= 83.37	= 83.37	NA	NA	
	c.681A>G	No effect on splicing	A	= 93.33	= 87.15 (-6.6%)	= 8.03	= 6.62 (-17.6%)	= 0.96	= 0.96 (+0.8%)	= 5.78	= 3.69 (-36.2%)	= 89.82	= 82.14 (-8.8%)	NA
	c.682-30A>C	r.632_681del	D	= 75.42	<70	= 8.56	= 3.25 (-62.1%)	= 0.93	= 0.1	= 82.62	= 70.61 (-14.5%)	NA	NA	
	c.1550A>G	No effect on splicing	D	= 88.05	= 7.05	= 0.18	= 0.18	= 0.18	= 0.18	= 4.31	= 4.26 (-1.0%)	= 86.68	= 86.68	NA
	c.1792A>G	No effect on splicing	D	nr	nr	nr	nr	nr	nr	nr	nr	NA	1 ESE disrupted	
	c.1863+1_1863-1_1875+1_1875-1dup	r.8633_8754dup	...	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
	c.8754-135G>A	No effect on splicing	D	nr	nr	nr	nr	nr	nr	nr	nr	NA	NA	
	c.8905G>A	No effect on splicing	D	= 100.00	= 10.86	= 1.00	= 1.00	= 1.00	= 1.00	= 4.24	= 4.27 (+0.7%)	= 100.00	= 100.00	NA
	c.8270C>T	No effect on splicing	D	= 91.58	= 13.32	= 0.98	= 0.98 (+0.4%)	= 10.29	= 10.07 (-2.2%)	= 87.14	= 87.14	NA	NA	
	c.9292T>C	No effect on splicing	A	= 91.58	= 13.32	= 0.98	= 0.98	= 0.98	= 0.98	= 10.29	= 10.80 (+4.9%)	= 87.14	= 87.14	NA

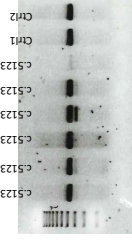
SS, splice site; A, consensus acceptor-splice site; D, consensus donor-splice site; NA, not applicable; nr, not recognized
 * The type of SS affected by the variant (A or D) is indicated. If the region tested by Alamut (199 bp centered in the variant) encompasses a whole exon, variations in both SS (A and D) are shown. In addition, we report new A or D SS when it was identified by more than one *in silico* program.
 ** In brackets, the range of scores used for each *in silico* program. In cells below, threshold used in Alamut.
 *** Wild-type scores are at the right side of the arrow and variant ones at the left side. In brackets, percentage of variation is shown. The symbol # indicates that no differences were found.

Online Resource 3. Table showing the *in silico* predictions for variants analyzed in this study

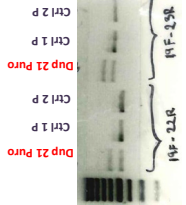
BRCA1	Variant (HGVS)	Exon affected	Forward	Reverse	T annealing	Size
	c.811T>C	Ex3	Ex2_5'-TTTATCGCTTGGCGTTGA	Ex7_5'-GACTTCACAGATGAAACCGGA	60	424
	c.181T>A	Ex5	Ex2_5'-TTTATCGCTTGGCGTTGA	Ex7_5'-GACTTCACAGATGAAACCGGA	60	424
	c.(441+1_442-1).(547+1_548-1)del	Ex8	Ex6_5'-CAGCTTACACAGGTTTCCA	Ex10_5'-CTCAGGAACACGGATGAAA	57	445
	c.(4357+1_4358-1).(4484+1_4485-1)del	Ex8	Ex13_5'-GACTCTTCTCGCCTTGAGGA	Ex16_5'-AGCAGGAGAAAGCGAATT	58	614
	c.4675+11A>G*	115	Ex14_5'-CCCTATAGCCAGAAATCCAGAA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	57	835
	c.4882G>C	Ex16	Ex14_5'-CCCTATAGCCAGAAATCCAGAA	Ex18_5'-TTGTGTGTGAACGGCACTG	58	715
	c.4956G>A	Ex16	Ex14_5'-CCCTATAGCCAGAAATCCAGAA	Ex18_5'-TTGTGTGTGAACGGCACTG	58	715
	c.4987-30_4987-11del*	116	Ex14_5'-CCCTATAGCCAGAAATCCAGAA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	57	835
	c.5075-49 del C	Ex18	Ex14_5'-CCCTATAGCCAGAAATCCAGAA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	59	299
	c.5100A>C	Ex18	Ex16_5'-GGGTCAACAAAGAATGTCCA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	59	299
	c.5170C>A	Ex18	Ex16_5'-GGGTCAACAAAGAATGTCCA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	59	299
	c.5124C>A #	Ex18	Ex17_5'-GTTGCCAGAAACACACAGA	Ex20_5'-GGAGATGTGTCATTTGGAAGA	59	232
	c.5154G>A	Ex19	Ex17_5'-GTTGCCAGAAACACACAGA	Ex24_5'-AGATGTGTAAGGCACCTGTG	60	496
	c.5175A>C	Ex19	Ex17_5'-GTTGCCAGAAACACACAGA	Ex24_5'-AGATGTGTAAGGCACCTGTG	60	496
	c.(193+1_194-1).(5278+1_5279-1)del	Ex20	Ex17_5'-GTTGCCAGAAACACACAGA	Ex24_5'-AGATGTGTAAGGCACCTGTG	60	496
	c.5467-9C>A	Ex23	Ex17_5'-GTTGCCAGAAACACACAGA	Ex24_5'-AGATGTGTAAGGCACCTGTG	60	496
	c.-39+12_39-10delTTCT	Ex2	Ex15-AGCTTACTTGGCGGCAAAA	Ex7_5'-GCTCACCCACCCCTTAGTTTC	58	578
	c.590C>T	Ex7	Ex4_5'-AGCAGATGATGTTTCTGTCTCC	Ex9_5'-CAGACAAGTGAACACAAATCAA	59	412
	c.632-3C>T	17	Ex6_5'-TGGTATGTGGAGTTTGTTTCA	Ex10_5'-CTCAAGACACCAATTGGAA	58	635
	c.681+5C>C	18	Ex6_5'-TGGTATGTGGAGTTTGTTTCA	Ex10_5'-CTCAAGACACCAATTGGAA	58	635
	c.682-30A>C	18	Ex6_5'-TGGTATGTGGAGTTTGTTTCA	Ex10_5'-CTCAAGACACCAATTGGAA	58	635
	c.1580A>G	Ex10	Ex9_5'-TCAAGAGAAAGCTGCAAGTCA	Ex11_5'-TTTTGGACAATTCGTAGGAA	59	1236
	c.1792A>G	Ex10	Ex9_5'-TCAAGAGAAAGCTGCAAGTCA	Ex11_5'-TTTTGGACAATTCGTAGGAA	59	1236
	c.8632+138G>A	Ex20	Ex18_5'-GATCTCCCTCTTACTGTCTTGA	Ex23_5'-TACTGAGTATTGGGTCCAT	58	682
	c.(8632+1_8633-1).(8754+1_8755-1)dup S	Ex21	Ex20_5'-GGAGAAAAGGAAAGCAGCAA	Ex23_5'-TACTGAGTATTGGGTCCAT	58	682
	c.8906G>C	Ex22	Ex20_5'-GGAGAAAAGGAAAGCAGCAA	Ex24_5'-CTTTCAGCCTCTGTTCTGAGG	61	678
	c.9270C>T	Ex25	Ex22_5'-CATGGAATCTCTGACAAAGG	Ex27_5'-CACTTTGTATGGCCAAAAGGA	60	860
	c.9292T>C	Ex25	Ex22_5'-CATGGAATCTCTGACAAAGG	Ex27_5'-CACTTTGTATGGCCAAAAGGA	60	860

* According to ENIGMA recommendations, forward and reverse primers must be at least 1 whole exon 5' or 3' of variant, respectively. For this reason, we performed a PCR using Primers: Ex13_5'-GACTCTTCTCGCCTTGAGGA and Ex20_5'-GGAGATGTGTCATTTGGAAGA. However, this PCR led to a constitutive alternative transcript lacking the first 3 bp of Exon 14. It made difficult the interpretation of the results, so we decided to use a Forward primer located in Exon 14, in order to avoid the constitutive transcript.

In first design, we used primers Ex16_5'-GGGTCAACAAAGAATGTCCA and Ex20_5'-GGGTCAACAAAGAATGTCCA and Ex20_5'-GGGTCAACAAAGAATGTCCA, according to ENIGMA recommendations. But in Online Resource 1 we showed an image of a PCR performed with Primers Ex17-Ex20 because the effect observed was the same but picture was of higher quality.



PCR Design: BRCA1 Ex16 – Ex21



\$ At first, we used primers Ex19_5'-TTCTCTTGCCCTTATCATCG and Ex23_5'-TACTGAGTATTGGGTCCAT, according to ENIGMA recommendations. However, in Figure 1, we showed the image of a PCR done with Primers Ex20-Ex23, because the effect was the same and we used 9 controls instead of 2.

Capítulo 2

ESTUDIO FUNCIONAL DE VSD EN LAS PROTEÍNAS BRCA1/2

ARTÍCULO 2

Artículo 2: Functional and structural analysis of C-terminal BRCA1 missense variants

Quiles F, Fernández-Rodríguez J, Mosca R, Feliubadaló L, Tornero E, Brunet J, Blanco I, Capellá G, Pujana MA, Aloy P, Monteiro A, Lázaro C.

PLoS One. 2013 Apr 17; 8(4)

Factor de Impacto (2013): 3,534

Functional and Structural Analysis of C-Terminal BRCA1 Missense Variants

Francisco Quiles¹, Juana Fernández-Rodríguez¹, Roberto Mosca², Lidia Feliubadaló¹, Eva Tornero¹, Joan Brunet¹, Ignacio Blanco¹, Gabriel Capellá¹, Miquel Àngel Pujana³, Patrick Aloy^{2,4}, Alvaro Monteiro⁵, Conxi Lázaro^{1*}

1 Hereditary Cancer Program, Catalan Institute of Oncology-(Bellvitge Institute for Biomedical Research; Girona Institute for Biomedical Research; Germans Trial i Pujol Research Institute) (ICO-IDIBELL, ICO-IdIBGI, ICO-IGTP), L'Hospitalet de Llobregat, Barcelona, Spain, **2** Institute for Research in Biomedicine (IRB) Barcelona, Joint IRB-BSC Program in Computational Biology, Barcelona, Spain, **3** Breast Cancer Unit, Translational Research Laboratory, Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain, **4** Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain, **5** Cancer Epidemiology Program, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida, United States of America

Abstract

Germline inactivating mutations in *BRCA1* and *BRCA2* genes are responsible for Hereditary Breast and Ovarian Cancer Syndrome (HBOCS). Genetic testing of these genes is available, although approximately 15% of tests identify variants of uncertain significance (VUS). Classification of these variants into pathogenic or non-pathogenic type is an important challenge in genetic diagnosis and counseling. The aim of the present study is to functionally assess a set of 7 missense VUS (Q1409L, S1473P, E1586G, R1589H, Y1703S, W1718L and G1770V) located in the C-terminal region of BRCA1 by combining *in silico* prediction tools and structural analysis with a transcription activation (TA) assay. The *in silico* prediction programs gave discrepant results making its interpretation difficult. Structural analysis of the three variants located in the BRCT domains (Y1703S, W1718L and G1770V) reveals significant alterations of BRCT structure. The TA assay shows that variants Y1703S, W1718L and G1770V dramatically compromise the transcriptional activity of BRCA1, while variants Q1409L, S1473P, E1586G and R1589H behave like wild-type BRCA1. In conclusion, our results suggest that variants Y1703S, W1718L and G1770V can be classified as likely pathogenic *BRCA1* mutations.

Citation: Quiles F, Fernández-Rodríguez J, Mosca R, Feliubadaló L, Tornero E, et al. (2013) Functional and Structural Analysis of C-Terminal BRCA1 Missense Variants. PLoS ONE 8(4): e61302. doi:10.1371/journal.pone.0061302

Editor: Paolo Peterlongo, IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Italy

Received: January 29, 2013; **Accepted:** March 7, 2013; **Published:** April 17, 2013

Copyright: © 2013 Quiles et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by: the Spanish Ministry of Health ISCIII FIS grants PI10/01422, PI09/02483 and PI12/01528, RETICC grants RD06/0020/1051, RD06/0020/1050, RD12/0036/008 and RD12/0036/0031; the AGAUR Catalan Government Agency grants 2009-SGR293 and 2009-SGR290; the Spanish Association Against Cancer (AECC 2010); and the Fundación Eugenio Rodríguez Pascual grant 2012. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: clazaro@iconcologia.net

Introduction

Between 10 and 20% of the breast cancer cases appearing in the general population present familial history of the disease [1]. Mutations in *BRCA1* and *BRCA2* genes confer high lifetime risks of breast and ovarian cancer, among other neoplasias [2,3]. Inactivating germline mutations in these genes account for 20–50% of familial cases, depending on the population [4]. Thus, genetic analysis of *BRCA1* and *BRCA2* is a cornerstone of genetic counseling practice. However, classification of genetic variants as pathogenic is challenging, particularly for missense changes and for silent or intronic variants that cannot be directly associated with increased cancer risk and are classified as variants of uncertain significance (VUS), which are found in 13% of *BRCA1* and *BRCA2* genetic tests [5]. Diverse multifactorial likelihood algorithms have been developed and applied for both *BRCA1* and *BRCA2* variants (reviewed in Spurdle et al., 2011) [6]. These models use the combination of a number of independent features (sequence conservation, type of amino-acid change, familial co-segregation, family and personal cancer history, tumor data, co-occurrence with a deleterious mutation, and severity of amino acid change) to establish the likelihood that a given VUS is pathogenic

or non-pathogenic. In many cases there is not enough data to classify such mutations for clinical purposes. Therefore, functional analyses that assess specific properties of BRCA1 or BRCA2 may help to classify VUS [7]. In particular, *BRCA1* encodes for a protein of 1,863 amino acids and with multiple functional domains [8]. Several functional assays have been conducted to evaluate VUS in BRCA1 at the level of its global and domain-based functions, including ubiquitin ligase activity assays, protease sensitivity assays, phosphopeptide binding assays, small colony phenotype assays, yeast localization phenotype assays, and embryonic stem cell-based functional assays (review in Millot et al., 2012) [9]. In the work presented here, we combine a functional assay - the transcription activation (TA) assay, which is based on the function of the BRCA1 carboxy-terminal region (aa 1396–1863) in transcriptional activation domain when linked to a sequence-specific DNA binding module - [10] with protein structural analyses [11] to assess the functional impact of seven BRCA1 C-terminal VUS. Our results indicate that three of them (Y1703S, W1718L, G1770V) have significant functional impact and may represent pathogenic *BRCA1* variants while the remaining four do not have a functional impact.

Materials and Methods

BRCA1 VUS

Seven missense BRCA1 C-terminal variants (Q1409L, S1473P, E1586G, R1589H, Y1703S, W1718L and G1770V), identified through genetic testing of patients with suspicion of Hereditary Breast and Ovarian Cancer Syndrome (HBOCS) as part of the Hereditary Cancer Program of the Catalan Institute of Oncology, were included in this study (Figure 1). Written informed consent was obtained from all subjects. The study received the approval of the Ethics Committee of IDIBELL. The pedigrees of these families are depicted in Figure S1. *BRCA1* and *BRCA2* genetic analyses consisted of screening of point mutations and large genomic rearrangements [12]. Genetic tests were carried out once patients had received appropriate genetic counseling and provided written informed consent.

In Silico Predictions

The *BRCA1* sequences containing the VUS were evaluated for their potential pathogenicity using the following algorithms: Align-GVGD [13], PolyPhen-2 [14], SIFT [15], Mutation Assessor [16], PhD-SNP [17], SNAP [18], I-Mutant [19], PON-P [20], and Condel [21]. The Align-GVGD algorithm was used at two depths of alignment (human to frog and human to sea urchin). The results of PolyPhen-2 were retrieved from the original webpage (version 2.2.2) but also from version 2.0.22 run by PON-P and version 1 run by Condel, which use them for weighted average scores.

Transcription Activation Assay

The assay was based on the mammalian expression of the firefly luciferase gene under the control of five GAL4 binding sites [10]. In this system, pcDNA3 constructs containing the GAL4 DNA binding domain (DBD) fused to the wild-type or VUS-containing BRCA1 C-terminal region were used. Pathogenic mutation controls (M1775R and Y1853X) were included in the assays. Constructs containing the VUS were generated by site-directed mutagenesis using the QuickChange II Site-Directed Mutagenesis Kit (Agilent Technologies, USA) and the wild-type sequence as a template (U14680). Human embryonic kidney (HEK) 293T cells were co-transfected (X-tremeGENE 9 DNA transfection Reagent, Roche Applied Science, Mannheim, Germany) with the pcDNA3 construct encoding either the wild-type or the VUS sequence, the pG5Luc plasmid that contains the luciferase reporter gene, and

the pHRG-TK plasmid that contains a *Renilla* luciferase gene under the control of a constitutive thymidine kinase promoter, which acts as an internal control. Luciferase was quantified using the Dual-Luciferase Reporter Assay System (Promega, Madison, USA).

Protein Structural Rationalizations

To predict the impact of the VUS on both the folding and phosphopeptide binding of the BRCT domains a structural analysis was carried out by mapping the position of the variants on the BRCA1 structures collected from the Protein Data Bank [22] (PDB, www.pdb.org) and analyzing them manually with Pymol (<http://www.pymol.org>). The structures were collected by BLAST-querying the PDB using the BRCA1 sequence (UniProt accession code P38398). Only the VUS mapped in the BRCT tandem domains (Y1703S, W1718L and G1770V) were analyzed using this approach, as there are no structural data for the upstream domain or regions.

Results

In Silico Predictions

Several algorithms designed to predict deleterious versus neutral missense changes were used (Table 1). Align-GVGD [13], PolyPhen-2 [14], SIFT [15], Mutation Assessor [16], PhD-SNP [17], SNAP [18] and I-Mutant [19] differ in the properties taken into account to generate the prediction and in the method and possible training for decision-making (reviewed in Thursberg, et al., 2010) [23]. PON-P [20] integrates PolyPhen2, SIFT, PhD-SNP, SNAP and I-Mutant using a random forest method, whereas Condel integrates PolyPhen2, SIFT and Mutation Assessor using a weighted average of the normalized scores of the individual methods. Both integration methods claim to outperform individual ones [20,21]. Table 1 shows the outcomes of all methods, for comparison with high discrepancy between different programs. The two integrative algorithms indicate that variants Q1409L, S1473P, E1586G and R1589H have a benign effect and that G1770V has a damaging effect, whereas they disagree in their predictions for variants Y1703S and W1718L. Furthermore the results of PON-P and Condel for Q1409L and G1770V strikingly diverge from those of Align-GVGD, which classify Q1409L as most likely to interfere with BRCA1 function and G1770V as less likely using the alignment from human to frog. Also SIFT,

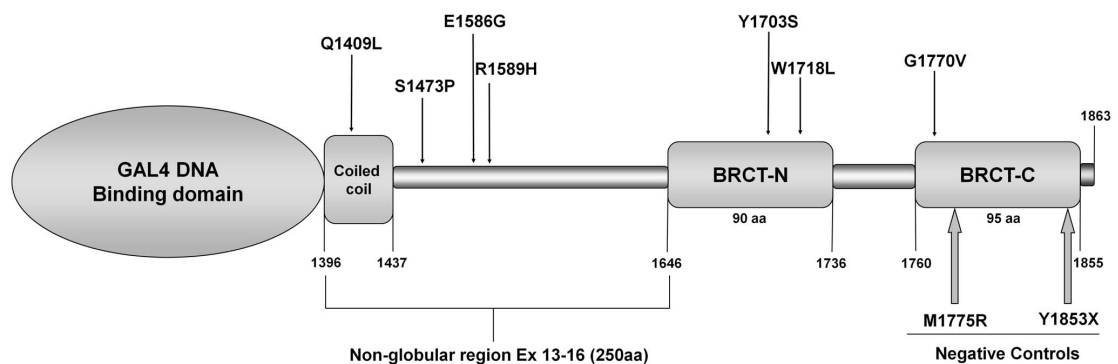


Figure 1. Illustration depicting the location of the C-terminal BRCA missense variants and negative controls. This Figure shows the position of missense variants analyzed in the present study (black arrows) and the negative controls (closed gray arrows, underlined) within the vector used in the TA experiments.

doi:10.1371/journal.pone.0061302.g001

Table 1. *In silico* predictions for pathogenicity of the variants presented in this study.

Predictor ¹		Align-GVGD ²		PolyPhen-2 v2.2.2 (Original)		PolyPhen-2 v2.0.22 (PON-P)		PolyPhen-2 v1 (Condel)		SIFT	
DNA variant	Protein variant	Human to frog	Human to sea urchin	Prob.	Pred.	Prob.	Class	Prob.	Pred.	Prob.	Class
c.4226A>T	Q1409L	C65	C15	0.015	benign	0.995	Del.	0.007	benign	0.00	damaging
c.4417T>C	S1473P	C0	C0	0.004	benign	0.960	Del.	0.124	benign	0.01	damaging
c.4757A>G	E1586G	C0	C0	0.003	benign	0.995	Del.	0.124	benign	0.01	damaging
c.4766G>A	R1589H	C0	C0	0.000	benign	0.000	N	0.000	benign	1.00	tolerated
c.5108A>C	Y1703S	C65	C65	0.430	benign	1.000	Del.	0.689	Possib. damaging	0.00	damaging
c.5153G>T	W1718L	C55	C55	0.086	benign	1.000	Del.	0.497	benign	0.00	damaging
c.5309G>T	G1770V	C0	C0	0.093	benign	0.999	Del.	0.964	possib. damaging	0.00	damaging

Mutation Assessor		PhD-SNP		SNAP		I-Mutant		PON-P ^o		Condel		
Funct. Impact Score	Funct. Impact	Pred.	Reliability	Pred.	Accur.	$\Delta\Delta G$	Pred.	Pred.	Class	Accur.	Prob.	Class
1.100	low	N	3	N	60	0.22	N	0.14	N	0.84	0.000	N
0.975	low	N	3	N	85	-2.07	Dest	0.18	N	0.76	0.018	N
1.040	low	N	8	P	58	-1.81	Dest	0.12	N	0.84	0.014	N
-1.040	neutral	N	8	N	92	-1.52	Dest	0.00	N	0.86	0.000	N
2.670	medium	P	9	P	58	-1.44	Dest	0.39	UV	0.60	0.746	Del
3.105	medium	P	4	P	58	-0.81	N	0.57	UV	0.46	0.065	N
2.135	medium	P	8	P	78	-1.20	Dest	0.98	P	0.96	0.906	Del

Each predictor yields its own type of results, usually quantitative and categorical. We have chosen the most representative of each. ¹Predictors PON-P and Condel integrate results from other predictors run in their own servers; as PolyPhen-2 versions and their results differ if they are run in the PolyPhen-2 website or the other 2, all of them are shown for comparison. ²Classification of variants in different classes according to Align-GVGD algorithm (Align-Grantham Variation Grantham Deviation; <http://agvgd.iarc.fr>), C65 means "Most likely to interfere with function, C0 means "least likely" (possible classes are: C65>C55>C45>C35>C25>C15>C0). Abbreviations: Prob.-Probability; Pred.-Prediction; Funct.-Functional; Accur.-Accuracy; $\Delta\Delta G$ - $\Delta\Delta G$ value (kcal/mol); Del-Deleterious; N-Neutral; Dest- Destabilizing; P-Pathogenic; UV-Unknown Variant. doi:10.1371/journal.pone.0061302.t001

Mutation Assessor, PhD-SNP, SNAP and I-Mutant show divergent tendencies.

Transcription Activation Assay

Using the transcriptional assay, three VUS (Y1703S, W1718L and G1770V) showed a significant decrease in reporter expression compared to wild-type *BRCA1* (Figure 2). The luciferase activity of these three mutants was very similar to that observed for the known pathogenic mutants used as controls: that is <5%, compared to the wild-type construct. Conversely, the remaining four VUS (Q1409L, S1473P, E1586G and R1589H) displayed similar transcriptional activity to that of the wild-type construct.

Protein Structural Rationalizations

A structural analysis was performed for those VUS located in the BRCT domains (Y1703S, W1718L and G1770V). Y1703 is located at the BRCT dual-repeat-interaction interface (Figure 3A), directed towards the linker. It forms a hydrogen bond with H1746 (Figure 3B). The change to a serine, by removal of the aromatic ring, is likely to disrupt this bond and the surrounding hydrophobic interactions and to affect the stability of the domain and the dual-repeat interaction. It may also affect the peptide binding through K1702 (Figure 3B), which interacts directly with the peptide's phosphoserine. W1718 is loosely packed in the core of the BRCT-N domain and is part of a highly conserved motif

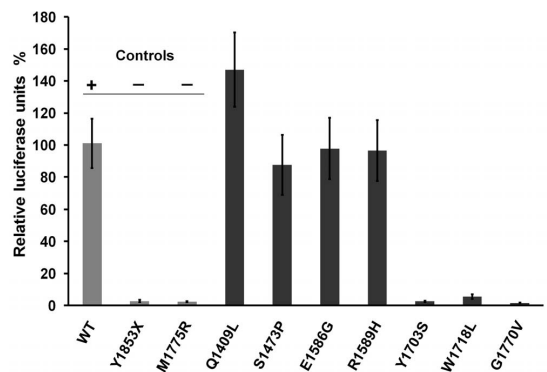


Figure 2. Results of functional assay of the studied missense variants. Percentage of transcriptional activity was expressed as percentage relative to the wild-type construct. Experiments were performed in triplicate and normalized against the *Renilla* luciferase internal control. Results are shown in bars: gray for controls and black for variants. doi:10.1371/journal.pone.0061302.g002

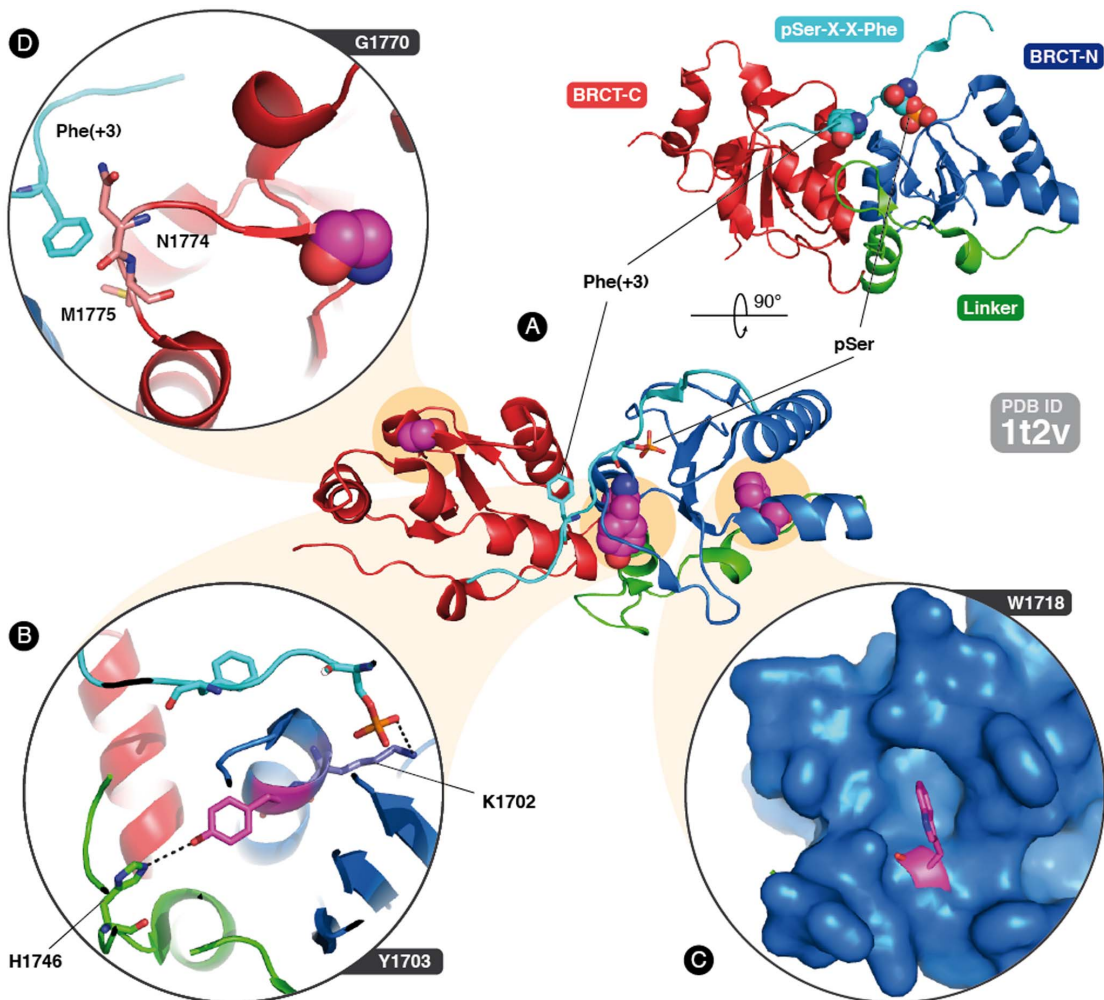


Figure 3. Structural analysis of missense variants lying in the BRCA1 BRCT domain. Representation of the structure of the BRCT tandem repeat from human BRCA1 in complex with a binding phosphopeptide (with the typical binding motif pSer-X-X-Phe, PDB ID 1t2v): in blue, the N-terminal repeat (BRCT-N); in red, the C-terminal repeat (BRCT-C); in green, the linker region between the two; in cyan, the phosphopeptide with the phosphoserine and the key binding residue Phe(+3); in magenta the three mutated residues. Y1703 (A) forms a hydrogen bond with H1746 and is adjacent to K1702, which binds directly to the pSer residue in the phosphopeptide. W1718 (B) fills the internal core and stabilizes the BRCT-N domain. G1770 (C) gives the loop connecting the β 1 strand with α 1-helix in the BRCT-C repeat the necessary flexibility to form a tight turn, bringing the residues N1774 and M1775 close to the phosphopeptide for interaction with the Phe(+3) key residue.
doi:10.1371/journal.pone.0061302.g003

within the α 3-helix (Figure 3C). The substitution of this bulky residue with a leucine likely destabilizes the folding of the domain, analogously to experimentally confirmed observations of substitution by a cysteine [24,25]. G1770 is solvent-exposed and located on a tight turn of the loop connecting the β 1 strand with the α 1 helix in the BRCT-C repeat (Figure 3D). It is highly conserved in the Pfam alignment of the BRCT domain and probably confers to the loop the necessary flexibility for phosphopeptide binding. Substitution by a valine would expose a hydrophobic side chain and confer rigidity to the loop, compromising the binding. In contrast to the other mutations, this change may not affect folding.

Discussion

This study combined a functional BRCA1 assay and structural analyses to predict whether specific missense variants in the BRCA1 C-terminal region have pathological significance, contributing to breast cancer risk. The variants selected were VUS identified in a clinical scenario, which are located between amino acid residues 1,396 and 1,863 of the full-length BRCA1 sequence and include part of a coiled-coil domain (aa 1,364–1,437) and the BRCT tandem domains (aa 1,646–1,863) (Figure 1). Notably, six of the VUS (Q1409L, S1473P, E1586G, Y1703S, W1718L and G1770V) had not been recorded in the locus-specific database for

mutations in *BRCA1* and *BRCA2*, BIC (Breast Cancer Information Core database, <http://research.nhgri.nih.gov/bic/>) (Figure 1 and Table 1).

The use of different *in silico* prediction algorithms gave contradictory results for some of the variants (Table 1) making it difficult to draw any clear conclusion. For example, there is a striking difference between the results for three versions of PolyPhen-2 as run from different sources; W1718L can yield probabilities of pathogenicity of 0.086, 0.497 and 1.000, while the probabilities for Q1409L, S1473P and E1586G range from values below 0.2 to values above 0.95 in the different versions. If we tried to order the variants according to their deleteriousness as predicted from each of the algorithms, they would yield as many orders as algorithms tested. In general terms, the two integrative programs gave quite similar results, predicting a damaging effect for one of the variants (G1770V) and a benign effect for four of them (Q1409L, S1473P, E1586G and R1589H), while variants Y1703S and W1718L were difficult to classify. Therefore, attending to the results obtained in this study, we consider the use of *in silico* programs not reliable enough to get an approximation of the pathogenicity of a given VUS. In a recent published paper by Li *et al.*, the predictive power of Condel and five individual methods (SIFT and PolyPhen2 among them) in discriminating between pathogenic nsSNVs and other rare nsSNVs, was examined [26]. Their results showed that the predictive power of the combined Condel methodology is not necessarily better than individual methods. Furthermore and in agreement with our data, they show that correlations in the scores between the predictive methods tested are weak or moderate. The observed discrepancies among programs clearly support the need of adding functional assay results in the process of classification of VUS.

The TA assay in mammalian cells was performed as described in previous studies [10]. This assay has been extensively validated, proving to be extremely robust for variants located in the BRCT domain [7,10,27,28]. In a previous study it has been benchmarked using 14 pathogenic and 10 not pathogenic variants, showing high sensitivity and specificity [28]. In this study two constructs containing previously classified pathogenic variants (IARC Class 5), M1775R and Y1853X, were used as negative controls and a construct with the wild-type BRCA1 C-terminal region was used as a positive control. This assay clearly identified a functional defect for three of the variants (Y1703S, W1718L and G1770V), whereas the remaining variants gave results very similar to those of the wild-type controls. It should be noted that residue W1718 appears to be highly prone to mutation, and although the mutation studied here has never been reported, mutations W1718S and W1718C have been described in HBOCS families; in addition, a panel of assays including protease sensitivity, phosphopeptide binding activity and specificity, as well as the TA assay, revealed that both of these variants have strong functional effects [28]. Both constitute large volume changes at a rigid position in a rigid region that is part of the BRCT signature motif WXXXX [29].

Next, to better interpret and understand the functional data obtained, a structural analysis was performed for those VUS

located in the BRCT domains. This analysis was limited to the BRCT domain as its structure has been solved [11], making it possible to rationalize whether a given variant alters the domain stability or the phosphopeptide-mediated binding to phosphorylated protein partners (Figure 3). The results indicated that the three studied VUS in the BRCT domains seem to damage proper BRCA1 structure. Interestingly, the experimental data obtained in the TA assay support these structural predictions.

The pedigree structure of patients with these missense mutations was, in general, not very informative for co-segregation analysis (see Figure S1), either due to the lack of relatives with breast (or ovarian) cancer or because no DNA samples for other relatives were available. Nevertheless, it is worth noting that variant Q1409L, which appears to have no effect on TA activity in the present study, was found in a patient with a pathogenic mutation in *BRCA2* (c.5720_5723delCTCT), and although a small number of carriers of two pathogenic mutations - one in *BRCA1* and one in *BRCA2* - have been described worldwide this event has a very low probability. In spite of this, and due to the fact that mutation Q1409L is mapped in the BRCA1/PALB2 interaction region, it would be worth to test if mutant L1409 is hampering the BRCA1/PALB2 interaction [30]. Moreover, mutation G1770V, which shows a clear impairment of TA activity, has been found in two families of Moroccan origin presumed to be independent. In both cases, patients suffer from breast or ovarian cancer at a very early age of onset (27 and 30 years of age, respectively).

In summary, in a clinical context, extensive worldwide analysis of *BRCA1* has identified a large number of VUS, most of them unique, which need to be classified for the purposes of diagnosis and genetic counseling. By combining functional analysis and structural rationalizations, this study shows that three *BRCA1* VUS (Y1703S, W1718L and G1770V) have profound functional impact and may be pathogenic variants. By contrast, the other four variants did not alter the function of BRCA1 in the TA assay and should remain classified as VUS. The implementation of combined approaches, such as the one presented here, is therefore of considerable clinical relevance, especially in those cases where co-segregation studies are not possible.

Supporting Information

Figure S1 Pedigrees of the families with the studied VUS. (DOC)

Acknowledgments

We thank Montserrat Soler from the Institute for Research in Biomedicine (IRB) for the fruitful discussion and suggestions. We would also like to thank the Spanish Association Against Cancer (AECC) for recognizing our group with one of its awards.

Author Contributions

Conceived and designed the experiments: GC MAP PA AM CL. Performed the experiments: FQ JFR LF ET RM. Analyzed the data: FQ JFR LF RM JB IB CL. Contributed reagents/materials/analysis tools: RM MAP PA AM. Wrote the paper: FQ RM MAP GA AL CL.

References

- Balmain A, Gray J, Ponder B (2003) The genetics and genomics of cancer. *nature genetics* 33: 238–244.
- Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, et al. (1994) BRCA1 mutations in primary breast and ovarian carcinomas. *Science* 266: 120–122.
- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266: 66–71.
- Kurian AW (2010) BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Curr Opin Obstet Gynecol* 22: 72–78.

5. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, et al. (2002) Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 20: 1480–1490.
6. Spurdle AB, Healey S, Devereau A, Hogervorst FB, Monteiro AN, et al. (2011) 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.
7. Carvalho MA, Couch FJ, Monteiro AN (2007) Functional assays for BRCA1 and BRCA2. *Int J Biochem Cell Biol* 39: 298–310.
8. Clark SL, Rodriguez AM, Snyder RR, Hankins GD, Boehning D (2012) Structure-Function Of The Tumor Suppressor BRCA1. *Comput Struct Biotechnol J* 1: e201204005.
9. Millot GA, Carvalho MA, Caputo SM, Vreeswijk MP, Brown MA, et al. (2012) A guide for functional analysis of BRCA1 variants of uncertain significance. *Hum Mutat* 33: 1526–1537.
10. Phelan CM, Dapic V, Tice B, Favis R, Kwan E, et al. (2005) Classification of BRCA1 missense variants of unknown clinical significance. *J Med Genet* 42: 138–146.
11. Williams RS, Green R, Glover JN (2001) Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nat Struct Biol* 8: 838–842.
12. Feliúbadaló L, Lopez-Doriga A, Castellsague E, del Valle J, Menéndez M, et al. (2012) Next-generation sequencing meets genetic diagnostics: development of a comprehensive workflow for the analysis of BRCA1 and BRCA2 genes. *European Journal of Human Genetics* In press.
13. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, et al. (2006) Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 43: 295–305.
14. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7: 248–249.
15. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4: 1073–1081.
16. Reva B, Antipin Y, Sander C (2011) Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* 39: e118.
17. Capriotti E, Calabrese R, Casadio R (2006) Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics* 22: 2729–2734.
18. Bromberg Y, Rost B (2007) SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res* 35: 3823–3835.
19. Capriotti E, Fariselli P, Rossi I, Casadio R (2008) A three-state prediction of single point mutations on protein stability changes. *BMC Bioinformatics* 9: S6.
20. Olatubosun A, Valiaho J, Harkonen J, Thusberg J, Vihinen M (2012) PON-P: integrated predictor for pathogenicity of missense variants. *Hum Mutat* 33: 1166–1174.
21. Gonzalez-Perez A, Lopez-Bigas N (2011) Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score. *Condel. Am J Hum Genet* 88: 440–449.
22. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.
23. Thusberg J, Olatubosun A, Vihinen M (2011) Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat* 32: 358–368.
24. Williams RS, Chasman DI, Hau DD, Hui B, Lau AY, et al. (2003) Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations. *J Biol Chem* 278: 53007–53016.
25. Glover J (2006) Insights into the Molecular Basis of Human Hereditary Breast Cancer from Studies of the BRCA1 BRCT Domain. *Fam Cancer*: 89–93.
26. Li M-X, Swan JSH, Bao S-YB, Yang W, Ho S-L, et al. (2013) Predicting Mendelian Disease-Causing Non-Synonymous Single Nucleotide Variants in Exome Sequencing Studies. *PLOS Genetics* 9(1): e1003143 doi:10.1371/journal.pgen.1003143.
27. Carvalho MA, Marsillac SM, Karchin R, Manoukian S, Grist S, et al. (2007) Determination of cancer risk associated with germ line BRCA1 missense variants by functional analysis. *Cancer Res* 67: 1494–1501.
28. Lee MS, Green R, Marsillac SM, Coquelle N, Williams RS, et al. (2010) Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays. *Cancer Res* 70: 4880–4890.
29. Mirkovic N, Marti-Renom MA, Weber BL, Sali A, Monteiro AN (2004) Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposition. *Cancer Res* 64: 3790–3797.
30. Zhang F, Fan Q, Ren K, Andreassen PR (2009) PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2. *Mol Cancer Res* 7: 1110–1118.

Resultados sometidos a publicación:***Identification of a founder BRCA1 mutation in Moroccan population.***

Francisco Quiles¹ (BSc), Àlex Teulé¹ (MD), Nora Martinussen Tandstad² (MD), Lídia Feliubadaló¹ (PhD), Eva Tornero¹ (BSc), Jesús del Valle¹ (PhD), Mireia Menéndez¹ (PhD), Mónica Salinas (BSc), Vibeke Wethe Rognlien² (MD), Angela Velasco (BSc), Angel Izquierdo¹ (PhD), Gabriel Capellá¹ (MD,MPH), Joan Brunet¹ (MD, PhD), and Conxi Lázaro¹ (PhD).

Clinical Genetics

Factor de Impacto (2015): 3,931

Full title:**Identification of a founder *BRCA1* mutation in the Moroccan population****Authors:**

Francisco Quiles¹ (BSc), Àlex Teulé¹ (MD), Nora Martinussen Tandstad² (MD), Lidia Feliubadaló¹ (PhD), Eva Tornero¹ (BSc), Jesús del Valle¹ (PhD), Mireia Menéndez¹ (PhD), Mónica Salinas (BSc), Vibeke Wethe Rognlien² (MD), Angela Velasco (BSc), Angel Izquierdo¹ (MD, MPH), Gabriel Capellá¹ (MD, PhD), Joan Brunet¹ (MD, PhD), and Conxi Lázaro¹ (PhD)

Authors' affiliations:

¹ Hereditary Cancer Program, Catalan Institute of Oncology (ICO-IDIBELL, ICO-IdIBGi, ICO-IGTP), L'Hospitalet de Llobregat, Barcelona, Spain.

² Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

Corresponding author:

Conxi Lázaro Ph.D., Hereditary Cancer Program, Molecular Diagnostics Unit, Catalan Institute of Oncology (ICO-IDIBELL), Hospital Duran i Reynals, Gran Via 199, L'Hospitalet de Llobregat, Barcelona 08908, Spain. Tel.: +34932607342; Fax: +34932607466; E-mail: clazaro@iconcologia.net

Conflict of interest statement

The authors of this paper certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial

1
2
3 interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the
4
5 subject matter or materials discussed in this manuscript.
6
7

8 9 **Acknowledgments**

10
11 We thank all patients who contributed to the study. The authors would like to particularly
12
13 acknowledge the support of the Asociación Española Contra el Cáncer (AECC). Contract grant
14
15 sponsor: Instituto de Salud Carlos III, organismo adscrito al Ministerio de Economía y
16
17 Competitividad and “Fondo Europeo de Desarrollo Regional (FEDER), una manera de hacer
18
19 Europa”, Catalan Health Institute and Autonomous Government of Catalonia. Contract grant
20
21 numbers: PI10/01422, PI13/00285, PIE13/00022, RD12/0036/0008, RD12/0036/0014,
22
23 RD12/0036/0031, 2009SGR290 and 2014SGR338.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Purpose: Breast cancer (BC) is the most frequent cancer among women in Morocco. However, the role of the most prevalent BC-predisposing genes, *BRCA1* and *BRCA2*, has been largely unexplored. To help define the role of *BRCA1* in breast cancer in Morocco, we characterized the first potential *BRCA1* founder mutation in this population.

Materials and Methods: Genetic testing of *BRCA1* and *BRCA2* in BC high-risk families identified mutation *BRCA1* c.5309G>T, p.(Gly1770Val) or G1770V in five independent families from Morocco, suggesting a founder effect. To confirm this hypothesis, haplotype construction was performed using seven intragenic and flanking *BRCA1* microsatellite markers. Clinical data were also compiled.

Results: Clinical data from carriers of mutation G1770V correspond to data from carriers of *BRCA1* pathogenic mutations. Microsatellite analysis showed a common haplotype for the five families in a region comprising 1.54Mb, confirming G1770V as the first specific founder *BRCA1* mutation in the Moroccan population.

Conclusion: Our findings contribute to a better understanding of BC genetics in the Moroccan population. Nevertheless, comprehensive studies of mutation G1770V in large series of BC patients from Morocco are needed to assess the real prevalence of this mutation and to improve genetic testing and risk assessment in this population.

Keywords: *BRCA1*, founder mutation, breast, cancer, Morocco, hereditary, genetics

INTRODUCTION

Between 10 and 20% of the breast cancer cases appearing in the general population present familial history of the disease(1). Mutations in *BRCA1* and *BRCA2* genes drastically increase lifetime risk of breast and ovarian cancer, among other neoplasias(2, 3). Inactivating germline mutations in these genes may account for 20-50% of familial cases, depending on the population(4). In North Africa (including Morocco, Algeria and Tunisia), the incidence of the disease has increased considerably in recent years, to the extent that breast cancer is now the most common malignancy among women(5). Interestingly, the average age of onset in this region is more than ten years below that of Western countries(6, 7). In Morocco, average onset occurs at 48 years, with 7% of cases diagnosed at less than 35 years(7, 8). In addition, the size and grade of breast tumors at diagnosis are higher in North Africa than elsewhere(6, 7). Such features suggest a strong influence of high-penetrance genetic factors in breast cancers appearing in this region. However, the contribution of mutations in breast and ovarian cancer-predisposing genes in this region remains largely unexplored. Few genetic studies of *BRCA1* have been reported for this population, examining less than 500 familial cases in total (reviewed in Laraqui et al., 2015)(7). An overview of such studies indicates that the North African population presents a spectrum of prevalent *BRCA1* mutations, some of which appear as recurrent or founder mutations in other populations. Founder mutations, understood as rare mutations which become frequent in certain populations due to isolation and consequent inbreeding, may account for a high percentage of familial cases in these populations(9). For example, three founder mutations (*BRCA1* c.68_69delAG, *BRCA1* c.5266dupC and *BRCA2* c.5946delT) account for 98-99% of mutations identified in breast and ovarian familial cancers associated with *BRCA* genes in the Ashkenazi Jewish population(9). The Arab population is genetically unique, characterized by a tribal lifestyle with very large families and frequent consanguinity(6). Therefore, the identification and geographic mapping of founder mutations can lead to cheaper and quicker mutational screening of high-risk patients. However, to date, only one founder *BRCA1* mutation (c.798_799delTT) specific to North African populations has been described(8).

1
2
3 Here, we report a new founder mutation in *BRCA1* gene specific to the Moroccan
4 population, c.5309G>T (G1770V). G1770V is located in the BRCT tandem domain of the
5
6 BRCA1 protein and has been identified as probably pathogenic on the basis of two independent
7
8 functional studies(10, 11). We found G1770V in five apparently unrelated Moroccan families
9
10 sharing a common haplotype, confirming it as a founder mutation. Case-control studies of
11
12 different North African populations are needed to clarify the contribution of this *BRCA1*
13
14 mutation to breast and ovarian cancer predisposition.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

MATERIAL AND METHODS

Patients

A total of 11 individuals, including eight carriers and three non-carriers of variant c.5309G>T (G1770V), were included in this study. They belonged to five unrelated families of Moroccan origin (Figure 1). Routine mutational analysis of *BRCA1* identified the presence of *BRCA1* c.5309G>T in the index case of these families. Three of the five families were high-risk families with suspected Hereditary Breast and Ovarian Cancer Syndrome (HBOCS) examined under the Hereditary Cancer Program at the Catalan Institute of Oncology (ICO) in Barcelona, Catalonia, Spain. The remaining two families were identified at Oslo University Hospital (Norway). Detailed family information from at least three generations and data on geographic origins were obtained. No relationships were identified between individuals from the different families.

Genetic counselors collected clinical and pathological data from affected carriers. Informed consent was obtained from all subjects evaluated in this study.

Microsatellite genotyping and haplotype analysis

Haplotypes were constructed manually from two intragenic and five *BRCA1* flanking microsatellite markers: *D17S486*, *D17S1793*, *D17S855*, *D17S1323*, *D17S1325*, *D17S951*, *D17S1183* [references and primer sequences described at USCS Genome Bioinformatics, (<https://genome.ucsc.edu>)], assuming the lowest number of recombinations. Amplifications were performed with one fluorescently labelled primer per set according to standard procedures. Fragment analysis was performed on an ABI 3130xl capillary sequencer using GeneMapper software (Applied Biosystems, Forster City, CA, USA).

RESULTS

Mutation screening of *BRCA1* identified the variant c.5309G>T in five breast/ovarian cancer patients belonging to five unrelated families of Moroccan origin. Direct sequencing of c.5309G>T in the patients' relatives identified three additional carriers and three non-carriers (DNA from non-carrier in Family C not available for haplotype analysis) (Figure 1). Five out of the eight carriers were diagnosed with breast tumors and one was diagnosed with an ovarian tumor; two of the carriers were asymptomatic at the time of the study (26 and 57 years old, respectively). The average age at first diagnosis was 36.5 years (SD=10.2, range: 27-57) (Table 1). Thirty-seven percent of carriers showed an age of onset below 35 years and 62.5% below 40 years. All breast tumors showed similar histopathological features: triple-negative invasive carcinomas, grade 3. The ovarian tumor was a serous invasive cystadenocarcinoma, grade 3 (Table 1). All families tested in this study were originally from the north coast of Morocco (Figure 1). Specifically, one family screened in Norway (Family C) and two families from Catalonia (A, B) were originally from the city of Nador. The other family identified in Norway was originally from Oudja. The remaining family, identified in Catalonia, was originally from Tangier (Table 1).

The identification of the same variant in five independent families from the same area of Morocco suggested a common ancestor. To confirm this hypothesis, two microsatellites markers in *BRCA1* and five flanking markers were analyzed in the ten individuals included in this study (Figure 1). In family B, DNA from two carriers and two non-carriers was available, and a haplotype spanning 2.62 Mb, linked to the c.5309G>T variant, was univocally constructed (Figure 1). In family C, DNA from three carriers was available (DNA from the non-carrier was not available), and the alleles of markers *DI7S855* and *DI7S1325* associated with c.5309G>T were ascertained and shared with family B; the remaining markers of the associated haplotype in family B could not be univocally defined in family C. For the remaining families, only DNA from the proband was available. All patients from families A, C, D and E carry the alleles

1
2
3 linked to c.5309C>T in family B, in the region extending from *D17S1793* to *D17S1171* (1.54
4 Mb), although the exact phase could not be assessed in all cases.
5
6

7 **DISCUSSION**

8
9 Breast cancer is the leading form of cancer among women in Morocco and has increased in
10 incidence in recent years(8). In fact, a recent study of a female population in Rabat showed that
11 breast cancer accounts for almost 40% of all new cancer cases(12). However, the contribution
12 of *BRCA1* mutations to breast and ovarian cancer in this region is limited to three studies,
13 covering a total of only 145 familial cases(7, 13, 14). Here we report a new *BRCA1* founder
14 mutation in the Moroccan population, c.5309G>T. This mutation was first reported in Spain in
15 two families of Moroccan origin and was classified as probably pathogenic on the basis of a
16 combination of functional and structural analyses(11). The study clearly showed that the
17 corresponding protein variant G1770V impairs the transcriptional activity of BRCA1 and also
18 predicted that the variant would affect the phosphor-peptide-mediated binding of BRCA1 to
19 phosphorylated protein partners(11). More evidence of the pathogenicity of this variant was
20 provided by a functional complementation assay that showed high sensitivity to cisplatin and
21 growth retardation for G1770V(10). Subsequently, this variant was identified twice in different
22 families from Norway and once more in Spain. In all cases, the families were of Moroccan
23 origin. Interestingly, the five families were from the north coast of Morocco, three of them from
24 the city of Nador, one from Tangier and one from Oujda. The co-occurrence of c.5309G>T in
25 five Moroccan families from the same region suggested that the variant has a founder effect. A
26 haplotype linked to c.5309G>T, constructed from five microsatellite markers and spanning
27 1.54Mb, was defined in family B. The alleles found in the other families are consistent with this
28 haplotype, supporting our hypothesis (Figure 1). While there are not enough patients/families to
29 calculate the age of the variant, the relatively short span of the haplotype shared by these
30 families suggests that c.5309G>T may be an old mutation. On the other hand, our study
31 suggests a high penetrance for mutation c.5309G>T, because six out of the eight identified
32 carriers are already affected, leading to tumor development before 40 years of age in more than
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 60% of cases. In addition, the clinical data collected from the breast tumors of patients included
4
5 in this study showed that all tumors were grade III and triple negative. Consequently, the
6
7 clinical and genetic features associated with c.5309G>T are consistent with those described for
8
9 *BRCA1* pathogenic mutations(15-18). Overall, our results suggest that c.5309G>T could be a
10
11 founder mutation in the Moroccan population, which may explain the hereditary breast and
12
13 ovarian cancer syndrome (HBOCS) observed in the five families presented here.
14

15 The fact that the five patients were screened outside Morocco could have two main
16
17 explanations: the scarcity of information on *BRCA* testing from Moroccan laboratories and the
18
19 extensive migration between Morocco and Europe(19) during the 20th century, following the
20
21 establishment of a French-Spanish protectorate in Morocco and the signing of migration
22
23 agreements with some European countries. As a result of this population movement and, in light
24
25 of our results, the variant G1770V could have spread among Moroccans settled in Europe. As
26
27 such, a comprehensive study of a large series of breast cancer patients from Morocco is needed
28
29 to ascertain the exact prevalence of this mutation in the Moroccan population and its
30
31 geographical distribution. The identification of recurrent mutations could facilitate the genetic
32
33 testing process for specific populations, with pre-screening for these mutations as a first step in
34
35 the testing algorithm. This could be a cost-effective approach, making it particularly suitable for
36
37 countries with limited budgets for genetic testing, such as Morocco.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

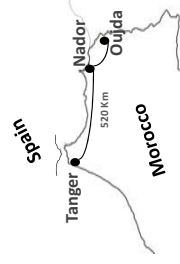
References

1. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *nature genetics* 2003; 33: 238-244.
2. Futreal PA, Liu Q, Shattuck-Eidens D et al. BRCA1 mutations in primary breast and ovarian carcinomas. *Science* (New York, NY 1994; 266 (5182): 120-122.
3. Miki Y, Swensen J, Shattuck-Eidens D et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* (New York, NY 1994; 266 (5182): 66-71.
4. Kurian AW. BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Current opinion in obstetrics & gynecology* 2010; 22 (1): 72-78.
5. Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer* 2014; 136 (5): E359-386.
6. Chouchane L, Boussen H, Sastry KS. Breast cancer in Arab populations: molecular characteristics and disease management implications. *The Lancet* 2013; 14 (10): e417-424.
7. Laraqui A, Uhrhammer N, Rhaffouli HE et al. BRCA genetic screening in Middle Eastern and North African: mutational spectrum and founder BRCA1 mutation (c.798_799delTT) in North African. *Disease markers* 2015: 194293.
8. Slaoui M, Razine R, Ibrahim A et al. Breast cancer in Morocco: a literature review. *Asian Pac J Cancer Prev* 2014; 15 (3): 1067-1074.
9. Ferla R, Calo V, Cascio S et al. Founder mutations in BRCA1 and BRCA2 genes. *Ann Oncol* 2007; 18 Suppl 6: vi93-98.
10. Bouwman P, van der Gulden H, van der Heijden I et al. A high-throughput functional complementation assay for classification of BRCA1 missense variants. *Cancer discovery* 2013; 3 (10): 1142-1155.
11. Quiles F, Fernandez-Rodriguez J, Mosca R et al. Functional and structural analysis of C-terminal BRCA1 missense variants. *PloS one* 2013; 8 (4): e61302.
12. Tazzite A, Jouhadi H, Saiss K et al. Relationship between family history of breast cancer and clinicopathological features in Moroccan patients. *Ethiopian journal of health sciences* 2013; 23 (2): 150-157.
13. Laarabi FZ, Jaouad IC, Ouldin K et al. Genetic testing and first presymptomatic diagnosis in Moroccan families at high risk for breast/ovarian cancer. *Oncology letters* 2011; 2 (2): 389-393.
14. Tazzite A, Jouhadi H, Nadifi S et al. BRCA1 and BRCA2 germline mutations in Moroccan breast/ovarian cancer families: novel mutations and unclassified variants. *Gynecologic oncology* 2012; 125 (3): 687-692.
15. Palacios J, Honrado E, Osorio A et al. Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast cancer research and treatment* 2005; 90 (1): 5-14.
16. Phillips KA. Immunophenotypic and pathologic differences between BRCA1 and BRCA2 hereditary breast cancers. *J Clin Oncol* 2000; 18 (21 Suppl): 107S-112S.
17. Sorlie T, Tibshirani R, Parker J et al. 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* 2003; 100 (14): 8418-8423.
18. Antoniou A, Pharoah PD, Narod S et al. 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* 2003; 72 (5): 1117-1130.
19. De Haas H. Morocco's Migration Experience: a Transitional Perspective. *International Migration* 2007; 45 (4): 39-70.

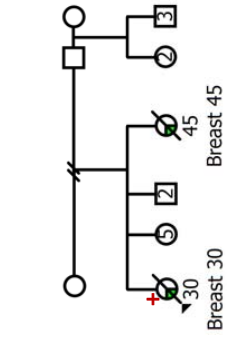
Figure Legends

Figure 1. The five family pedigrees are shown. The upper left corner shows a map of the cities of origin of the five families. In the bottom right corner the common haplotype is depicted, showing the minimal common region among the five families. As can be seen, all carriers of the c.5309G>T variant shared a common haplotype comprising the region between *D17S1793* and *D17S951* (1.54 Mb). In bold, the alleles linked to c.5308G>T are labeled. Mb, Megabases. +, carrier; -, non-carrier.

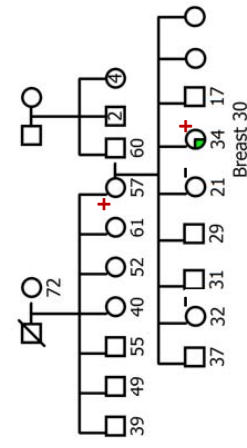
For Review Only



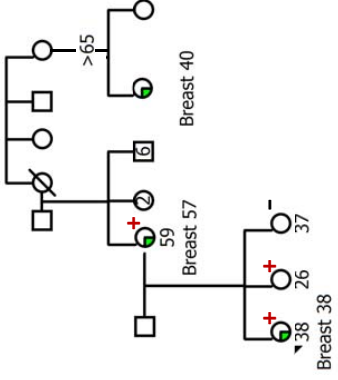
Family A



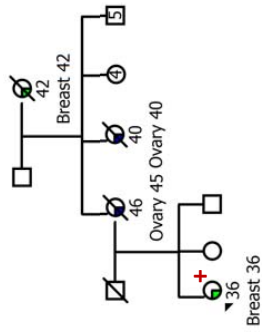
Family B



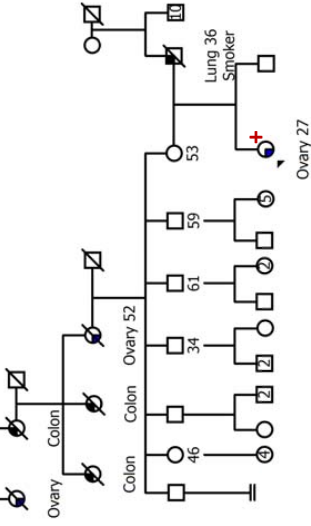
Family C



Family D



Family E



Family	Relationship	Carrier	D17S1846 (41.73 Mb)	D17S1793 (42.20 Mb)	c.5309C>T (43.05 Mb)	D17S865 (43.05 Mb)	D17S1323 (43.09 Mb)	D17S1325 (43.45 Mb)	D17S951 (43.74 Mb)	D17S183 (44.35 Mb)				
A	Proband	+	229	233	196	196	153	155	195	216	169	167	125	143
	Mother	+	229	233	196	196	144	153	155	195	189	169	141	125
B	Proband	+	229	233	196	200	146	140	153	157	195	169	169	128
	Sister	-	233	237	196	200	144	140	155	157	193	197	169	125
	Sister	-	233	233	196	196	144	146	155	151	193	197	167	137
C	Mother	+	225	249	196	194	146	144	153	151	195	189	169	171
	Proband	+	225	237	196	194	146	146	153	151	195	195	169	169
	Sister	+	225	257	196	194	146	146	153	151	195	195	169	181
D	Proband	+	225	225	196	196	146	138	153	151	195	199	169	165
	Proband	+	229	233	196	200	146	146	153	151	195	195	169	177

Shared Haplotype (1.54 Mb)

Table 1. Clinical and histopathological information of breast and ovarian tumors from *BRCA1* c.5309G>T carriers.

Family	Country	Origin	ID	G1770V	Sex	Age Dx	Current Age	Cancer	Histopathology	Grade	ER	PR	HER2
A	Spain	Nador	I	Carrier	F	30	Dead	BC	Invasive Ductal Carcinoma	3	-	-	-
B	Spain	Nador	I	Carrier	F	31	34	BC	Invasive Ductal Carcinoma	3	-	-	-
			II	Carrier	F	NA	57	NA					
			III	Non-Carrier	F	NA	32	NA					
			IV	Non-Carrier	F	NA	21	NA					
C	Norway	Nador	I	Carrier	F	57	60	BC	Invasive Ductal Carcinoma	3	-	-	-
			II	Carrier	F	38	39	BC	Invasive Ductal Carcinoma	3	-	-	-
			III	Carrier	F	NA	26	NA					
			IV	Non-Carrier*	F	NA	37	NA					
D	Norway	Oudja	I	Carrier	F	36	38	BC	Invasive Medullary Carcinoma	3	-	-	-
E	Spain	Tanger	I	Carrier	F	27	31	OC	Invasive Adenocarcinoma Serous	3	NA	NA	NA

*DNA from non-carrier in Family C not available for haplotype analysis. BC, breast cancer. OC, ovarian cancer. NA, not applicable

Resultados no publicados:

Análisis funcional de variantes de significado desconocido en BRCA2 mediante un ensayo basado en células embrionarias de ratón.

Resumen

El análisis mutacional de los genes *BRCA1* y *BRCA2* identifica un gran número de variantes raras cuya baja frecuencia poblacional impide su clasificación mediante modelos epidemiológicos, genéticos o multifactoriales. En estos casos, determinar el efecto de dichas variantes en la función de la proteína es de gran relevancia para su clasificación. En este estudio, se seleccionaron 5 VSD en *BRCA2* (T200I, K331I, K1434I, V1532F, R2645G) identificadas en el Programa de Cáncer Hereditario del Instituto Catalán de Oncología (ICO). El impacto funcional de estas cinco variantes se evaluó mediante un ensayo fundamentado en células madre embrionarias de ratón modificadas genéticamente, denominadas PI2F7^{227, 228, 234}. La viabilidad de estas células es dependiente de tener una proteína *Brca2* funcional. Por tanto, el papel funcional de una determinada variante se puede evaluar en base a la capacidad de la misma de rescatar dichas células cuando es transfectada en ellas y *Brca2* endógeno es posteriormente *knocked out*. Las variantes transfectadas son generadas en un cromosoma artificial de bacteria que contiene *BRCA2*. Así pues, en el presente estudio, se analizó la capacidad de las variantes seleccionadas de rescatar las células PI2F7 *brca2*^{ko/ko}. Además, se analizó la sensibilidad a distintos agentes genotóxicos [Mitomicina C³⁴²; Cisplatino³⁴³; Metil metanosulfonato (MMS)³⁴⁴ y radiación ionizante (IR)] en células transgénicas PI2F7 *brca2*^{ko/ko} transfectadas por estas variantes. Todas las variantes estudiadas mostraron defectos en al menos uno de los ensayos utilizados, sugiriendo un comportamiento cuanto menos hipomórfico de las citadas mutaciones.

Metodología

Selección de Variantes

El Programa de Cáncer Hereditario del Instituto Catalán de Oncología (ICO), realiza estudios mutacionales de los genes *BRCA1* y *BRCA2*, a todos aquellos pacientes de alto riesgo que cumplen los criterios clínicos de síndrome de cáncer de mama y ovario hereditario (SCMOH) establecidos por la OncoGuía de Consejo Genético del Cáncer de Cataluña. En el año 2012, se habían identificado 35 VSD del tipo cambio de sentido (*missense*) en *BRCA2*. De entre estas variantes, para el presente estudio se seleccionaron 12 variantes que presentaban mayor agregación familiar. A continuación, se utilizó el programa Alamut Visual 2.7 (*Interactive Biosoftware*), que integra diferentes programas de predicción *Align-GVGD*¹³³, *SIFT*¹⁶⁷, *Mutation Taster*¹⁶⁸ y *PolyPhen 2*¹⁶⁹, para clasificar las variantes con mayor probabilidad de afectar la función de la proteína. De esta manera, se seleccionaron las cinco variantes que mostraron una mayor probabilidad de tener un

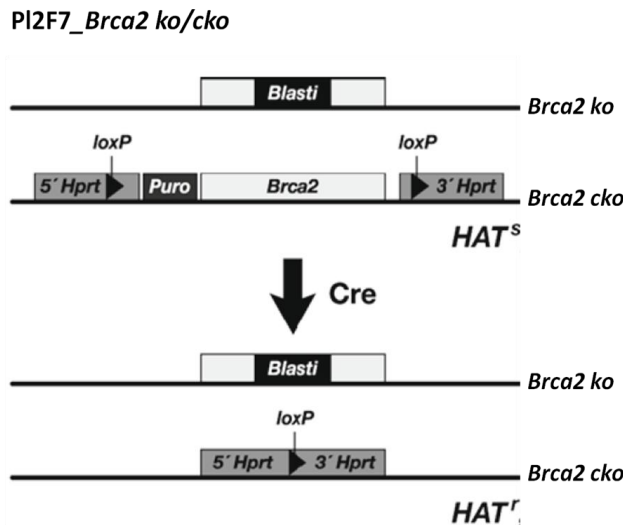
impacto funcional (T200I, K331I, K1434I, V1532F, R2645G) (Tabla 1) para la realización de este estudio funcional.

Células embrionarias de ratón PI2F7

Las células utilizadas para realizar este ensayo son las denominadas PI2F7. Estas células, generadas por el grupo del Dr. Shyam Sharan²³⁴, son células embrionarias de ratón que contienen un alelo de *Brca2* knock out (*Brca2^{ko}*) y otro condicional (*Brca2^{cko}*) (Fig. 1). Proviene de la modificación genética de las células AB2.2 (*129S7/SvEvBrd-Hprt^{b-m2}*), las cuales son deficientes en *hprt*, por lo que son inviables en medios que contengan hipoxantina-aminopterina-timidina (HAT)³⁴⁵.

Para generar las células PI2F7 a partir de las AB2.2, se insertó en el exón 11 de uno de los alelos de *Brca2* un gen de resistencia a la blasticina, dejando este alelo inactivo (*brca2^{ko}*). En el segundo alelo de *Brca2* se integró, en cada uno de sus extremos, una de las dos mitades del gen *HPRT* fusionada con un sitio *loxP*. De esta forma, la depleción de este alelo condicional (*brca2^{cko}*) es inducible mediante la expresión transitoria de *Cre* (Fig. 1)³⁴⁵. Las células recombinantes expresarán *HPRT* funcional y serán resistentes al tratamiento con HAT.

Figura 1. Células PI2F7 con *Brca2^{ko/cko}*



ko, knock out; cko, condicional; Puro, Puromicina; *Blasti*, blasticina, HAT^r, resistente a HAT. (Figura adaptada de Kuznestov et. al, 2008)

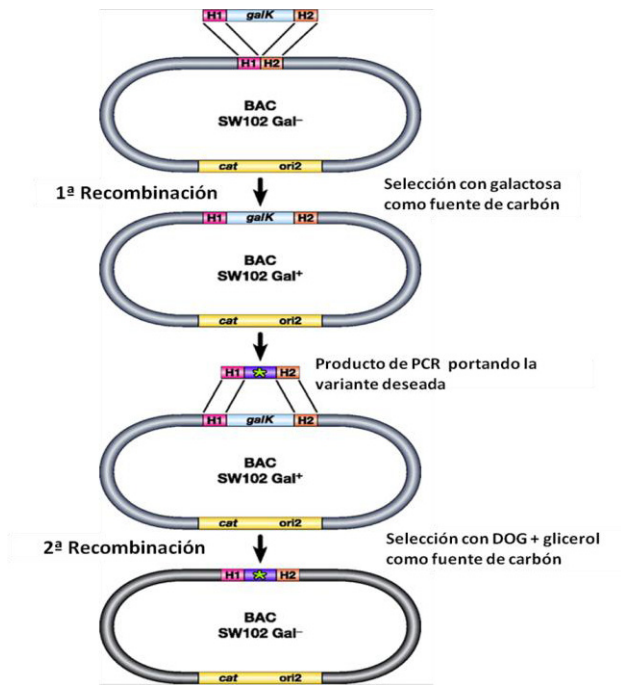
El cultivo y mantenimiento de las células PI2F7, y sus derivados transgénicos, se realizó como había sido descrito previamente²³⁴. Fueron cultivadas sobre una monocapa

de células SNL 76/7 STO mitóticamente inactivadas mediante el tratamiento con Mitomicina C. La línea celular SNL 76/7 STO es resistente a la gentamicina (G418r) y expresa el factor inhibidor de leucemia (LIF), necesario para mantener las células embrionarias de ratón indiferenciadas^{346, 347}.

Generación de variantes en el BAC de BRCA2

Las variantes seleccionadas fueron introducidas en el cromosoma artificial de bacteria (BAC) CTD-2342K5, el cual contenía el gen *BRCA2* (para simplificar BAC *BRCA2*). El BAC *BRCA2* fue generado por el grupo del Dr. Shyam Sharan a partir del vector pBACe3.6, al que se le introdujo un gen de resistencia a la gentamicina bajo el control de un promotor constitutivo *Pkg*. Además, se le fusionó el tag c-myc en el extremo N-terminal de *BRCA2*^{234, 348, 349}. El gen de resistencia a la gentamicina permite la selección de las PL2F7 transgénicas (PL2F7 *BRCA2 Brca2^{ko/cko}*) mediante tratamiento con el antibiótico G418.

Para generar las variantes en el BAC *BRCA2*, se utilizó un sistema de selección positiva/negativa basado en la expresión del gen galactokinasa (*galk*)³⁵⁰(Fig. 2). La cepa bacteriana *Escherichia coli* SW102 se utilizó como huésped²³⁴. SW102 deriva de la cepa *Escherichia coli* DY380, que contiene el sistema de recombinación del profago λ . El sistema profago λ puede ser regulado por temperatura para inducir la recombinación homóloga³⁵¹. Además, las SW102, contienen un operón galactosa completamente funcional, excepto por la deleción de gen *galk*. De este modo, la expresión de *galk* en *trans* en las SW102, restablecerá la función del operón galactosa y éstas serán capaces de crecer en un medio con galactosa como única fuente de carbono.

Figura 2. Esquema de la generación de variantes en el BAC *BRCA2*

galk, galactokinasa; H1 y H2 colas de homología con la región del BAC donde se quiere generar la variante (Figura Adaptada de Warming et. al, 2005)

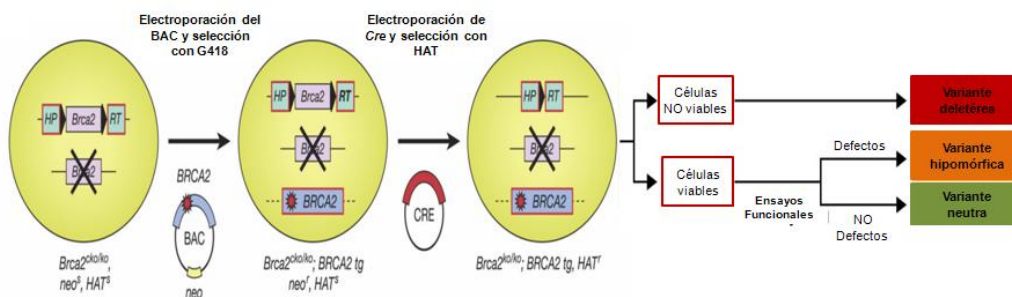
El sistema de generación de variantes en el BAC *BRCA2*, consta de una primera etapa en la que *galk* fue insertado en el BAC *BRCA2* mediante recombinação homóloga inducida por el sistema profago λ (Fig. 2). *Galk* se insertó en el sitio donde se quería generar la variante por homología con la región colindante a ésta. Para ello, se amplificó *galk*, a partir de un plásmido *galk*, utilizando cebadores con colas de 60 pb de homología con el BAC *BRCA2* en cada uno de los lados del nucleótido que se quería modificar, excluyendo dicho nucleótido (Fig. 2). Las SW102 *galk*⁺ fueron seleccionadas positivamente con un medio con galactosa como única fuente de carbono. En la segunda etapa de selección, *galk* fue substituido por un oligonucleótido (producto de PCR), en el que se había generado la variante deseada. Para generar dicho oligonucleótido, se realizó una PCR utilizando dos cebadores de 100 pb que contenían la variante deseada y estaban formados por 90pb de homología con el BAC *BRCA2* a ambos lados del sitio donde se quería introducir la variante, y 10 pb más al otro lado, de manera que compartían 20 pb de complementariedad entre ellos. En esta segunda etapa, la selección de las colonias que contenían la variante se realizó utilizando 2-deoxy-galactosa (DOG) y glicerol como fuente de carbono (selección negativa de *galk*). DOG es inocuo para las SW102 si no es fosforilado por *galk*. En cambio, DOG fosforilado da lugar a 2-deoxy-galactosa-sulfonato,

el cual no es metabolizable por las SW102 y, por tanto, tóxico para ellas (Fig. 2). Las mutaciones generadas en el BAC *BRCA2* se validaron mediante secuenciación Sanger.

Ensayo funcional basado en células embrionarias de ratón

El ensayo funcional basado en células de ratón embrionarias (RE) está fundamentado en la observación de que *BRCA2* es esencial para la viabilidad de estas células cuando *Brca2* endógeno es inactivado²⁴⁴. En base a esta observación, la capacidad de que una variante en *BRCA2* transfectada en células-RE deficientes en *brca2* (PI2F7 *Brca2*^{ko/ko}) tiene para restablecer la viabilidad de estas células, se utiliza para evaluar su impacto funcional (Fig. 3). Las variantes que no logran rescatar las células PI2F7 *Brca2*^{ko/ko} se consideran deletéreas. Las variantes que consiguen restablecer la viabilidad de estas células, se testarán para diferentes funciones conocidas de *BRCA2*, como son: proliferación *in vitro*, formación del RAD51 *foci*, sensibilidad a agentes genotóxicos, sensibilidad a radiación IR, recombinación homóloga o estabilidad genómica. Si no muestran ninguna alteración en estas funciones se clasificarán como neutras. En caso contrario, se clasificarán como hipomórficas. En el presente estudio se evaluó el efecto de las variantes seleccionadas en la sensibilidad a distintos agentes genotóxicos y a la radiación ionizante.

Figura 3. Esquema del ensayo funcional basado en células embrionarias de ratón para variantes en *BRCA2*



ko, knock down; cko, condicional, neo^r, resistente a la gentamicina; HAT^r resistente a HAT. (Figura adaptada de Kuznestov et. al, 2008 y Biswas et. al, 2012).

Para realizar este ensayo, las variantes generadas en el BAC *BRCA2* y un control positivo *wild-type* fueron electroparadas en células PI2F7 *Brca2*^{ko/cko}. Las células transgénicas PI2F7 *BRCA2* *Brca2*^{ko/cko} fueron seleccionadas mediante tratamiento con gentamicina G418. La expresión de *BRCA2* en las colonias G418^r obtenidas se analizó mediante PCR de retro-transcripción, utilizando oligonucleótidos específicos de *BRCA2* localizados en el exón 11 (*Forward*: 5'-ACATGTCCCAAATGAGGA) y el exón 18 (*Reverse*: 5'-GCCGATCTTCTGCTTCTATCA). Seis clones G418^r independientes, que expresaban *BRCA2*, fueron electroparadas con un plásmido *Pgk-Cre*³⁴⁵ para deletar el alelo

Brca2^{cko}. Posteriormente, 10^5 células electroporadas fueron sembradas en una placa de cultivo de 10 cm y fueron seleccionadas con un medio con HAT. Cuando se obtuvieron colonias HAT^r, 48 de ellas fueron analizadas mediante *Southern Blot* para confirmar la pérdida del alelo *brca2^{cko}*. Las colonias HAT^r que permanecieron en la placa, fueron teñidas con azul de metileno, o cristal de violeta, para ser cuantificadas y comparadas con las colonias obtenidas para el alelo *BRCA2 wt*. Para confirmar la pérdida del alelo *Brca2^{cko}* mediante *Southern Blot*, se utilizó una sonda de 1,5 Kb específica del exón 11 de *Brca2* (nt 5208-6710 del NM 009765). Esta sonda detectaba un fragmento de 2,2 kb en el alelo *brca2^{ko}* y de 4,8 kb en el alelo *Brca2^{cko}*. Finalmente, dos colonias HAT^r independientes de cada variante, fueron analizadas para evaluar el efecto dicha variante en la función reparadora del DNA de BRCA2.

Ensayos funcionales realizados en este estudio

Es conocida la hipersensibilidad de células deficientes en *BRCA2* a diferentes agentes *crosslinkers* y metilantes del DNA, así como a la radiación IR^{242, 259, 262}. Por ello, en este estudio se evaluó la sensibilidad de las variantes estudiadas a estos agentes que dañan el DNA.

Sensibilidad a agentes genotóxicos

Para cada dosis de genotóxico utilizado, se sembraron, por triplicado, 8.000 células en una placa de cultivo de 96 pozos. El tratamiento se empezó 18h después de la siembra y se mantuvo durante 3 días sin cambiar el medio. La viabilidad se midió mediante un ensayo de MTT a las 72h del comienzo del tratamiento como se había descrito previamente³⁵².

Los agentes genotóxicos utilizados y las distintas dosis que se emplearon de cada uno de ellos, se muestran a continuación:

Mitomicina C (ng/ul): 0, 5, 10, 20, 40, 60, 80.

MMS (Metil-metanosulfonato) (μg/ml): 0, 5, 10, 15, 20, 30, 40.

Cisplatino (Cis-diaminodicloroplatino (II)) (μM): 0 - 0,1 - 0,2 - 0,3 - 0,4 - 0,6 - 1.

Sensibilidad a la radiación IR

Para testar la sensibilidad de las variantes estudiadas a la IR, se sembraron, por triplicado, 8.000 células en diferentes pozos de en una placa de 96. Se utilizó una placa distinta para cada dosis de radiación. A las 18h de la siembra, sin cambiar el medio de cultivo, se irradiaron las placas con una fuente de ¹³⁷Cs, utilizando diferentes tiempos

para lograr las dosis deseadas. Las dosis utilizadas fueron 0, 1, 2, 4, 6 Gy. La viabilidad celular se midió 72 horas después del tratamiento mediante un ensayo de MTT como se había descrito previamente³⁵².

Análisis de resultados

Se realizaron un mínimo de tres experimentos para cada ensayo funcional (genotóxicos o IR), utilizando dos colonias independientes de cada variante en cada experimento. Como las dos colonias independientes mostraron fenotipos muy similares y comparables estadísticamente (Test de Wilcoxon, p-valor>0,05), se trataron como un único grupo y se calculó la media entre ambas para futuras comparaciones. Esta media se comparó contra la media de cada control utilizado en dichos experimentos (un clon wt y un clon de la variante hipomórfica G25R²²⁷). Debido al número reducido de repeticiones de cada experimento (n=3-5), el test de Wilcoxon se vio muy afectado por pequeñas variaciones entre experimentos y careció de poder estadístico suficiente para realizar comparaciones consistentes entre controles y variantes. Por ello, se decidió interpretar los resultados cualitativamente basándose en las gráficas obtenidas a partir de la media de los resultados de cada experimento realizado y la desviación estándar entre ellos.

Modelaje estructural

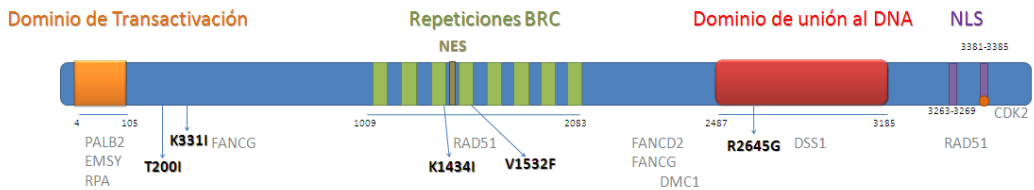
La información estructural sobre BRCA2 es escasa. Si bien, modelos estructurales representativos de la región comprendida entre los residuos 2479-3190, han podido ser definidos a partir de sus ortólogos en rata (PDB: 1IYJ) y ratón (PDB: 1MJE), cuya identidad con BRCA2 en esta región es del 78% y 77%, respectivamente. En esta región se produce la interacción BRCA2-DSS1³⁵³, que puede ser modelada utilizando los citados modelos con una identidad de secuencia en la región de interacción mayor del 60%. El efecto de la variante R2645G en la interacción BRCA2-DSS1 fue analizada utilizando estos modelos (análisis realizado por Carles Pons perteneciente al *Structural Bioinformatics & Network Biology Lab*, IRB Barcelona).

Resultados

Descripción de las variantes seleccionadas

Cinco VSD en *BRCA2* fueron seleccionadas para este estudio: T200I, K331I, K1434I, V1532F, R2645G (Fig. 4, Tabla 1). Las variantes T200I y K331I, se encontraban localizadas en una región de *BRCA2* sin estructura tridimensional definida en la que no se ha sido descrito ningún dominio funcional. Las variantes K1434I y V1532F estaban localizadas en el dominio BRC de *BRCA2*³⁵⁴, concretamente en los subdominios BRC3 y BRC4. El dominio BRC consta de 8 repeticiones y es el encargado principal de mediar la interacción *BRCA2*-*RAD51*, la cuál es de gran importancia para un correcto funcionamiento del mecanismo de recombinación homóloga¹⁰⁸. El subdominio BRC4 exhibe la interacción más fuerte con *RAD51*³⁵⁵. Estudios estructurales y funcionales previos revelaron que el aminoácido V1532 está fuertemente involucrado en ésta interacción^{356, 357}. Además, Ochiai et. al, mostraron mediante un estudio *in vitro*, que el cambio V1532F podría romper dicha interacción³⁵⁶. El papel de BRC3 en la interacción *BRCA2*-*RAD51* es menos conocida. Hasta ahora, no existe ninguna información sobre la participación del aminoácido K1434 en la citada interacción. Por último, la variante R2645G está localizada en el dominio altamente conservado de unión al DNA de *BRCA2*¹⁰³.

Figura 4. Localización de las variantes analizadas en este estudio en *BRCA2*



Las variantes K331I y V1532F habían sido previamente descritas como VSD en la base de datos *Breast Cancer Information Core* (BIC). La variante K331I había sido identificada 3 veces, dos en población española y otra en población no definida. La variante V1532F había sido descrita 2 veces, una vez en población española y la otra en población no concretada. El resto de variantes analizadas en este estudio no habían sido reportadas previamente en BIC. No obstante, la variante K1434I había sido clasificada como probablemente neutra (*odds* a favor de su neutralidad= 0,009) en un estudio basado en la conservación evolutiva (*Align-GVGD*) y la histopatología del tumor¹⁵¹. Ninguna de las variantes descritas ha sido reportada en la base de datos *Exome Sequencing Project* (ESP) (<http://evs.gs.washington.edu/EVS/>), donde se recoge información de las variantes encontradas en la secuenciación de 6503 exomas de individuos de población normal europea y afro-americana.

El análisis de cosegregación de las variantes descritas en las familias afectadas se realizó cuando fue posible (Fig. 5). Para la variante T200I no fue posible tal análisis (Fig. 5, Familia A). La variante K331I fue identificada en la probando, diagnosticada de cáncer de mama (CM) a los 28 años, en su hermano y en su madre, no afectos hasta el momento (Fig. 5, Familia B). La variante K1434I fue identificada en la probando, y en sus dos hermanas, todas ellas afectas y diagnosticadas a los 51, 49 y 56 años, respectivamente (Fig. 5, Familia C). La variante V1532F fue encontrada en una paciente que sufrió dos CM, uno a los 26 años y el otro a los 44 años, y un cáncer de riñón a los 44 años (Fig. 5, Familia D). La variante V1532F fue encontrada en una paciente que sufrió dos CM, uno a los 26 años y el otro a los 44 años, y un cáncer de riñón a los 44 años (Fig. 5, Familia D). Además, también fue detectada en su hermana y en su madre, no afectas hasta ahora. Esta paciente también era portadora de la VSD R841W en BRCA1, la cual afecta a un aminoácido poco conservado y ha sido reportada en ESP con una frecuencia del 0,4%. Por último, la variante R2645G fue identificada en una paciente que sufrió cáncer de ovario a los 58 años. Su hermana no portadora sufrió CM a los 74 años (Fig. 5, Familia E).

El posible efecto de dichas variantes en la función de BRCA2 fue evaluado mediante el programa Alamut 2.7 el cual integra los programas de predicción *Align-GVGD*¹³³, *SIFT*¹⁶⁷, *Mutation Taster*¹⁶⁸ y *PolyPhen 2* (ref.¹⁶⁹)(Tabla 1). Las variantes T200I y K331I fueron consideradas, respectivamente, deletérea y neutra de forma inequívoca por todos los programas utilizados. Las variantes K1434I y R2645G fueron consideradas deletéreas por todos los programas excepto por *Align-GVGD*, que las clasificó como Clase 15 o de impacto funcional incierto. Por último, la variante V1532F fue considerada neutra por los programas *Align-GVGD* y *Mutation Taster* y deletérea por *SIFT* y *Polyphen 2*. Además, el programa *Align-GVGD* permitió analizar el grado de conservación evolutiva de los aminoácidos afectados. Para ello se utilizó un alineamiento basado en 12 especies, de humano a erizo de mar. En base a dicho alineamiento distinguimos entre aminoácidos no conservados, poco conservados, moderadamente conservados y altamente conservados. Las variantes T200I, V1532F y R2645G afectaban a aminoácidos altamente conservados y las variantes K331I y K1434I afectaban a aminoácidos moderadamente conservados. Adicionalmente, el efecto de la variante R2645G en la interacción BRCA2-DSS1 se analizó utilizando modelos estructurales disponibles realizados a partir de los ortólogos de BRCA2 en rata (PDB: 1IYJ) y ratón (PDB: 1MJE). En ambos modelos estructurales, la variante R2645G se encuentra en la interfaz de la interacción, probablemente desestabilizándola (datos no mostrados).

La variante G25R, de patogenicidad conocida, fue utilizada como control en los ensayos de sensibilidad a genotóxicos y a la IR. G25R estaba localizada en el dominio de transactivación de BRCA2, concretamente en la región de interacción BRCA2-PALB2

(Tabla 1). Esta variante afectaba a un aminoácido altamente conservado. Los programas de predicción de impacto funcional sugirieron que dicha variante era deletérea. Estas predicciones fueron confirmadas por los resultados obtenidos en un estudio previo basado también en el ensayo funcional en células-RE donde mostró sensibilidad moderada a agentes genotóxicos, a la IR, redujo la recombinación homóloga e incrementó la inestabilidad genómica^{105, 227}. En base a esto, G25R fue clasificada como probablemente patogénica con un riesgo asociado moderado.

Figura 5. Pedigríes de las familias con las variantes analizadas en este estudio.

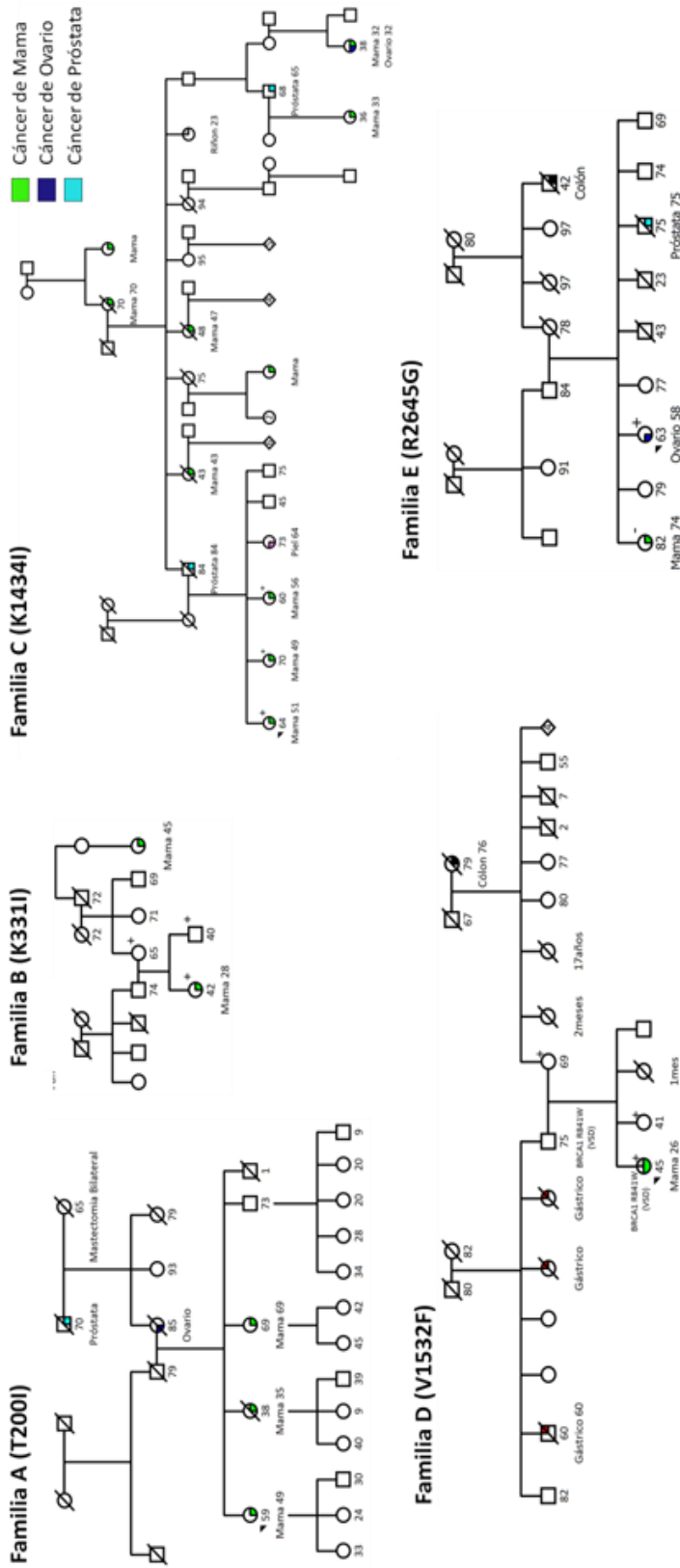


Figura 5. Cabezas de punta de color negro se utilizaron para señalar al probando. La edad de defunción, según corresponda (cuándo se dispuso de ella) se encuentra debajo del símbolo del individuo. El tipo de cáncer con la edad de diagnóstico a la derecha, están escritos debajo de la edad. El cáncer de mama se marco en verde, el cáncer de ovario en azul oscuro y el cáncer de próstata en azul claro. El resto de cánceres en distintos colores. +, portador; -, no portador

Tabla 1. Descripción y predicciones funcionales de las variantes analizadas en este estudio.

Alamut 2.7										
Variante Proteína (BIC)	Variante cDNA (HGVS)	BIC	ESP	Clasificación	Dominio Funcional	Align-GVGD*	SIFT	Mutation Taster	Polyphen 2	Conservación Evolutiva (Align-GVGD)#
T200I	c.599C>T	NR	NR	VSD	Ninguno	C65	Deletérea	Prob. Delet.	Prob. Pat.	Alta
K331I	c.922A>T	3 VSD, 2 pob. española	NR	VSD	Ninguno	C0	Tolerada	Polimorfismo	Pos. Pat.	Moderada
K1434I	c.4301A>T	2 VSD, 1 pob. española	NR	Prob. no pat.	BRC3	C15	Deletérea	Prob. Delet	Prob. Pat.	Moderada
V1532F	c.4594G>T	NR	NR	VSD	BRC4	C0	Deletérea	Polimorfismo	Prob. Pat.	Moderada
R2645G	c.7933A>G	NR	NR	VSD	Dominio de unión al DNA	C15	Deletérea	Prob. Delet	Prob. Pat.	Alta
G25R	c.73G>A	NR	NR	Prob. Pat.	Dominio de Transcativación	C65	Deletérea	Prob. Delet	Prob. Pat.	Alta

BIC (*Breast Information Core database*) y en ESP (*Exome Sequencing Project*), HGVS, *Nomenclatura Human Genome Variation Society*; NR, no reportado; VSD, variante de significado desconocido, Prob. Pat., Probablemente Patogénica; Pos. Pat, Posiblemente Patogénica.

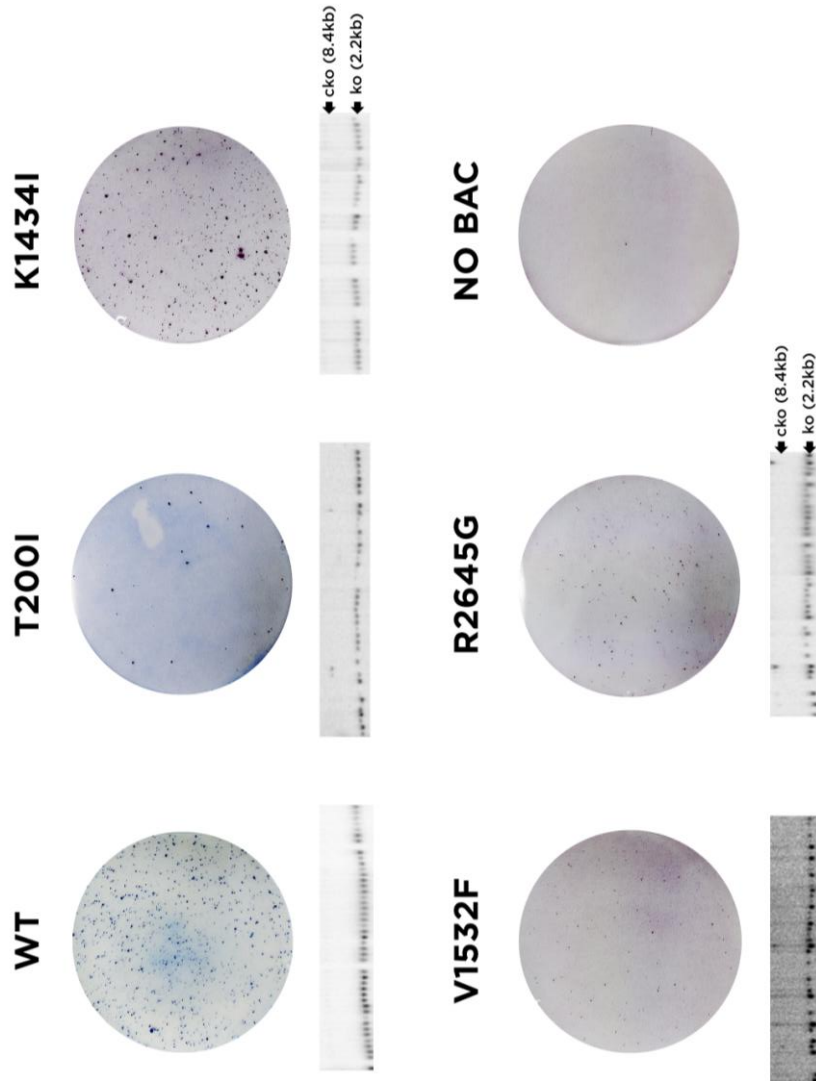
* *Align-GVGD* distingue cinco clases (C0, C15, C25, C35, C45, C55 y C65) en relación a la probabilidad de que una variante afecte a la función de la proteína, siendo la Clase 0 la menos probable y la Clase 65 la más probable. Los cambios de aminoácidos clasificados como C0 se consideran neutros, los clasificados como C15 y C25 se consideran de efecto intermedio y los clasificados como clase C35 o superior se consideran probablemente deletéreos.

El grado de conservación evolutiva se obtuvo a partir del programa *Align-GVGD* y se basó en el alineamiento de 12 especies, de humano a erizo de mar. *Align-GVGD* considera 4 grados de conservación evolutiva: No conservado, Baja, Moderada, Alta.

Evaluación funcional de las variantes seleccionadas en base a su capacidad de restablecer la viabilidad de las células PI2F7 $Brca2^{ko/ko}$

El impacto funcional de las variantes seleccionadas fue evaluado utilizando el ensayo basado en células-RE. Cada variante fue generada en un BAC que contenía el gen completo de *BRCA2* y un gen de resistencia a la G418. Los BAC portando las distintas variantes fueron individualmente electroporados en células PI2F7, células-RE que contienen un alelo de *Brca2 knock out* y otro condicional ($Brca2^{ko/cko}$). Los clones PI2F7 ($BRCA2 Brca2^{ko/cko}$) fueron seleccionados con un medio de G418, para eliminar aquellos clones que no han incorporado el BAC con la variante a analizar. La expresión de *BRCA2* fue analizada mediante PCR de retro-transcripción en 48 colonias independientes (datos no presentados). Seis de las colonias expresando *BRCA2* fueron escogidas para el ensayo de viabilidad. La capacidad de restablecer la viabilidad de las células PI2F7 cuando se delecta su alelo $Brca^{cko}$ mediante la expresión de *Cre* sirvió como aproximación para evaluar el impacto funcional de las variantes en *BRCA2* que fueron transfectadas en dichas células. Por ello, se expresó *Cre* de forma transitoria en los 6 clones independientes PI2F7 ($BRCA2 Brca2^{ko/cko}$) de cada variante. La delectación del alelo $Brca^{cko}$ generó un alelo funcional del gen *HPRT* que permitió la selección de los clones recombinantes PI2F7 ($BRCA2 Brca2^{ko/ko}$) utilizando un medio con HAT. Cuarenta y ocho colonias HAT^r de cada variante fueron genotipadas mediante *Southern Blot* para confirmar la pérdida del alelo $Brca2^{cko}$ y validar el experimento de *pgk-Cre* (Fig. 6). El resto de colonias HAT^r fueron teñidas y cuantificadas. La variante K1434I dio lugar a un número de colonias HAT^r comparable al obtenido para el *wt* (Fig. 6). Las colonias T200I, V1532F y R2645G dieron lugar a un número de colonias HAT^r muy reducido en comparación al *wt*, sugiriendo un defecto en la función de *BRCA2* (Fig. 5). Para la variante K331I, más del 70% de colonias de las 48 que se utilizaron para confirmar la pérdida del alelo $brca2^{cko}$, mantenían dicho alelo. Debido a este alto *background*, no se pudo comparar el número de colonias HAT^r con el *wt*.

Figura 6. Ensayo de viabilidad en células P1ZF7 *Brca2*^{ko/ko}



En esta figura se muestran las colonias HAT^r obtenidas para cada variante después de inducir la delección del alelo *Brca2*^{ko} en las células P1ZF7 y seleccionar con HAT. Encima de cada placa se indica el nombre de la variante correspondiente. Para cada variante, debajo de la foto de la placa, se muestra el resultado del *Southern Blot*, realizado en 48 colonias HAT^r para detectar la pérdida del alelo *Brca2*^{ko}. La banda inferior corresponde con el alelo *Brca2*^{ko} (2,2 kb) y la banda de arriba al alelo *Brca2*^{cko} (4,8 kb).

Efecto de las variantes seleccionadas en las funciones reparadoras de BRCA2

BRCA2 tiene un papel clave en la reparación de roturas de doble cadena en el DNA. Células *BRCA2* deficientes han mostrado hipersensibilidad a varios agentes genotóxicos (*crosslinkers* y metilantes), lo que sugiere que la pérdida de función de BRCA2 origina esta hipersensibilidad^{242, 259, 262}. Por ello, el posible impacto funcional de las variantes que restablecieron la viabilidad de las PI2F7 *Brca2*^{ko/ko}, fue evaluado mediante ensayos de sensibilidad a agentes que dañan el DNA. Los agentes genotóxicos utilizados fueron la Mitomicina C (MitC), el cisplatino, el metanosulfonato (MMS) y la radiación ionizante (IR) (Fig. 7). En cada experimento, los resultados obtenidos para las variantes fueron comparados contra los obtenidos para PI2F7 transfectadas con el alelo *wt* y con la variante probablemente patogénica G25R. El análisis cualitativo indicó que las variantes T200I, V1532F mostraban sensibilidad a la Mit C, cisplatino y a la IR muy similar a la variante G25R y un comportamiento intermedio para el MMS. Las variantes K331I y R2645G mostraron una sensibilidad similar a G25R para la Mit C y el cisplatino y una sensibilidad intermedia al MMS y la IR. Por último, La variante K1434I mostró una sensibilidad muy leve a la Mit C, al MMS y al IR y una sensibilidad al cisplatino comparable a la G25R.

Figura 7. Resultados de los ensayos de sensibilidad a diferentes agentes genotóxicos

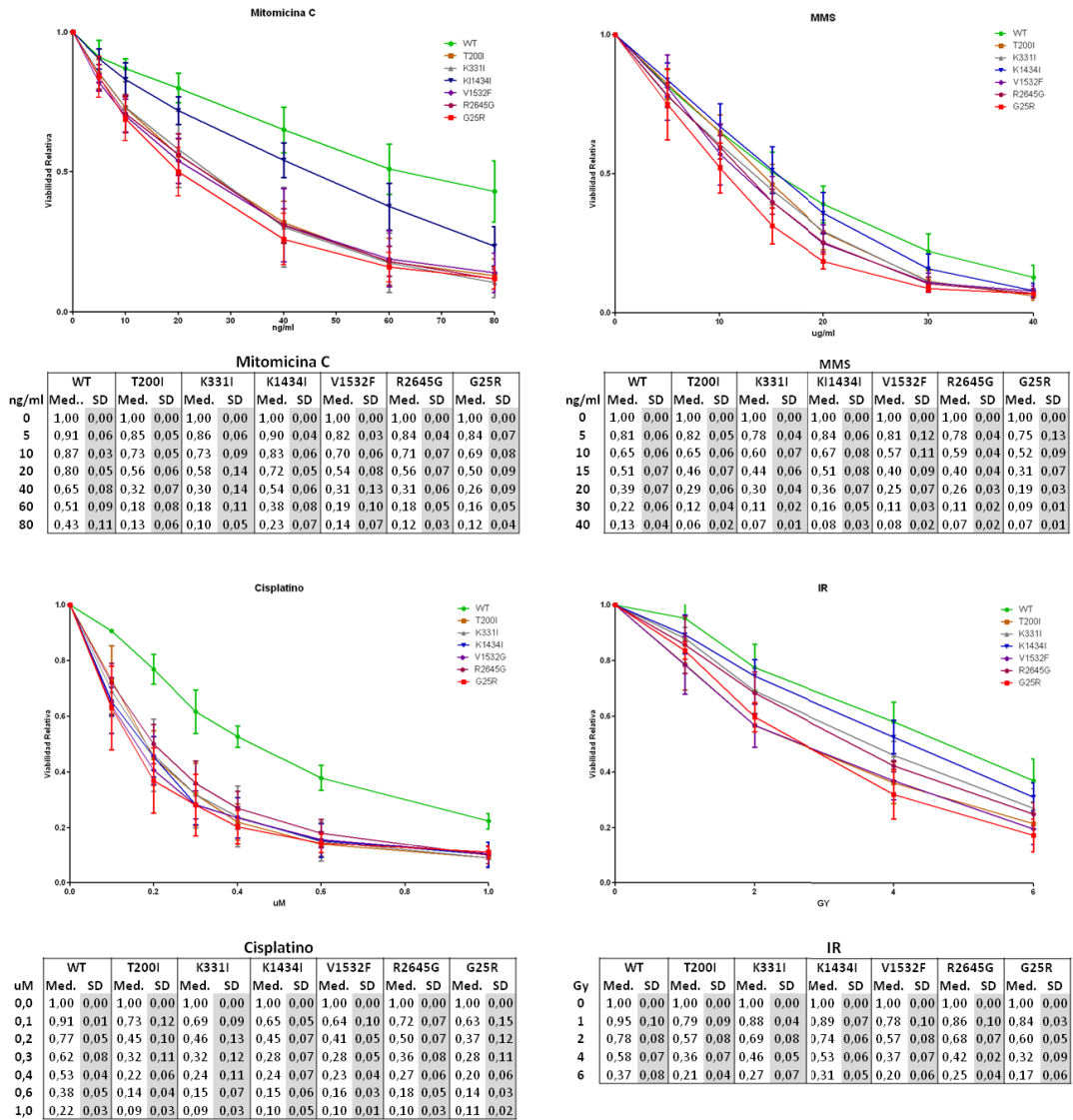


Figura 7. Resultados de los ensayos de sensibilidad a la Mitomicina C, el metil metanosulfonato (MMS), el cisplatin y la radiación ionizante. Los distintos colores utilizados para representar cada variante, se muestran en la leyenda situada en la esquina superior de cada gráfica. En cada ensayo se utilizaron dos controles, células BRCA2 wt (verde) y células con la variante probablemente patogénica G25R (rojo). Debajo de cada gráfica se muestra una tabla con los valores de viabilidad relativa de cada variante y de los controles utilizados para la representación gráfica. Las medias (Med.) y las SD han sido calculados a partir de los resultados obtenidos para dos colonias independientes de una misma variante en tres o más experimentos.

Discusión

Interpretar el riesgo asociado a las variantes identificadas en los pacientes que se someten al test genético de *BRCA1* y *BRCA2* es de gran importancia desde el punto de vista del asesoramiento clínico. Sin embargo, muchas de las variantes de cambio de sentido o missense identificadas en estos genes son únicas o muy raras, lo que impide que su riesgo asociado se pueda estimar a partir de datos clínicos o genéticos (modelos de cosegregación, concurrencia, modelos multifactoriales, entre otros). En estos casos, determinar el impacto funcional que pueden tener dichas variantes es determinante para poder clasificarlas. Con tal fin, en este estudio, se ha evaluado el impacto funcional de 5 VSD (T200I, K331I, K1434I, V1532F, R2645G) en *BRCA2* identificadas en nuestra población, mediante el ensayo basado en células-RE. El estudio funcional se complementó con otras evidencias obtenidas a partir del estudio de cosegregación, las predicciones de impacto funcional y de conservación evolutiva. Estas evidencias nos añaden más información para la clasificación de estas variantes de una forma más robusta, sobre todo en el caso de variantes hipomórficas.

Como primera aproximación, cuando fue posible, se realizó el análisis de cosegregación de las variantes seleccionadas en las familias afectadas (Fig. 5). En las familias afectadas por las variantes K331I y V1532F se detectó, en cada una de ellas, dos portadores no afectados, siendo uno de ellos mayor de 60 años, lo que no se ajusta a la mayoría de patrones observados para mutaciones de alto riesgo en *BRCA2*. Por el contrario, para la variante K1434I se observó cosegregación en tres hermanas afectadas. Este fuerte patrón de cosegregación sugiere que esta variante podría ser la causante de la enfermedad en esta familia. En la familia afectada por la variante R2645G, no se detectó la variante en un paciente afecto de CM. No obstante, dado el importante número de fenocopias asociados al CM debido a mutaciones en los genes *BRCA*, esta observación no excluye que esta variante sea responsable del cáncer en la paciente portadora de la misma. A continuación, se utilizaron varios programas de predicción para evaluar el posible efecto funcional que dichas variantes podían tener (Tabla 1). Las predicciones obtenidas sugirieron que las variantes T200I, K1434I y R2645G podrían tener un efecto en la función de *BRCA2*. Por el contrario, las predicciones obtenidas para la variante K331I indicaban que era muy poco probable que tuviera un impacto funcional. Para la variante V1532F, se obtuvieron resultados contradictorios. Por tanto, el análisis conjunto de los resultados de cosegregación y de las predicciones funcionales sugirió una posible patogenicidad asociada a la variante K1434I y la posible neutralidad de la variante K331I, siendo no concluyente para el resto de variantes.

Tabla 2. Tabla resumen de los resultados de las variantes estudiadas

Variante	Cosegregación	Predicción <i>In silico</i> / Grado Conservación	Observaciones Previas	Ensayo Viabilidad PI2F7ko/cko	Sensibilidad a Genotóxicos				Clasificación
					Mit C	MMS	Cisplatino	IR	
T200I	NA	Deletérea/Alta	No	+++	+++	++	+++	+++	Riesgo Moderado
K331I	Detectada en no afectados (uno > 65)	Neutra/Moderada	No	NA	+++	++	+++	++	Riesgo Bajo
K1434I	SI	Deletérea/Moderada	Prob. Neutra	ND	+	ND/+	+++	ND/+	Riesgo Moderado
V1532F	Detectada en no afectados (uno > 65)	No concluyente/Moderado	Rompe interacción BRCA2-RAD51	+++	+++	++	+++	+++	Riesgo Bajo/Moderado
R2645G	NO	Deletérea/Alta	No	+++	+++	++	+++	++	Riesgo Bajo/Moderado

En esta tabla se muestra una visión global de los resultados obtenidos en los ensayos funcionales realizados en este estudio junto con el análisis de cosegregación y las predicciones *in silico*, así como la valoración final de la patogenicidad asociada a estas variantes. En caso de que las haya, también se muestran las observaciones previas reportadas para estas variantes, a modo de comparación con los resultados obtenidos en este estudio. NA, no aplicable; ND, no diferencias. Los resultados de los ensayos funcionales, han sido jerarquizados de una forma cualitativa en 4 clases: ND, no diferencias con el wild-type (wt); +, efecto muy moderado, sólo a concentraciones muy altas de genotóxico; ++, efecto intermedio entre el wt y el control negativo, G25R; +++, igual efecto que el observado para la variante G25R.

Con el fin de verificar las predicciones funcionales y obtener más evidencias para clasificar las variantes seleccionadas, estas variantes se analizaron mediante el ensayo basado en células-RE. Este ensayo ha sido previamente utilizado para testar 34 variantes en *BRCA2*, mostrando una gran sensibilidad y especificidad a la hora de discriminar entre variantes neutras y patogénicas conocidas^{227, 228, 234}. En primer lugar, se testó la capacidad que las citadas variantes tenían para rescatar las células PI2F7 *Brca2*^{ko/ko} (Fig. 6). Posteriormente, las variantes que resultaron en colonias PI2F7 *Brca2*^{ko/ko} viables, se analizaron mediante ensayos de sensibilidad a diferentes agentes genotóxicos, para así determinar su efecto en la función reparadora del DNA de *BRCA2* (Fig. 7). La variante hipomórfica G25R se utilizó como control en dichos ensayos de sensibilidad. Esta variante había sido previamente testada por el ensayo basado en células-RE, mostrando una sensibilidad moderada a los agentes genotóxicos utilizados en el presente estudio. La variante K1434I no mostró diferencias con *BRCA2 wt* en la capacidad de rescatar las células PI2F7 *Brca2*^{ko/ko}. Tampoco mostró sensibilidad al MMS y a la IR. En cambio, mostró una sensibilidad muy leve a la Mit C (fenotipo intermedio entre *wt* y G25R) y una sensibilidad al cisplatino comparable a G25R, indicando una pérdida parcial de la función de *BRCA2*. Esta variante había sido previamente clasificada como probablemente no patogénica mediante un modelo probabilístico basado en información histopatológica (*odds* favor de su neutralidad=0,027; *odd* límite a favor de la neutralidad = 0,01)¹⁵¹. Esta clasificación fue poco consistente debido a la poca información clínica que se disponía de la variante. Además, los tumores asociados a *BRCA2* presentan rasgos muy heterogéneos, haciendo que los datos patológicos asociados a variantes en *BRCA2* no sean de gran valor predictivo. Las nuevas evidencias funcionales y de cosegregación aportadas por el presente estudio rebaten la clasificación previa de esta variante, ya que indican que puede estar asociada a la enfermedad en la familia afectada (Fig. 5, Familia C). No

obstante, un estudio más amplio de cosegregación en esta familia ayudaría a determinar su riesgo patogenicidad. Las variantes T200I, V1532F y R2645G resultaron en colonias PI2F7 *Brca2*^{ko/ko} viables, aunque en un número significativamente menor que *BRCA2 wt* (Fig. 6), indicando un claro defecto en la función de BRCA2. Estas tres variantes mostraron un fenotipo para la sensibilidad a agentes genotóxicos similar a la variante G25R, lo que confirmó un defecto en la función de BRCA2 (Fig. 7). Para las variantes T200I y R2546G, estos resultados concuerdan con el hecho de que estén afectando a aminoácidos altamente conservados. Particularmente, para la variante R2546G, este resultado también concuerda con su posible efecto sobre la interacción BRCA2-DSS1, que es de vital importancia para estabilizar BRCA2 y asegurar su localización celular donde desarrolla su papel en la RH^{108, 109}. De hecho, se han descrito varias mutaciones asociadas al CM que afectaban esta interacción^{110, 353}. Para el caso de V1532F nuestros resultados están acorde con la observación realizada en un estudio previo, que mostró que esta variante podría estar afectando a la interacción BRCA2-RAD51, esencial para un correcto funcionamiento de la RH³⁵⁶. Cuando fue analizada, la variante G25R no mostró diferencias con *BRCA2 wt* en el ensayo de viabilidad de las células PI2F7 *Brca2*^{ko/ko} con lo que cabría esperar un fenotipo más severo de nuestras 3 variantes en el ensayo de sensibilidad a agentes genotóxicos, similar al de otras variantes patogénicas que mostraron una reducción de la viabilidad de las células PI2F7 *Brca2*^{ko/ko} e hipersensibilidad a los agentes genotóxicos (E3309X, R2336H o L2510P)²²⁷. No obstante, para las variantes hipomórficas, el fenotipo observado depende de la función o funciones afectadas por la variante y puede ser muy heterogéneo. Por ejemplo, la variante delta3-4, mostró una reducción del 30% de la viabilidad y, en cambio, no mostró ninguna sensibilidad a agentes genotóxicos²²⁸. Nuestros resultados para las variantes T200I, V1532F y R2645G sugiere un fenotipo intermedio, menos severo que las variantes patogénicas citadas, indicando que son variantes asociadas a un riesgo de sufrir CM bajo o moderado. Por último, para variante K331I no fue posible realizar el ensayo de viabilidad en células PI2F7 *Brca2*^{ko/ko}. Sin embargo, mostró una sensibilidad a agentes genotóxicos similar, o un poco más leve, que la variante G25R. En base a ello, y a falta de determinar su efecto en el ensayo de viabilidad de células PI2F7 *Brca2*^{ko/ko}, consideramos que la variante K331I podría estar ligada a un aumento de riesgo bajo. Los resultados funcionales obtenidos para esta variante como para V1532F, concuerdan con sus patrones de cosegregación, los cuales correspondería con un patrón asociado a una variante de penetrancia moderado/baja más que al de una variante alta penetrancia en *BRCA2*.

El programa de predicción *Align-GVGD* es, hoy en día, el referente para predecir el efecto de una variante en la función de BRCA1 y BRCA2. De hecho, los resultados de este programa están siendo integrados en los actuales modelos multifactoriales de clasificación de variantes. Sin embargo, para BRCA2 sólo ha sido correctamente validado

para variantes localizadas en los dominios funcionales de transactivación y de unión al DNA¹²⁴. Nuestros resultados funcionales para la variante R2645G, localizada en el dominio de unión al DNA, concuerdan con la predicción de *Align-GVGD* (Clase 15 o de efecto intermedio), corroborando la validez de este programa para predecir el efecto de variantes localizadas en este dominio. No obstante, *Align-GVGD* predijo como Clase 0 o neutras las variantes K331 y V1532F, las cuales mostraron una pérdida parcial de la función de BRCA2 en nuestro estudio. Además, la variante K331I fue unívocamente clasificada como neutra por todos los programas de predicción utilizados en este estudio. De aquí, se deduce la necesidad de obtener información funcional de un número mayor de variantes localizadas en las regiones no conservadas de la proteína BRCA2 para una mayor validación estos programas.

En conclusión, el análisis funcional y de cosegregación de cinco variantes en *BRCA2* nos ha permitido determinar el efecto de estas variantes en la función de BRCA2. Gracias a ello, cuatro variantes (T200I, K1434I, V1532F, R1645G) han podido ser clasificadas como variantes de riesgo moderado, mientras que la variante K331I puede ser clasificada como variantes de bajo riesgo. Sin embargo, los distintos fenotipos observados para variantes que hipomórficas, dificulta en gran medida la interpretación exacta de su riesgo asociado. El estudio funcional de un amplio número de variantes de este tipo, clasificadas mediante modelos no funcionales, nos permitirá determinar correctamente el riesgo asociados a los distintos fenotipos observados. Con ello, se podrá establecer probabilidades de patogenicidad para estas variantes hipomórficas y, como consecuencia, podrán ser incluidas en los modelos multifactoriales de clasificación de variantes, dándole así una mayor significancia a los resultados funcionales en el contexto clínico.

Capítulo 3

ESTUDIO MUTACIONAL DE NUEVOS GENES IDENTIFICADOS EN LA VÍA DE ANEMIA DE FANCONI/BRCA EN FAMILIAS BRCA1

ARTÍCULO 3

Artículo 3: Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families

Fernández-Rodríguez J*, Quiles F*, Blanco I, Teulé A, Feliubadaló L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capellá G, Brunet J, Lázaro C, Pujana MA.

BMC Cancer. 2012 Mar 8;12:84.

(*) Estos autores han contribuido equitativamente en este trabajo.

Factor de Impacto (2012): 3,333

RESEARCH ARTICLE

Open Access

Analysis of *SLX4/FANCP* in non-*BRCA1/2*-mutated breast cancer families

Juana Fernández-Rodríguez^{1†}, Francisco Quiles^{1†}, Ignacio Blanco¹, Alex Teulé¹, Lúdia Feliubadaló¹, Jesús del Valle¹, Mónica Salinas¹, Àngel Izquierdo², Esther Darder², Detlev Schindler³, Gabriel Capellá¹, Joan Brunet², Conxi Lázaro^{1*} and Miguel Angel Pujana^{4*}

Abstract

Background: Genes that, when mutated, cause Fanconi anemia or greatly increase breast cancer risk encode for proteins that converge on a homology-directed DNA damage repair process. Mutations in the *SLX4* gene, which encodes for a scaffold protein involved in the repair of interstrand cross-links, have recently been identified in unclassified Fanconi anemia patients. A mutation analysis of *SLX4* in German or Byelorussian familial cases of breast cancer without detected mutations in *BRCA1* or *BRCA2* has been completed, with globally negative results.

Methods: The genomic region of *SLX4*, comprising all exons and exon-intron boundaries, was sequenced in 94 Spanish familial breast cancer cases that match a criterion indicating the potential presence of a highly-penetrant germline mutation, following exclusion of *BRCA1* or *BRCA2* mutations.

Results: This mutational analysis revealed extensive genetic variation of *SLX4*, with 21 novel single nucleotide variants; however, none could be linked to a clear alteration of the protein function. Nonetheless, genotyping 10 variants (nine novel, all missense amino acid changes) in a set of controls (138 women and 146 men) did not detect seven of them.

Conclusions: Overall, while the results of this study do not identify clearly pathogenic mutations of *SLX4* contributing to breast cancer risk, further genetic analysis, combined with functional assays of the identified rare variants, may be warranted to conclusively assess the potential link with the disease.

Background

A functionally coherent network of gene and/or protein interactions, altered in Fanconi anemia (FA) and breast cancer (BrCa), has emerged in recent years [1]. Fifteen genes that, when mutated, cause FA (*FANC* genes) and several genes that may harbor mutations of high, moderate or low penetrance for BrCa risk encode for proteins that converge on a homology-directed DNA damage repair process [2]. As further evidence of a fundamental common causal basis between these diseases, germline bi- and mono-allelic loss-of-function mutations

in four of these genes cause FA and BrCa, respectively: *FANCD1/BRCA2* [3,4], *FANCF/BRIP1* [5-8], *FANCG/PALB2* [9-11] and *FANCI/RAD51C* [12,13] (more recent data suggests that mutations in *RAD51C* may be primarily linked to ovarian cancer risk [14]). This evidence marks any novel gene involved in the aforementioned network or process as a candidate to harbor mutations in unclassified FA and/or BrCa patients.

Interstrand DNA cross-link agents, such as mitomycin-C used in diagnostic tests for FA, block replication forks and may therefore cause genome instability. Homologs of *SLX4* in model organisms were initially identified as necessary for replication fork restart following exposure to DNA-damaging agents [15]. Subsequently, *SLX4* homologs have been shown to play a key role as docking molecules for the repair of interstrand cross-links [16,17]. These observations pointed to the human *SLX4* gene as a *FANC* candidate for unclassified

* Correspondence: clazaro@iconcologia.net; mapujana@iconcologia.net

† Contributed equally

¹Hereditary Cancer Program, Catalan Institute of Oncology (ICO), Hospital Duran i Reynals, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet, Barcelona, Catalonia, Spain

⁴Translational Research Laboratory, Biomedical Research Center Network for Epidemiology and Public Health (CIBERESP), ICO, IDIBELL, L'Hospitalet, Barcelona, Catalonia, Spain

Full list of author information is available at the end of the article

patients and, as a result, two groups have recently described mutations (renamed *FANCP* gene or FA-P subtype) [18,19]. The potential link to BrCa risk has been examined, to date, in 52 German or Byelorussian patients with familial breast cancer [20]: the study has not revealed truncating or clearly pathogenic mutations, but has identified four unclassified missense variants. Here, we conducted a more detailed study of the *SLX4* gene in 94 index BrCa cases from Spanish families negative for *BRCA1* and *BRCA2* mutations. As recently reported [20], our results do not show truncating or clearly pathogenic mutations, although they do describe seven missense variants of unknown biological significance that are not found in controls.

Methods

Study samples

Since its creation in 1999, the Hereditary Cancer Program at the ICO has identified a set of high-risk families with suspected hereditary breast and/or ovarian cancer syndrome. Following the Catalan Consensus Onco Guidelines on genetic testing for this condition, patients are analysed for mutations in the *BRCA1* or *BRCA2* genes after receiving appropriate genetic counselling and providing written informed consent. This genetic analysis consists of screening for point mutations and large rearrangements affecting those genes. For the present study, a total of 94 affected individuals belonging to 94 unrelated families negative for *BRCA1* or *BRCA2* mutations were selected. In addition to negativity for mutations in *BRCA1* or *BRCA2*, the inclusion criteria were: at least three first-degree relatives affected by breast or ovarian cancer; or at least two first-degree female relatives affected by breast cancer (at least one of them diagnosed before the age of 50); or at least one case of female breast cancer plus at least one case of either ovarian, female bilateral breast, or male breast cancer. Among the selected cases, 10 families were represented with an elevated prior risk of harboring a high-penetrance mutation, as calculated with the BRCAPRO [21] algorithm (briefly, 94 families mean = 0.40, standard deviation = 0.26, 95% confidence interval 0.11-0.99). Control samples, consisting of 138 women and 146 men, were taken from a hospital-based cancer association study (a detailed description of the study population, composition and interviews has been given elsewhere [22]). Specifically, these individuals were randomly enrolled from non-cancer patients admitted to the same general hospital as the BrCa cases. To avoid selection bias, the inclusion criterion for controls was that the current admission to the hospital should be for a new disease (not previously diagnosed). The studies were approved by the IDIBELL ethics committee and participants gave

written informed consent for their participation, and for the genetic analysis of their biological samples, according to the Declaration of Helsinki.

Mutation analysis

The *SLX4* exons and exon-intron boundaries were sequenced from polymerase chain reactions using previously defined primers and conditions [18]; for exon 7, AmpliTaq Gold DNA Polymerase (Applied Biosystems) and 10% of dimethyl sulfoxide were used. The reaction products were purified from remaining primers using ExoSAP-IT (GE Healthcare) and sequencing reactions performed following a standard Sanger method with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Samples harboring DNA variants were re-sequenced at least once using an independent DNA aliquot from the first-pass analysis. The guidelines of the Human Genome Variation Society and the reference sequences NM_032444.2 and NP_115820.2 of the National Center for Biotechnology Information were used for nomenclature.

Genotyping

Assays based on the KASPar technology were performed following the manufacturer's instructions (KBioscience). Reactions were carried out in a 384-well format with 2% of duplicates, and negative and positive (i.e. BrCa patient carrier) sample controls present in each plate.

Results and discussion

The genomic region of *SLX4*, comprising all exons and exon-intron boundaries, was sequenced in 94 BrCa familial cases that match a criterion indicating the potential presence of a highly-penetrant germline mutation, following exclusion of *BRCA1* or *BRCA2* mutations (see Methods). This mutational analysis revealed 49 variants: 21 novel and 28 which are currently annotated in the single nucleotide polymorphism database (dbSNP [23]) (Table 1). Of the 49 variants, 21 were found only once, which include three changes identified by the 1,000 Genomes Project [24] (Table 1): rs72778139-T has no known frequency data; rs76488917-A has an allele frequency of 0.02 in Caucasians; and rs115694169-A has an allele frequency of 0.03 in the Yoruba people of Ibadan (there is no data for Caucasians). Excluding these from the set of 21 with low frequency revealed eight missense and five silent changes at the protein level, and five intronic changes (Table 1). A neural network splicing prediction [25] model did not strongly support alteration by any of the identified intronic variants (data not shown). Together, these results suggest extensive genetic variation at the *SLX4* locus among individuals in our population, but provide no obvious link to BrCa risk.

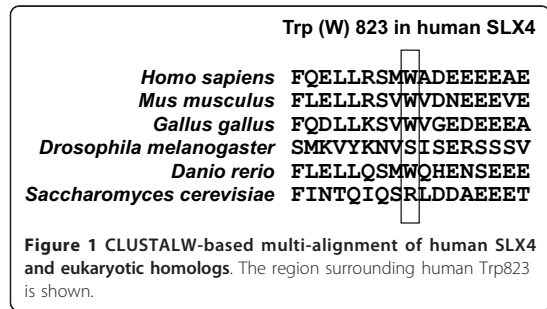
Table 1 SLX4 variants found in non-*BRCA1/2*-mutated familial BrCa cases

Exon	Nucleotide change	Change type	Amino acid change	Number of carriers*			dbSNP†
				Het (%)	Hom (%)	Total	
2	c.248G > C	Missense	p.Gly83Ala	1 (1.1)	0	94	NA
2	c.339T > C	Silent	p. =	1 (1.1)	0	94	NA
2	c.421G > T	Missense	p.Gly141Trp	1 (1.1)	0	94	NA
3	c.555C > T	Silent	p. =	9 (9.5)	0	94	rs74640850
3	c.610C > T	Missense	p.Arg204Cys	10 (10.6)	0	94	rs79842542
3	c.678C > T	Silent	p. =	3 (3.2)	0	94	rs28516461
3	c.590T > C	Missense	p.Val197Ala	1 (1.1)	0	94	NA
3	c.710G > A	Missense	p.Arg237Gln	2 (2.1)	0	94	NA
3	c.753G > A	Silent	p. =	24 (25.5)	2 (2.1)	94	rs8061528
4	c.761-32T > G	Intronic	p. =	2 (2.1)	0	94	NA
5	c.1152A > G	Silent	p. =	11 (11.7)	0	94	rs112511042
5	c.1153C > A	Missense	p.Pro385Thr	1 (1.1)	0	94	rs115694169
5	c.1156A > G	Missense	p.Met386Val	11 (11.7)	0	94	rs113490934
5	c.1163 + 10C > T	Intronic	p. =	11 (11.7)	0	94	rs80116508
6	c.1164-16T > C	Intronic	p. =	1 (1.1)	0	94	NA
6	c.1164-40C > A	Intronic	p. =	1 (1.1)	0	94	NA
6	c.1164-66T > A	Intronic	p. =	2 (2.1)	0	94	NA
6	c.1164-75C > G	Intronic	p. =	11 (11.7)	0	94	rs59622164
6	c.1366 + 11T > C	Intronic	p. =	12 (12.8)	0	94	rs76350200
7	c.1371T > G	Missense	p.Asn457Lys	10 (10.6)	0	94	rs74319927
7	c.1419C > T	Silent	p. =	1 (1.1)	0	94	NA
8	c.1846G > A	Missense	p.Val616Met	1 (1.1)	0	94	NA
9	c.2012T > C	Missense	p.Leu671Ser	11 (11.8)	0	93	rs77985244
9	c.2013 + 23G > A	Intronic	p. =	11 (11.7)	0	94	rs112226642
9	c.2013 + 137G > C	Intronic	p. =	11 (11.7)	0	94	rs801186343
10	c.2160 + 50C > T	Intronic	p. =	10 (10.6)	0	94	rs75762935
12	c.2346C > T	Silent	p. =	1 (1.1)	0	94	NA
12	c.2469G > C	Missense	p.Trp823Cys	1 (1.1)	0	94	NA
12	c.2854G > A	Missense	p.Ala952Thr	8 (8.5)	0	94	rs59939128
12	c.2855C > T	Missense	p.Ala952Val	8 (8.5)	0	94	rs78637028
12	c.3162G > A	Silent	p. =	1 (1.1)	0	94	rs76488917
12	c.3365C > T	Missense	p.Pro1122Leu	12 (12.8)	1 (1.1)	94	rs714181
12	c.3662C > T	Missense	p.Ala1221Val	10 (10.6)	0	94	rs3827530
12	c.3812C > T	Missense	p.Ser1271Phe	4 (4.2)	0	94	rs3810813
12	c.3872C > T	Missense	p.Thr1291Met	1 (1.1)	0	94	NA
12	c.4261A > T	Missense	p.Ile1421Phe	1 (1.1)	0	94	NA
12	c.4409C > T	Missense	p.Pro1470Leu	1 (1.1)	0	94	rs72778139
12	c.4500T > C	Silent	p. =	42 (44.7)	21 (22.3)	94	rs3810812
12	c.4530G > T	Silent	p. =	1 (1.1)	0	94	NA
13	c.4637-125C > T	Intronic	p. =	1 (1.1)	0	94	NA
13	c.4637-227C > T	Intronic	p. =	9 (9.6)	0	94	rs75693937
13	c.4739 + 10C > T	Intronic	p. =	1 (1.1)	0	94	NA
13	c.4739 + 24G > T	Intronic	p. =	20 (21.3)	2 (2.1)	94	rs12933120
14	c.5072A > G	Missense	p.Asn1691Ser	1 (1.1)	0	94	NA
15	c.5389C > T	Silent	p. =	1 (1.1)	0	93	NA
15	c.5501A > G	Missense	p.Asn1834Ser	2 (2.2)	0	93	rs111738042
15	c.*8A > G	Intronic	p. =	9 (9.7)	0	93	rs3751839
15	c.*102G > A	Intronic	p. =	1 (1.1)	0	93	NA
15	c.*113C > T	Intronic	p. =	8 (8.6)	0	93	rs76661336

*Het, heterozygous; Hom, homozygous
†Build 133; NA, not applicable

Having identified rare variants in BrCa familial cases, we next assessed the presence of 10 of these variants in a cohort of controls collected at the same hospital as the cases (see Methods). The selection of these variants was based on the observed low frequency in the 94 BrCa cases and on their identification as missense variations. In addition, a causative prediction was obtained using two algorithms (PolyPhen-2 [26] and SIFT [27]), plus a weighted average of scores (Condel [28]). Seven of these variants were not found in controls and, intriguingly, five of them were predicted to be “deleterious” (Table 2). Among this group, only one amino acid position (Trp823) showed some evolutionary conservation (Figure 1), and the substitution may be disfavored (Trp to Cys) [29]; tumor samples were not available for any case that would have allowed assessment of the existence of loss of heterozygosity at the *SLX4* locus. Nonetheless, predictions of a deleterious effect should be taken with caution as neutral polymorphisms can frequently be misclassified (from ~15-50% depending on the method [28]). On the other hand, extensive genetic variation in *SLX4* might reflect an unknown evolutionary pressure or could be related to a similar observation made for other DNA repair-related genes [30].

While *SLX4* serves as a scaffold for multiple proteins involved in the DNA damage response [16,17], the functional involvement of the Trp823 position, and of the other rare variants not found in controls in this study, is unknown. None of the identified variants changes a



critical amino acid residue and there is no data that could suggest an alteration of protein interactions or complexes; however, the Pro1470Leu variant might disrupt a mitotic phosphorylation site at Ser1469 [31,32]. In this context, a *SLX4* pathological variant linked to BrCa should consist of a hypomorphic mutation that would cause genome instability. Accordingly, *SLX4* is a key regulator of the function of structure-specific endonucleases involved in the repair of DNA damage; in particular, proper function of *SLX4* is fundamental for repair during replication and for resolving Holliday junctions formed during homologous recombination [16,17,33,34]. Overall, the results of this study do not support the existence of loss-of-function mutations of *SLX4* associated with BrCa risk; nonetheless, further genetic analysis in patients and controls, combined with

Table 2 Pathological prediction and frequency in controls of selected *SLX4* missense variants

Exon	Nucleotide change	Amino acid change	dbSNP†	Pathological prediction				Controls tested (n)	Number of control carriers (%)
				SIFT (score < 0.05, deleterious)	PolyPhen-2 (false positive rate)	Condel (weighted average of scores)	Condel prediction		
2	c.248G > C	p.Gly83Ala	NA	0.14	0.15	0.15	Neutral	283	0
2	c.421G > T	p.Gly141Trp	NA	0.00	0.86	0.86	Deleterious	284	2 (0.7)
3	c.590T > C	p.Val197Ala	NA	0.48	0.01	0.00	Neutral	284	1 (0.4)
3	c.710G > A	p. Arg237Gln	NA	0.49	0.00	0.38	Neutral	284	4 (1.4)
8	c.1846G > A	p. Val616Met	NA	0.17	0.62	0.80	Deleterious	281	0
12	c.2469G > C	p. Trp823Cys	NA	0.01	1.00	0.97	Deleterious	282	0
12	c.3872C > T	p. Thr1291Met	NA	0.05	0.98	0.99	Deleterious	283	0
12	c.4261A > T	p. Ile1421Phe	NA	0.08	0.77	0.77	Deleterious	285	0
12	c.4409C > T	p. Pro1470Leu	rs72778139	0.02	0.99	0.96	Deleterious	283	0
14	c.5072A > G	p. Asn1691Ser	NA	0.56	0.00	0.01	Neutral	285	0

†Build 133; NA, not applicable

functional assays of specific rare variants, may be warranted.

Conclusions

The mutational analysis of *SLX4* in 94 familial BrCa index cases without mutations in *BRCA1* or *BRCA2* has revealed extensive genetic variation. Twenty-nine novel single nucleotide variants have been detected, 21 of them showing relatively low allele frequencies: however, none can be linked to a clear alteration of the protein function. Nonetheless, analysis of 10 of these variants failed to detect seven of them in a set of controls. While the results of this study do not support the common existence of *SLX4* mutations contributing to BrCa risk, additional studies may be warranted.

Abbreviations

BrCa: Breast cancer; *BRCA1/2*: Breast cancer 1/2: early onset genes; BRCAPRO: *BRCA* carrier prediction model; Condel: Consensus deleteriousness score; dbSNP: Single nucleotide polymorphism database; DNA: Deoxyribonucleic acid; FA: Fanconi anemia; Het: Heterozygous; Hom: Homozygous; PolyPhen: Polymorphism phenotyping; SIFT: Sorting intolerant from tolerant; *SLX4*: Structure-specific endonuclease subunit homolog (*S. cerevisiae*) gene.

Acknowledgements

We wish to thank all study participants, clinicians, centers and members of the case-control study at the Catalan Institute of Oncology for their valuable contributions. The authors recognize the following grants provided to support this study: Generalitat de Catalunya 2009SGR290 to IB and CL, and 2009-SGR283 to MAP; Spanish Association Against Cancer (AECC groups) to GC; Spanish Ministry of Health "Instituto de Salud Carlos III" 10/31488, 10/01422, and 09/02483 to IB, CL, and MAP, respectively; 2009 Spanish Society of Medical Oncology to JB; and CIBERESP group 55 and Ramón Areces Foundation XV to MAP. JF-R was supported by a research support contract (CA08/00248) from the "Instituto de Salud Carlos III".

Author details

¹Hereditary Cancer Program, Catalan Institute of Oncology (ICO), Hospital Duran i Reynals, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet, Barcelona, Catalonia, Spain. ²Hereditary Cancer Program, ICO, Hospital Josep Trueta, Girona Biomedical Research Institute (IdIBGi), Girona, Catalonia, Spain. ³Department of Human Genetics, University of Würzburg, Biozentrum, Würzburg, Germany. ⁴Translational Research Laboratory, Biomedical Research Center Network for Epidemiology and Public Health (CIBERESP), ICO, IDIBELL, L'Hospitalet, Barcelona, Catalonia, Spain.

Authors' contributions

The project was conceived by DS, CL and MAP. The experiments and data analyses were coordinated by JF-R, CL and MAP. The mutational analysis was performed by JF-R and FQ. The clinical and *BRCA1/2* annotations were performed by IB, AT, LF, JV, MS, AI, ED, GC, JB and CL. The selection of index cases was performed by IB. The manuscript was written by MAP. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 13 September 2011 Accepted: 8 March 2012

Published: 8 March 2012

References

1. Wang W: Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet* 2007, **8**:735-748.
2. D'Andrea AD: Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med* 2010, **362**:1909-1919.

3. Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, et al: Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 2002, **297**:606-609.
4. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G: Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* 1995, **378**:789-792.
5. Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, et al: Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 2006, **38**:1239-1241.
6. Levran O, Attwooll C, Henry RT, Milton KL, Neveling K, Rio P, Batish SD, Kalb R, Velleuer E, Barral S, et al: The *BRCA1*-interacting helicase *BRIP1* is deficient in Fanconi anemia. *Nat Genet* 2005, **37**:931-933.
7. Levitus M, Waisfisz Q, Godthelp BC, de Vries Y, Hussain S, Wiegant WW, Elghalbzouri-Maghrani E, Steltenpool J, Roomians MA, Pals G, et al: The DNA helicase *BRIP1* is defective in Fanconi anemia complementation group J. *Nat Genet* 2005, **37**:934-935.
8. Litman R, Peng M, Jin Z, Zhang F, Zhang J, Powell S, Andreassen PR, Cantor SB: *BACH1* is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCF. *Cancer Cell* 2005, **8**:255-265.
9. Erkkö H, Xia B, Nikkila J, Schleutker J, Syrjäkoski K, Mannermaa A, Kallioniemi A, Pylkas K, Karppinen SM, Rapakko K, et al: A recurrent mutation in *PALB2* in Finnish cancer families. *Nature* 2007, **446**:316-319.
10. Reid S, Schindler D, Hanenberg H, Barker K, Hanks S, Kalb R, Neveling K, Kelly P, Seal S, Freund M, et al: Biallelic mutations in *PALB2* cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet* 2007, **39**:162-164.
11. Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, et al: *PALB2*, which encodes a *BRCA2*-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 2007, **39**:165-167.
12. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H, et al: Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nat Genet* 2010, **42**:410-414.
13. Vaz F, Hanenberg H, Schuster B, Barker K, Wiek C, Erven V, Neveling K, Endt D, Kesterton I, Autore F, et al: Mutation of the *RAD51C* gene in a Fanconi anemia-like disorder. *Nat Genet* 2010, **42**:406-409.
14. Pelttari LM, Heikkinen T, Thompson D, Kallioniemi A, Schleutker J, Holli K, Blomqvist C, Aittomäki K, Butzow R, Nevanlinna H: *RAD51C* is a susceptibility gene for ovarian cancer. *Hum Mol Genet* 2011, **20**:3278-3288.
15. Rites TM, Kobor MS, Bastin-Shanover SA, Li M, Horte SA, Gin JW, Emili A, Rine J, Brill SJ, Brown GW: *Slx4* regulates DNA damage checkpoint-dependent phosphorylation of the BRCT domain protein Rtt107/Esc4. *Mol Biol Cell* 2006, **17**:539-548.
16. Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ, Harper JW: Mammalian *BTBD12/SLX4* assembles a Holliday junction resolvase and is required for DNA repair. *Cell* 2009, **138**:63-77.
17. Fekairi S, Scaglione S, Chahwan C, Taylor ER, Tissier A, Coulon S, Dong MQ, Ruse C, Yates JR, Russell P, et al: Human *SLX4* is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* 2009, **138**:78-89.
18. Stoepker C, Hain K, Schuster B, Hillhorst-Hofstee Y, Roomians MA, Steltenpool J, Oostra AB, Eirich K, Korthof ET, Nieuwint AW, et al: *SLX4*, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet* 2011, **43**:138-141.
19. 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.
20. Landwehr R, Bogdanova NV, Antonenkova N, Meyer A, Bremer M, Park-Simon TW, Hillemanns P, Karstens JH, Schindler D, Dork T: Mutation analysis of the *SLX4/FANCP* gene in hereditary breast cancer. *Breast Cancer Res Treat* 2011, **130**:1021-1028.
21. Euhus DM, Smith KC, Robinson L, Stucky A, Olopade OI, Cummings S, Garber JE, Chittenden A, Mills GB, Rieger P, et al: Pretest prediction of *BRCA1* or *BRCA2* mutation by risk counselors and the computer model BRCAPRO. *J Natl Cancer Inst* 2002, **94**:844-851.
22. Landi S, Moreno V, Gioia-Patricola L, Guino E, Navarro M, de Oca J, Capella G, Gazian F: Association of common polymorphisms in inflammatory genes interleukin (IL) 6, *IL8*, tumor necrosis factor alpha,

- NFKB1**, and peroxisome proliferator-activated receptor gamma with colorectal cancer. *Cancer Res* 2003, **63**:3560-3566.
23. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29**:308-311.
 24. Consortium GP: **A map of human genome variation from population-scale sequencing.** *Nature* 2010, **467**:1061-1073.
 25. Reese MG, Eeckman FH, Kulp D, Haussler D: **Improved splice site detection in Genie.** *J Comput Biol* 1997, **4**:311-323.
 26. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: **A method and server for predicting damaging missense mutations.** *Nat Methods* 2010, **7**:248-249.
 27. Kumar P, Henikoff S, Ng PC: **Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.** *Nat Protoc* 2009, **4**:1073-1081.
 28. González-Pérez 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**:440-449.
 29. Betts MJ, Russell RB: **Amino acid properties and consequences of substitutions.** In *Bioinformatics for Geneticists*. Edited by: Barnes MR, Gray IC. Chichester: John Wiley and Sons; 2003:289-316.
 30. Mohrenweiser HW, Wilson DM, Jones IM: **Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes.** *Mutat Res* 2003, **526**:93-125.
 31. Cantin GT, Yi W, Lu B, Park SK, Xu T, Lee JD, Yates JR: **Combining protein-based IMAC, peptide-based IMAC, and MudPIT for efficient phosphoproteomic analysis.** *J Proteome Res* 2008, **7**:1346-1351.
 32. Dephoure N, Zhou C, Villen J, Beauleil SA, Bakalarski CE, Elledge SJ, Gygi SP: **A quantitative atlas of mitotic phosphorylation.** *Proc Natl Acad Sci USA* 2008, **105**:10762-10767.
 33. Crossan GP, van der Weyden L, Rosado IV, Langevin F, Gaillard PH, McIntyre RE, Gallagher F, Kettunen MI, Lewis DY, Brindle K, *et al*: **Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia.** *Nat Genet* 2011, **43**:147-152.
 34. Wechsler T, Newman S, West SC: **Aberrant chromosome morphology in human cells defective for Holliday junction resolution.** *Nature* 2011, **471**:642-646.

Pre-publication history

The pre-publication history for this paper can be accessed here:
<http://www.biomedcentral.com/1471-2407/12/84/prepub>

doi:10.1186/1471-2407-12-84

Cite this article as: Fernández-Rodríguez *et al*: **Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families.** *BMC Cancer* 2012 **12**:84.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



ARTÍCULO 4

Artículo 4: Evaluation of rare variants in the new fanconi anemia gene ERCC4 (FANCC) as familial breast/ovarian cancer susceptibility alleles.

Osorio A¹, Bogliolo M, Fernández V, Barroso A, de la Hoya M, Caldés T, Lasa A, Ramón y Cajal T, Santamariña M, Vega A, **Quiles F**, Lázaro C, Díez O, Fernández D, González-Sarmiento R, Durán M, Piqueras JF, Marín M, Pujol R, Surrallés J, Benítez J.

Human Mutation. 2013 Dec; 34(12):1615-8.

Factor de Impacto (2013): 5,050

Evaluation of Rare Variants in the New Fanconi Anemia Gene *ERCC4* (*FANQ*) as Familial Breast/Ovarian Cancer Susceptibility Alleles

Ana Osorio,^{1,2†} Massimo Bogliolo,^{2,3†} Victoria Fernández,¹ Alicia Barroso,¹ Miguel de la Hoya,⁴ Trinidad Caldés,⁴ Adriana Lasa,⁵ Teresa Ramón y Cajal,⁶ Marta Santamariña,^{2,7} Ana Vega,^{2,8} Francisco Quiles,⁹ Conxi Lázaro,⁹ Orland Díez,¹⁰ Daniel Fernández,¹¹ Rogelio González-Sarmiento,¹¹ Mercedes Durán,¹² José Fernández Piqueras,^{2,13} María Marín,³ Roser Pujol,^{2,3} Jordi Surrallés,^{2,3*} and Javier Benítez^{1,2,**}

¹Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, CNIO, Madrid, Spain; ²Centre for Biomedical Network Research on Rare Diseases (CIBERER), Spain; ³Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain; ⁴Molecular Oncology Laboratory, Hospital Clínico San Carlos. Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Barcelona, Spain; ⁵Genetics Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ⁶Oncology Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ⁷Grupo de Medicina Xenómica -USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain; ⁸Fundación Pública Galega de Medicina Xenómica-SERGAS Santiago de Compostela, Spain; ⁹Hereditary Cancer Program, Catalan Institute of Oncology (ICO), Hospital Duran i Reynalds, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet, Barcelona, Spain; ¹⁰Oncogenetics Laboratory, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, and University Hospital of Vall d'Hebron, Barcelona, Spain; ¹¹Instituto de Biología Molecular y Celular del Cáncer (IBMCC), laboratory 14, Universidad de Salamanca-CSIC, Salamanca, Spain; ¹²Institute of Biology and Molecular Genetics, Universidad de Valladolid (IBGM-UVA), Valladolid, Spain; ¹³Centro de Biología Molecular Severo Ochoa (CBM), Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid, Spain

Communicated by Christopher G. Mathew

Received 31 May 2013; accepted revised manuscript 28 August 2013.

Published online 11 September 2013 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22438

ABSTRACT: Recently, it has been reported that biallelic mutations in the *ERCC4* (*FANQ*) gene cause Fanconi anemia (FA) subtype FA-Q. To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, as occurs with other FA genes, we screened the 11 coding exons and exon–intron boundaries of *ERCC4* in 1573 index cases from high-risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for *BRCA1* and *BRCA2* mutations and 854 controls.

The frequency of *ERCC4* mutation carriers does not differ between cases and controls, suggesting that *ERCC4* is not a cancer susceptibility gene. Interestingly, the prevalence of *ERCC4* mutation carriers (one in 288) is similar to that reported for *FANCA*, whereas there are approximately 100-fold more FA-A than FA-Q patients, indicating that most biallelic combinations of *ERCC4* mutations are embryo lethal. Finally, we identified additional bone-fide FA *ERCC4* mutations specifically disrupting interstrand cross-link repair.

Hum Mutat 34:1615–1618, 2013. © 2013 Wiley Periodicals, Inc.

KEY WORDS: Fanconi anemia; breast cancer; *ERCC4*; *FANQ*; XPF

[†]These authors contributed equally to this work.

*Correspondence to: Jordi Surrallés, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. E-mail: jordi.surralles@uab.es

**Correspondence to: Javier Benítez, Human Genetics Programme Spanish National Cancer Centre (CNIO) C/Melchor Fernández Almagro 3, 28029 Madrid, Spain. E-mail: jbenitez@cnio.es

Contract grant sponsors: Mutua Madrileña Foundation (FMMA); Intrasalud PI12/00070; Acción Cooperativa y Colaborativa Intramural-CIBER2012; SAF2010-20493; Generalitat de Catalunya (SGR0489-2009), the ICREA-Academia award, the Spanish Ministry of Economy and Competitiveness (CIBERER CB06/07/0023, SAF2009-11936 and SAF2012-31881); the European Regional Development FEDER Funds; Fondo de Investigación Sanitaria (FIS) (PI 12/00539); European Regional Development FEDER Funds; Asociación Española Contra el Cáncer; Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute and Autonomous Government of Catalonia (ISCIII/RETIC RD06/0020/1051, RD12/0036/008, PI10/01422, PI10/00748, and 2009SGR290). Acción Cooperativa y Colaborativa Intramural-CIBERER 2012, the OncoCycle Programme (S2011/BMD-2470) and SAF2012-36556; and Fondo de Investigación Sanitaria (PI10/00219)

The *ERCC4* (*FANQ*) gene (MIM #133520) encodes for a DNA repair endonuclease (XPF) that plays essential roles in nucleotide excision repair (NER) and interstrand cross-link repair (ICLR) [Gregg et al., 2011]. Biallelic mutations in *ERCC4* have been linked to Xeroderma Pigmentosum (XP; MIMs #278700, #610651, #278720, #278730, #278740, #278760, #278780, and #278750) [Sijbers et al., 1996] and progeria (XFE; MIM #610965) [Niedernhofer et al., 2006] diseases. Very recently, it has been published that biallelic mutations in the gene are also causative of disease in previously unclassified Fanconi anemia (FA) patients subtype FA-Q and, therefore, the use of *FANQ* as an alias for *ERCC4* was approved by the HUGO Gene Nomenclature Committee [Bogliolo et al., 2013]. Individuals harboring mutations in *ERCC4* show high variability on their clinical manifestations, ranging from mild XP symptoms (sun sensitivity, freckling of the skin, and basal or squamous cell carcinomas) to the dramatic accelerated aging symptoms of a progeroid

Table 1. Unique Genetic Variants Identified in the *ERCC4* Gene in 1573 Cases and 854 Controls

Exon	Nucleotide change ^a	Protein change	Previously described	MAF reported ^c	Phenotype	Predicted to affect protein function ^c	Functional assay ^f
3	c.448C>T ^b	p.Arg150Cys	rs145402255	0.0	Case	Yes	Mild effect on ICLR
3	c.540_541delAG	p.Arg180SerfsX30	—	—	Control ^d	Yes	—
IVS3	c.584+1G>A ^b	p.Ile131SerfsX2	—	—	Case	Yes	—
5	c.800G>A	p.Arg267His	rs143479220	0.0001	Case	No	—
9	c.1861A>G	p.Thr621Ala	—	—	Case	No	—
11	c.2065C>A	p.Arg689Ser	rs149364215	0.0001	Control	Yes	Disrupts ICLR
11	c.2291delG	p.Ser764ThrfsX53	—	—	Control	Yes	—
11	c.2357C>T ^b	p.Ser786Phe	rs143081574	—	Case	Yes	Disrupts ICLR

^aMutation nomenclature listed uses GenBank reference sequences NM_005236.2 with numbering starting at the A of the first ATG, following the HGVS guidelines (www.hgvs.org/mutnomen).

^bApart from the 854 controls in which *XPF/ERCC4* was fully sequenced, these variants were specifically analyzed in 300 additional controls. The mutation gave rise to skipping of exon 3 (r.389_584del) that was confirmed at the cDNA level (data not shown).

^cMAF (minor allele frequency) in European American Population as described in the NHLBI Exome Sequencing Project (ESP) (<http://evs.gs.washington.edu/EVS/>).

^dMutation found in two independent controls.

^eAll mutations giving rise to a PTC were considered as predicted to affect protein function. Missense mutations were previously evaluated using the program CONDEL that predicts the pathogenicity of nonsynonymous variants using a consensus deleteriousness score that combines various tools such as SIFT, Polyphen2, and MutationAssessor (<http://bg.upf.edu/condel/home>). Those missense variants predicted to be deleterious by condel were considered as predicted to affect protein function and this was confirmed by functional analysis; those predicted by condel to be neutral were considered as such and not further evaluated.

^fDetailed in the text and in Figure 1.

syndrome, and finally, as recently shown [Bogliolo et al., 2013] in FA patients, to malformations and progressive bone marrow failure. It has also been demonstrated that a balance between NER and ICLR activities due to different effects caused by distinct mutations in *ERCC4* determined the final clinical outcome [Bogliolo et al., 2013]. On the other hand, given that FA-Q patients are normal in FANCD2 monoubiquitination, it has been suggested that monoallelic mutations in *ERCC4* could be linked to the breast and ovarian hereditary cancer syndrome as it occurs with other FA genes such as *BRCA2/FANCD1* [Howlett et al., 2002], *BRIP1/FANCF* [Seal et al., 2006], *PALB2/FANCF* [Rahman et al., 2007], and *RAD51C/FANCO* [Meindl et al., 2010; Osorio et al., 2012] (MIMs #300514, #610832, #614082, and #614083, respectively) [Levy-Lahad, 2010].

To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, we screened, by DHPLC (denaturing high-performance liquid chromatography) and direct sequencing, the 11 coding exons and exon–intron boundaries of the *ERCC4* gene in 1573 index cases from high-risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for mutations in *BRCA1* and *BRCA2* and 854 controls without personal or familial antecedents of cancer. Criteria for inclusion of cases and controls, and methods of screening for mutations in *BRCA1/2* have been previously published [Bonache et al., 2013; Fernandez-Rodriguez et al., 2012; Osorio et al., 2012; Romero et al., 2011]. We identified five and four unique variants among cases and controls among which three and four, respectively, were considered as putatively deleterious. Missense mutations were evaluated using the program CONDEL that predicts the pathogenicity of nonsynonymous variants using a consensus deleteriousness score that combines various tools such as SIFT, Polyphen2, and MutationAssessor (<http://bg.upf.edu/condel/home>). Those missense variants predicted to be deleterious by CONDEL were considered as predicted to affect protein function and this was studied later by functional analysis; those predicted to be neutral were considered as such and not further evaluated (Table 1). Regarding cases, one of the mutations found, c.584+1G>A in intron 3, was confirmed to cause skipping of exon 3 of the gene and a premature stop codon (PTC) (data not shown). The other two mutations, p.Arg150Cys in exon 3 and p.Ser786Phe in exon 11, were later functionally investigated. Regarding controls, we identified two different frameshift mutations c.540_541delAG in exon 3 (found in two independent controls) and c.2291delG in exon 11, both predicted to cause PTCs. Even though

this later PTC is very C-terminal and could potentially result in a shorter but partially functional protein, this mutation results in a truncated XPF protein that lacks the double helix-hairpin-helix (HhH2) domain involved in heterodimerization with ERCC1 and DNA binding [de Laat et al., 1998], very similar to a pathogenic *ERCC4* mutation (c.2371_2398dup28; p.Ile800Thrfs*24) that functionally disrupts NER and ICLR activities [Bogliolo et al., 2013]. The last deleterious variant found in controls was the missense p.Arg689Ser in exon 11 previously found in a FA patient, and demonstrated to cause abnormal nuclease activity and to specifically disrupt ICLR [Bogliolo et al., 2013]. No differences were found regarding the localization of mutations in the gene among cases and controls (Table 1). All variants reported have been submitted to the Leiden Open Variation Database (LOVD).

To evaluate the functional impact of the missense variants, we cloned a HA-tagged wild type (WT) *ERCC4* cDNA in a pBABE-puro retroviral vector (Addgene plasmid 14430, kindly shared by Dr. L.M. Martins) in internal ribosome entry site (IRES) with the GFP cDNA, and the c.448C>T, c.2065C>A, and c.2357C>T variants were introduced by site directed mutagenesis [Bogliolo et al., 2013]. The resulting constructs were transduced in NER and ICLR deficient *Ercc4* KO mouse embryonic fibroblasts (MEFs) and, after puromycin selection, the green cells were sorted to achieve a purity of over 98% by fluorescence activated cell sorting [Bogliolo et al., 2013]. Due to the bicistronic nature of the IRES construct, we were able to assess the stability of the mutant XPF proteins using GFP as a reference, since both proteins are encoded by the same mRNA [Pelletier and Sonenberg, 1988]. The p.Arg689Ser variant reduced by 40% the stability of XPF, whereas the p.Arg150Cys and p.Ser786Phe variants had no impact on protein stability (Fig. 1A and B). Ultraviolet radiation subtype C (UVC) sensitivity of Xpf KO MEFs was complemented with the expression of both p.Arg150Cys and p.Ser786Phe-XPF (Fig. 1C), and only the Xpf KO MEFs expressing p.Ser786Phe-XPF or p.Arg689Ser showed a FA phenotype in terms of mitomycin C (MMC) sensitivity (Fig. 1D), MMC-induced cell cycle arrest at the G2/M phase (Fig. 1E), and DEB-induced chromosome fragility (Fig. 1F). These data confirm that, resembling p.Arg689Ser, p.Ser786Phe specifically disrupts ICLR and, therefore, is a bone-fide FA mutation. Interestingly, both mutations are located in the nuclease domain of XPF. Despite a mild MMC sensitivity (Fig. 1D and E), *Ercc4* KO MEFs expressing p.Arg150Cys-XPF did not show DEB-induced chromosome fragility (Fig. 1F). These data,

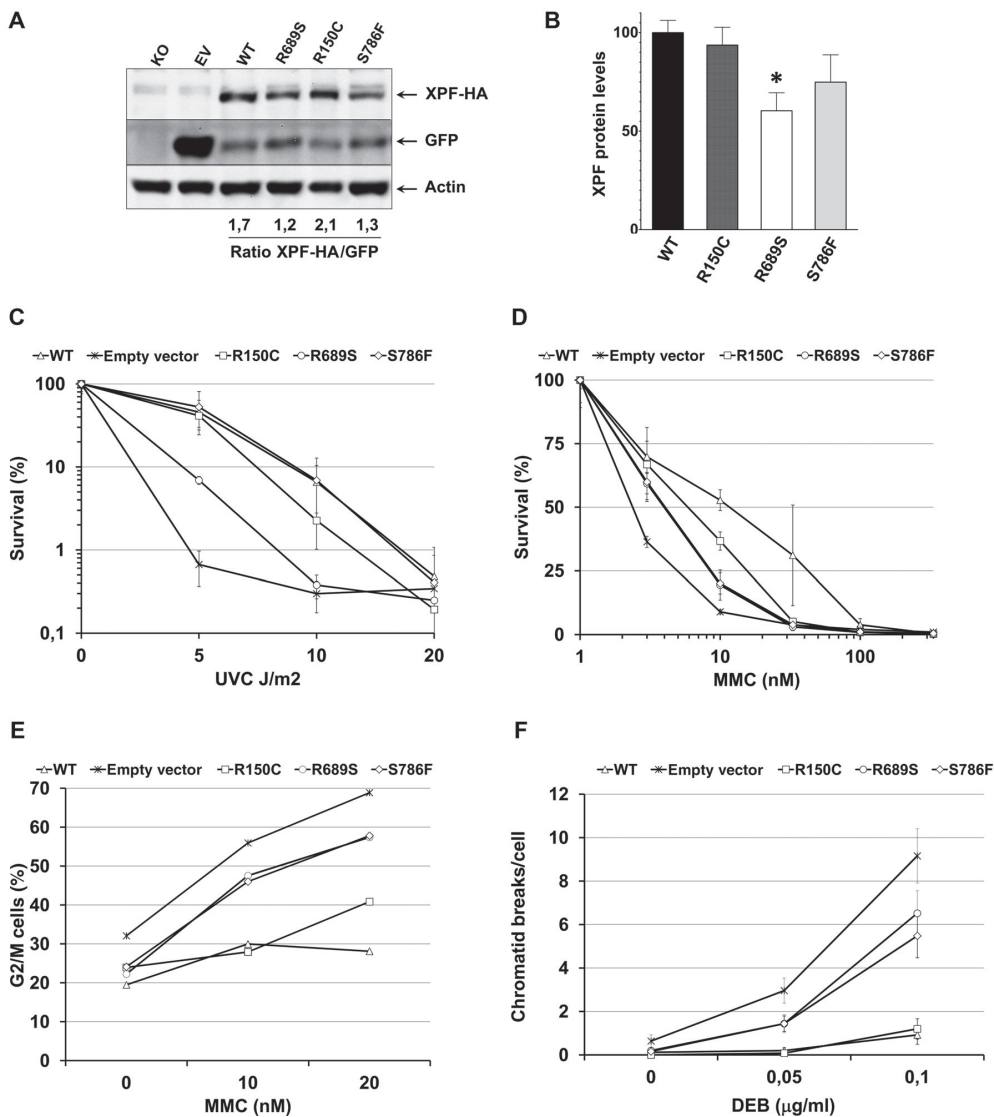


Figure 1. Functional studies of *ERCC4* missense variants. **A:** Western blot analysis showing HA-XPF and GFP expressions in the transduced MEFs as a measure of XPF protein stability. XPF levels were quantified with ImageJ software and are expressed as a ratio relative to the GFP levels. **B:** XPF levels were quantified as in (A) and expressed as a percentage of the WT. Mean and SEM of six experiments are shown (* $p = 0.005$ student's t -test). **C:** UVC sensitivity of *Ercc4* KO MEFs expressing the different alleles of *ERCC4*. Cells were analyzed after 72 hr of irradiation. Data represent means and SD of three independent experiments. **D:** MMC sensitivity in the same cells as in (C). Data represent means and SD of two independent experiments. **E:** MMC-induced G2/M cell cycle arrest in the same cells as in (C). **F:** DEB-induced chromosome fragility test arrest in the same cells as in (C). Chromatid breaks were scored in 25 metaphase per point. Methods for western blot, site-directed mutagenesis, cDNA transduction, antibodies used, and the experiments presented in (D), (E), and (F) were performed as reported earlier [Bogliolo et al., 2013; Trujillo et al., 2012].

together with the protein stability and UVC sensitivity data, indicate a null impact of the c.448C>T variant on XPF NER functions and a mild effect on ICLR activity.

In conclusion, the frequency of Spanish individuals heterozygous for pathogenic mutations in the *ERCC4* gene is approximately 0.3%, and it does not differ between familial breast/ovarian cancer patients and healthy controls ($p = 0.251$), suggesting that monoallelic muta-

tions in *ERCC4* are not linked to cancer susceptibility in the general population. Similar results were found with *SLX4* that, like *ERCC4*, acts downstream FANCD2 monoubiquitination but upstream the homologous recombination step of ICLR [Fernandez-Rodriguez et al., 2012]. The prevalence of *ERCC4* mutation carriers (one in 288) is similar to that reported for *FANCA*. However, there are approximately 100-fold more FA-A than FA-Q individuals, suggesting

that over 90% of biallelic combinations of *ERCC4* mutations are embryo lethal in humans. All reported XP patients subtype XPF worldwide have at least one missense mutation disrupting NER, whereas all missense mutations found in 2.427 Spanish individuals have substantial NER activity explaining why there are no reported XPF families in Spain.

Acknowledgments

This study was approved by the Institutional Committee on Ethical Research in Human Samples and proper informed consent was obtained. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain. We wish to thank all the members of the ICO Hereditary Cancer Program team led by Dr. Gabriel Capella.

All authors of this manuscript made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data, collaborated in drafting the article or revising it critically for important intellectual content and gave final approval of the version to be published.

Disclosure statement: The authors declare no conflict of interest.

References

- Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillon J, Ramirez MJ, Pujol R, Casado JA, Banos R, et al. 2013. Mutations in *ERCC4*, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* 92:800–806.
- Bonache S, de la Hoya M, Gutierrez-Enriquez S, Tenes A, Masas M, Balmana J, Diez O. 2013. Mutation analysis of the *SHFM1* gene in breast/ovarian cancer families. *J Cancer Res Clin Oncol* 139:529–532.
- de Laat WL, Sijbers AM, Odijk H, Jaspers NG, Hoeijmakers JH. 1998. Mapping of interaction domains between human repair proteins *ERCC1* and *XPF*. *Nucleic Acids Res* 26:4146–4152.
- Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capella G, Brunet J, et al. 2012. Analysis of *SLX4/FANCP* in non-*BRCA1/2*-mutated breast cancer families. *BMC Cancer* 12:84.
- Gregg SQ, Robinson AR, Niedernhofer LJ. 2011. Physiological consequences of defects in *ERCC1-XPF* DNA repair endonuclease. *DNA Repair (Amst)* 10:781–791.
- Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, et al. 2002. Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 297:606–609.
- Levy-Lahad E. 2010. Fanconi anemia and breast cancer susceptibility meet again. *Nat Genet* 42:368–369.
- Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H, Ramsr J, Honisch E, et al. 2010. Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nat Genet* 42:410–414.
- Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, van Leeuwen W, Theil AF, Vermeulen W, et al. 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* 444:1038–1043.
- Osorio A, Endt D, Fernandez F, Eirich K, de la Hoya M, Schmutzler R, Caldes T, Meindl A, Schindler D, Benitez J. 2012. Predominance of pathogenic missense variants in the *RAD51C* gene occurring in breast and ovarian cancer families. *Hum Mol Genet* 21:2889–2898.
- Pelletier J, Sonenberg N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334:320–325.
- Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, et al. 2007. *PALB2*, which encodes a *BRCA2*-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 39:165–167.
- Romero A, Perez-Segura P, Tosar A, Garcia-Saenz JA, Diaz-Rubio E, Caldes T, 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 Res Treat* 129:939–946.
- Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, et al. 2006. Truncating mutations in the Fanconi anemia *J* gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 38:1239–1241.
- Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, et al. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 86:811–822.
- Trujillo JP, Mina LB, Pujol R, Bogliolo M, Andrieux J, Holder M, Schuster B, Schindler D, Surralles J. 2012. On the role of *FAN1* in Fanconi anemia. *Blood* 120: 86–89.

Capítulo 4

SECUENCIACIÓN DE EXOMAS EN FAMILIAS BRCA1

Resultados no publicados:

Búsqueda de nuevos genes de predisposición al SCMOH mediante secuenciación de exomas de pacientes pertenecientes a familias BRCA1.

Resumen

Cerca de un 25% de los pacientes diagnosticados de cáncer de mama (CM) y ovario muestran historia familiar de esta enfermedad. Mutaciones en los genes *BRCA1* y *BRCA2* son detectadas aproximadamente en un 15% de estos casos. Por otra parte, mutaciones de alta, moderada o baja penetrancia se han identificado en otros genes (como *PALB2*, *BRIP1*, *CHEK2*...) y han podido explicar una pequeña porción de los casos restantes. No obstante, hoy en día, más de un 50% de los casos familiares permanecen sin causa genética conocida. Son los casos comúnmente conocidos como casos *BRCA*. En este contexto, las nuevas técnicas de secuenciación masiva se presentaron como una potente herramienta para la identificación de nuevos genes de predisposición al cáncer. Estas técnicas nos permiten interrogar exomas completos de una forma coste-efectiva. En el presente estudio, se utilizaron las técnicas de secuenciación masiva para la secuenciación de los exomas de varios pacientes pertenecientes a 12 familias *BRCA*, con un aparente patrón de herencia autosómico dominante, con la finalidad de identificar mutaciones en nuevos genes de predisposición al SCMOH. El análisis de las variantes obtenidas en la secuenciación, condujo a la selección de 20 genes candidatos que están siendo actualmente validados en un panel de 1368 pacientes *BRCA* pertenecientes a familias independientes.

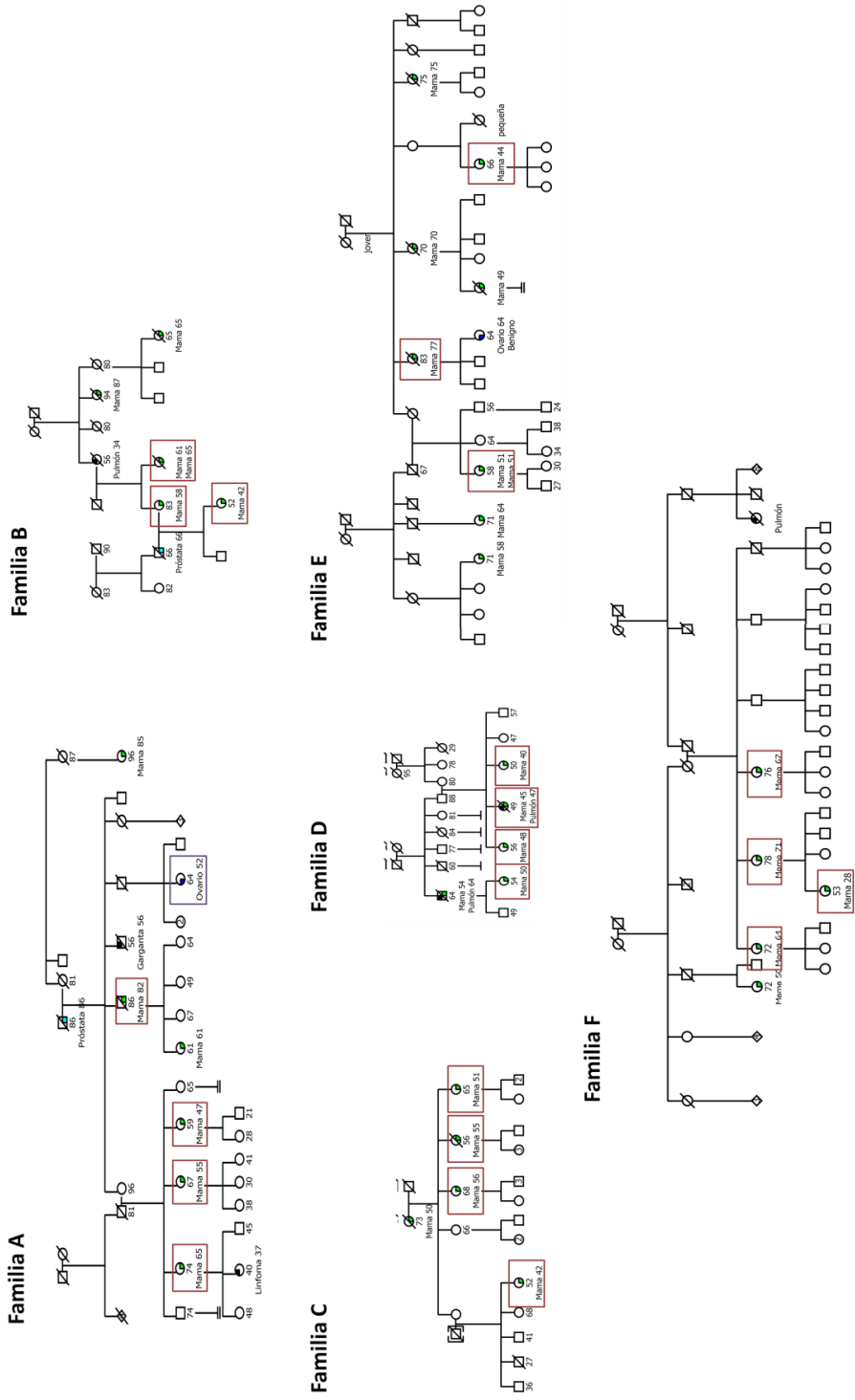
Metodología

Selección de familias

Doce familias con sospecha de síndrome de cáncer de mama y ovario hereditario (SCMOH) fueron seleccionadas para este estudio. Se seleccionaron de entre todas las familias *BRCA* identificadas en nuestro centro, el Programa de Cáncer Hereditario, Instituto Catalán de Oncología (ICO), y en centros colaboradores de la Comunidad Valenciana.

Desde su creación en 1998, el Programa de Cáncer Hereditario del ICO ha identificado un conjunto de 2402 familias con sospecha de sufrir SCMOH. Siguiendo el modo de actuación establecido por la OncoGuía de Consejo Genético de Cataluña (http://www.gencat.net/salut/depsan/pdf/og0601esgrr_din4.pdf), se les realizó el estudio mutacional de *BRCA1* y *BRCA2* a los casos índices de dichas familias. El test genético de estos genes resultó positivo para 430 familias (18% de los casos) y negativo para 1976 familias (Familias *BRCA*) (1743 de ellas sin variante detectada y 233 portadoras de variantes de significado desconocido). De entre estas familias, se seleccionaron las que mostraban un patrón de herencia aparentemente autosómico dominante y presentaban tres o más personas afectas en un mínimo de dos-tres generaciones diferentes, con disponibilidad de muestra de DNA de sangre de al menos 3 afectos. Además, con el fin de aumentar el poder de identificación de la estrategia planteada, cuando fue posible, se escogieron familias que presentarían casos de bilateralidad y/o con DNA disponible de familiares afectos en ramas alejadas^{297, 298, 358}. En base a este criterio, se seleccionaron por consenso, entre un equipo multidisciplinar formado por miembros de la Unidades de Consejo Genético y de Diagnóstico Molecular del Programa de Cáncer Hereditario, las 7 familias *BRCA* más representativas del ICO. En base al mismo criterio, 5 familias más fueron seleccionadas por los centros colaboradores de la Comunidad Valenciana, para así completar la serie de las 12 familias *BRCA* que se analizaron en este estudio (Fig. 1).

Figura 1. Genealogías de las 12 familias BRCAx analizadas en este estudio



Secuenciación de exomas de pacientes pertenecientes a familias BRCA1 y selección de variantes candidatas

El exoma de varios pacientes de cada familia seleccionada fue secuenciado haciendo uso de la técnica de secuenciación masiva. Las familias se secuenciaron en dos etapas, primero se secuenciaron 5 familias del ICO y después se secuenciaron 7 más, 2 del ICO y 5 de la Comunidad Valenciana. Se secuenciaron entre 3-5 individuos de cada familia, en base a la disponibilidad de muestras. Las librerías de exomas se obtuvieron mediante la fragmentación, captura y enriquecimiento exónico del DNA de los pacientes, utilizando el *SureSelect Human All Exon 50 Mb Kit* (Agilent) junto con el *Paired-End Sample Preparation Kit* (Illumina), siguiendo las instrucciones del fabricante. Estas librerías fueron secuenciadas con el secuenciador *HiSeq 2000* (Illumina), utilizando dos carriles y 52 ciclos por muestra, lo que resultó en más de 60 millones de lecturas emparejadas por muestra. Tanto la obtención de librerías como la secuenciación masiva fueron realizadas por el Centro de Análisis Genómicos (CNAG). La selección inicial (*Variant calling*) de variantes candidatas fue realizada por el grupo de investigación del Dr. Xose Ponte (Universidad de Oviedo) haciendo uso de una metodología de filtraje bioinformático que ellos mismo habían descrito previamente³⁵⁹. Dicha selección se hizo en base a la hipótesis de que las variantes candidatas debían responder a un patrón de herencia mendeliana autosómica dominante. De este modo, se seleccionaron las variantes en heterocigosis comunes a todos los pacientes de una misma familia y que podían alterar la función de la proteína. Las variantes frecuentes presentes en la base de datos dbSNP o con una frecuencia mayor al 2% en una base de datos local obtenida por dicho grupo fueron excluidas. Más del 90% del exoma fue cubierto por al menos diez lecturas.

Identificación de mutaciones germinales en “pools” de muestras

Los 20 genes seleccionados como candidatos están en proceso de ser validados en un set de 1368 pacientes BRCA1 del ICO. Para ello, se adaptó la metodología SMIPS (*Somatic Mutation Identification in Pooled Samples*), desarrollada previamente por el grupo del Dr. Ponte³⁶⁰, para la búsqueda de mutaciones germinales. Esta metodología permite acelerar y facilitar el proceso de cribado gracias a la posibilidad de amplificar y secuenciar simultáneamente hasta 96 muestras de DNA mezcladas en un único tubo (*pool*). La metodología seguida fue la siguiente, en un placa de 96 pozos se prepararon diluciones equimolares de cada DNA que sería incluido en el *pool* (100ng/ μ l). Se añadió la misma cantidad de cada uno de estos DNA al *pool*. Las regiones exónicas e intrónicas flanqueantes de cada gen candidato fueron amplificadas con la *Read-proof Platinum Pfx DNA Polymerase (Invitrogen)* utilizando el DNA del *pool*. Los productos de PCR fueron purificados con el kit *QIAquick 96 PCR Purification (Qiagen)* y cuantificados en Nanodrop ND-1000. Todos los amplicones generados fueron mezclados de forma equimolar en un nuevo *pool* para ser ligados y secuenciados utilizando un secuenciador HiSeq de Illumina en el CNAG. La ligación, las librerías de DNA y la obtención del listado de las variantes se realizó por el grupo del Dr. Ponte (para más información técnica véase Ponte et. al, 2013; *Supplementary Material*). La secuenciación se realizó en varias tandas de 40 kb cada una, obteniéndose una cobertura por base entre 40.000X-50.000X. La alta cobertura obtenida por base permitió diferenciar entre variante reales y fallos derivados de la secuenciación, de la construcción de librerías o de los errores introducidos por la DNA polimerasa. En una situación ideal, al utilizar cantidades

equimolares de DNA en el *pool*, la frecuencia alélica de una determinada variante en dicho *pool* debería ser 0.52%. No obstante, en una situación real, ciertos DNA pueden estar menos representados debido a su mala calidad o a errores humanos en la preparación del *pool*. Para reducir el riesgo de perder variantes por este motivo, en el proceso de *Variant calling*, se decidió establecer una frecuencia alélica límite de 0,3% para considerar una variante como real.

Resultados

La secuenciación de exomas y la selección de genes candidatos se realizó en dos etapas consecutivas en el tiempo, una primera con 5 familias (Familias de la A-E) y la segunda con las 7 restantes (Familias de la F-L). El proceso de selección de variantes fue el mismo en ambas etapas, con la única diferencia que, gracias a la mayor disponibilidad de tiempo, en la primera etapa, las variantes candidatas fueron genotipadas en una cohorte de de 288 controles de población española. Por otra parte, también en la primera etapa, se decidió validar el uso de la metodología SMIPS en nuestro laboratorio. Para ello, se genotiparon, mediante esta técnica, varios de los genes seleccionados en esta etapa en una cohortes de 262 casos BRCA. Dicha validación se realizó en vistas a futuros estudios a mayor escala, en los que se incluiría todos los genes candidatos seleccionados en el presente estudio. Los resultados de la validación se muestran al final del apartado de resultados.

Selección de variantes candidatas

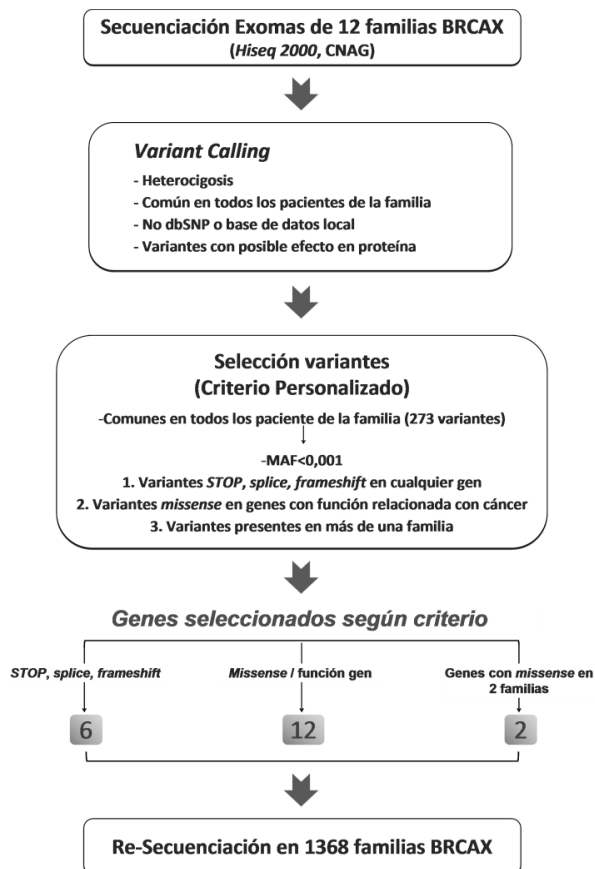
Los exomas completos de varios pacientes pertenecientes a 12 familias BRCA independientes (Fig. 1), fueron secuenciados utilizando la tecnología de secuenciación masiva en paralelo. Entre 3 y 5 pacientes fueron analizados en cada familia. Dado que este estudio fue diseñado para identificar nuevos genes dominantes de predisposición al CM de alta penetrancia, las 12 familias seleccionadas presentaban una fuerte agregación de CM, considerándose 2-3 generaciones, con un aparente patrón de herencia autosómico dominante. Además, por este mismo motivo, en el filtraje de los datos crudos de la secuenciación de exomas (*Variant calling* inicial), se seleccionaron las variantes en heterocigosis comunes en todos los pacientes de una misma familia y que podían alterar la función de la proteína (*STOP*, *frameshift*, *indels*, *splicing* (+/- 1,2) y *missense*). Este *Variant Calling* inicial resultó en 261 variantes distribuidas en las diferentes familias y los diferentes tipos de variantes como se muestra en la Tabla 1.

Tabla 1. Variantes seleccionadas después de *Variant Calling* inicial distribuidas según familias y tipo de variante.

	Familia A	Familia B	Familia C	Familia D	Familia E	Familia F	Familia G	Familia H	Familia I	Familia J	Familia K	Familia L	TOTAL
Misense	5	33	18	11	8	32	19	11	18	0	14	29	198
STOP	4	0	1	0	0	0	1	1	0	0	0	2	9
Splice	0	0	2	0	0	0	0	0	1	0	0	0	3
Frameshift	1	4	0	1	0	2	0	3	2	1	0	3	17
Indels	4	1	1	4	3	6	1	5	2	0	2	5	34
TOTAL	14	38	22	16	11	40	21	20	23	1	16	39	261

Algunas de las variantes que pasaron el filtro inicial de selección fueron posteriormente descartadas por afectar a genes que están frecuentemente mutados en población general). Estos genes suelen aparecer mutados en los proyectos de secuenciación de exomas pero la probabilidad de ser causales de la enfermedad es muy baja. Son genes tales como los de la familia de los *DNAH*, *MUC16* o los que codifican para receptores olfativos^{361, 362}. Un esquema global sobre la metodología seguida en la selección de variantes se muestra en la Figura 2.

Figura 2. Metodología seguida para la selección de genes candidatos



Para la selección de las variantes candidatas, de entre las resultantes del *Variant Calling* inicial, se estableció un criterio personalizado basado en una serie de premisas:

1. Se quiso atender a la heterogeneidad genética de las familias BRCA1/2. Por ello, de entrada, ningún gen quedó descartado por su función o por no haber sido relacionado con el cáncer.

2. En primera estancia, se seleccionaron variantes que podían generar un codón de terminación prematuro (CTP) [*STOP* (st), *splice* (sp) o *frameshift* (fs)]. Este tipo de variantes son

consideradas potencialmente deletéreas y han sido ampliamente relacionadas con el desarrollo de enfermedad³⁶³.

3. Conocimiento previo de la prevalencia a la susceptibilidad al CM en familias BRCA.

Además de estas premisas, el criterio de selección se ajustó a la hipótesis de que los posibles nuevos genes de predisposición al CM de alta penetrancia, son mucho menos prevalentes que los genes *BRCA1/2*, ya que estudios previos así lo sugieren^{329, 330}. La prevalencia de portadores de mutaciones en *BRCA1/BRCA2*, en la población general, se estima del orden de 1/1000 individuos³⁶⁴⁻³⁶⁶. La frecuencia de la mutación recurrente que más veces ha sido reportada en BIC, la mutación fundadora en judíos Azkenazi, la c.68_69delAG, corresponde al 13% del total de variantes reportadas en esta base datos. Por tanto, la frecuencia poblacional de la mutación más frecuente en *BRCA1* es del 0,013%. Así, considerando la hipótesis planteada, se decidió excluir aquellas variantes que aparecieran con una *Minor Allele Frequency (MAF)* mayor que 0,001% (10 veces menor que él de la variante más recurrente en *BRCA1*). La frecuencia poblacional de las variantes seleccionadas se comprobó en las bases de datos *Exome Sequencing Project (ESP)* (<http://evs.gs.washington.edu/>) y *1000 Genomes* (<http://www.1000genomes.org/>).

- Los genes de susceptibilidad descritos hasta ahora en estas familias han sido genes involucrados en procesos de reparación de DNA, regulación del ciclo celular, diferenciación o migración celular.

En base a estas premisas, se jerarquizaron las variantes en tres tipos:

Variantes raras (MAF<0,001%) que podían causar un CTP y afectaban a todos los pacientes de una misma familia, independientemente del gen afectado.

Todas las variantes que cumplían este criterio (n=26) fueron confirmadas mediante secuenciación Sanger. Siete variantes resultaron verdaderos positivos: cuatro de ellas, en los genes *ARHGEF38 (sp)*, *ECL2 (st)*, *MUTYH (fs)*, *PIPOX (fs)*, fueron identificadas en la primera etapa y las tres restantes en la segunda etapa en los genes *KIR3DL3 (st)*, *POLE (fs)*, *ST6GALNAC2 (sp)*.

Variantes *missense* o *indels in frame* raras (MAF<0,001), comunes en todos los pacientes de una misma familia y que afectaban a genes con funciones relacionadas con el desarrollo del cáncer (reparación de DNA, regulación del ciclo celular, diferenciación o migración celular). Variantes de este tipo fueron seleccionadas en los genes *ACTR8*, *BEX2*, *FGFR2*, *MCM7*, *PHF20L1*, *PRDM2*, *SEPT9*, *TET1* y, *ZFH3* (primera etapa) y *DVL1*, *DVL2* y *MCM2* (segunda etapa) (Fig 3). Todas las variantes *missense* seleccionadas también fueron confirmadas por secuenciación Sanger, resultando todas en verdaderos positivos. No se identificó ninguna variante indels que afectara a genes relacionados con cáncer.

Missense o *indels in frame* raras (MAF<0,001%) identificadas en un gen mutado en más de una familia, independientemente de la función del gen afectado. Dos genes, *GRM6*, *ZSWIM6*, fueron seleccionados siguiendo este criterio (Fig. 3). Las variantes presentes en estos genes también fueron confirmadas por Sanger.

A continuación, se evaluó, mediante varios programas de predicción *in silico*, el posible efecto funcional de las variantes *missense* seleccionadas (Tabla 2). Las predicciones fueron consistentes para 10 de las 16 variantes identificadas, siendo seis de ellas consideradas deletéreas por todos los programas (genes *DVL2*, *FGFR2*, *GRM6*, *MCM2*, *MCM7*, *SEPT9*) y cuatro consideradas probablemente neutras (genes *ACTR8*, *PHF20L1*, *PRDM2*, *ZSWIM6*). Para cinco de las variantes los resultados no fueron concluyentes (genes *BEX2*, *DVL1*, *GRM6*, *TET1*, *ZHX3*, *ZSWIM6*). Dado la alta variabilidad de resultados obtenidos por estos programas, las predicciones *in silico* no fueron utilizadas para descartar ningún gen y se realizaron únicamente a modo informativo. Si bien, nos sirvieron como evidencias secundarias para interpretar los resultados obtenidos^{134, 164-166}.

Tabla 2. Predicciones *in silico* para las variantes *missense* seleccionadas en este estudio

Familia	Genes	Variante	SIFT	Mutation	
				PolyPhen2	Taster
A	FGFR2	V533M	Deletérea	Deletérea	Deletérea
	PHF20L1	K561R	Neutra	Neutra	Neutra
B	ACTR8	S70N	Neutra	Neutra	Neutra
	BEX2	R141W	Deletérea	Deletérea	Neutra
	GRM6	R141W	Neutra	Deletérea	Deletérea
	TET1	P1345L	Neutra	Posible Deletérea	Neutra
	ZFH3	S1648R	Neutra	NR	NR
D	MCM7	R561C	Deletérea	Deletérea	Deletérea
	SEPT_9	R289H	Deletérea	Deletérea	Deletérea
	ZSWIM6	D875E	Neutra	Deletérea	Deletérea
E	PRDM2	S1505G	Neutra	Neutra	Neutra
F	DVL1	R289H	Neutra	Neutra	Deletérea
I	GRM6	R274W	Deletérea	Deletérea	Deletérea
	DVL2	Y553T	Deletérea	Deletérea	Deletérea
K	MCM2	R122W	Deletérea	Deletérea	Deletérea
	ZSWIM6	I238V	Neutra	Neutra	Neutra

NR, no reconocido

Por último, las variantes seleccionadas en la primera etapa de este estudio fueron genotipadas en un set de 288 controles sanos de población española. Únicamente la variante de *STOP* en el gen *ECTL2* fue identificada con una frecuencia superior al 1% (1,1%, 3 individuos), lo que llevó a la exclusión de este gen para futuros análisis.

Cabe destacar que, durante el proceso de selección de variantes, apareció un nuevo caso de cáncer de ovario en la familia A (Fig. 1). Se estudió la cosegreación de las variantes seleccionadas en esta nueva paciente. Solo la variante en *PHF20L1* fue identificada en esta paciente. Paralelamente, se consiguió muestra tumoral de varios pacientes de esta familia. Se realizó el estudio de LOH de los genes seleccionados. Ninguno de los dos genes seleccionados mostró LOH en ninguno de los pacientes analizados.

En resumen, se seleccionaron 20 genes candidatos a partir de la secuenciación de exomas de 12 familias BRCA (Fig. 3). El conjunto de genes seleccionados está siendo actualmente validado en una serie de 1384 pacientes BRCA independientes mediante una adaptación de la metodología SMIPS.

Validación de la Metodología de secuenciación SMIPS

Tras el análisis de los exomas de las cinco primeras familias surgieron cuatro genes candidatos con los criterios comentado anteriormente. De esta forma, se decidió genotipar los genes seleccionados (*BEX2*, *FGFR2*, *MCM7*, *PHF20L1*, *SEPT9*) en una cohorte de 262 pacientes BRCA1/2 utilizando una adaptación de la metodología SMIPS, a modo de validación de dicha metodología. Entre los DNA que se secuenciaron, se encontraba el DNA del portador de la mutación en *BEX2* identificada en la secuenciación de exomas. Este DNA pudo ser utilizado como control positivo y fue correctamente identificado. Este análisis SIMPS identificó dos nuevas mutaciones *missense* raras en *BEX2*, *MCM7* y *SEPT9*, una en *PHF20L1* y ninguna en *FGFR2*. Es interesante destacar que las dos nuevas mutaciones en *MCM7* estaban localizadas en la misma región en la que se identificó la primera mutación en el análisis de los exomas. En esta región se produce la interacción con *Rad17*, la cual es crítica para el correcto funcionamiento del punto de control en la fase S del ciclo celular como respuesta al daño celular³⁶⁷. La mutación en *PHF20L1* afectaba su dominio funcional Tudor, involucrado en la mayoría de las funciones de este gen³⁶⁸. *BEX2* y *SEPT9* no tienen dominios funcionales definidos con lo que no se pudo interpretar el efecto de las variantes identificadas.

Las 262 muestras utilizadas en esta prueba piloto han sido incluidas también en la cohorte de 1384 pacientes en las que se están genotipando los genes candidatos seleccionados en este estudio, lo que servirá para validar nuestros resultados.

Figura 3. Genes candidatos distribuidos en las familias en las que han sido identificados.

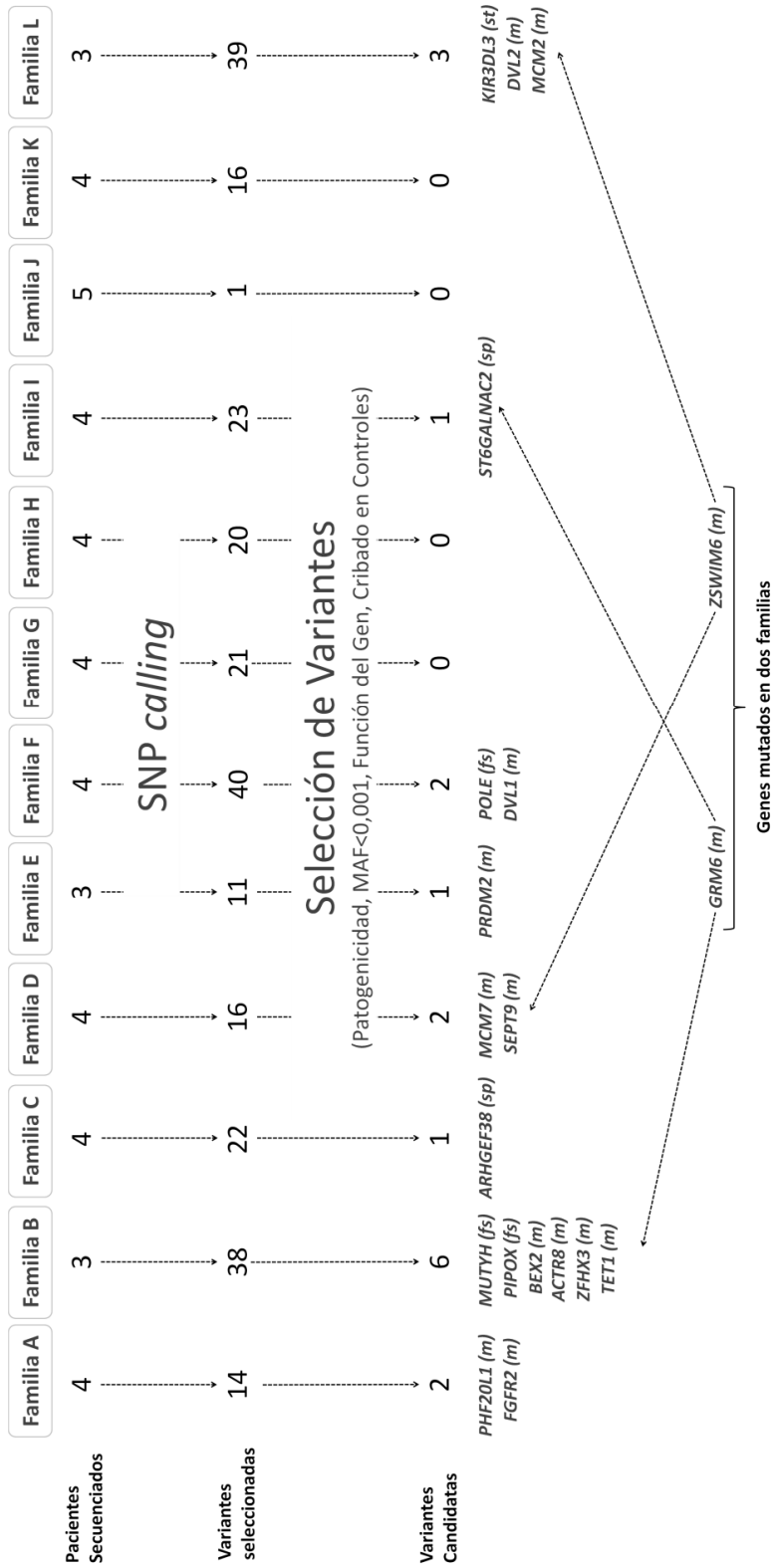


Figura 3. En esta figura se muestran los genes candidatos seleccionados en este estudio distribuidos entre las 12 familias analizadas. Después del Variant calling inicial, se aplicaron diferentes criterios para la selección de genes candidatos. Primero se seleccionaron los genes con variantes consideradas potencialmente patogénicas (truncantes, frameshift, splice). A continuación se seleccionaron los genes con funciones relacionadas con el desarrollo del cáncer y que portaban una variante missense común en todos los pacientes de la misma familia. Por último, se seleccionaron genes que estaban mutados en más de una familia. Sólo se consideraron variantes con frecuencia poblacional menor del 0,001%. Como resultado final, se obtuvo un conjunto de 20 genes candidatos distribuidos en varias familias. m, missense; fs, frameshift, sp, splice.

Discusión

El presente estudio fue planteado con el objetivo de identificar nuevos genes de predisposición al SCMOH, el diseño fue pensado para la identificación de genes con un patrón de herencia autosómico dominante y con alta penetrancia. Se utilizó un criterio de selección de variantes candidatas que quiso atender a la heterogeneidad genética de los casos BRCA. Siguiendo este criterio, se seleccionaron 20 genes candidatos. Una breve descripción de la función de estos genes y de su relación con el cáncer se puede ver en la Tabla 3 situada al final de este apartado.

La interpretación de los resultados obtenidos, así como la determinación de la causalidad de las diferentes variantes identificadas es una tarea de gran complejidad. Seis de los genes candidatos presentaban mutaciones de *STOP*, *splice*, o *frameshift*. A priori, este tipo de variantes suelen considerarse patogénicas y por tanto, asociadas a enfermedad^{16, 55, 297}. No obstante, estudios recientes han demostrado que la identificación de una mutación de este tipo en un determinado gen, aunque ésta cosegregue con la enfermedad, no es suficiente para probar su causalidad, debido a que este tipo de variantes pueden aparecer en población sana con una frecuencia considerable o conjuntamente con mutaciones causales conocidas^{334, 369}. Esta observación se ve confirmada en nuestro estudio con la detección de la variante de *STOP* identificada en el gen *ECTL2* en tres controles sanos. Ahora bien, el conocimiento previo de la función y el espectro mutacional del gen afectado puede ayudar a determinar la causalidad de este tipo de variantes. En este sentido, la conocida susceptibilidad al cáncer ligada a mutaciones truncantes en los genes *MUTYH* y *POLE* permite interpretar los resultados obtenidos en este estudio. Mutaciones en el gen *MUTYH* (*A/G-specific adenine DNA glycosylase*), predisponen a cáncer colorrectal de forma autosómica recesiva^{370, 371}. La variante detectada en este estudio, p.Glu410Glyfs*43, es una de las variantes patogénicas más frecuentes en este gen y está descrito que está presente en un 1% de la población general en heterocigosis. En un estudio previo³⁷², esta mutación había sido identificada en homocigosis en dos pacientes que sufrieron, además de cáncer colorrectal, CM, a los 50 y 59 años^{373, 374}. Dado que el modelo de predisposición al cáncer descrito para este gen es recesivo, cabría esperar la identificación de una segunda mutación en este gen para asociar el fenotipo tumoral al mismo. Se ha revisado los datos de los exomas en la muestra de estos pacientes y no se ha identificado ninguna mutación patogénica adicional. Evidentemente, no podemos descartar que estos pacientes presenten una segunda mutación que escape el método de análisis utilizado por lo que haría falta un estudio complementario de RNA o de grandes reordenamientos para descartar el papel de *MUTYH* en esta familia. La estructura familiar y la edad avanzada de los pacientes analizados, sin que ninguno de ellos haya desarrollado cáncer colorrectal, parece descartar a *MUTYH* como responsable del SCMOH en esta familia³⁷⁵. Sin embargo también puede existir la posibilidad de que mutaciones en heterocigosis en *MUTYH* predispongan al CM de forma dominante. De hecho, varios estudios demuestran un aumento significativo en la frecuencia de variantes patogénicas en heterocigosis en *MUTYH* en sujetos BRCA, estimándose un posible aumento del riesgo asociado a este gen entre 1-3,4 (ref^{320, 376, 377}), similar al reportado para *CHEK2* (ref³⁵⁸). Si bien, no todos los estudios publicados encuentran evidencias en este sentido, por lo que el papel de *MUTYH* en el CM está aún por determinar en análisis de grandes cohortes de pacientes^{320, 376-381}.

Mutaciones en el gen *POLE*, que codifica para la DNA polimerasa ϵ , predisponen a un nuevo síndrome de predisposición al cáncer colorrectal, denominado *polymerase proofreading-associated polyposis* (PPAP)^{382, 383}. Además, estudios publicados recientemente indican que mutaciones en este gen predisponen a varios tipos de cáncer, entre ellos al cáncer de ovario^{384, 385}. Sin embargo, hasta ahora no ha sido reportada ninguna mutación en *POLE* asociada a la predisposición al CM. Todas las mutaciones de predisposición al cáncer descritas para este gen afectan su dominio exonucleasa (aminoácidos 86-426), sugiriendo que este dominio es el responsable de tal predisposición. La mutación identificada en el presente estudio, p.Gly395Glyfs*16, causa la pérdida de los últimos 31 aminoácidos de dicho dominio. Esta variante está presente en tres hermanas y una sobrina con cáncer de mama, existe una prima paterna de las tres hermanas con diagnóstico de cáncer de mama de la cual no se ha podido obtener muestra para ampliar el análisis de cosegregación (Fig. 1, Familia F). No obstante, las evidencias aportadas por este estudio sugieren que *POLE* actúa como gen de predisposición al CM de alta penetrancia en la familia analizada. Por otra parte, se ha descrito que los tumores asociados a *POLE* corresponden con tumores hipermutados³⁸². El estudio del espectro mutacional de los tumores de estas pacientes podrían aportar mayores evidencias que confirmaran nuestros resultados. Lamentablemente, este estudio no se ha podido realizar hasta el momento. Finalmente, el estudio SIMPS que se está realizando ayudará a elucidar la prevalencia de mutaciones en *POLE* en familias BRCA.

Más difícil es determinar la susceptibilidad asociada al resto de genes identificados en este estudio. Muchas de las familias analizadas presentaron mutaciones en varios genes que, por su función relacionada con el desarrollo del cáncer, son potencialmente candidatos por sí solos. Esta observación no es inesperada, ya que tanto las evidencias aportadas por estudios de asociación como la identificación de numerosos genes de moderada penetrancia indican que parte de las familias BRCA podrían estar asociadas a un modelo de herencia poligénico^{297, 298, 307, 312, 358}. Además, la presencia de mutaciones en varios genes en una misma familia sugiere que en muchos casos, la susceptibilidad puede ser familia-específica. Cabe destacar que, varios estudios previos, diseñados con el mismo objetivo que el nuestro, encontraron resultados similares, indicando que este es un escenario habitual en el estudio de familias BRCA^{310, 334, 335}. Por otra parte, no se puede excluir que en algunas de estas familias coexistan mutaciones en un gen dominante junto con alelos modificadores del riesgo, como sucede con los pacientes portadores de mutaciones patogénicas en *BRCA1/2* que además portan SNPs modificadores de riesgo (el consorcio internacional CIMBA, del que nuestro equipo forma parte, ha identificado varios de estos modificadores) (ver bibliografía del consorcio CIMBA, a modo de ejemplo las referencias 58-63). Por ejemplo, se ha identificado una mutación en *FGFR2* (Familia A) que es un gen modificador de riesgo asociado a *BRCA2*, pero que por sí solo únicamente confiere un pequeño incremento de riesgo al CM. En esta familia podría ser que la mutación identificada en *FGFR2* fuera un modificador del riesgo de una mutación dominante en el gen *PH20L1*, identificada en esta familia^{312, 386-388}. También podría ser que algunas de las variantes seleccionadas en este estudio no tengan ninguna repercusión sobre la predisposición al cáncer, como sugieren algunas de las predicciones *in silico* (Tabla 2). En este sentido, se espera que el estudio SIMPS que se está realizando actualmente con nuestros genes candidatos, ayude a determinar cuáles de ellos son verdaderos genes de susceptibilidad. No obstante, la asignación de una mutación específica como una verdadera variante de

predisposición al SCMOH requerirá de más evidencias clínicas, genéticas y funcionales (análisis *in vitro*, análisis de tumores, análisis en grandes cohortes, modelos animales, etc.).

El hecho de que no se haya identificado ningún gen candidato evidente en varias de las familias estudiadas (Familias G, H, J, K), aún cuando presentan una alta agregación de CM, sugiere que quizás el algoritmo de clasificación utilizado no ha sido suficiente específico. Estos genes pueden no haber sido identificado por varios motivos. Una revisión crítica a posteriori nos sugiere que los criterios de selección utilizados quizás fueron muy restrictivos y presentaban ciertas limitaciones. Así, sólo se seleccionaron aquellas variantes comunes en todos los pacientes de una misma familia excluyendo así la posibilidad de fenocopias en las familias estudiadas. A día de hoy sabemos que las fenocopias en casos de familias de CM asociadas a *BRCA1/2* se sitúan alrededor del 15% (ref^{139, 389}) por lo que estamos pensando en realizar un reanálisis de los resultados teniendo en cuenta esta posibilidad. Además, la estrategia de selección fue diseñada para detectar genes dominantes de alta penetrancia, por lo que se seleccionaron únicamente variantes en heterocigosis muy raras ($MAF < 0,001$). Con esto se excluyó la posibilidad de un modelo poligénico o un modelo recesivo de herencia. Subir el punto de corte, por ejemplo, a $MAF < 0,01$, podría conducir a la identificación de nuevas variantes de moderada o baja penetrancia. Por otra parte, en algunas de las familias se podría modelar un patrón de herencia recesivo (Familias A, C, D) con fenocopias asociadas. Por último, sólo se han considerado variantes que puedan afectar a la función de la proteína de forma clara. Sin embargo, existen estudios que demuestran que variantes en regiones reguladoras o intrónicas también pueden ser causantes de enfermedad^{265, 266}. Destacar que, en la familia K, en la que se secuenciaron 5 individuos, sólo se identificaron dos variantes que cosegregaban con la enfermedad, una *deep intronic* y otra en una región 3' UTR. Estas variantes no fueron consideradas para futuros estudios por la dificultad de clasificación debido a su ubicación. Con todo ello, estamos convencidos que el reanálisis de los datos obtenidos desde nuevas perspectivas, probablemente, podría identificar nuevos genes de predisposición en las familias analizadas.

Como resumen de este estudio, se concluye que la secuenciación de exomas de pacientes pertenecientes a familias BRCA seleccionadas por su alta carga familiar de CM es una buena estrategia para identificar nuevos genes de predisposición, aunque probablemente la diversidad clínica de estas familias limita la obtención de resultados concluyentes y quizás la realización de una clasificación clínica más homogénea previa a los estudios genéticos hubiera sido más útil para la interpretación de los resultados obtenidos. Aunque la identificación de genes de alta penetrancia es aún posible, como sugiere la identificación de la mutación en el gen *POLE* en una de las familias, parece que la susceptibilidad en los casos BRCA es muy heterogénea y podría estar asociada a diferentes modelos de herencia, específicos de cada familia. Tales observaciones se deberán tener en cuenta en el diseño e interpretación de resultados de futuros experimentos, siendo conveniente el análisis de resultados desde diferentes perspectivas, independientemente del diseño planteado. Este escenario indica que, serán necesarios estudios a gran escala para conseguir evidencias claras de la susceptibilidad asociada a muchos de los nuevos genes de predisposición identificados.

Tabla 2. Descripción de los genes seleccionados. Función y relación con cáncer.

Gen	Nombre (inglés)	Mut.	Función	Relación con Cáncer	Ref.
ACTR8	Actin-related protein 8	m	Reparación de DNA mediante el reclutamiento del complejo de remodelación de la cromatina INO80	frecuentemente fusionado en células cancerígenas de próstata y leucemia	390-392
ARHGEF38	Rho guanine nucleotide exchange factor 38	sp	Liberación de guaninas	...	393
BEX2	Brain-Expressed X-linked gene 2	m	Regulación ciclo celular	Necesario para la progresión del ciclo celular en la fase G1 en CMT (Oncogen)	394
DVL1, DVL2	Segment polarity protein dishevelled homolog proteins	m	Papel importante en la vía de <i>Wnt</i> , involucrada en la proliferación, diferenciación y migración celular	La pérdida de expresión de estos genes puede dar lugar a la desregulación de esta vía <i>Wnt</i> y conducir a la tumorigénesis	395-397
FGFR2	Fibroblast growth factor receptor 2	m	Tirosina quinasa; regulación de la proliferación celular, la diferenciación, la migración, apoptosis	Modificador del riesgo asociado a <i>BRCA2</i> , mantenimiento de las células madre iniciadoras de tumor (sobreeexpresado puede inducir el desarrollo de varios tipos de cáncer)	312, 387, 394, 398-400
GRM6	glutamate receptor, metabotropic 6	m	Receptor de glutamato y transporte de calcio	...	401
KIR3DL3	Killer cell immunoglobulin-like receptor 3DL3	st	Control de las infecciones víricas y eliminación de células tumorales	Mutaciones asociadas con incremento de riesgo a neoplasias de cervix. Genes de su familia posibles oncogenes o protectores en CM	402, 403
MCM2 MCM7	, mini-chromosome maintenance protein 2	m	Iniciación replicación del DNA. Dianas de <i>ATR</i> como respuesta al estrés genotóxico. <i>ATR foci</i> en sitios de daño al DNA en fase S dependiente de <i>MCM7</i> .	Disminución de su expresión en la fase G0-->G1 causa defectos en la reparación del DNA y aberraciones cromosómicas.	367, 404-407
MUTYH	A/G-specific adenine DNA glycosylase	fs	Reparación DNA por escisión de bases	Predispone a cáncer colorrectal de forma autosómica recesivo	320, 370, 371, 376-379, 408
PHF20L1	PHD finger protein 20-like protein 1	m	Posible factor de transcripción (por similitud)	Potencial supresor tumoral en CO, frecuentemente fusionado (posible <i>driver</i> en líneas celulares de CM)	409-411

Tabla 2 (Continuación). Descripción de los genes seleccionados. Función y relación con cáncer.

Gen	Nombre	Mut.	Función	Relación con Cáncer	Ref.
PIPOX	Peroxisomal sarcosine oxidase	fs	Metaboliza sarcosina, ácido L-pipecólico o y L-prolina.	Reduce el potencial oncogénico de células cancerígenas de próstata. Su expresión está asociada a mal pronóstico en CM.	412
POLE	DNA polymerase epsilon catalytic subunit A	fs	Reparación y replicación DNA	Predispone al cáncer colorrectal y a otros tipos de cáncer; autosómico dominante	28, 29
PRDM2	PR domain zinc finger protein	m	Regulación negativa de la proliferación celular	En CMT, comúnmente su expresión está reducida o eliminada. Reconstitución su expresión en estas células causa parada de ciclo celular (G2-M) y/o apoptosis.	413-416
SEPT9	Septin-9	m	Actividad GTPasa, involucrada en la citoquinesis y la división celular	Localizado en una región del cromosoma 17q comúnmente pérdida en tumores de mama y ovario (posible supresor tumoral). Identificación de tumores de mama con línea germinal mutada.	417-419
ST6GALNAC2	Alpha-N-acetylgalactosaminide 2,6-sialyltransferase 2	sp	Glicosilación de proteínas	Supresor de metástasis en tumores de mama ER-negativos	420, 421
TET1	Methylcytosine dioxygenase TET1	m	Dioxigenasa involucrada en la regulación de cromatina, la metilación del DNA y la represión transcripcional	Supresor tumoral en varios tipos de cáncer. Desregulación de su expresión, activa la vía de señalización TNFa-p38-MAPK, resultando en producción de CMT iniciadoras	422-426
ZFX3	Zinc finger homeobox protein 3	m	Regulador transcripcional. Regulador de la glándula mamaria durante la pubertad mediante la inhibición de la función proliferativa del receptor de estrógenos	Frecuentemente inactivado en células tumorales de mama (mal pronóstico). Gen supresor tumoral en cáncer de próstata	427-431
ZSWIM6	Zinc finger SWIM domain-containing protein 6	m	Regulación de la morfogénesis y la migración de células neuronales	...	432

Mut.: Tipo de mutación; m, missense; sp, splice; st, STOP; CMT, células de mama tumorales; CM: Cáncer de Mama

RESUMEN DE RESULTADOS

La presente tesis se planteó con el objetivo de mejorar el diagnóstico y el consejo genético de pacientes afectados por el CM familiar. Para ello, se plantearon dos estrategias, la clasificación de VSD en *BRCA1/2* identificadas en pacientes con sospecha de SCMOH y la búsqueda de nuevos genes de predisposición a este síndrome.

Para la clasificación de VSD en los genes *BRCA1/2* se utilizaron dos aproximaciones. Por una parte, se analizaron 28 VSD a nivel transcripcional para determinar su efecto en el *splicing* del RNA de los citados genes. Este análisis identificó seis variantes que producían un patrón aberrante de *splicing*. Como consecuencia de ello, estas 6 variantes pudieron ser clasificadas como probablemente patogénicas. Previamente, nuestro grupo había participado en el trabajo *Quality Control Project*, diseñado por el *Splicing Working Group (SWG)* del consorcio ENIGMA, a modo de validación de la metodología utilizada por nuestro equipo en análisis de RNA de este tipo de variantes. Los resultados obtenidos por nuestro grupo en este estudio, mostraron que nuestra metodología de análisis era válida, y por tanto, nuestros resultados robustos.

Por otra parte, se analizaron 12 VSD en *BRCA1/2* a nivel funcional. Siete de estas variantes, localizadas en la región C-terminal de *BRCA1*, fueron testadas mediante el ensayo de actividad transcripcional (TA) diseñado para este gen (Q1409L, S1473P, E1586G, R1589H, Y1703S, W1718L, G1770V). Tres de ellas, Y1703S, W1718L y G1770V pudieron ser clasificadas como probablemente patogénicas en base a este estudio. Otras seis VSD en *BRCA2* (T200I, K331I, K1434I, V1532F, R2645G, F3146S) fueron analizadas mediante el ensayo funcional basado en células embrionarias de ratón. Estas variantes mostraron un efecto hipomórfico, lo que sugiere que pueden estar asociadas al SCMOH con un riesgo bajo/moderado.

Durante el desarrollo de esta tesis, la variante G1770V fue identificada en cinco familias de origen Marroquí, lo que hizo pensar en un posible efecto fundador de esta mutación en dicha población. Para confirmar esta hipótesis se realizó un estudio de haplotipos en la región de *BRCA1*. El haplotipo de riesgo asociado a esta mutación se identificó en las cinco familias portadoras, confirmando que la variante G1770V es la primera mutación fundadora en población Marroquí.

En la búsqueda de nuevos genes de predisposición al SCMOH, también se plantearon dos aproximaciones. La primera, basada en evidencias previas aportadas por otros grupos, que indican que los genes pertenecientes a la vía de FA/*BRCA* o que interactúan con *BRCA1/2*, son buenos genes candidatos de predisposición al SCMOH. En base a esta observación, se secuenciaron dos nuevos genes de la vía de FA, *ERCC4* y

SLX4. Este estudio identificó varias variantes *missense* raras nuevas en cada gen. Algunas de ellas fueron clasificadas como potencialmente deletéreas por varios programas de predicción *in silico*. No obstante, la validación funcional de estas variantes es necesaria para elucidar su asociación con el SCMOH. Como segunda aproximación, se secuenció el exoma de varios pacientes pertenecientes a familias BRCA1 con un aparente patrón de herencia autosómico dominante. Como resultado, se identificaron 20 genes potencialmente candidatos por su función o por portar mutaciones que conllevan un CTP prematuro. Entre ellos, *POLE* destaca como firme candidato.

DISCUSIÓN

Los resultados de la presente tesis doctoral se han presentado como un compendio de trabajos publicados, o sometidos a publicación actualmente, más resultados no publicados a día de hoy, estos últimos resumidos en formato de artículo científico redactados en castellano. En cada uno de los artículos existe un apartado de discusión de los hallazgos de cada trabajo. Por tanto, en este apartado de discusión global de la tesis, se ha decidido realizar una presentación breve de los resultados obtenidos y se ha intentado profundizar en las limitaciones y beneficios del trabajo realizado así como su contextualización en el marco actual del CM hereditario. Al igual que en la sección de resultados, la discusión se ha dividido en apartados según los diferentes proyectos realizados. Finalmente se presenta una visión global del aporte de este trabajo en su conjunto.

Clasificación de Variantes de Significado Desconocido (VSD) en los genes BRCA1/2

El estudio mutacional de de los genes *BRCA1/2* detecta un alto porcentaje de variantes de significado biológico desconocido (VSD)^{53, 56, 134}. Concretamente, en la unidad de diagnóstico molecular del programa de cáncer hereditario del Instituto Catalán de Oncología esta proporción es de aproximadamente un 8-10%. La identificación de este tipo de variantes dificulta el asesoramiento genético de los portadores de las mismas y en algunos casos puede generar una sensación de ansiedad. Por este motivo, uno de los principales objetivos de nuestro laboratorio es el desarrollo e implementación de nuevas metodologías para la clasificación de estas variantes. En esta tesis doctoral se utilizaron dos aproximaciones distintas para lograr este objetivo. Por una parte, se analizó el efecto de una serie de VSD en el procesamiento del mRNA. Por otra, se realizaron ensayos funcionales a nivel de proteína para algunas de las variantes identificadas en nuestra población.

Primera Aproximación: Evaluación del efecto de VSD identificadas en nuestro centro en el procesamiento del mRNA de los genes BRCA1/2.

El análisis de VSD mediante ensayos de RNA ha demostrado ser una técnica robusta para la clasificación de variantes^{129, 215, 433, 434}. Además, un resultado alterado del *splicing* normal de estos genes, permite la clasificación casi directa de la variante estudiada como patogénica o probablemente patogénica, en base al esquema de clasificación de variantes propuesto por la *Agency for Research on Cancer (IARC)*¹⁸⁴ y por diferentes consorcios internacionales como el consorcio internacional ENIGMA^{129, 149, 215}. Por este motivo, el análisis de RNA es una de las técnicas comúnmente utilizadas en nuestro laboratorio para clasificar VSD.

Validación de la metodología de análisis utilizada en nuestro laboratorio

Uno de los objetivos del presente trabajo fue la validación de la metodología seguida al realizar este tipo de análisis, para así, asegurar la calidad y fiabilidad de nuestros resultados. Para ello, se participó en un proyecto que tenía como finalidad evaluar y poner en común las metodologías seguidas por 23 laboratorios en los ensayos de RNA para evaluar VSD en los genes *BRCA1/2*. Este trabajo estuvo dirigido por el consorcio *ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles)* englobado en su proyecto *Quality Control Project* diseñado por el *Splicing Working Group*. En este trabajo, cada grupo analizó las mismas 8 variantes siguiendo su propia metodología. Los resultados obtenidos para cada variante se compararon, de una forma cualitativa, con los resultados obtenidos en una serie de 11 controles. A partir de aquí, siguiendo unos criterios establecidos por ENIGMA, las variantes se clasificaron en una de las cinco clases establecidas en el esquema de clasificación de variantes propuestos por la *IARC*. A continuación, el consorcio ENIGMA, recopiló los resultados obtenidos por cada grupo y realizó un análisis comparativo de éstos. La metodología utilizada en nuestro laboratorio** fue capaz de identificar todos los transcritos aberrantes producidos por esta variantes, lo que indica que tal metodología es completamente válida para el análisis de VSD detectadas en *BRCA1/2* en pacientes del SCMOH²²⁴. Como consecuencia de este trabajo, se detectaron una serie de puntos críticos que deben ser considerados en el diseño de los ensayos de RNA. Se determinó que tanto la localización de los oligonucleótidos como el tiempo de elongación utilizados en la PCR de amplificación son los puntos más críticos para la detección los transcritos aberrantes y la interpretación de los resultados. En cuanto a la localización de los oligonucleótidos, se concluyó que el conocimiento previo de los transcritos constitutivos y la predicción de los transcritos aberrantes que se esperan son los factores más importantes. La utilización de cebadores que eviten la amplificación de transcritos constitutivos y amplifiquen específicamente los asociados a enfermedad, aumentará el poder de detección y de interpretación de los resultados. En este sentido, los esfuerzos realizados por algunos grupos, en el contexto de ENIGMA, para la identificación de todos los transcritos constitutivos de *BRCA1/2* presentes en sangre se espera que mejore el poder interpretativo de estos ensayos.

** La metodología utilizada en nuestro laboratorio consiste en la extracción de RNA a partir de linfocitos del portador, previamente tratados con puromicina (con la finalidad de evitar el fenómeno del *Non-Sense Mediated Decay*), generación de cDNA mediante PCR de retro-transcripción, amplificación por PCR de la región afectada por la variante y posterior detección de transcritos alternativos mediante la combinación de electroforesis en gel de agarosa y secuenciación del producto de PCR. El resultado obtenido para la variante se compara con, al menos, seis controles sanos (Menendez et. al, 2012)

Estudio de RNA de 28 Variantes identificadas en pacientes con sospecha de sufrir SCMOH.

En un trabajo previo realizado por nuestro equipo, se estudiaron 26 VSD en *BRCA1/2* identificadas en pacientes ICO y se identificó que ocho variantes producían un *splicing* aberrante del gen implicado¹⁷⁶. En la presente tesis doctoral se planteó una continuación del citado estudio con el análisis de 28 VSD más identificadas en el ICO. Ninguna de las variantes seleccionadas, excepto la variante patogénica c.5123C>A, había sido estudiada previamente a nivel de RNA. En estudios previos, la variante c.5123C>A había mostrado diferentes patrones de *splicing* dependiendo de la fuente de RNA utilizada, mostrando un patrón aberrante de *splicing* cuando el RNA era extraído a partir del linfocitos frescos y no mostrando ningún efecto cuando se extraía a partir de cultivos de linfocitos tratados con puromicina. Se decidió incluir esta variante en el estudio para aportar nuevas evidencias sobre su efecto y a modo de control de calidad de nuestra metodología.

Dieciséis de las variantes seleccionadas estaban localizadas en *BRCA1* y las 12 restantes en *BRCA2*. Diez de ellas eran variantes intrónicas, 4 grandes reordenamientos genómicos (LGR), detectados por análisis de *Multiplex Ligation-dependent Probe Amplification* (MLPA) y que afectaban a un solo exón, y las 14 restantes eran variantes *missense*. Las predicciones *in silico* de las variantes puntuales seleccionadas, a excepción de dos variantes intrónicas identificada posteriormente al primer estudio, no indicaron un efecto claro en el *splicing* para ninguna de ellas. Las dos variantes que mostraron una reducción drástica del reconocimiento de los sitios *splice* consenso fueron las variantes *BRCA1* c.4987-20_4987-11del y *BRCA2* c.681+5G>C. El resultado obtenido en las predicciones *in silico* está en concordancia con que las VSD con mayor probabilidad de afectar al *splicing* habían sido seleccionadas para el primer estudio. No obstante, los programas de predicción *in silico* han mostrado ciertas limitaciones para identificar sitios de *splicing* crípticos o *de novo* o alteraciones en las secuencias reguladoras de *splicing*^{129, 176}. Por ello, el análisis de RNA es necesario como prueba inequívoca de un defecto en el *splicing* de toda variante puntual, o pequeña indel, fuera de los sitios invariables de *splicing*. Por otra parte, aunque el efecto de los LGR sobre el *splicing* es probable, es necesario analizar tal efecto por varios motivos. En primer lugar, en el caso de deleciones, el hecho de que hayan sido detectados mediante MLPA implica que no se pueda asegurar si el exón afectado está completamente o parcialmente delecionado. Por tanto, no se puede predecir si el transcrito producido generará una proteína *in frame* o *frameshift* con codón de terminación prematuro (CTP). Además, suponiendo que se pierde completamente el exón afectado y, a priori, se espera que se produzca un transcrito con CTP no se puede descartar la posibilidad de un segundo evento durante el *splicing* que dé lugar a una transcrito *in frame*. En tal caso, si esta deleción *in frame* no está afectando a

un dominio funcional de conocida importancia, no se podrá clasificar como patogénica. En caso de que lo afecte, si no afecta a aminoácidos de conocida importancia clínica, es recomendable realizar ensayos funcionales a nivel de proteína para determinar su patogenicidad^{215, 228}. De los LGR estudiados, únicamente uno en *BRCA1*, c.(5193+1_5194-1)_(5278+1_5279-1)del, está localizado en uno de estos dominios, el dominio BRCT, afectando a un aminoácido con relevancia clínica conocida (p.G1738). Por otra parte, en el caso de las duplicaciones, el análisis de MLPA no detecta si dicha duplicación ha ocurrido en tándem o el exón duplicado se encuentra en alguna otra región del genoma. En este caso el estudio de RNA es necesario para determinar si realmente está en tándem y produce un transcrito con CTP.

Siete de las 28 variantes analizadas mostraron un patrón aberrante de *splicing*. La variante c.5123C>A produjo un *skipping* parcial del exón 18 cuando se utilizó RNA recién extraído de linfocitos congelados y mostró un patrón de *splicing* normal cuando se extrajo el RNA de linfocitos en cultivo tratados con puomicina. Por tanto, los resultados obtenidos para c.5123C>A concuerdan con los estudios previamente reportados para esta variante, validando así la metodología utilizada en nuestro estudio. Las otras seis variantes que produjeron un *splicing* aberrante fueron, los cuatro LRG, y las variantes intrónicas *BRCA1* c.4987-20_4987-11del y *BRCA2* c.681+5G>C. Todas ellas produjeron únicamente la pérdida del exón afectado, excepto el LGR *BRCA1* c.(5193+1_5194-1)_(5278+1_5279-1)del, que afectaba el exón 20 y produjo el *skipping* de los exones 18, 19, 20. En todos los casos el transcrito aberrante resultó en una proteína *frameshift* con CTP. Dado que la expresión de los transcritos aberrantes producidos por estas seis variantes no pudo ser cuantificada, se decidió, siendo conservadores, clasificarlas como probablemente patogénicas. Las 21 variantes restantes no mostraron un patrón de *splicing* aberrante en base al diseño utilizado, por lo que es muy poco probable que afecten al *splicing*. Estos resultados concuerdan con las predicciones *in silico* para estas variantes.

La clasificación de estas seis variantes como probablemente patogénicas tiene una serie de implicaciones clínicas. Estas variantes podrán ser utilizadas como valor predictivos en las familias afectadas, de manera que los portadores se podrán beneficiar de las estrategias de seguimiento y prevención establecidas, mientras que los no portadores verán su estrés aliviado. Por otra parte, este estudio proporciona nuevas evidencias sobre el resto de variantes analizadas. Tales evidencias pueden ayudar a su clasificación mediante los modelos multifactoriales actuales, lo cuales integran los resultados de *splicing*. Además, pueden ser utilizadas para mejorar el poder predictivo de los programas *in silico* en el reconocimiento de lugares exónicos activadores e inhibidores del *splicing* funcionales. En conclusión, nuestro estudio demuestra la utilidad de realizar ensayos de RNA para aquellas VSD identificadas en pacientes de alto riesgo,

especialmente cuando los programas *in silico* sugieren un defecto en el reconocimiento de los sitios de *splice* consensos.

Segunda Aproximación: Estudio funcional de diferentes VSD en las proteínas BRCA1/2

En la actualidad, han sido desarrollados diferentes ensayos funcionales para analizar variantes en las proteínas *BRCA1/2* (ref ^{130, 233}). En esta tesis doctoral utilizaron dos de ellos para clasificar en una serie de VSD en *BRCA1/2* identificadas en pacientes del ICO con sospecha de SCMOH. El ensayo utilizado para analizar variantes en *BRCA1* fue el ensayo de actividad transcripcional y el utilizado para *BRCA2* fue el ensayo funcional basado en células madre de ratón embrionarias. La selección de estos estudios estuvo basado en que su valor predictivo había sido ampliamente contrastado^{234, 275}.

Hasta la actualidad han sido desarrollados diferentes ensayos funcionales para analizar variantes en las proteínas *BRCA1/2* (ref ^{130, 233}). En esta tesis doctoral se utilizaron dos de ellos para clasificar una serie de VSD en *BRCA1/2* identificadas en pacientes del ICO con sospecha de SCMOH. El ensayo utilizado para analizar variantes en *BRCA1* fue el ensayo de actividad transcripcional y el utilizado para *BRCA2* fue el ensayo funcional basado en células madre de ratón embrionarias. La selección de estos estudios estuvo basado en que su valor predictivo había sido ampliamente contrastado^{234, 275}.

Ensayo de Actividad Transcripcional

El ensayo de actividad transcripcional (TA), fue diseñado para analizar variantes localizadas en la región C-terminal de *BRCA1*, en la que está situado su dominio BRCT. El dominio BRCT (aa 1650-1863) o dominio de activación transcripcional, es un dominio común entre las proteínas involucradas en la reparación del DNA^{95, 96}. Se ha demostrado que este dominio es requerido para la mayoría de funciones de *BRCA1*, muchas de ellas, incluida la actividad transcripcional, involucradas en su función como supresor tumoral^{97, 273, 274, 435}. Esta observación se ha visto reflejada en el hecho de que una gran parte de las mutaciones asociados al SCMOH afectan a este dominio, comprometiendo su función^{97, 282, 436}. Por ello, evaluar el efecto que una determinada variante puede tener sobre la función del dominio BRCT de *BRCA1*, puede aportar una fuerte evidencia sobre su patogenicidad.

El ensayo TA consiste en un ensayo de complementación fundamentado en la capacidad que tiene la región C-terminal de *BRCA1* (1.396-1.863) cuando está fusionada a un dominio de unión al DNA GAL4, para activar la expresión de un gen reportero de *Luciferasa*, cuya expresión se encuentra regulada por el control de 5 sitios de unión GAL4.

Para ello, un plásmido conteniendo BRCA1:GAL4 y otro conteniendo el gen *Luciferasa*, son cotransfectados en células mamarias (HEK293T)^{††}. La capacidad de activación transcripcional de la región C-terminal de BRCA1, se puede medir mediante la cuantificación de la expresión del gen *Luciferasa*. Un plásmido *Luciferasa Renilla*, bajo el control de un promotor constitutivo, se utiliza como control de transfección y viabilidad..

En nuestro estudio, se combinó el ensayo TA con un análisis estructural para analizar siete VSD identificadas en pacientes con SMCOH del ICO. Estas variantes estaban localizadas en la región C-terminal de BRCA1, cuatro de ellas en la región colindante al dominio BRCT (Q1409L, S1473P, E1586G, R1589H) y las tres restantes dentro del citado dominio (Y1703S, W1718L y G1770V). El análisis estructural de las variantes Y1703S, W1718L y G1770V, predijo que estas tres variantes podían alterar bien la estructura del dominio BRCT o su unión a motivos fosfopeptídicos (o ambas) causando así la pérdida de funcionalidad de dicho dominio. Estas predicciones fueron confirmadas por el resultado obtenido en el ensayo TA, en el cual se observó la reducción drástica de la actividad transcripcional para estas tres mutaciones. Por otra parte, las cuatro variantes restantes, localizadas en la región precedente al dominio BRCT, no pudieron ser analizadas estructuralmente, debido a que están localizadas en una región sin estructura definida. No obstante, el ensayo TA permite evaluar el efecto de variantes situadas en esta región, concretamente en la región comprendida entre los aminoácidos 1.396-1.863. Ninguna de estas variantes evidenció reducción de la actividad transcripcional de BRCA1 en este estudio.

Nuestros resultados para las variantes Y1703S, W1718L y G1770V indican que la actividad transcripcional del dominio BRCT sirve como medida indirecta de la estabilidad estructural de dicho dominio, un hecho que ya había quedado evidenciado en estudios previos que se observó una gran correlación entre los resultados de este ensayo y el ensayo de sensibilidad a proteasas^{231, 274, 275, 437}. Con ello el resultado obtenido para las otras cuatro variantes, puede ser debido a que están localizadas en regiones desordenadas y por ello, mucho menos sensibles a cambios estructurales.

Previamente a nuestro estudio, había sido demostrada la fuerte correlación existente entre los resultados de este ensayo y la patogenicidad de las variantes analizadas. Esta observación fue ampliamente validada mediante un set de variantes previamente clasificadas por modelos genéticos^{274, 275, 437}. Esta validación permitió establecer un valor límite de expresión de *Luciferasa* para determinar la patogenicidad de

^{††} Una descripción detallada de la metodología y los aspectos técnicos de este ensayo se puede encontrar en la Introducción, Apartado10

las variantes analizadas²⁷⁵. Las variantes patogénicas mostraban una expresión de *Luciferasa* menor del 45%, mientras que las neutras mayor de 50%, con respecto al alelo *wild-type*. Este resultado sugiere que no existen variantes de riesgo moderado que afecten a este dominio. En base a esta observación, y a los resultados obtenidos en nuestro estudio, las variantes Y1703S, W1718L y G1770V fueron clasificadas como probablemente patogénicas, mientras que las variantes Q1409L, S1473P, E1586G, R1589H siguen siendo consideradas VSD hasta que futuros estudios genéticos, epidemiológicos o funcionales aporten nuevas evidencias de su patogenicidad.

La clasificación de estas variantes como probablemente patogénicas es de interés para el manejo clínico de los pacientes de estas familias aunque la situación ideal sería el poder clasificarlas como patogénicas añadiendo evidencias genéticas, histopatológicas y/o epidemiológicas que acompañen al resultado funcional. De ahí, la importancia de incluir los resultados funcionales en los modelos multifactoriales de clasificación de variantes. Con ello, se le dará mayor relevancia clínica a dichos resultados a la vez que se mejorará el poder predictivo de dichos modelos. En este sentido, posteriormente a nuestro estudio, se diseñó un modelo estadístico que integraba los resultados del ensayo TA con evidencias aportadas por la historia familiar y personal, cosegregación y datos histopatológicos¹³². Este hecho, junto con el gran valor predictivo de este ensayo, así como la baja complejidad técnica y el tiempo relativamente corto requerido para testar una VSD mediante él, hacen de éste uno de los ensayos más robustos para la translación clínica de los resultados obtenidos.

Como inciso, destacar que para varias de las variantes analizadas en este estudio, las predicciones funcionales obtenidas no correspondieron con los resultados experimentales obtenidos. Además, en muchos casos, las predicciones para la misma variante divergían entre diferentes programas. Los programas Polyphen2 (1 falso positivo^{††}) y *Align-GVGD* (1 falso positivo y un falso negativo), mostraron ser los más consistentes con los resultados funcionales. El falso negativo obtenido en *Align-GVGD* (variante G1770V) es de especial interés, ya que las predicciones de éste programa son incluidas actualmente como una de las evidencias de peso en los modelos multifactoriales de clasificación de variantes actuales. Este programa se considera correctamente validado para los dominios conservados RING y BRCT, por ello, en los citados modelos solo se consideran las predicciones para las variantes situadas en estos dominios (toda variante que esté fuera de ellos es considerada como CO (o sin evidencias de patogenicidad))¹³³. En este sentido, nuestros resultados evidencian la necesidad de una mayor validación de

^{††} Falso positivo; variante neutra predicha como patogénica; Falso negativo, variante deletérea predicha como neutra.

estos programas *in silico*, especialmente en el caso de *Align-GVGD*, incluso en los dominios que se considera validado.

G1770V, nueva variante fundadora de origen marroquí

En el momento del análisis de la variante G1770V mediante el ensayo TA, esta variante había sido detectada en dos familias de origen marroquí, ambas pertenecientes a la serie del ICO. Más tarde, dicha variante fue identificada en una tercera familia de alto riesgo, también de origen marroquí, identificada en el ICO. Posteriormente, y como consecuencia de la publicación de nuestros resultados del estudio TA, se puso en contacto con nosotros la Dra. Nora Martinussen (*Department of Medical Genetics, Oslo University Hospital*) para notificarnos la identificación de esta variante en dos familias de origen marroquí estudiadas en su centro en Noruega. La información clínica recogida por los consejeros genéticos sobre estas cinco familias indicó que no existía relación aparente entre ellas. Además, dicha información indicaba que el fenotipo de esta variante correspondía con el de portadores de variantes de alto riesgo en el gen *BRCA1* (Tumores tripla negativos, grado 3 y edad de diagnóstico <40 años en el 60% de los casos). Las citadas familias pertenecían todas a ciudades localizadas en la costa norte de Marruecos. Concretamente, tres de ellas pertenecían a Nador, una a Tánger y la otra a Oujda. La identificación de esta variante en cinco familias independiente de origen marroquí, planteaba la posibilidad de un efecto fundador asociado a esta variante. Para confirmar esta hipótesis se analizaron dos marcadores intragénicos en *BRCA1* (*D17S855*, *D17S1323*) y 5 en la región flanqueante de este gen (*D17S486*, *D17S1793*, *D17S1325*, *D17S951*, *D17S1183*). En una de las familias analizadas (Familia B), la disponibilidad de DNA de dos portadores y de dos no portadores, permitió definir inequívocamente el haplotipo *BRCA1* de riesgo asociado a la mutación G1770V. Los portadores de la variante G1770V en el resto de familias estudiadas compartían los alelos asociados a esta variante en la familia B, en la región que se comprendida entre los marcadores *D17S1793* y *D17S1171* (1,54 Mb). Aunque, la fase y la asociación de estos alelos con G1770V no pudo ser demostrada en estas familias, los resultados obtenidos sugieren que G1770V es una mutación fundadora de origen marroquí, confirmando nuestra hipótesis. El hecho de que esta mutación no haya sido identificada antes en Marruecos puede ser debido a la escasa información sobre estudios mutacionales en *BRCA1* en esta región población (menos de 500 familias han sido estudiadas en todo el norte de África y sólo 125 en Marruecos)⁴³⁸. Además, el hecho de que esta mutación no se haya encontrado en cribados mutacionales de *BRCA1* en otros países del Norte de África, sugiere que es una mutación específica de población marroquí. Por otra parte, el que haya sido identificada en varias familias marroquíes asentadas en Europa, plantea la hipótesis de que esta variante se haya diseminado por la población marroquí asentada en este continente, consecuencia de la larga tradición migratoria entre este país y varios países Europeos⁴³⁹. Un estudio

exhaustivo en una gran serie de pacientes con sospecha de SCMOH de población de origen marroquí sería necesario para determinar la prevalencia exacta de esta mutación en esta población, así como su distribución geográfica.

La identificación de este tipo de mutaciones puede facilitar el proceso de cribado genético en las poblaciones afectadas, ya que se puede proceder al cribado selectivo de las mismas como primer paso en el algoritmo de pruebas genéticas. En países con presupuesto limitado para las pruebas genéticas, como podría ser Marruecos, este hecho podría tener repercusiones económicas muy favorables.

Por otra parte, este estudio aporta evidencias adicionales sobre la patogenicidad de esta variante, validando así los resultados previos obtenidos en el estudio funcional TA.

Ensayo funcional basado en células de ratón embrionarias para evaluar VSD en BRCA2

El ensayo de basado en células de ratón embrionarias (RE) está fundamentado en la observación de que estas células no son viables si no tienen un *brca2* funcional y que *BRCA2* funcional transfectado en células *brca2*^{ko/ko} es capaz de rescatarlas²⁴⁴. Por tanto, para testar VSD mediante este ensayo, en primer lugar se han de generar las variantes sujetas a estudio utilizando un BAC que contiene *BRCA2* humano completo y se transfectan en células-RE que contienen un alelo *brca2* funcional condicional y otro *knock out*²³⁴. De esta forma tras la transfección con el BAC portador de la variante, se deleciona el alelo *brca2* condicional y se analiza la capacidad de *BRCA2* transfectado de hacer viable estas células y así evaluar el impacto funcional de dicha variante⁵⁵. Las variantes que no dan lugar a células *brca2*^{ko/ko} viables, son consideradas patogénicas. Las variantes que dan lugar a células *brca2*^{ko/ko} viables pueden ser testadas para ver su efecto en diferentes funciones de conocidas de *BRCA2* (sensibilidad a agentes genotóxicos, formación del *RAD51 foci*, recombinación homóloga y estabilidad cromosómica) para determinar si producen un efecto hipomórfico o son neutras. En el presente estudio, por razones técnicas y temporales, solo se pudo analizar la sensibilidad a agentes genotóxicos.

La selección de este ensayo se basó en las grandes ventajas de que aporta con respecto a otros ensayos funcionales existentes para *BRCA2*. Por una parte, ha sido ampliamente validado en una serie de variantes neutras y patogénicas conocidas, y se ha utilizado de forma satisfactoria para la clasificación de una batería de variantes localizadas en la zona de interacción *BRCA2*-*PALB2* así como para el estudio de variantes identificadas en pacientes con anemia de Fanconi^{227, 228, 234}. Una de las ventajas de este

⁵⁵ Una extensa descripción de la metodología y los matices técnicos de este ensayo ha sido redactada en el apartado de resultados correspondiente.

ensayo es el uso de células embrionarias, ya que están libres de la inestabilidad genómica asociada a la mayoría de líneas celulares cancerígenas utilizadas en el resto de ensayos funcionales. Por otra parte, el uso de un BAC permite testar las variantes en el contexto de expresión de la proteína completa y evita la variabilidad de expresión asociada al uso de plásmidos con fracciones de la proteína. Además, el BAC permite analizar variantes en regiones intrónicas y reguladoras, lo que supone una gran ventaja en un contexto en el que se espera que la implementación de las nuevas técnicas de secuenciación masiva detecten un alto número de variantes de este tipo. Por otra parte, este ensayo permite el análisis de varias funciones de BRCA2. Estas ventajas permiten obtener resultados más robustos que los ensayos de complementación basados en fragmentos de cDNA y que testan un único dominio/función de la proteína.

En nuestro estudio este ensayo se utilizó para testar 5 variantes en BRCA2 (T200I, K331I, K1434I, V1532F, R2645G) identificadas en familias con sospecha de sufrir SCMOH. Fueron seleccionadas aquellas variantes en las que se observaba una mayor agregación familiar. Cuatro de las variantes (T200I, K1434I, V1532F, R2645G) fueron capaces de rescatar las células *brca2*^{ko/ko}, aunque, a excepción de K1434I, el número de colonias viables obtenidas fue drásticamente menor que las obtenidas en el control con BRCA2 *wild-type*, sugiriendo un posible defecto funcional de estas variantes, aunque no completo. Para la variante K331I problemas técnicos impidieron la cuantificación objetivo aunque se obtuvieron colonias viables para realizar ensayos funcionales. En el ensayo de sensibilidad a agentes genotóxicos todas las variantes mostraron sensibilidad. Las variantes T200I, V1532F y R2645G mostraron una sensibilidad similar al control negativo utilizado (G25R, una variante clasificada como probablemente patogénica) mientras que K331I y K1434I mostraron un fenotipo más leve. Adicionalmente, cuando fue posible, se realizó el análisis de cosegregación de estas variantes siendo sólo informativo para la variante K1434I, que mostró cosegregación con la enfermedad en tres hermanas afectas. Por último destacar que para la variante R2645G fue posible realizar un estudio de modelaje estructural. Los datos estructurales sugirieron que esta variante podría afectar la función de la proteína, acorde con los resultados obtenidos en el presente ensayo.

En base a nuestros resultados, y a observaciones previas aportadas para otras variantes analizadas mediante este estudio, como la G25R, las variantes T200I, K1434I, V1532F, R2645G podrían clasificarse como de riesgo moderado y la variante K331I se podría considerar de riesgo bajo. Hoy en día, la detección de variantes hipomórficas de riesgo moderado o bajo conlleva un problema a nivel clínico ya que es difícil de trasladar estos resultados al manejo clínico del paciente. Es necesaria la realización de estudios complementarios para estimar el riesgo asociado a cada variante. Es importante que los modelos multifactoriales se modifiquen para contemplar resultados intermedios,

asociados a riesgos moderados ya que los modelos actuales parten de la base de que una variante es patogénica o es neutra, y no consideran fenotipos intermedios.

Es interesante remarcar, que la variante K1434I había sido previamente clasificada como probablemente neutra¹⁵¹, en base a resultados de un modelo multifactorial que utilizaba información clínica y datos histopatológicos. No obstante, esta clasificación fue poco consistente por varios motivos. Por una parte, el uso de datos histopatológico de tumores para clasificar variantes en *BRCA2* no es un método robusto debido a la gran heterogeneidad que presentan los tumores asociados a mutaciones en este gen. Además, la información clínica disponible sobre esta variante era muy limitada. Por último, esta variante fue clasificada como clase C15 por *Align-GVGD*, pero fue considerada C0 por estar fuera de los dominios de conocida importancia de *BRCA2*. En este sentido, las predicciones de *Align-GVGD* estuvieron en concordancia con los experimentales para las variantes localizadas en el dominio de unión al DNA, mientras que clasificó las variantes K313I y V1532F como C0 cuando realmente mostraron un efecto funcional, lo que remarca de nuevo la necesidad de una mayor validación de este programa.

Aunque, como se ha explicado anteriormente, este ensayo aporta una serie de ventajas respecto a otros, y proporciona resultados muy fiables, requiere una gran demanda técnica y tiene una serie de limitaciones económicas. Por ejemplo, el mantenimiento de las células embrionarias, requiere de cuidado diario y un gran gasto económico en material fungible. El empleo de *Southern Blots* para la identificación de colonias, requiere del uso de reactivos radioactivos. Además, se requiere de casi dos meses para analizar una variante y, en base a nuestra experiencia, no es aconsejable testar más de 5 variantes simultáneamente. Por todo lo expuesto es inviable su uso en laboratorios de diagnóstico rutinario y es por eso que se están desarrollando métodos de cribado más coste-efectivos, como, por ejemplo, el desarrollado por Bouwman et al, 2013 para *BRCA1* que permite testar variantes de una forma rápida y robusta. Este ensayo, basado en la recombinación dirigida de *BRCA1* completo, transfectado mediante un pcDNA3, en células embrionarias de ratón, sortea las limitaciones temporales y técnicas asociadas a la generación y análisis de variantes haciendo uso de BAC, mientras que asegura unos niveles de expresión de *BRCA1* igual a éstos⁴⁴⁰. No obstante, el ensayo basado en células-RE es, a día de hoy, una de las pocas opciones para testar variantes situadas en regiones intrónicas o reguladoras. Sin embargo, se espera que el empleo de las nuevas tecnologías de ingeniería genética, como la basada en el sistema *CRISP/Cas9* (ref ^{236, 441}), que permiten generar las variantes endógenamente, mejorarán las características de este tipo de ensayos obteniéndose resultados más representativos y acortando tiempos y gastos, mejorando así su uso como herramienta de clasificación de VSD en estos genes.

Búsqueda de nuevos genes relacionados con el SCMOH

El segundo gran objetivo de esta tesis doctoral se centró en la búsqueda de nuevos genes de predisposición al SMOH. Este objetivo se planteó como consecuencia de la clara necesidad de ahondar en el conocimiento de las causas genéticas de SMOH. Más del 50% de las familias de alto riesgo siguen sin causa genética conocida, son las denominadas familias BRCAX^{***57}. En este sentido, se han desarrollado varias estrategias para la búsqueda de nuevo genes de predisposición a esta enfermedad^{†††}. En nuestro estudio, se plantearon dos aproximaciones independientes, la primera basada en el conocimiento previo de la susceptibilidad al CM en familias BRCAX y la segunda en la implementación de las nuevas técnicas de secuenciación masiva en la búsqueda de nuevos de genes de susceptibilidad.

Primera Aproximación: Análisis mutacional de nuevos genes identificados en la vía de Anemia de Fanconi/BRCA en familias BRCAX

La primera aproximación se fundamentó en el hecho de que todos los genes de predisposición al CM identificados tras la era de *BRCA1/2*, exceptuando aquellos que predisponen a otros síndromes raros con alta tasa de CM (*TP53*, *PTEN*, *SKT11*, *CDH1*), forman parte de una vía de reparación de DNA común a *BRCA1/2*, la vía de BRCA/FA (*Introducción Apartado 4*), o son interactores de éstos^{85, 111, 297-299, 358}. Tales evidencias sugieren que los nuevos genes identificados en esta vía o en procesos comunes con los genes *BRCA1/2*, sean buenos candidatos para estar involucrados en la predisposición al CM. Durante el desarrollo de la presente tesis doctoral se habían identificado dos nuevos genes pertenecientes a la vía FA/BRCA, *ERCC4* y *SLX4* (ref⁴⁴²⁻⁴⁴⁴). Estos dos genes interactúan entre ellos, siendo *SLX4* regulador de la actividad de *ERCC4* (ref⁴⁴⁵). Siguiendo la hipótesis planteada se decidió evaluar la posible relación de estos genes en el SCMOH, *SLX4* dentro de un proyecto propio y *ERCC4* formando parte de un proyecto colaborativo^{446, 447}.

El gen *SLX4 (FANCP)*, codifica para una proteína que funciona como de andamio para otras proteínas reparadoras que está involucrada en la reactivación de las horquillas de replicación después de la reparación del daño al DNA^{443, 448, 449}. Mutaciones bilalélicas en este gen fueron identificadas como un nuevo subtipo de Anemia de Fanconi, denominado tipo P^{442, 443}. En nuestro estudio, se secuenció la región codificante completa más las regiones intrónicas colindantes en 94 familias BRCAX. Se identificaron 49

*** Familias BRCAX entendidas como aquellas familias negativas para el test genético de *BRCA1/2* y que cumplen los criterios establecidos por la Oncoguía de Cataluña para ser consideradas de alto riesgo (*Introducción Apartado 3*)

††† Una explicación detallada de la susceptibilidad al CM y la situación actual de los casos BRCAX se puede encontrar en el *Apartado 10* de la introducción de esta tesis.

variantes, 21 nuevas y 28 reportadas en dbSNP. Estas variantes se dividían en 21 variantes del tipo *missense*, nueve de ellas reportadas por primer vez, 11 sinónimas y 17 intrónicas⁴⁴⁶. El efecto de las variante intrónicas en el *splicing* fue evaluado utilizando varios programas de predicción *in silico*. Para ninguna de ellas se predijo un efecto sobre el *splicing*. Las variantes *missense* nuevas y una previamente reportada en dbSNP, que no disponía de datos de frecuencia poblacional, fueron genotipadas en un set de controles formado por DNA de 138 mujeres y 146 hombres. Tres de ellas fueron identificadas en el grupo control. De las 7 restantes, 5 fueron consideradas como deletéreas por el programa de predicción CONDEL⁴⁵⁰, que integra los *scores* de los programas SIFT¹⁶⁷ y Polyphen 2¹⁶⁹. El efecto funcional de estas variantes no pudo ser analizado. Por ello, ninguna de las mutaciones identificadas pudo ser asociada con una pérdida clara de la función del gen y, por tanto, nuestro estudio no pudo determinar la posible susceptibilidad al CM ligada a *SLX4*. Si bien, no se puede descartar que algunas de las mutaciones *missense* identificadas en este estudio tengan un efecto hipomórfico suficiente para causar la inestabilidad genómica, necesaria para el inicio de la tumorigénesis, como consecuencia de horquillas de replicación estancadas^{451, 452}. La variante W783G, una de las variantes *missense* detectadas, afecta a un aminoácido muy conservado y la sustitución de W a G es potencialmente desfavorable⁴⁵³. Otra variante detectada, P1470L, está situada al lado de un sitio de fosforilación mitótico S1469, pudiendo afectar su fosforilación (ref⁴⁵⁴). Sin embargo, para determinar su efecto es necesario analizarlas a nivel funcional. *SLX4* es una proteína multidominio que interacciona con múltiples exonucleasas que ejercen diferentes funciones en la célula. Hoy en día, se desconoce con exactitud cuáles de estas interacciones o funciones de *SLX4* están involucradas en la supresión tumoral. Sin embargo, es conocido que células deficientes en *SLX4* expresan topoisomerasa I y son hipersensibles a agentes genotóxicos que causan entrecruzamientos de doble cadena (ICL) así como a inhibidores de PARP^{442, 443, 455}. Estudios posteriores al nuestro han evaluado funcionalmente el efecto de varias *missense* a estos tres niveles. Ninguna de estas variantes ha mostrado sensibilidad a estas drogas o han expresado topoisomerasa I^{455, 456}. Varias de estas variantes afectaban a sitios potencialmente críticos, como la región de interacción con *ERCC4* esencial para la reactivación de las horquillas de replicación, o el dominio BTB, en el que fueron identificadas variantes en 4 familias de Anemia de Fanconi asociadas este gen^{442, 443}. Varios estudios han analizado *SLX4* en familias BRCA, sumando un total de 2625 familias. Estos estudios han detectado numerosas variantes *missense* raras. Sin embargo, únicamente se identificaron 3 variantes de pérdida de función de la proteína asociadas a la generación de un CTP (c.2496G>A, c.2013+2T>A, c.4549G>T), sugiriendo que aunque *SLX4* sea un gen de predisposición al CM, la prevalencia de mutaciones de este tipo en los casos BRCA será muy baja (0,038%)^{446, 448, 455-460}. El análisis conjunto de estos datos sugiere que mutaciones patogénicas en *SLX4* podrían dar un incremento del riesgo al CM de hasta el

doble que el estimado para la población general⁴⁵⁹. No obstante, recientemente, Kurian et al. incluyeron este gen en un panel constituido por 42 genes de predisposición al CM y testaron 140 pacientes BRCA. Dos nuevas mutaciones de pérdida de función en *SLX4* fueron identificadas (c.5233_5234delGC, c.5229insG), mostrando una prevalencia del 1,14% en este estudio. Es interesante destacar que en este estudio el número de VSD identificadas en *SLX4* fue bastante alta en comparación con el resto de genes incluidos en el panel (21 VSD identificadas en *SLX4* respecto a una media de VSD por gen de 8). Los datos clínicos reportados para los tumores asociados a las cinco mutaciones claramente patogénicas identificadas en *SLX4* se recogen en la Tabla 1.

Tabla 1. Datos clínicos y histopatológicos de los tumores asociados a mutaciones patogénicas en *SLX4*

HGVS	Efecto Proteína	Dx	Tipo histológico	ER	PR	HER2	LOH	Historia Familiar	Ref
c.2496G>A	p.W823*	40	CLI	+	+	-	Alelo mut.	2 CM (Madre 60a, Abuela Paterna 30a); 1 CP (Padre 60a)	Shah et. al
c.2013+2T>A		52	CDI	-	-	-	No	1 CM (Madre, 85a), 1 CG (Padre, 85a), Hija portadora no afecta 42a	Bakker et. al.
c.4549G>T	p.E1517*	60	NA	NA	NA	NA	NR	4 CM (Dos hermanas, 48a y 49a; Prima, 48a; Hija Prima, 35), 1 CR (Padre 69a)	de Garibay et. al
c.5233_5234delGC	Ala1745Serfs*32	NR	NR	NR	NR	NR	NR	3 CM (NR, 40a), 1CR (NR, 75a), 1CCR (NR, 85)	Kurian et.al
c.5229insG	Gln1744Alafs*34	61	NR	+	+	-	NR	3CM (NR, 40a), CCR (NR, 39 y 73a)	Kurian et. al

Los pocos datos patológicos reportados sobre los tumores asociados a estas mutaciones no permiten establecer un perfil histopatológico definido asociado a ellas. Aún así, parece que están ligadas a un fenotipo HER2-. Además, la concurrencia de varios tipos de cáncer en las familias afectadas, sugiere que mutaciones en *SLX4* podrían predisponer a varios tipos de cáncer, similar a lo que ocurre con otros genes como *BRCA1/2* o *POLE*^{384, 385, 461, 462}. Por otra parte, no se identificó LOH en el tumor asociado a la mutación c.2013+2T>A y se identificó la pérdida del alelo mutante para la mutación c.2496G>A^{455, 456}. Lo que sugiere que la inactivación de *SLX4* en tumor puede ser debida a otros mecanismos como, por ejemplo, mecanismos epigenéticos, etc.⁴⁵⁵ No obstante, el estudio de LOH en tumor no es suficiente para determinar la predisposición al CM, ya que varios estudios de LOH realizados para mutaciones en otros genes de predisposición al CM han encontrado resultados similares a estos⁴⁶³⁻⁴⁶⁵.

Con todo esto, la contribución de *SLX4* al CM hereditario parece estar limitada a unos pocos casos. No obstante, como ha quedado demostrado en el estudio realizado por Kurian et. al 2015, *SLX4* es un buen candidato para ser incorporado en los paneles genéticos de susceptibilidad al SCMOH. De todas formas, puede que el papel de este gen en la susceptibilidad al SCMOH esté infraestimado, tanto en los estudios de genotipado descritos como en el estudio realizado por Kurian et. al, se han identificado un gran

número de VSD *missense*. Si bien, el estudio funcional de varias de ellas ha demostrado no tener efecto en la función de la proteína, será necesario testar un número mayor de estas variantes para determinar su contribución real al SCMOH. Por otra parte, los datos clínicos aportados hasta ahora sugieren que podría estar asociado a perfiles histológicos específicos (Her2 +) o podría contribuir a la predisposición a otros tipos de cáncer. La identificación de más mutaciones patogénicas correctamente anotadas, podría aportar más evidencias a favor de esta hipótesis. A partir de ahí, se podrían definir nuevos grupos de cribado acotados a estas características para determinar una posible susceptibilidad específica asociada a este gen.

El otro gen de la vía analizado fue *ERCC4* (*FANCG*; *XPF*) codifica para una endonucleasa que, mediante la formación de un heterodímero con *ERCC1*, está involucrada en la reparación del DNA por escisión de nucleótidos (NER) y en la reparación de los ICL⁴⁶⁶. Mutaciones bialélicas en este gen causan *Xeroderma Pigmentosum* y progeria^{467, 468}. Más recientemente, también se identificaron mutaciones bialélicas en *ERCC4* asociadas a un nuevo subtipo de FA, el subtipo Q⁴⁴⁴. Además, ciertos SNP comunes en este gen habían sido asociados con un aumento de riesgo al CM. Si bien, estos estudios son poco consistente y los SNP de riesgo identificados están principalmente ligados a ciertas poblaciones asiáticas o han sido identificados en poblaciones definidas por otros factores de riesgo⁴⁶⁹⁻⁴⁷⁴. En base a las evidencias previas, se decidió genotipar este gen en una serie de 1573 casos BRCA y 853 controles sanos sin historia familiar de cáncer, para determinar su posible asociación con el CM familiar⁴⁴⁷. Este estudio fue un estudio de colaboración estatal dirigido por el grupo de los doctores Ana Osorio y Javier Benítez (Grupo de Genética Humana de Centro de Investigación de Genética Oncológica, CNIO). Nuestro grupo contribuyó con la secuenciación de *ERCC4* en 96 casos BRCA, los mismos que habían sido utilizados para el gen *SLX4* (ref⁴⁴⁶). En el cribado realizado por nuestro grupo, se identificaron tres mutaciones *missense* nuevas (p.M12V, p.S786F) y 7 ya reportadas en dbSNP, entre ellas, únicamente 2 con MAF<0,001 o sin frecuencia poblacional descrita (p.S521R, p.I706Thr). Destacar que la mutación p.S786F, que había sido reportada pero sin datos de frecuencia poblacional, apareció 3 veces en nuestros casos. Estas cuatro variantes fueron consideradas patogénicas por el programa de predicción CONDEL⁴⁵⁰. Además, nuestro estudio detectó 6 variantes sinónimas, dos de ellas nuevas, y 11 intrónicas, cuatro de ellas nuevas. Ninguna de las variantes identificadas por nuestro grupo mostró tener un efecto sobre el *splicing*, en base a los resultados de los programas de predicción *in silico* utilizados. La variante p.S786F no fue identificada en controles y fue considerada para futuros estudios funcionales.

En total, en todo el estudio, este gen fue estudiado en 1573 casos y 854 controles de población española. Fueron identificadas 5 variantes únicas en casos (1 *frameshift* y 4 *missense*) y 4 variantes únicas en controles (2 *frameshift* y 2 *missense*). Las variantes

frameshift identificadas tanto en los casos como en controles, se consideraron como supuestamente deletéreas por crear un codón prematuro de terminación que afectaba a dominios funcionales de conocida importancia^{445, 475}. Las variantes *missense* se analizaron mediante CONDEL, considerándose supuestamente deletéreas dos de las identificadas en los casos (p.R150C y p.S786F, esta última identificada por nuestro grupo) y una identificada en controles (p.R689S), que había sido previamente identificada en casos de FA y causaba pérdida de la actividad NER y de reparación de ICL⁴⁴⁴. Estas tres variantes *missense* fueron analizadas funcionalmente para determinar su efecto en los citados procesos de reparación. Los resultados funcionales mostraron que R689S, R786S tenían un fenotipo asociado a la FA, alterando específicamente la reparación de ICL. La variante R150C no tuvo ningún impacto sobre la actividad NER y un efecto leve sobre la actividad reparadora de ICL. En vista a estos resultados se concluyó que la frecuencia de mutaciones patogénicas en heterocigosis en población española no difiere entre los pacientes de CM y CO familiares y controles sanos en este gen (0,3%, p-valor=0,251), lo que sugiere que mutaciones monoalélicas en *ERCC4* no están vinculados a la susceptibilidad al cáncer en población general. Por otra parte, la prevalencia de portadores de mutaciones monoalélicas en *ERCC4* (uno en 288) es similar a la reportada para *FANCA*. Sin embargo, hay aproximadamente 100 veces más de casos de FA-A que de FA-Q, lo que sugiere que más del 90% de mutaciones bialélicas en *ERCC4* son letales a nivel embrionario.

Un único estudio posterior al nuestro fue realizado en el mismo sentido. En dicho estudio genotiparon *ERCC4* en 75 casos de población Alemana y Bielorrusa⁴⁷⁶. En este estudio no fue identificada ninguna mutación de pérdida de función y se concluyó que *ERCC4*, en caso de que contribuya al CM familiar en población europea, tendrá un papel menor, en concordancia con nuestros resultados. En este estudio, también se estudio el riesgo asociado al polimorfismo p.R415G, un polimorfismo común que había sido asociado en estudio anteriores a un incremento de riesgo bajo en población europea^{469, 473}. En su estudio, p.R415G no mostró estar asociado a un incremento de riesgo. Aunque está presente en población española, este polimorfismo no fue identificado en nuestro estudio. Esta observación, refuerza la hipótesis de que la contribución de este gen al SCMOH, si la hay, es mínima, al menos en población Caucásica.

En líneas generales, aunque nuestros resultados no han sido concluyentes en la definición de *ERCC4* y *SLX4* como nuevos genes de predisposición al SCMOH, los estudios posteriores realizados para *SLX4* indican que la estrategia seguida en nuestro estudio para la identificación de nuevos genes de susceptibilidad es una estrategia válida. Si bien, nuestro resultados pueden ser consecuencia de la cada vez más plausible hipótesis de que un gran número de familias BRCA1/2 podrían estar asociadas a un modelo de herencia poligénico^{307, 477}, lo que puede dificultar la identificación de genes de predisposición utilizando estrategias de este tipo.

Segunda Aproximación: Búsqueda de nuevos genes de predisposición al SCMOH mediante secuenciación de exomas en familias BRCA.

La segunda aproximación utilizada para identificar nuevos genes de predisposición al SCMOH se fundamentó en la secuenciación de exomas completos de pacientes pertenecientes a familias BRCA. El proyecto se diseñó con la finalidad de identificar genes de predisposición dominantes y de alta penetrancia. Un esquema detallado de la metodología seguida para la selección de genes candidatos se puede ver en la Figura 2 del Apartado de Resultados correspondiente. En resumen, se seleccionaron 12 familias con los siguientes criterios.

- Alta agregación familiar, considerando 2-3 generaciones, mostrando un aparente patrón de herencia autosómico dominante.
- Disponibilidad de DNA de, al menos, 3 afectados.
- Casos de bilateralidad o disponibilidad de de DNA de afectados en ramas distantes de la familia.

Se secuenciaron los exomas de 3-5 pacientes de cada familia. Las variantes obtenidas fueron filtradas mediante un algoritmo bioinformático diseñado acorde a nuestra hipótesis. Se seleccionaron las variantes en heterocigosis comunes a todos los pacientes de una misma familia que se presumía que podían afectar la función de la proteína. La priorización de las variantes identificadas siguió los siguientes criterios:

1. Variantes raras (MAF<0,001%) presentes en todos los pacientes de una misma familia.

2. Variantes claramente patogénicas con los estándares de clasificación canónicos, es decir variantes del tipo: *STOP* (st), *splice site* $\pm 1,2$ (sp) o *frameshift* (fs), independientemente del gen afectado. Las variantes de este tipo seleccionadas se encontraban en los siguientes genes: *ARHGEF38* (sp), *KIR3DL3* (st), *MUTYH* (fs), *PIPOX* (fs), *POLE* (fs), *ST6GALNAC2* (sp).

3. Variantes del tipo *Missense* o *indels in frame* que afectaban a genes con funciones relacionadas con el desarrollo del cáncer (reparación del DNA, regulación ciclo celular, migración y diferenciación celular). Las variantes de este tipo seleccionadas se encontraban en los siguientes genes: *ACTR8*, *BEX2*, *DVL1*, *DVL2*, *FGFR2*, *MCM2*, *MCM7*, *PHF20L1*, *PRDM2*, *SEPT9*, *TET1* y *ZFH3*.

4. Missense o *indels in frame* identificadas en genes mutados en más de una familia, independientemente de la función del gen afectado. Los genes seleccionados atendiendo a este criterio fueron *GRM6* y *ZSWIM6*.

Actualmente, los 20 genes seleccionados están siendo validados en un conjunto de 1384 pacientes BRCA1 independientes. Entre los genes seleccionados, *POLE* destaca como firme candidato. La identificación de una mutación *frameshift* en este gen que conduce a la generación de un CTP, junto con su relación con la susceptibilidad a otros tipos de cáncer así lo indican. Otros genes candidatos, bien por su función o por portar una mutación truncante también son buenos candidatos para ser genes de predisposición al SCMOH. Sin embargo, la contribución de los mismos a este síndrome no puede ser determinada únicamente a partir de los datos obtenidos. Se espera que la validación SMIPS que se está realizando actualmente aporte mayores evidencias sobre la posible contribución de todos estos genes al SCMOH. No obstante, para determinar de forma inequívoca su relación con este síndrome, se requerirá de una validación funcional, especialmente en aquellos genes con funciones menos conocidas. Por otra parte, en algunas de las familias analizadas no se ha identificado ningún gen candidato con los criterios utilizados en este estudio, por lo que está previsto un reanálisis de los mismos teniendo en cuenta otras aproximaciones como la posibilidad de fenocopias, otros modelos de herencia mendeliana o análisis poligénicos (ver más detalle de este aspecto en el apartado de Discusión del Capítulo 4 de resultados).

Al igual que en nuestro estudio, estudios previos que han seguido estrategias similares a la nuestra, han podido identificar nuevos genes de predisposición al SCMOH (*BLM*, *FANCC*, *REQCL* o *XRCC2*)^{330-332, 335, 478-481}. Este hecho refleja las ventajas que aporta la secuenciación de exomas frente a otras técnicas como el análisis de ligamiento o de asociación que no han dado un rendimiento demasiado alto en el pasado. En el contexto actual, con la aparición de paneles de NGS que contienen un gran número de genes asociados a predisposición al cáncer, se podrá dilucidar el papel de alguno de estos en el desarrollo del SCMOH (como *POLE*, entre otros). El uso de estos paneles se está convirtiendo en una práctica habitual en el diagnóstico genético gracias al avance de las nuevas tecnologías de secuenciación y a la gran mejoría en los métodos bioinformáticos de análisis y filtraje de datos, que han hecho de ellos una herramienta técnicamente factible, al mismo tiempo que coste/efectiva para el diagnóstico clínico^{128, 322, 482}. Estudios recientes, como el de Walsh et. al, 2010, donde se evaluó un panel de 21 genes de susceptibilidad al CM, o el de Kurian et. al, donde se testó un panel de 42 genes demuestran firmemente las ventajas del uso de estas técnicas en el diagnóstico genético^{128, 320}. De hecho, algunos laboratorios, como *Myriad* con su panel *MyRisk* o *Ambry Genetics* con sus paneles asociados a diferentes tipos tumorales, ya los están

utilizando de forma rutinaria en el diagnóstico clínico. La inclusión de nuevos genes de predisposición en estos paneles ayudará a aumentar el conocimiento en este campo.

Al igual que en nuestro estudio, estudios previos que han seguido estrategias similares a la nuestra, han podido identificar nuevos genes de predisposición al SCMOH (*BLM*, *FANCC*, *REQL* o *XRCC2*)^{330-332, 335, 478-481}. Este hecho refleja las ventajas que aporta la secuenciación de exomas frente a otras técnicas como el análisis de ligamiento o de asociación que no han dado un rendimiento demasiado alto en el pasado. En el contexto actual, con la aparición de paneles de NGS que contienen un gran número de genes asociados a predisposición al cáncer, se podrá dilucidar el papel de alguno de estos en el desarrollo del SCMOH (como *POLE*, entre otros). El uso de estos paneles se está convirtiendo en una práctica habitual en el diagnóstico genético gracias al avance de las nuevas tecnologías de secuenciación y a la gran mejoría en los métodos bioinformáticos de análisis y filtraje de datos, que han hecho de ellos una herramienta técnicamente factible, al mismo tiempo que coste/efectiva para el diagnóstico clínico^{128, 322, 482}. Estudios recientes, como el de Walsh et. al, 2010, donde se evaluó un panel de 21 genes de susceptibilidad al CM, o el de Kurian et. al, donde se testó un panel de 42 genes demuestran firmemente las ventajas del uso de estas técnicas en el diagnóstico genético^{128, 320}. De hecho, algunos laboratorios, como *Myriad* con su panel *MyRisk* o *Ambry Genetics* con sus paneles asociados a diferentes tipos tumorales, ya los están utilizando de forma rutinaria en el diagnóstico clínico. La inclusión de nuevos genes de predisposición en estos paneles ayudará a aumentar el conocimiento en este campo.

Aunque es muy probable la existencia de nuevos genes de predisposición que sigan un modelo de herencia autosómico dominante, los resultados de los estudios de ligamiento realizados hasta ahora indican que éstos explicarán muy pocas familias por sí solos. En este sentido se espera que estudios futuros confirmen a algunos de los candidatos aquí presentados como nuevos genes de este grupo. No obstante, tanto nuestro estudio como estudios previos basados en la secuenciación de exomas, reflejan la gran heterogeneidad genética que parece estar asociada a las familias BRCA y no se puede descartar que algunos de los genes identificados en nuestro trabajo se ajusten a otros patrones de herencia no contempladas en nuestro estudio piloto^{310, 334, 481}. Esto implica que probablemente se deban considerar nuevas estrategias cuando se diseñen estudios para la identificación de nuevos genes. Por ejemplo, se deberán considerar modelos de herencia recesivos, o asociados a oncogenes, se deberán definir grupos de estudio específicos basado en perfiles de expresión o histopatológicos para aumentar el poder de identificación unificando el tipo de paciente/familia^{311, 330, 483-485, 486-488}. Otra alternativa, a más largo plazo, con el abaratamiento de los costes y la mayor capacidad de análisis e interpretación de datos, será la secuenciación de genomas enteros, que nos permitirá la identificación de variantes en zonas no investigadas en la secuenciación de

exomas. Algunas pacientes/familias se explicarán siguiendo modelos poligénicos, en este escenario, la secuenciación de un gran número de exomas y la creación de bases de datos con información de grandes cohortes será crucial. Por último, será necesario el estudio funcional de los nuevos genes involucrados, tanto por separado como en conjunto, para una correcta interpretación del papel de cada uno y de las vías de interacción entre ellos. Así, las técnicas de ingeniería genética, como las basadas en el sistema CRISPR/Cas9, que permiten la generación simultánea de varias mutaciones de forma endógena, serán de gran utilidad en estos estudios^{489, 490}. Es de esperar que la identificación de estos genes y sus interacciones causales, permitirá estratificar mejor el riesgo, incluso dentro de una misma familia, aumentando el poder predictivo de los actuales paneles genéticos^{322, 491, 492}.

Como resumen de este capítulo de resultados, en este estudio se han identificado varios genes candidatos de predisposición al SMOH. El conocimiento actual de estos genes nos permite sugerir que *POLE* podría ser un buen candidato para la familia portadora de la mutación, tanto por el tipo de mutación identificada (truncante), como por la función del gen (relacionado con cáncer hereditario) más los datos de segregación familiar obtenidos. En el caso de que estos resultados se demuestren, nuestros hallazgos podrán ser utilizados a nivel clínico para el seguimiento y monitorización de los miembros portadores de esta familia. Además, si se confirma la presencia de mutaciones en este gen en el estudio SMIPS que se está realizando actualmente, *POLE* podría ser un buen candidato para ser introducido en los paneles de genes utilizados actualmente en diagnóstico genético de este síndrome. Por otra parte, la identificación de otros posibles candidatos potenciales y la consideración de diferentes modelos de herencia es una aproximación interesante para mejorar la comprensión de las bases genéticas de las familias BRCA. Ensayos funcionales, poblacionales y de segregación familiar serán necesarios para determinar el papel de cada uno de estos candidatos, por si solos o interaccionando entre ellos, en la predisposición al SMOH.

SUMARIO

La presente tesis se planteó con el objetivo de mejorar el diagnóstico y el consejo genético de pacientes con sospecha de SCMOH. Para ello, se abordaron dos de las principales problemáticas en el diagnóstico genético de estos pacientes: la clasificación clínica de VSD identificadas en los *BRCA1/2* y la identificación de las causas genéticas asociadas a la predisposición en las familias BRCA.

Para el estudio de las VSD en los genes *BRCA1/2*, se propusieron diferentes aproximaciones metodológicas, algunas ya utilizadas en nuestro laboratorio (estudios de RNA) y otras nuevas pero ya validadas en la literatura, como el ensayo TA y el ensayo funcional basado en células-RE. En este sentido, mientras que el ensayo de TA ha demostrado ser un ensayo óptimo para el uso clínico por su gran valor predictivo y su facilidad técnica; una serie de limitaciones económicas, técnicas y temporales asociadas al ensayo basado en células-RE hacen que éste quede relegado a un uso más científico-investigacional. Es importante destacar que este método tiene la ventaja de poder ser utilizado específicamente para analizar variantes intrónicas y en regiones reguladoras, lo que es de gran valor en un contexto en el que se espera que haya un aumento en el número de este tipo de variantes debido a la implantación de las técnicas de secuenciación masiva en el diagnóstico clínico. Por otra parte, una de las técnicas comúnmente utilizada en nuestro laboratorio para la clasificación de VSD, los análisis de RNA, fue validada con éxito mediante la participación en el Proyecto de Control de Calidad dirigido por el consorcio ENIGMA. Esto nos permite asegurar la calidad de nuestros resultados en este tipo de ensayos. Por tanto, como resultado conjunto de este trabajo, se ha conseguido mejorar las prestaciones de nuestro laboratorio para clasificar VSD en los genes *BRCA1/2*. En un contexto en el que la rareza de las VSD dificulta su clasificación mediante otro tipo de modelos, la implantación de estas técnicas en un laboratorio de diagnóstico tiene una importante relevancia clínica, lo que le da un gran valor al trabajo realizado en esta tesis.

Utilizando las tres técnicas anteriormente citadas, se analizaron 40 VSD identificadas en nuestro centro. En total, 9 de ellas pudieron ser asociadas con el desarrollo del cáncer y fueron clasificadas como probablemente patogénicas. La clasificación de estas variantes como probablemente patogénicas facilita el manejo clínico de los pacientes de estas familias y alivia el estrés ligado a la incerteza de la VSD en los portadores y familiares. Otras cinco variantes, respondieron a un fenotipo hipomórfico y fueron asociadas a un riesgo bajo/moderado a sufrir el SCMOH. Si bien, la identificación de variantes hipomórficas no tiene cabida si las evidencias funcionales no van acompañadas de evidencias adicionales genéticas, histopatológicas y/o epidemiológicas. Es lógico pensar que, el creciente uso de ensayos funcionales para analizar VSD

identificará un mayor número de variantes de este tipo. Por tanto, como se ha dicho anteriormente, nuestros esfuerzos futuros deberían ir dirigidos al desarrollo de modelos multifactoriales que integren los resultados funcionales y puedan ser utilizados para interpretar fenotipos hipomórficos. Con ello, a su vez, se conseguirá clasificar variantes de una forma más robusta y rápida. No obstante, los modelos actuales deberán ser validados para su uso en la clasificación de variantes hipomórficas.

Por otra parte, en esta tesis se han utilizado una serie de programas de predicción *in silico* tanto a nivel de *splicing* como a nivel funcional. Los programas de predicción *in silico* de *splicing* han mostrado ser consistentes en la identificación de los sitios de *splicing* consenso, lo que hace de ellos una buena herramienta para priorizar variantes cuando no todas ellas puedan ser analizadas mediante ensayos de RNA. Estos programas han mostrado mayores limitaciones en la identificación de secuencias reguladoras funcionales. Por otra parte, los diferentes programas de predicción *in silico* de cambios de aminoácido en la proteína mostraron discrepancias para la mayoría de las variantes analizadas. Estas discrepancias apoyan la necesidad de utilizar ensayos funcionales en el proceso de clasificación de VSD. Además, nuestros resultados sugieren la necesidad de una mayor validación de estos programas, tanto los de *splicing* como los de proteína, en un contexto en el que algunos de ellos están siendo integrados actualmente en los citados modelos multifactoriales o se espera sean integrados en un futuro cercano. En este sentido, nuestros resultados pueden ayudar a la validación de los citados programas.

La segunda parte de este proyecto se centró en mejorar el conocimiento de las bases genéticas de predisposición asociadas a las familias BRCAX. Para ello, se propuso identificar nuevos genes de predisposición al SCMOH. Con este fin, por una parte, se secuenciaron genes candidatos en familias BRCAX y por otra se secuenciaron exomas completos de pacientes pertenecientes a familias BRCAX en base a un patrón de herencia autosómico dominante. El cribado mutacional en familias BRCAX de los genes de anemia de fanconi, *SLX4* y *ERCC4*, identificó una serie de variantes *missense* que podrían causar un impacto funcional. No obstante, el estudio funcional de estas variantes es necesario para demostrar la asociación de estos genes con el SCMOH. Si bien, una serie de estudios posteriores al nuestro demostraron una contribución pequeña de *SLX4* al SCMOH. Además, uno de estos estudios, mostró que su inclusión en los paneles de genes utilizados para el diagnóstico genético de pacientes de este síndrome podría ser beneficiosa. En cambio, estudios posteriores indican que *ERCC4* no parece estar asociado al SCMOH, al menos en población Caucásica. El cribado de mutaciones de este gen en tamaños muestrales más grandes es necesario para la correcta estimación de su asociación al SCMOH.

Por otra parte, la secuenciación de exomas de pacientes pertenecientes a familias BRCAX ha identificado 20 genes candidatos de predisposición al SCMOH. El conocimiento actual de los genes candidatos permite sugerir a *POLE* como un buen candidato para la familia en la que ha sido identificada la mutación en este gen. Se espera, que el análisis de este gen en una serie amplia de familias, como se está haciendo actualmente, certifique su relación con el SCMOH. El resto de candidatos, bien por su función o por el tipo de mutación identificada, también pueden considerarse buenos candidatos. No obstante, en estos casos se necesitará de más validaciones funcionales y genéticas para determinar su contribución al SCMOH. Si bien, al igual que *POLE*, se espera que el estudio de estos genes en nuestras familias BRCAX, aporte evidencias que permitan determinar cuáles de ellos se postulan como candidatos reales. En su conjunto, nuestros resultados demuestran que el uso de las nuevas técnicas de secuenciación masiva para la identificación de nuevos genes de predisposición autosómicos dominantes al SCMOH es una buena estrategia. Los resultados obtenidos por nuestro estudio, junto con otros estudios previos sugieren que plantear nuevas aproximaciones de búsqueda (modelos recesivos, oncogenes, modelos poligénicos), puede ser beneficioso para la identificación de nuevos genes. Nuestros resultados contribuyen a un mayor conocimiento de la susceptibilidad asociada a los casos BRCAX y aportan nuevas evidencias sobre la heterogeneidad genética asociada en dicha susceptibilidad. En este sentido, nuestros resultados pueden servir como punto de partida para la identificación de nuevos genes o nuevas interacciones genéticas ligadas a la susceptibilidad al SCMOH, que permitan individualizar el riesgo asociado a cada caso.

CONCLUSIONES

- La metodología utilizada en nuestro laboratorio para analizar VSD a nivel transcripcional ha podido ser validada con éxito mediante la participación en el Proyecto de Control de Calidad dirigido por el consorcio ENIGMA.
- El uso de la citada metodología para el análisis de 28 VSD localizadas en los genes *BRCA1/2* identificó un *splicing* aberrante en seis de estas variantes demostrando su asociación con el SCMOH.
- El ensayo de activación transcripcional (TA) combinado con el modelaje estructural de las variantes Y1703S, W1718L y G1770V, ubicadas en el dominio BRCT de *BRCA1*, indica que estas variantes pueden ser clasificada como mutaciones probablemente patogénicas. Esta clasificación es de gran relevancia para el manejo clínico de los pacientes de estas familias.
- La mutación G1770V fue posteriormente identificada como la primera mutación fundadora en población de origen Marroquí.
- El ensayo funcional basado en células embrionarias de ratón sugiere que las variantes T200I, K331I, K1434I, V1532R y G2645G en *BRCA2*, producen un fenotipo hipomórfico y podrían estar asociadas al SCMOH con un riesgo moderado/bajo.
- Nuestros estudios evidencian que los resultados de análisis funcionales han de ser integrados en modelos multifactoriales de clasificación de variantes. De este modo se obtendrán clasificaciones robustas que podrán ser utilizadas y trasladadas al manejo clínico individualizado de acuerdo con el riesgo personal de cada paciente.
- El cribado mutacional de *SLX4* en una serie de pacientes BRCA1 de nuestro centro identificó 21 variantes del tipo *missense*, 9 de ellas nunca antes reportadas y no identificadas en una serie de 284 controles. Si bien, estudios funcionales de estas variantes son necesarios para demostrar su asociación al SCMOH, estudios posteriores de otros grupos de investigación han identificado variantes de pérdida de función en este gen en pacientes con SCMOH.
- El cribado mutacional de *ERCC4* en la misma serie de pacientes BRCA1, identificó 10 mutaciones *missense*, dos de ellas nunca antes reportadas. El estudio de *ERCC4* en 1573 casos y 854 controles de población española evidenció que este gen no contribuye de forma relevante al SCMOH en nuestra población.
- La secuenciación de exomas de pacientes pertenecientes a familias BRCA1 seleccionadas en base a un patrón de herencia autosómico dominante con alta

agregación familiar ha identificado 20 genes candidatos de predisposición al SCMOH.

- El conocimiento actual de los genes candidatos permite sugerir que la mutación identificada en *POLE* podría ser la responsable del SCMOH en los pacientes de esta familia. Para el resto de candidatos, considerados también buenos candidatos, bien por su función o por el tipo de mutación identificada, se necesitan más datos poblacionales, familiares y funcionales para determinar su papel en el SCMOH de las familias estudiadas.
- Nuestros hallazgos sugieren que el considerar nuevas aproximaciones de análisis de los resultados (modelos recesivos, introducción de fenocopias, etc.), podría identificar nuevos genes de predisposición al SCMOH que han pasado inadvertidos con los criterios de análisis utilizados en la presente tesis doctoral.
- Como resumen final consideramos que el presente trabajo, en su conjunto, contribuye a la mejora del diagnóstico y consejo genético del SCMOH gracias a la clasificación clínica de variantes VSD en los genes *BRCA1/2* así como a la identificación de nuevos genes candidatos para el SCMOH.

BIBLIOGRAFÍA

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer* 2014;136:E359-86.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: a cancer journal for clinicians* 2011;61:69-90.
3. Youlden DR, Cramb SM, Dunn NA, Muller JM, Pyke CM, Baade PD. The descriptive epidemiology of female breast cancer: an international comparison of screening, incidence, survival and mortality. *Cancer epidemiology* 2012;36:237-48.
4. Hery C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. *Ann Oncol* 2008;19:1187-94.
5. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International journal of cancer* 2004;132:1133-45.
6. Dumitrescu RG, Cotarla I. Understanding breast cancer risk -- where do we stand in 2005? *Journal of cellular and molecular medicine* 2005;9:208-21.
7. Breast CGoHFi. 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 (London, England)* 2001;358:1389-99.
8. Garcia-Closas M, Brinton LA, Lissowska J, Chatterjee N, Peplonska B, Anderson WF, Szeszenia-Dabrowska N, Bardin-Mikolajczak A, Zatonski W, Blair A, Kalaylioglu Z, Rymkiewicz G, et al. Established breast cancer risk factors by clinically important tumour characteristics. *British journal of cancer* 2006;95:123-9.
9. Cancer. 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. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet (London, England)* 1997;350:1047-59.
10. Breast CGoHFi. 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 (London, England)* 2002;360:187-95.
11. Boyd NF, Stone J, Vogt KN, Connelly BS, Martin LJ, Minkin S. Dietary fat and breast cancer risk revisited: a meta-analysis of the published literature. *British journal of cancer* 2003;89:1672-85.
12. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ (Clinical research ed)* 2007;335:1134.
13. Stuckey A. Breast cancer: epidemiology and risk factors. *Clinical obstetrics and gynecology* 2011;54:96-102.
14. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *The New England journal of medicine* 2000;343:78-85.
15. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. *International journal of cancer* 1997;71:800-9.
16. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (New York, N.Y)* 1994;266:66-71.
17. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;378:789-92.
18. Farhat GN, Walker R, Buist DS, Onega T, Kerlikowske K. Changes in invasive breast cancer and ductal carcinoma in situ rates in relation to the decline in hormone therapy use. *J Clin Oncol* 2010;28:5140-6.
19. Li CI, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *Jama* 2003;289:1421-4.
20. Ward EM, DeSantis CE, Lin CC, Kramer JL, Jemal A, Kohler B, Brawley OW, Gansler T. Cancer statistics: Breast cancer in situ. *CA: a cancer journal for clinicians* 2015;65:481-95.
21. Diab SG, Clark GM, Osborne CK, Libby A, Allred DC, Elledge RM. Tumor characteristics and clinical outcome of tubular and mucinous breast carcinomas. *J Clin Oncol* 1999;17:1442-8.
22. Rosen PP, Groshen S, Kinne DW. Prognosis in T2N0M0 stage I breast carcinoma: a 20-year follow-up study. *J Clin Oncol* 1991;9:1650-61.
23. Anderson WF, Schairer C, Chen BE, Hance KW, Levine PH. Epidemiology of inflammatory breast cancer (IBC). *Breast disease* 2005;22:9-23.
24. Caliskan M, Gatti G, Sosnovskikh I, Rotmensz N, Botteri E, Musmeci S, Rosali dos Santos G, Viale G, Luini A. Paget's disease of the breast: the experience of the European Institute of Oncology and review of the literature. *Breast cancer research and treatment* 2008;112:513-21.
25. Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. *British journal of cancer* 2005;93:1046-52.
26. Cancer AJCo. *AJCC Cancer Staging Manual*. 7th ed. New York, NY: Springer 2010.
27. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98:10869-74.

28. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, Bordoni A. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. *Ann Oncol* 2009;20:628-35.
29. Karp SE, Tonin PN, Begin LR, Martinez JJ, Zhang JC, Pollak MN, Foulkes WD. Influence of BRCA1 mutations on nuclear grade and estrogen receptor status of breast carcinoma in Ashkenazi Jewish women. *Cancer* 1997;80:435-41.
30. Lynch BJ, Holden JA, Buys SS, Neuhausen SL, Gaffney DK. Pathobiologic characteristics of hereditary breast cancer. *Human pathology* 1998;29:1140-4.
31. Palacios J, Honrado E, Osorio A, Cazorla A, Sarrio D, Barroso A, Rodriguez S, Cigudosa JC, Diez O, Alonso C, Lerma E, Dopazo J, et al. Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast cancer research and treatment* 2005;90:5-14.
32. Eisinger F, Nogues C, Birnbaum D, Jacquemier J, Sobol H. BRCA1 and medullary breast cancer. *Jama* 1998;280:1227-8.
33. Lakhani SR, Van De Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L, Easton DF. 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. *J Clin Oncol* 2002;20:2310-8.
34. Phillips KA, Andrulis IL, Goodwin PJ. Breast carcinomas arising in carriers of mutations in BRCA1 or BRCA2: are they prognostically different? *J Clin Oncol* 1999;17:3653-63.
35. Eisinger F, Stoppa-Lyonnet D, Longy M, Kerangueven F, Noguchi T, Bailly C, Vincent-Salomon A, Jacquemier J, Birnbaum D, Sobol H. Germ line mutation at BRCA1 affects the histoprognostic grade in hereditary breast cancer. *Cancer research* 1996;56:471-4.
36. Johannsson OT, Idvall I, Anderson C, Borg A, Barkardottir RB, Egilsson V, Olsson H. Tumour biological features of BRCA1-induced breast and ovarian cancer. *Eur J Cancer* 1997;33:362-71.
37. Foulkes WD, Chappuis PO, Wong N, Brunet JS, Vesprini D, Rozen F, Yuan ZQ, Pollak MN, Kuperstein G, Narod SA, Begin LR. Primary node negative breast cancer in BRCA1 mutation carriers has a poor outcome. *Ann Oncol* 2000;11:307-13.
38. Loman N, Johannsson O, Bendahl PO, Borg A, Ferno M, Olsson H. Steroid receptors in hereditary breast carcinomas associated with BRCA1 or BRCA2 mutations or unknown susceptibility genes. *Cancer* 1998;83:310-9.
39. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, et al. 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* 2003;100:8418-23.
40. Foulkes WD, Stefansson IM, Chappuis PO, Begin LR, Goffin JR, Wong N, Trudel M, Akslen LA. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *Journal of the National Cancer Institute* 2003;95:1482-5.
41. Syrjakoski K, Vahteristo P, Eerola H, Tamminen A, Kivinummi K, Sarantaus L, Holli K, Blomqvist C, Kallioniemi OP, Kainu T, Nevanlinna H. Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. *Journal of the National Cancer Institute* 2000;92:1529-31.
42. Honrado E, Osorio A, Palacios J, Milne RL, Sanchez L, Diez O, Cazorla A, Syrjakoski K, Huntsman D, Heikkila P, Lerma E, Kallioniemi A, et al. Immunohistochemical expression of DNA repair proteins in familial breast cancer differentiate BRCA2-associated tumors. *J Clin Oncol* 2005;23:7503-11.
43. Palacios J, Robles-Frias MJ, Castilla MA, Lopez-Garcia MA, Benitez J. The molecular pathology of hereditary breast cancer. *Pathobiology* 2008;75:85-94.
44. Robson M, Rajan P, Rosen PP, Gilewski T, Hirschaut Y, Pressman P, Haas B, Norton L, Offit K. BRCA-associated breast cancer: absence of a characteristic immunophenotype. *Cancer research* 1998;58:1839-42.
45. Armes JE, Trute L, White D, Southey MC, Hammet F, Tesoriero A, Hutchins AM, Dite GS, McCredie MR, Giles GG, Hopper JL, Venter DJ. Distinct molecular pathogenesis of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study. *Cancer research* 1999;59:2011-7.
46. Marcus JN, Watson P, Page DL, Narod SA, Lenoir GM, Tonin P, Linder-Stephenson L, Salerno G, Conway TA, Lynch HT. Hereditary breast cancer: pathobiology, prognosis, and BRCA1 and BRCA2 gene linkage. *Cancer* 1996;77:697-709.
47. Lakhani SR, Gusterson BA, Jacquemier J, Sloane JP, Anderson TJ, van de Vijver MJ, Venter D, Freeman A, Antoniou A, McGuffog L, Smyth E, Steel CM, et al. The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. *Clin Cancer Res* 2000;6:782-9.
48. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *nature genetics* 2003;33:238-44.
49. Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. *American journal of human genetics* 1991;48:232-42.
50. Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *American journal of epidemiology* 1990;131:961-72.
51. NCCN Clinical Practice Guidelines in Oncology: Genetic/Familial High Risk Assessment: Breast and Ovarian. Version 1.2010. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed 13 December 2013. 2013.
52. de la Hoya M, Diez O, Perez-Segura P, Godino J, Fernandez JM, Sanz J, Alonso C, Baiget M, Diaz-Rubio E, Caldes T. Pre-test prediction models of BRCA1 or BRCA2 mutation in breast/ovarian families attending familial cancer clinics. *Journal of medical genetics* 2003;40:503-10.

53. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpfer KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol* 2002;20:1480-90.
54. Malone KE, Daling JR, Doody DR, Hsu L, Bernstein L, Coates RJ, Marchbanks PA, Simon MS, McDonald JA, Norman SA, Strom BL, Burkman RT, et al. Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years. *Cancer research* 2006;66:8297-308.
55. Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, et al. BRCA1 mutations in primary breast and ovarian carcinomas. *Science (New York, N.Y.)* 1994;266:120-2.
56. Kurian AW. BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. *Current opinion in obstetrics & gynecology* 2010;22:72-8.
57. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science (New York, N.Y.)* 2014;343:1466-70.
58. Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, Soucy P, Fredericksen Z, Barrowdale D, Dennis J, Gaudet MM, Dicks E, et al. Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS genetics* 2013;9:e1003212.
59. Gaudet MM, Kuchenbaecker KB, Vijai J, Klein RJ, Kirchoff T, McGuffog L, Barrowdale D, Dunning AM, Lee A, Dennis J, Healey S, Dicks E, et al. Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk. *PLoS genetics* 2013;9:e1003173.
60. Antoniou AC, Beesley J, McGuffog L, Sinilnikova OM, Healey S, Neuhausen SL, Ding YC, Rebbeck TR, Weitzel JN, Lynch HT, Isaacs C, Ganz PA, et al. Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction. *Cancer research* 2010;70:9742-54.
61. Antoniou AC, Sinilnikova OM, Simard J, Leone M, Dumont M, Neuhausen SL, Struwing JP, Stoppa-Lyonnet D, Barjhoux L, Hughes DJ, Coupier I, Belotti M, et al. RAD51 135G-->C modifies breast cancer risk among BRCA2 mutation carriers: results from a combined analysis of 19 studies. *American journal of human genetics* 2007;81:1186-200.
62. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, Ghousaini M, Barrowdale D, et al. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nature genetics* 2010;42:885-92.
63. Wang X, Pankratz VS, Fredericksen Z, Tarrell R, Karas M, McGuffog L, Pharaoh PD, Ponder BA, Dunning AM, Peock S, Cook M, Oliver C, et al. Common variants associated with breast cancer in genome-wide association studies are modifiers of breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Human molecular genetics* 2010;19:2886-97.
64. Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE. An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). *Breast Cancer Res* 2007;9:104.
65. Easton DF. How many more breast cancer predisposition genes are there? *Breast Cancer Res* 1999;1:14-7.
66. Rahman N. Realizing the promise of cancer predisposition genes. *Nature* 2014;505:302-8.
67. Apostolou P, Fostira F. Hereditary breast cancer: the era of new susceptibility genes. *BioMed research international* 2013;2013:747318.
68. Collins N, McManus R, Wooster R, Mangion J, Seal S, Lakhani SR, Ormiston W, Daly PA, Ford D, Easton DF, et al. Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13. *Oncogene* 1995;10:1673-5.
69. Boulton SJ. Cellular functions of the BRCA tumour-suppressor proteins. *Biochemical Society transactions* 2006;34:633-45.
70. Smith SA, Easton DF, Evans DG, Ponder BA. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nat Genet* 1992;2:128-31.
71. Narod SA, Foulkes WD. BRCA1 and BRCA2: 1994 and beyond. *Nat Rev Cancer* 2004;4:665-76.
72. Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes & development* 2000;14:927-39.
73. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, Grompe M, D'Andrea AD. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell* 2001;7:249-62.
74. Moynahan ME, Chiu JW, Koller BH, Jasin M. Brca1 controls homology-directed DNA repair. *Molecular cell* 1999;4:511-8.
75. Moynahan ME, Pierce AJ, Jasin M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Molecular cell* 2001;7:263-72.
76. Holloman WK. Unraveling the mechanism of BRCA2 in homologous recombination. *Nature structural & molecular biology* 2011;18:748-54.
77. Yuan R, Fan S, Wang JA, Meng Q, Ma Y, Schreiber D, Goldberg ID, Rosen EM. Coordinate alterations in the expression of BRCA1, BRCA2, p300, and Rad51 in response to genotoxic and other stresses in human prostate cancer cells. *The Prostate* 1999;40:37-49.

Bibliografia

78. Chen J, Silver DP, Walpita D, Cantor SB, Gazdar AF, Tomlinson G, Couch FJ, Weber BL, Ashley T, Livingston DM, Scully R. Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells. *Molecular cell* 1998;2:317-28.
79. Wang W. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet* 2007;8:735-48.
80. D'Andrea AD, Grompe M. The Fanconi anaemia/BRCA pathway. *Nat Rev Cancer* 2003;3:23-34.
81. Garaycochea JI, Patel KJ. Why does the bone marrow fail in Fanconi anemia? *Blood* 2014;123:26-34.
82. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* 2013;493:356-63.
83. Pang Q, Keeble W, Christianson TA, Faulkner GR, Bagby GC. FANCC interacts with Hsp70 to protect hematopoietic cells from IFN-gamma/TNF-alpha-mediated cytotoxicity. *The EMBO journal* 2001;20:4478-89.
84. Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, et al. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science (New York, N.Y)* 2002;297:606-9.
85. Sawyer SL, Tian L, Kahkonen M, Schwartzentruber J, Kircher M, Majewski J, Dymant DA, Innes AM, Boycott KM, Moreau LA, Moilanen JS, Greenberg RA. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. *Cancer discovery* 2015;5:135-42.
86. Chandler J, Hohenstein P, Swing DA, Tessarollo L, Sharan SK. Human BRCA1 gene rescues the embryonic lethality of Brca1 mutant mice. *Genesis* 2001;29:72-7.
87. Jasin M. Homologous repair of DNA damage and tumorigenesis: the BRCA connection. *Oncogene* 2002;21:8981-93.
88. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 2002;108:171-82.
89. Larson JS, Tonkinson JL, Lai MT. A BRCA1 mutant alters G2-M cell cycle control in human mammary epithelial cells. *Cancer research* 1997;57:3351-5.
90. Joukov V, Groen AC, Prokhorova T, Gerson R, White E, Rodriguez A, Walter JC, Livingston DM. The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly. *Cell* 2006;127:539-52.
91. Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet* 1998;19:254-6.
92. Hashizume R, Fukuda M, Maeda I, Nishikawa H, Oyake D, Yabuki Y, Ogata H, Ohta T. The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *The Journal of biological chemistry* 2001;276:14537-40.
93. Lorick KL, Jensen JP, Fang S, Ong AM, Hatakeyama S, Weissman AM. RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proceedings of the National Academy of Sciences of the United States of America* 1999;96:11364-9.
94. Rodriguez JA, Henderson BR. Identification of a functional nuclear export sequence in BRCA1. *The Journal of biological chemistry* 2000;275:38589-96.
95. Bork P, Hofmann K, Bucher P, Neuwald AF, Altschul SF, Koonin EV. A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *Faseb J* 1997;11:68-76.
96. Callebaut I, Mornon JP. From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. *FEBS letters* 1997;400:25-30.
97. Chapman MS, Verma IM. Transcriptional activation by BRCA1. *Nature* 1996;382:678-9.
98. Yu X, Chini CC, He M, Mer G, Chen J. The BRCT domain is a phospho-protein binding domain. *Science (New York, N.Y)* 2003;302:639-42.
99. Aprelikova ON, Fang BS, Meissner EG, Cotter S, Campbell M, Kuthiala A, Bessho M, Jensen RA, Liu ET. BRCA1-associated growth arrest is RB-dependent. *Proceedings of the National Academy of Sciences of the United States of America* 1999;96:11866-71.
100. Scully R, Chen J, Plug A, Xiao Y, Weaver D, Feunteun J, Ashley T, Livingston DM. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* 1997;88:265-75.
101. Zhong Q, Chen CF, Li S, Chen Y, Wang CC, Xiao J, Chen PL, Sharp ZD, Lee WH. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science (New York, N.Y)* 1999;285:747-50.
102. Clark SL, Rodriguez AM, Snyder RR, Hankins GD, Boehning D. Structure-Function Of The Tumor Suppressor BRCA1. *Computational and structural biotechnology journal* 2012;1.
103. Shamoo Y. Structural insights into BRCA2 function. *Current opinion in structural biology* 2003;13:206-11.
104. Siddique H, Zou JP, Rao VN, Reddy ES. The BRCA2 is a histone acetyltransferase. *Oncogene* 1998;16:2283-5.
105. Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, Christ N, Liu X, Jasin M, Couch FJ, Livingston DM. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Molecular cell* 2006;22:719-29.
106. Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF, Milner J, Brown LA, Hsu F, Gilks B, Nielsen T, Schulzer M, et al. EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* 2003;115:523-35.
107. Wong JM, Ionescu D, Ingles CJ. Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. *Oncogene* 2003;22:28-33.
108. Rajendra E, Venkitaraman AR. Two modules in the BRC repeats of BRCA2 mediate structural and functional interactions with the RAD51 recombinase. *Nucleic acids research* 2010;38:82-96.

109. Wong AK, Pero R, Ormonde PA, Tavtigian SV, Bartel PL. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *brca2*. *The Journal of biological chemistry* 1997;272:31941-4.
110. Jeyasekharan AD, Liu Y, Hattori H, Pisupati V, Jonsdottir AB, Rajendra E, Lee M, Sundaramoorthy E, Schlachter S, Kaminski CF, Ofir-Rosenfeld Y, Sato K, et al. A cancer-associated BRCA2 mutation reveals masked nuclear export signals controlling localization. *Nature structural & molecular biology* 2013;20:1191-8.
111. Hussain S, Wilson JB, Medhurst AL, Hejna J, Witt E, Ananth S, Davies A, Masson JY, Moses R, West SC, de Winter JP, Ashworth A, et al. Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Human molecular genetics* 2004;13:1241-8.
112. Thorslund T, Esashi F, West SC. Interactions between human BRCA2 protein and the meiosis-specific recombinase DMC1. *The EMBO journal* 2007;26:2915-22.
113. Esashi F, Christ N, Gannon J, Liu Y, Hunt T, Jasin M, West SC. CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* 2005;434:598-604.
114. Tamura K, Utsunomiya J, Iwama T, Furuyama J, Takagawa T, Takeda N, Fukuda Y, Matsumoto T, Nishigami T, Kusuhashi K, Sagayama K, Nakagawa K, et al. Mechanism of carcinogenesis in familial tumors. *International journal of clinical oncology* 2004;9:232-45.
115. Lindor NM, Goldgar DE, Tavtigian SV, Plon SE, Couch FJ. BRCA1/2 sequence variants of uncertain significance: a primer for providers to assist in discussions and in medical management. *The oncologist* 2013;18:518-24.
116. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *The oncologist* 2004;9:361-77.
117. Konishi H, Mohseni M, Tamaki A, Garay JP, Croessmann S, Karnan S, Ota A, Wong HY, Konishi Y, Karakas B, Tahir K, Abukhdeir AM, et al. Mutation of a single allele of the cancer susceptibility gene BRCA1 leads to genomic instability in human breast epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* 2011;108:17773-8.
118. Santarosa M, Ashworth A. Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. *Biochimica et biophysica acta* 2004;1654:105-22.
119. Szabo C, Masiello A, Ryan JF, Brody LC. The breast cancer information core: database design, structure, and scope. *Human mutation* 2000;16:123-31.
120. Montagna M, Dalla Palma M, Menin C, Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E. Genomic rearrangements account for more than one-third of the BRCA1 mutations in northern Italian breast/ovarian cancer families. *Human molecular genetics* 2003;12:1055-61.
121. Welch PL, King MC. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Human molecular genetics* 2001;10:705-13.
122. Judkins T, Rosenthal E, Arnell C, Burbidge LA, Geary W, Barrus T, Schoenberger J, Trost J, Wenstrup RJ, Roa BB. Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer* 2012;118:5210-6.
123. Thompson D, Easton D. The genetic epidemiology of breast cancer genes. *Journal of mammary gland biology and neoplasia* 2004;9:221-36.
124. Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro AN, Iversen ES, Couch FJ, Goldgar DE. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *American journal of human genetics* 2007;81:873-83.
125. Janavicius R. Founder BRCA1/2 mutations in the Europe: implications for hereditary breast-ovarian cancer prevention and control. *The EPMA journal* 2010;1:397-412.
126. Fackenthal JD, Lee Y, Olopade OI. Hidden dangers: a cryptic exon disrupts BRCA2 mRNA. *Clin Cancer Res* 2012;18:4865-7.
127. Gambino G, Tancredi M, Falaschi E, Aretini P, Caligo MA. Characterization of three alternative transcripts of the BRCA1 gene in patients with breast cancer and a family history of breast and/or ovarian cancer who tested negative for pathogenic mutations. *International journal of molecular medicine* 2015;35:950-6.
128. Walsh T, Lee MK, Casadei S, Thornton AM, Stray SM, Pennil C, Nord AS, Mandell JB, Swisher EM, King MC. 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* 2010;107:12629-33.
129. Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM. Prediction and assessment of splicing alterations: implications for clinical testing. *Human mutation* 2008;29:1304-13.
130. Guidugli L, Carreira A, Caputo SM, Ehlen A, Galli A, Monteiro AN, Neuhausen SL, Hansen TV, Couch FJ, Vreeswijk MP. Functional assays for analysis of variants of uncertain significance in BRCA2. *Human mutation* 2014;35:151-64.
131. 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. A classification model for BRCA2 DNA binding domain missense variants based on homology-directed repair activity. *Cancer research* 2013;73:265-75.
132. Iversen ES, Jr., Couch FJ, Goldgar DE, Tavtigian SV, Monteiro AN. A computational method to classify variants of uncertain significance using functional assay data with application to BRCA1. *Cancer Epidemiol Biomarkers Prev* 2011;20:1078-88.

133. Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. *Human mutation* 2008;29:1342-54.
134. Eggington JM, Bowles KR, Moyes K, Manley S, Esterling L, Sizemore S, Rosenthal E, Theisen A, Saam J, Arnell C, Pruss D, Bennett J, et al. A comprehensive laboratory-based program for classification of variants of uncertain significance in hereditary cancer genes. *Clinical genetics* 2014;86:229-37.
135. Ready K, Gutierrez-Barrera AM, Amos C, Meric-Bernstam F, Lu K, Hortobagyi G, Arun B. Cancer risk management decisions of women with BRCA1 or BRCA2 variants of uncertain significance. *The breast journal* 2011;17:210-2.
136. Murray ML, Cerrato F, Bennett RL, Jarvik GP. Follow-up of carriers of BRCA1 and BRCA2 variants of unknown significance: variant reclassification and surgical decisions. *Genet Med* 2011;13:998-1005.
137. Pruss D, Morris B, Hughes E, Eggington JM, Esterling L, Robinson BS, van Kan A, Fernandes PH, Roa BB, Gutin A, Wenstrup RJ, Bowles KR. Development and validation of a new algorithm for the reclassification of genetic variants identified in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 2014;147:119-32.
138. Heyn H, Sayols S, Moutinho C, Vidal E, Sanchez-Mut JV, Stefansson OA, Nadal E, Moran S, Eyfjord JE, Gonzalez-Suarez E, Pujana MA, Esteller M. Linkage of DNA methylation quantitative trait loci to human cancer risk. *Cell reports* 2014;7:331-8.
139. Smith A, Moran A, Boyd MC, Bulman M, Shenton A, Smith L, Iddenden R, Woodward ER, Laloo F, Maher ER, Evans DG. Phenocopies in BRCA1 and BRCA2 families: evidence for modifier genes and implications for screening. *Journal of medical genetics* 2007;44:10-5.
140. Kang PC, Phuah SY, Sivanandan K, Kang IN, Thirthagiri E, Liu JJ, Hassan N, Yoon SY, Thong MK, Hui M, Hartman M, Yip CH, et al. Recurrent mutation testing of BRCA1 and BRCA2 in Asian breast cancer patients identify carriers in those with presumed low risk by family history. *Breast cancer research and treatment* 2014;144:635-42.
141. Friedman LS, Szabo CI, Ostermeyer EA, Dowd P, Butler L, Park T, Lee MK, Goode EL, Rowell SE, King MC. Novel inherited mutations and variable expressivity of BRCA1 alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. *American journal of human genetics* 1995;57:1284-97.
142. Ferla R, Calo V, Cascio S, Rinaldi G, Badalamenti G, Carreca I, Surmacz E, Colucci G, Bazan V, Russo A. Founder mutations in BRCA1 and BRCA2 genes. *Ann Oncol* 2007;18 Suppl 6:vi93-8.
143. Neuhausen S, Gilewski T, Norton L, Tran T, McGuire P, Swensen J, Hampel H, Borgen P, Brown K, Skolnick M, Shattuck-Eidens D, Jhanwar S, et al. Recurrent BRCA2 6174delT mutations in Ashkenazi Jewish women affected by breast cancer. *Nat Genet* 1996;13:126-8.
144. Roa BB, Boyd AA, Volcik K, Richards CS. Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nat Genet* 1996;14:185-7.
145. Castilla LH, Couch FJ, Erdos MR, Hoskins KF, Calzone K, Garber JE, Boyd J, Lubin MB, Deshano ML, Brody LC, et al. Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. *Nat Genet* 1994;8:387-91.
146. Garcia C, Lyon L, Littell RD, Powell CB. Comparison of risk management strategies between women testing positive for a BRCA variant of unknown significance and women with known BRCA deleterious mutations. *Genet Med* 2014;16:896-902.
147. Lindor NM, Guidugli L, Wang X, Vallee MP, Monteiro AN, Tavtigian S, Goldgar DE, Couch FJ. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Human mutation* 2012;33:8-21.
148. Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Human mutation* 2008;29:1265-72.
149. 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. 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* 2012;33:2-7.
150. Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *American journal of human genetics* 2004;75:535-44.
151. Spearman AD, Sweet K, Zhou XP, McLennan J, Couch FJ, Toland AE. Clinically applicable models to characterize BRCA1 and BRCA2 variants of uncertain significance. *J Clin Oncol* 2008;26:5393-400.
152. Goh G, Choi M. Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics & informatics* 2012;10:214-9.
153. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, Quinlan AR, Nickerson DA, Eichler EE. Copy number variation detection and genotyping from exome sequence data. *Genome research* 2012;22:1525-32.
154. Mohammadi L, Vreeswijk MP, Oldenburg R, van den Ouweland A, Oosterwijk JC, van der Hout AH, Hoogerbrugge N, Ligtenberg M, Ausems MG, van der Luijt RB, Dommering CJ, Gille JJ, et al. A simple method for co-segregation analysis to evaluate the pathogenicity of unclassified variants; BRCA1 and BRCA2 as an example. *BMC cancer* 2009;9:211.
155. Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *American journal of human genetics* 2003;73:652-5.

156. Yurgelun MB, Allen B, Kaldate RR, Bowles KR, Judkins T, Kaushik P, Roa BB, Wenstrup RJ, Hartman AR, Syngal S. Identification of a Variety of Mutations in Cancer Predisposition Genes in Patients With Suspected Lynch Syndrome. *Gastroenterology* 2015;149:604-13 e20.
157. Abkevich V, Zharkikh A, Deffenbaugh AM, Frank D, Chen Y, Shattuck D, Skolnick MH, Gutin A, Tavtigian SV. Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation. *Journal of medical genetics* 2004;41:492-507.
158. Wagner JE, Tolar J, Levran O, Scholl T, Deffenbaugh A, Satagopan J, Ben-Porat L, Mah K, Batish SD, Kutler DJ, MacMillan ML, Hanenberg H, et al. Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood* 2004;103:3226-9.
159. Domchek SM, Tang J, Stopfer J, Lilli DR, Hamel N, Tischkowitz M, Monteiro AN, Messick TE, Powers J, Yonker A, Couch FJ, Goldgar DE, et al. Biallelic deleterious BRCA1 mutations in a woman with early-onset ovarian cancer. *Cancer discovery* 2013;3:399-405.
160. Tavtigian SV, Samollow PB, de Silva D, Thomas A. An analysis of unclassified missense substitutions in human BRCA1. *Familial cancer* 2006;5:77-88.
161. Lovelock PK, Spurdle AB, Mok MT, Farrugia DJ, Lakhani SR, Healey S, Arnold S, Buchanan D, Couch FJ, Henderson BR, Goldgar DE, Tavtigian SV, et al. Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants? *Breast Cancer Res* 2007;9:R82.
162. Spurdle AB, Whiley PJ, Thompson B, Feng B, Healey S, Brown MA, Pettigrew C, Van Asperen CJ, Ausems MG, Kattentidt-Mouravieva AA, van den Ouweland AM, Lindblom A, et al. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. *Journal of medical genetics* 2012;49:525-32.
163. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *The EMBO journal* 1986;5:823-6.
164. Dorfman R, Nalpathamkalam T, Taylor C, Gonska T, Keenan K, Yuan XW, Corey M, Tsui LC, Zielenski J, Durie P. Do common in silico tools predict the clinical consequences of amino-acid substitutions in the CFTR gene? *Clinical genetics* 2010;77:464-73.
165. Quiles F, Fernandez-Rodriguez J, Mosca R, Feliubadalo L, Tornero E, Brunet J, Blanco I, Capella G, Pujana MA, Aloy P, Monteiro A, Lazaro C. Functional and structural analysis of C-terminal BRCA1 missense variants. *PloS one* 2013;8:e61302.
166. Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB. In silico analysis of missense substitutions using sequence-alignment based methods. *Human mutation* 2008;29:1327-36.
167. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome research* 2001;11:863-74.
168. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nature methods* 2012;7:575-6.
169. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nature methods* 2010;7:248-9.
170. De Nicolo A, Parisini E, Zhong Q, Dalla Palma M, Stoeckert KA, Domchek SM, Nathanson KL, Caligo MA, Vidal M, Cusick ME, Garber JE. Multimodal assessment of protein functional deficiency supports pathogenicity of BRCA1 p.V1688del. *Cancer research* 2009;69:7030-7.
171. Karchin R, Agarwal M, Sali A, Couch F, Beattie MS. Classifying Variants of Undetermined Significance in BRCA2 with protein likelihood ratios. *Cancer informatics* 2008;6:203-16.
172. Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, de Wind N. Assessment of functional effects of unclassified genetic variants. *Human mutation* 2008;29:1314-26.
173. Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic acids research* 1987;15:7155-74.
174. Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. *Journal of medical genetics* 2005;42:737-48.
175. Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Leone M, 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. *Human mutation* 2012;33:1228-38.
176. Menendez M, Castellsague J, Mirete M, Pros E, Feliubadalo L, Osorio A, Calaf M, Tornero E, del Valle J, Fernandez-Rodriguez J, Quiles F, Salinas M, et al. Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes. *Breast cancer research and treatment* 2012;132:979-92.
177. Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, Devilee P. Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. *Human mutation* 2009;30:107-14.
178. Thomassen M, Blanco A, Montagna M, Hansen TV, Pedersen IS, Gutierrez-Enriquez S, Menendez M, Fachal L, Santamarina M, Steffensen AY, Jonson L, Agata S, et al. Characterization of BRCA1 and BRCA2 splicing variants: a collaborative report by ENIGMA consortium members. *Breast cancer research and treatment* 2012;132:1009-23.
179. Chenevix-Trench G, Healey S, Lakhani S, Waring P, Cummings M, Brinkworth R, Deffenbaugh AM, Burbidge LA, Pruss D, Judkins T, Scholl T, Bekessy A, et al. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer research* 2006;66:2019-27.

180. Osorio A, de la Hoya M, Rodriguez-Lopez R, Martinez-Ramirez A, Cazorla A, Granizo JJ, Esteller M, Rivas C, Caldes T, Benitez J. Loss of heterozygosity analysis at the BRCA loci in tumor samples from patients with familial breast cancer. *International journal of cancer* 2002;99:305-9.
181. Osorio A, Milne RL, Honrado E, Barroso A, Diez O, Salazar R, de la Hoya M, Vega A, Benitez J. Classification of missense variants of unknown significance in BRCA1 based on clinical and tumor information. *Human mutation* 2007;28:477-85.
182. Spurdle AB, Lakhani SR, Healey S, Parry S, Da Silva LM, Brinkworth R, Hopper JL, Brown MA, Babikyan D, Chenevix-Trench G, Tavtigian SV, Goldgar DE. Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis--a report from the kConFab Investigators. *J Clin Oncol* 2008;26:1657-63.
183. Tavtigian SV, Greenblatt MS, Goldgar DE, Boffetta P. Assessing pathogenicity: overview of results from the IARC Unclassified Genetic Variants Working Group. *Human mutation* 2008;29:1261-4.
184. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human mutation* 2008;29:1282-91.
185. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, Lyon E, Ward BE. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet Med* 2008;10:294-300.
186. Vallee MP, Francy TC, Judkins MK, Babikyan D, Lesueur F, Gammon A, Goldgar DE, Couch FJ, Tavtigian SV. Classification of missense substitutions in the BRCA genes: a database dedicated to Ex-UVs. *Human mutation* 2012;33:22-8.
187. Berget SM, Sharp PA. A spliced sequence at the 5'-terminus of adenovirus late mRNA. *Brookhaven symposia in biology* 1977:332-44.
188. Chow LT, Gelinis RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 1977;12:1-8.
189. Alberts B. *Molecular Biology of the Cell*, 5e. New York: Garland Science 2008.
190. Darnell JE, Jr. Implications of RNA-RNA splicing in evolution of eukaryotic cells. *Science (New York, N.Y)* 1978;202:1257-60.
191. Blencowe BJ. Alternative splicing: new insights from global analyses. *Cell* 2006;126:37-47.
192. Tarn WY, Steitz JA. Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends in biochemical sciences* 1997;22:132-7.
193. Aebi M, Hornig H, Weissmann C. 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. *Cell* 1987;50:237-46.
194. Blencowe BJ. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends in biochemical sciences* 2000;25:106-10.
195. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002;3:285-98.
196. Lamond AI. The spliceosome. *Bioessays* 1993;15:595-603.
197. Nilsen TW. RNA-RNA interactions in the spliceosome: unraveling the ties that bind. *Cell* 1994;78:1-4.
198. Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. *Cell* 2002;108:501-12.
199. Panicola M. Catholic teaching on prolonging life: setting the record straight. *The Hastings Center report* 2001;31:14-25.
200. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009;136:701-18.
201. Patel AA, Steitz JA. Splicing double: insights from the second spliceosome. *Nature reviews* 2003;4:960-70.
202. Izquierdo JM, Valcarcel J. A simple principle to explain the evolution of pre-mRNA splicing. *Genes & development* 2006;20:1679-84.
203. Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular cell* 2007;27:435-48.
204. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008;456:470-6.
205. Colombo M, Blok MJ, Whiley P, Santamarina M, Gutierrez-Enriquez S, Romero A, Garre P, Becker A, Smith LD, De Vecchi G, Brandao RD, Tserpelis D, et al. Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Human molecular genetics* 2014;23:3666-80.
206. Romero A, Garcia-Garcia F, Lopez-Perolio I, Ruiz de Garibay G, Garcia-Saenz JA, Garre P, Ayllon P, Benito E, Dopazo J, Diaz-Rubio E, Caldes T, de la Hoya M. BRCA1 Alternative splicing landscape in breast tissue samples. *BMC cancer* 2015;15:219.
207. Liu HX, Zhang M, Krainer AR. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & development* 1998;12:1998-2012.
208. Honig A, Auboeuf D, Parker MM, O'Malley BW, Berget SM. Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. *Molecular and cellular biology* 2002;22:5698-707.
209. Stickeler E, Fraser SD, Honig A, Chen AL, Berget SM, Cooper TA. The RNA binding protein YB-1 binds A/C-rich exon enhancers and stimulates splicing of the CD44 alternative exon v4. *The EMBO journal* 2001;20:3821-30.

210. Martinez-Contreras R, Cloutier P, Shkreta L, Fiset JF, Revil T, Chabot B. hnRNP proteins and splicing control. *Advances in experimental medicine and biology* 2007;623:123-47.
211. Sterne-Weiler T, Howard J, Mort M, Cooper DN, Sanford JR. Loss of exon identity is a common mechanism of human inherited disease. *Genome research* 2011;21:1563-71.
212. Ars E, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Human molecular genetics* 2000;9:237-47.
213. Pros E, Gomez C, Martin T, Fabregas P, Serra E, Lazaro C. Nature and mRNA effect of 282 different NF1 point mutations: focus on splicing alterations. *Human mutation* 2008;29:E173-93.
214. Lopez-Bigas N, Audit B, Ouzounis C, Parra G, Guigo R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS letters* 2005;579:1900-3.
215. 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. 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* 2013;34:1424-31.
216. Buratti E, Baralle M, Baralle FE. Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic acids research* 2006;34:3494-510.
217. Crispino JD, Mermoud JE, Lamond AI, Sharp PA. Cis-acting elements distinct from the 5' splice site promote U1-independent pre-mRNA splicing. *RNA (New York, N.Y)* 1996;2:664-73.
218. Villate O, Rastrojo A, Lopez-Diez R, Hernandez-Torres F, Aguado B. Differential splicing, disease and drug targets. *Infectious disorders drug targets* 2008;8:241-51.
219. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic acids research* 2003;31:3568-71.
220. Whiley PJ, Pettigrew CA, Brewster BL, Walker LC, Spurdle AB, Brown MA. Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts. *BMC medical genetics* 2010;11:80.
221. Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornero L, Esteban-Cardenosa E, Lastra E, Pagani F, Miner C, Velasco EA. A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clin Cancer Res*;16:1957-67.
222. Brandao RD, van Roozendaal K, Tserpelis D, Gomez Garcia E, Blok MJ. Characterisation of unclassified variants in the BRCA1/2 genes with a putative effect on splicing. *Breast cancer research and treatment* 2010;129:971-82.
223. 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. Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity. *Human mutation* 2011;31:E1484-505.
224. Whiley PJ, de la Hoya M, Thomassen M, Becker A, Brandao R, Pedersen IS, Montagna M, Menendez M, Quiles F, Gutierrez-Enriquez S, De Leeneer K, Tenes A, et al. Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing. *Clinical chemistry* 2014;60:341-52.
225. Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Current opinion in cell biology* 2004;16:293-9.
226. Vibe-Pedersen K, Kornblihtt AR, Baralle FE. Expression of a human alpha-globin/fibronectin gene hybrid generates two mRNAs by alternative splicing. *The EMBO journal* 1984;3:2511-6.
227. Biswas K, Das R, Egginton JM, Qiao H, North SL, Stauffer S, Burkett SS, Martin BK, Southon E, Sizemore SC, Pruss D, Bowles KR, et al. Functional evaluation of BRCA2 variants mapping to the PALB2-binding and C-terminal DNA-binding domains using a mouse ES cell-based assay. *Human molecular genetics* 2011;21:3993-4006.
228. Biswas K, Das R, Alter BP, Kuznetsov SG, Stauffer S, North SL, Burkett S, Brody LC, Meyer S, Byrd RA, Sharan SK. A comprehensive functional characterization of BRCA2 variants associated with Fanconi anemia using mouse ES cell-based assay. *Blood* 2012;118:2430-42.
229. Lee MS, Green R, Marsillac SM, Coquelle N, Williams RS, Yeung T, Foo D, Hau DD, Hui B, Monteiro AN, Glover JN. Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays. *Cancer research* 2010;70:4880-90.
230. Jhurany A, Velkova A, Johnson RC, Kessing B, Carvalho RS, Whiley P, Spurdle AB, Vreeswijk MP, Caputo SM, Millot GA, Vega A, Coquelle N, et al. BRCA1 Circos: a visualisation resource for functional analysis of missense variants. *Journal of medical genetics* 2015;52:224-30.
231. Carvalho M, Pino MA, Karchin R, Beddor J, Godinho-Netto M, Mesquita RD, Rodarte RS, Vaz DC, Monteiro VA, Manoukian S, Colombo M, Ripamonti CB, et al. Analysis of a set of missense, frameshift, and in-frame deletion variants of BRCA1. *Mutation research* 2009;660:1-11.
232. Tischkowitz M, Hamel N, Carvalho MA, Birrane G, Soni A, van Beers EH, Joosse SA, Wong N, Novak D, Quenneville LA, Grist SA, Nederlof PM, et al. Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. *Eur J Hum Genet* 2008;16:820-32.
233. Millot GA, Carvalho MA, Caputo SM, Vreeswijk MP, Brown MA, Webb M, Rouleau E, Neuhausen SL, Hansen T, Galli A, Brandao RD, Blok MJ, et al. A guide for functional analysis of BRCA1 variants of uncertain significance. *Human mutation* 2012;33:1526-37.
234. Kuznetsov SG, Liu P, Sharan SK. Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2. *Nature medicine* 2008;14:875-81.

235. Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, Kuperwasser N, Motola DL, Meissner TB, Hendriks WT, Trevisan M, Gupta RM, et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell stem cell* 2013;12:238-51.
236. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y)* 2013;339:823-6.
237. Chen Y, Farmer AA, Chen CF, Chen PL, Lee WH. BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer research* 1996;56:3168-72.
238. Spain BH, Larson CJ, Shihabuddin LS, Gage FH, Verma IM. Truncated BRCA2 is cytoplasmic: implications for cancer-linked mutations. *Proceedings of the National Academy of Sciences of the United States of America* 1999;96:13920-5.
239. Feng Z, Kachnic L, Zhang J, Powell SN, Xia F. DNA damage induces p53-dependent BRCA1 nuclear export. *The Journal of biological chemistry* 2004;279:28574-84.
240. Au WW, Henderson BR. The BRCA1 RING and BRCT domains cooperate in targeting BRCA1 to ionizing radiation-induced nuclear foci. *The Journal of biological chemistry* 2005;280:6993-7001.
241. Henderson BR. Regulation of BRCA1, BRCA2 and BARD1 intracellular trafficking. *Bioessays* 2005;27:884-93.
242. Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer research* 2005;65:417-26.
243. Gowen LC, Johnson BL, Latour AM, Sulik KK, Koller BH. Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat Genet* 1996;12:191-4.
244. Sharan SK, Morimatsu M, Albrecht U, Lim DS, Regel E, Dinh C, Sands A, Eichele G, Hasty P, Bradley A. Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* 1997;386:804-10.
245. Chang S, Biswas K, Martin BK, Stauffer S, Sharan SK. Expression of human BRCA1 variants in mouse ES cells allows functional analysis of BRCA1 mutations. *The Journal of clinical investigation* 2009;119:3160-71.
246. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes & development* 1999;13:2633-8.
247. Farrugia DJ, Agarwal MK, Pankratz VS, Deffenbaugh AM, Pruss D, Frye C, Wadum L, Johnson K, Mentlick J, Tavtigian SV, Goldgar DE, Couch FJ. Functional assays for classification of BRCA2 variants of uncertain significance. *Cancer research* 2008;68:3523-31.
248. Moynahan ME, Cui TY, Jasin M. Homology-directed dna repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation. *Cancer research* 2001;61:4842-50.
249. Ransburgh DJ, Chiba N, Ishioka C, Toland AE, Parvin JD. Identification of breast tumor mutations in BRCA1 that abolish its function in homologous DNA recombination. *Cancer research* 2010;70:988-95.
250. Balia C, Galli A, Caligo MA. Effect of the overexpression of BRCA2 unclassified missense variants on spontaneous homologous recombination in human cells. *Breast cancer research and treatment* 2011;129:1001-9.
251. Guidugli L, Rugani C, Lombardi G, Aretini P, Galli A, Caligo MA. A recombination-based method to characterize human BRCA1 missense variants. *Breast cancer research and treatment* 2010;125:265-72.
252. Caligo MA, Bonatti F, Guidugli L, Aretini P, Galli A. A yeast recombination assay to characterize human BRCA1 missense variants of unknown pathological significance. *Human mutation* 2009;30:123-33.
253. Spugnoli L, Balia C, Collavoli A, Falaschi E, Quercioli V, Caligo MA, Galli A. Effect of the expression of BRCA2 on spontaneous homologous recombination and DNA damage-induced nuclear foci in *Saccharomyces cerevisiae*. *Mutagenesis* 2013;28:187-95.
254. Kais Z, Chiba N, Ishioka C, Parvin JD. Functional differences among BRCA1 missense mutations in the control of centrosome duplication. *Oncogene* 2011;31:799-804.
255. Starita LM, Machida Y, Sankaran S, Elias JE, Griffin K, Schlegel BP, Gygi SP, Parvin JD. BRCA1-dependent ubiquitination of gamma-tubulin regulates centrosome number. *Molecular and cellular biology* 2004;24:8457-66.
256. Connor F, Bertwistle D, Mee PJ, Ross GM, Swift S, Grigorieva E, Tybulewicz VL, Ashworth A. Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation. *Nat Genet* 1997;17:423-30.
257. Cortez D, Wang Y, Qin J, Elledge SJ. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science (New York, N.Y)* 1999;286:1162-6.
258. Shen MR, Zdzienicka MZ, Mohnweiser H, Thompson LH, Thelen MP. Mutations in hamster single-strand break repair gene XRCC1 causing defective DNA repair. *Nucleic acids research* 1998;26:1032-7.
259. Morimatsu M, Donoho G, Hasty P. Cells deleted for Brca2 COOH terminus exhibit hypersensitivity to gamma-radiation and premature senescence. *Cancer research* 1998;58:3441-7.
260. Rahden-Staron I, Szumilo M, Grosicka E, Kraakman van der Zwet M, Zdzienicka MZ. Defective Brca2 influences topoisomerase I activity in mammalian cells. *Acta biochimica Polonica* 2003;50:139-44.
261. Tomlinson GE, Chen TT, Stastny VA, Virmani AK, Spillman MA, Tonk V, Blum JL, Schneider NR, Wistuba II, Shay JW, Minna JD, Gazdar AF. Characterization of a breast cancer cell line derived from a germ-line BRCA1 mutation carrier. *Cancer research* 1998;58:3237-42.
262. Godthelp BC, Wiegant WW, Waisfisz Q, Medhurst AL, Arwert F, Joenje H, Zdzienicka MZ. Inducibility of nuclear Rad51 foci after DNA damage distinguishes all Fanconi anemia complementation groups from D1/BRCA2. *Mutation research* 2006;594:39-48.

263. Kote-Jarai Z, Matthews L, Osorio A, Shanley S, Giddings I, Moreews F, Locke I, Evans DG, Eccles D, Williams RD, Girolami M, Campbell C, et al. Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage. *Clin Cancer Res* 2006;12:3896-901.
264. Waddell N, Ten Haaf A, Marsh A, Johnson J, Walker LC, Gongora M, Brown M, Grover P, Girolami M, Grimmond S, Chenevix-Trench G, Spurdle AB. BRCA1 and BRCA2 missense variants of high and low clinical significance influence lymphoblastoid cell line post-irradiation gene expression. *PLoS genetics* 2008;4:e1000080.
265. Brown MA, Lo LJ, Catteau A, Xu CF, Lindeman GJ, Hodgson S, Solomon E. Germline BRCA1 promoter deletions in UK and Australian familial breast cancer patients: Identification of a novel deletion consistent with BRCA1:psiBRCA1 recombination. *Human mutation* 2002;19:435-42.
266. Signori E, Bagni C, Papa S, Primerano B, Rinaldi M, Amaldi F, Fazio VM. A somatic mutation in the 5'UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. *Oncogene* 2001;20:4596-600.
267. Wardrop SL, Brown MA. Identification of two evolutionarily conserved and functional regulatory elements in intron 2 of the human BRCA1 gene. *Genomics* 2005;86:316-28.
268. Morris JR, Keep NH, Solomon E. Identification of residues required for the interaction of BARD1 with BRCA1. *The Journal of biological chemistry* 2002;277:9382-6.
269. Brzovic PS, Rajagopal P, Hoyt DW, King MC, Klevit RE. Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nature structural biology* 2001;8:833-7.
270. Brzovic PS, Meza JE, King MC, Klevit RE. BRCA1 RING domain cancer-predisposing mutations. Structural consequences and effects on protein-protein interactions. *The Journal of biological chemistry* 2001;276:41399-406.
271. Morris JR, Pangon L, Boutell C, Katagiri T, Keep NH, Solomon E. Genetic analysis of BRCA1 ubiquitin ligase activity and its relationship to breast cancer susceptibility. *Human molecular genetics* 2006;15:599-606.
272. Dapic V, Monteiro AN. Functional implications of BRCA1 for early detection, prevention, and treatment of breast cancer. *Critical reviews in eukaryotic gene expression* 2006;16:233-52.
273. Monteiro AN, August A, Hanafusa H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93:13595-9.
274. Vallon-Christersson J, Cayanan C, Haraldsson K, Loman N, Bergthorsson JT, Brondum-Nielsen K, Gerdes AM, Moller P, Kristofferson U, Olsson H, Borg A, Monteiro AN. Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. *Human molecular genetics* 2001;10:353-60.
275. Carvalho MA, Marsillac SM, Karchin R, Manoukian S, Grist S, Swaby RF, Urmenyi TP, Rondinelli E, Silva R, Gayol L, Baumbach L, Sutphen R, et al. Determination of cancer risk associated with germ line BRCA1 missense variants by functional analysis. *Cancer research* 2007;67:1494-501.
276. Botuyan MV, Nomine Y, Yu X, Juranic N, Macura S, Chen J, Mer G. Structural basis of BACH1 phosphopeptide recognition by BRCA1 tandem BRCT domains. *Structure* 2004;12:1137-46.
277. Williams RS, Lee MS, Hau DD, Glover JN. Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. *Nature structural & molecular biology* 2004;11:519-25.
278. Williams RS, Chasman DI, Hau DD, Hui B, Lau AY, Glover JN. Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations. *The Journal of biological chemistry* 2003;278:53007-16.
279. Williams RS, Glover JN. Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. *The Journal of biological chemistry* 2003;278:2630-5.
280. Williams RS, Green R, Glover JN. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nature structural biology* 2001;8:838-42.
281. Coyne RS, McDonald HB, Edgemon K, Brody LC. Functional characterization of BRCA1 sequence variants using a yeast small colony phenotype assay. *Cancer biology & therapy* 2004;3:453-7.
282. Monteiro AN, Humphrey JS. Yeast-based assays for detection and characterization of mutations in BRCA1. *Breast disease* 1998;10:61-70.
283. Millot GA, Berger A, Lejour V, Boule JB, Bobo C, Cullin C, Lopes J, Stoppa-Lyonnet D, Nicolas A. Assessment of human Nter and Cter BRCA1 mutations using growth and localization assays in yeast. *Human mutation*;32:1470-80.
284. Millot GA, Berger A, Lejour V, Boule JB, Bobo C, Cullin C, Lopes J, Stoppa-Lyonnet D, Nicolas A. Assessment of human Nter and Cter BRCA1 mutations using growth and localization assays in yeast. *Human mutation* 2011;32:1470-80.
285. Brough R, Bajrami I, Vatcheva R, Natrajan R, Reis-Filho JS, Lord CJ, Ashworth A. APRIN is a cell cycle specific BRCA2-interacting protein required for genome integrity and a predictor of outcome after chemotherapy in breast cancer. *The EMBO journal* 2012;31:1160-76.
286. Hucl T, Rago C, Gallmeier E, Brody JR, Gorospe M, Kern SE. A syngeneic variance library for functional annotation of human variation: application to BRCA2. *Cancer research* 2008;68:5023-30.
287. Richter S, Haroun I, Graham TC, Eisen A, Kiss A, Warner E. Variants of unknown significance in BRCA testing: impact on risk perception, worry, prevention and counseling. *Ann Oncol* 2013;24 Suppl 8:viii69-viii74.
288. Petrucelli N, Lazebnik N, Huelsman KM, Lazebnik RS. Clinical interpretation and recommendations for patients with a variant of uncertain significance in BRCA1 or BRCA2: a survey of genetic counseling practice. *Genetic testing* 2002;6:107-13.
289. van Dijk S, van Asperen CJ, Jacobi CE, Vink GR, Tibben A, Breuning MH, Otten W. Variants of uncertain clinical significance as a result of BRCA1/2 testing: impact of an ambiguous breast cancer risk message. *Genetic testing* 2004;8:235-9.

290. Vos J, Gomez-Garcia E, Oosterwijk JC, Menko FH, Stoel RD, van Asperen CJ, Jansen AM, Stiggelbout AM, Tibben A. Opening the psychological black box in genetic counseling. The psychological impact of DNA testing is predicted by the counselees' perception, the medical impact by the pathogenic or uninformative BRCA1/2-result. *Psycho-oncology* 2012;21:29-42.
291. Vos J, Oosterwijk JC, Gomez-Garcia E, Menko FH, Jansen AM, Stoel RD, van Asperen CJ, Tibben A, Stiggelbout AM. Perceiving cancer-risks and heredity-likelihood in genetic-counseling: how counselees recall and interpret BRCA 1/2-test results. *Clinical genetics* 2011;79:207-18.
292. Wu CC, Shete S, Amos CI, Strong LC. Joint effects of germ-line p53 mutation and sex on cancer risk in Li-Fraumeni syndrome. *Cancer research* 2006;66:8287-92.
293. Bubien V, Bonnet F, Brouste V, Hoppe S, Barouk-Simonet E, David A, Edery P, Bottani A, Layet V, Caron O, Gilbert-Dussardier B, Delnatte C, et al. High cumulative risks of cancer in patients with PTEN hamartoma tumour syndrome. *Journal of medical genetics* 2013;50:255-63.
294. Hearle N, Schumacher V, Menko FH, Olschwang S, Boardman LA, Gille JJ, Keller JJ, Westerman AM, Scott RJ, Lim W, Trimbath JD, Giardiello FM, et al. Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. *Clin Cancer Res* 2006;12:3209-15.
295. Casadei S, Norquist BM, Walsh T, Stray S, Mandell JB, Lee MK, Stamatoyannopoulos JA, King MC. Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer. *Cancer research* 2011;71:2222-9.
296. Pelttari LM, Heikkinen T, Thompson D, Kallioniemi A, Schleutker J, Holli K, Blomqvist C, Aittomaki K, Butzow R, Nevanlinna H. RAD51C is a susceptibility gene for ovarian cancer. *Human molecular genetics* 2011;20:3278-88.
297. Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature genetics* 2007;39:165-7.
298. Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nature genetics* 2006;38:1239-41.
299. Kiiski JI, Pelttari LM, Khan S, Freysteinsdottir ES, Reynisdottir I, Hart SN, Shimelis H, Vilske S, Kallioniemi A, Schleutker J, Leminen A, Butzow R, et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2014;111:15172-7.
300. Economopoulou P, Dimitriadis G, Psyrrri A. Beyond BRCA: new hereditary breast cancer susceptibility genes. *Cancer treatment reviews* 2015;41:1-8.
301. Ahmed M, Rahman N. ATM and breast cancer susceptibility. *Oncogene* 2006;25:5906-11.
302. Morrell D, Cromartie E, Swift M. Mortality and cancer incidence in 263 patients with ataxia-telangiectasia. *J Natl Cancer Inst* 1986;77:89-92.
303. Roeb W, Higgins J, King MC. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. *Hum Mol Genet* 2012;21:2738-44.
304. Bogdanova N, Feshchenko S, Schurmann P, Waltes R, Wieland B, Hillemanns P, Rogov YI, Dammann O, Bremer M, Karstens JH, Sohn C, Varon R, et al. Nijmegen Breakage Syndrome mutations and risk of breast cancer. *Int J Cancer* 2008;122:802-6.
305. Heikkinen K, Rapakko K, Karpainen SM, Erkkö H, Knuutila S, Lundan T, Mannermaa A, Borresen-Dale AL, Borg A, Barkardottir RB, Petrini J, Winqvist R. RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. *Carcinogenesis* 2006;27:1593-9.
306. Ratajska M, Antoszewska E, Piskorz A, Brozek I, Borg A, Kusmierek H, Biernat W, Limon J. Cancer predisposing BARD1 mutations in breast-ovarian cancer families. *Breast Cancer Res Treat* 2012;131:89-97.
307. Maxwell KN, Nathanson KL. Common breast cancer risk variants in the post-COGS era: a comprehensive review. *Breast Cancer Res* 2013;15:212.
308. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;45:353-61, 61e1-2.
309. Bahcall OG. iCOGS collection provides a collaborative model. Foreword. *Nat Genet* 2013;45:343.
310. Gracia-Aznarez FJ, Fernandez V, Pita G, Peterlongo P, Dominguez O, de la Hoya M, Duran M, Osorio A, Moreno L, Gonzalez-Neira A, Rosa-Rosa JM, Sinilnikova O, et al. Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. *PloS one* 2013;8:e55681.
311. Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. *Annual review of genomics and human genetics* 2008;9:321-45.
312. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087-93.
313. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007;39:865-9.

314. Wacholder S, Han SS, Weinberg CR. Inference from a multiplicative model of joint genetic effects for [corrected] ovarian cancer risk. *Journal of the National Cancer Institute* 2010;103:82-3.
315. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *The New England journal of medicine* 2014;372:2243-57.
316. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM, Roeb W, Agnew KJ, Stray SM, Wickramanayake A, Norquist B, Pennington KP, et al. 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* 2011;108:18032-7.
317. Couch FJ, Hart SN, Sharma P, Toland AE, Wang X, Miron P, Olson JE, Godwin AK, Pankratz VS, Olswold C, Slettedahl S, Hallberg E, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. *J Clin Oncol* 2014;33:304-11.
318. Cybulski C, Lubinski J, Wokolorczyk D, Kuzniak W, Kashyap A, Sopik V, Huzarski T, Gronwald J, Byrski T, Szwiec M, Jakubowska A, Gorski B, et al. Mutations predisposing to breast cancer in 12 candidate genes in breast cancer patients from Poland. *Clinical genetics* 2015;88:366-70.
319. Hirotsu Y, Nakagomi H, Sakamoto I, Amemiya K, Oyama T, Mochizuki H, Omata M. Multigene panel analysis identified germline mutations of DNA repair genes in breast and ovarian cancer. *Molecular genetics & genomic medicine* 2015;3:459-66.
320. Kurian AW, Hare EE, Mills MA, Kingham KE, McPherson L, Whittemore AS, McGuire V, Ladabaum U, Kobayashi Y, Lincoln SE, Cargill M, Ford JM. Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol* 2014;32:2001-9.
321. Tung N, Battelli C, Allen B, Kaldete R, Bhatnagar S, Bowles K, Timms K, Garber JE, Herold C, Ellisen L, Krejdivsky J, DeLeonardis K, et al. Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. *Cancer* 2014;121:25-33.
322. Ku CS, Cooper DN, Iacopetta B, Roukos DH. Integrating next-generation sequencing into the diagnostic testing of inherited cancer predisposition. *Clinical genetics* 2012;83:2-6.
323. Curtin N. PARP inhibitors for anticancer therapy. *Biochemical Society transactions* 2014;42:82-8.
324. Pennington KP, Walsh T, Harrell MI, Lee MK, Pennil CC, Rendi MH, Thornton A, Norquist BM, Casadei S, Nord AS, Agnew KJ, Pritchard CC, et al. Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. *Clin Cancer Res*;20:764-75.
325. Drost R, Bouwman P, Rottenberg S, Boon U, Schut E, Klarenbeek S, Klijn C, van der Heijden I, van der Gulden H, Wientjens E, Pieterse M, Cateau A, et al. BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. *Cancer cell* 2011;20:797-809.
326. Lerner-Ellis J, Khalouei S, Sopik V, Narod SA. Genetic risk assessment and prevention: the role of genetic testing panels in breast cancer. *Expert review of anticancer therapy* 2015;15:1315-26.
327. Ruark E, Snape K, Humburg P, Loveday C, Bajrami I, Brough R, Rodrigues DN, Renwick A, Seal S, Ramsay E, Duarte Sdel V, Rivas MA, et al. Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. *Nature*;493:406-10.
328. Stadler ZK, Vijai J, Thom P, Kirchhoff T, Hansen NA, Kauff ND, Robson M, Offit K. Genome-wide association studies of cancer predisposition. *Hematology/oncology clinics of North America* 2010;24:973-96.
329. Thompson D, Szabo CI, Mangion J, Oldenburg RA, Odefrey F, Seal S, Barfoot R, Kroeze-Jansema K, Teare D, Rahman N, Renard H, Mann G, et al. 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* 2002;99:827-31.
330. Hedenfalk I, Ringner M, Ben-Dor A, Yakhini Z, Chen Y, Chebil G, Ach R, Loman N, Olsson H, Meltzer P, Borg A, Trent J. Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100:2532-7.
331. Gonzalez-Neira A, Rosa-Rosa JM, Osorio A, Gonzalez E, Southey M, Sinilnikova O, Lynch H, Oldenburg RA, van Asperen CJ, Hoogerbrugge N, Pita G, Devilee P, et al. 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* 2007;8:299.
332. Smith P, McGuffog L, Easton DF, Mann GJ, Pupo GM, Newman B, Chenevix-Trench G, Szabo C, Southey M, Renard H, Odefrey F, Lynch H, et al. A genome wide linkage search for breast cancer susceptibility genes. *Genes, chromosomes & cancer* 2006;45:646-55.
333. Aloraifi F, Boland MR, Green AJ, Geraghty JG. Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. *Surgical oncology* 2015;24:100-9.
334. Snape K, Ruark E, Tarpey P, Renwick A, Turnbull C, Seal S, Murray A, Hanks S, Douglas J, Stratton MR, Rahman N. Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast cancer research and treatment* 2012;134:429-33.
335. Thompson ER, Doyle MA, Ryland GL, Rowley SM, Choong DY, Tothill RW, Thorne H, Barnes DR, Li J, Ellul J, Philip GK, Antill YC, et al. Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS genetics* 2012;8:e1002894.

336. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM, Roeb W, Agnew KJ, Stray SM, Wickramanayake A, Norquist B, Pennington KP, et al. 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* 2012;108:18032-7.
337. Honrado E, Osorio A, Milne RL, Paz MF, Melchor L, Cascon A, Urioste M, Cazorla A, Diez O, Lerma E, Esteller M, Palacios J, et al. Immunohistochemical classification of non-BRCA1/2 tumors identifies different groups that demonstrate the heterogeneity of BRCA families. *Mod Pathol* 2007;20:1298-306.
338. Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genetic epidemiology* 2001;21:1-18.
339. Cui J, Antoniou AC, Dite GS, Southey MC, Venter DJ, Easton DF, Giles GG, McCredie MR, Hopper JL. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *American journal of human genetics* 2001;68:420-31.
340. Aloraifi F, Alshehhi M, McDevitt T, Cody N, Meany M, O'Doherty A, Quinn CM, Green AJ, Bracken A, Geraghty JG. Phenotypic analysis of familial breast cancer: comparison of BRCAx tumors with BRCA1-, BRCA2-carriers and non-familial breast cancer. *Eur J Surg Oncol* 2015;41:641-6.
341. Didraga MA, van Beers EH, Joosse SA, Brandwijk KI, Oldenburg RA, Wessels LF, Hogervorst FB, Ligtenberg MJ, Hoogerbrugge N, Verhoef S, Devilee P, Nederlof PM. A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients. *Breast cancer research and treatment* 2015;130:425-36.
342. Mao Y, Varoglu M, Sherman DH. Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564. *Chemistry & biology* 1999;6:251-63.
343. Wang WD, Li R, Chen ZT, Li DZ, Duan YZ, Cao ZH. Cisplatin-controlled p53 gene therapy for human non-small cell lung cancer xenografts in athymic nude mice via the CARg elements. *Cancer science* 2005;96:706-12.
344. Jiang Y, Zhang XY, Sun L, Zhang GL, Duerksen-Hughes P, Zhu XQ, Yang J. Methyl methanesulfonate induces apoptosis in p53-deficient H1299 and Hep3B cells through a caspase 2- and mitochondria-associated pathway. *Environmental toxicology and pharmacology* 2012;34:694-704.
345. Ramirez-Solis R, Liu P, Bradley A. Chromosome engineering in mice. *Nature* 1995;378:720-4.
346. Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* 1988;336:684-7.
347. Friedrich G, Soriano P. Insertional mutagenesis by retroviruses and promoter traps in embryonic stem cells. *Methods in enzymology* 1993;225:681-701.
348. Court DL, Swaminathan S, Yu D, Wilson H, Baker T, Bubunenko M, Sawitzke J, Sharan SK. Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene* 2003;315:63-9.
349. Sharan SK, Pyle A, Coppola V, Babus J, Swaminathan S, Benedict J, Swing D, Martin BK, Tessarollo L, Evans JP, Flaws JA, Handel MA. BRCA2 deficiency in mice leads to meiotic impairment and infertility. *Development (Cambridge, England)* 2004;131:131-42.
350. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. *Nucleic acids research* 2005;33:e36.
351. Serra-Moreno R, Acosta S, Hernalsteens JP, Jofre J, Muniesa M. Use of the lambda Red recombinase system to produce recombinant prophages carrying antibiotic resistance genes. *BMC molecular biology* 2006;7:31.
352. Plumb JA, Milroy R, Kaye SB. Effects of the pH dependence of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide-formazan absorption on chemosensitivity determined by a novel tetrazolium-based assay. *Cancer research* 1989;49:4435-40.
353. Li J, Zou C, Bai Y, Wazer DE, Band V, Gao Q. DSS1 is required for the stability of BRCA2. *Oncogene* 2006;25:1186-94.
354. Bignell G, Micklem G, Stratton MR, Ashworth A, Wooster R. The BRC repeats are conserved in mammalian BRCA2 proteins. *Human molecular genetics* 1997;6:53-8.
355. Davies AA, Masson JY, McIlwraith MJ, Stasiak AZ, Stasiak A, Venkitaraman AR, West SC. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Molecular cell* 2001;7:273-82.
356. Ochiai K, Yoshikawa Y, Yoshimatsu K, Oonuma T, Tomioka Y, Takeda E, Arikawa J, Mominoki K, Omi T, Hashizume K, Morimatsu M. Valine 1532 of human BRC repeat 4 plays an important role in the interaction between BRCA2 and RAD51. *FEBS letters* 2011;585:1771-7.
357. Pellegrini L, Yu DS, Lo T, Anand S, Lee M, Blundell TL, Venkitaraman AR. Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 2002;420:287-93.
358. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, Elstrodt F, van Duijn C, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 2002;31:55-9.
359. Puente XS, Quesada V, Osorio FG, Cabanillas R, Cadinanos J, Fraile JM, Ordonez GR, Puente DA, Gutierrez-Fernandez A, Fanjul-Fernandez M, Levy N, Freije JM, et al. Exome sequencing and functional analysis identifies BANF1 mutation as the cause of a hereditary progeroid syndrome. *American journal of human genetics* 2011;88:650-6.

360. Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, Escaramis G, Jares P, Bea S, González-Díaz M, Bassaganyas L, Baumann T, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* 2011;475:101-5.
361. Fuentes Fajardo KV, Adams D, Mason CE, Sincan M, Tift C, Toro C, Boerkoel CF, Gahl W, Markello T. Detecting false-positive signals in exome sequencing. *Human mutation* 2012;33:609-13.
362. Shyr C, Tarailo-Graovac M, Gottlieb M, Lee JJ, van Karnebeek C, Wasserman WW. FLAGS, frequently mutated genes in public exomes. *BMC medical genomics* 2014;7:64.
363. Feng BJ, Tavtigian SV, Southey MC, Goldgar DE. Design considerations for massively parallel sequencing studies of complex human disease. *PLoS one* 2011;6:e23221.
364. Balmana J, Diez O, Castiglione M. BRCA in breast cancer: ESMO clinical recommendations. *Ann Oncol* 2009;20 Suppl 4:19-20.
365. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nature reviews* 2007;7:937-48.
366. John EM, Miron A, Gong G, Phipps AI, Felberg A, Li FP, West DW, Whittemore AS. Prevalence of pathogenic BRCA1 mutation carriers in 5 US racial/ethnic groups. *Jama* 2007;298:2869-76.
367. Tsao CC, Geisen C, Abraham RT. Interaction between human MCM7 and Rad17 proteins is required for replication checkpoint signaling. *The EMBO journal* 2004;23:4660-9.
368. Esteve PO, Terragni J, Deepti K, Chin HG, Dai N, Espejo A, Correa IR, Jr., Bedford MT, Pradhan S. Methyllysine reader plant homeodomain (PHD) finger protein 20-like 1 (PHF20L1) antagonizes DNA (cytosine-5) methyltransferase 1 (DNMT1) proteasomal degradation. *The Journal of biological chemistry* 2014;289:8277-87.
369. MacArthur DG, Tyler-Smith C. Loss-of-function variants in the genomes of healthy humans. *Human molecular genetics* 2010;19:R125-30.
370. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR, Cheadle JP. Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors. *Nature genetics* 2002;30:227-32.
371. van Loon B, Hubscher U. An 8-oxo-guanine repair pathway coordinated by MUTHYH glycosylase and DNA polymerase lambda. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106:18201-6.
372. Isidro G, Laranjeira F, Pires A, Leite J, Regateiro F, Castro e Sousa F, Soares J, Castro C, Gira J, Brito MJ, Medeira A, Teixeira R, et al. Germline MUTHYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas. *Human mutation* 2004;24:353-4.
373. Gomez-Fernandez N, Castellvi-Bel S, Fernandez-Rozadilla C, Balaguer F, Munoz J, Madrigal I, Mila M, Grana B, Vega A, Castells A, Carracedo A, Ruiz-Ponte C. Molecular analysis of the APC and MUTHYH genes in Galician and Catalanian FAP families: a different spectrum of mutations? *BMC medical genetics* 2009;10:57.
374. Nielsen M, Franken PF, Reinards TH, Weiss MM, Wagner A, van der Klift H, Kloosterman S, Houwing-Duistermaat JJ, Aalfs CM, Ausems MG, Brocker-Vriends AH, Gomez Garcia EB, et al. Multiplicity in polyp count and extracolonic manifestations in 40 Dutch patients with MYH associated polyposis coli (MAP). *Journal of medical genetics* 2005;42:e54.
375. Samadder NJ, Jasperson K, Burt RW. Hereditary and common familial colorectal cancer: evidence for colorectal screening. *Digestive diseases and sciences* 2015;60:734-47.
376. Out AA, Wasielewski M, Huijts PE, van Minderhout IJ, Houwing-Duistermaat JJ, Tops CM, Nielsen M, Seynaeve C, Wijnen JT, Breuning MH, van Asperen CJ, Schutte M, et al. MUTHYH gene variants and breast cancer in a Dutch case-control study. *Breast cancer research and treatment* 2012;134:219-27.
377. Rennert G, Lejbkovic F, Cohen I, Pinchev M, Rennert HS, Barnett-Griness O. MUTHYH mutation carriers have increased breast cancer risk. *Cancer* 2012;118:1989-93.
378. Wasielewski M, Out AA, Vermeulen J, Nielsen M, van den Ouweland A, Tops CM, Wijnen JT, Vasen HF, Weiss MM, Klijn JG, Devilee P, Hes FJ, et al. Increased MUTHYH mutation frequency among Dutch families with breast cancer and colorectal cancer. *Breast cancer research and treatment* 2010;124:635-41.
379. Ellingson MS, Hart SN, Kalari KR, Suman V, Schahl KA, Dockter TJ, Felten SJ, Sinnwell JP, Thompson KJ, Tang X, Vedell PT, Barman P, et al. Exome sequencing reveals frequent deleterious germline variants in cancer susceptibility genes in women with invasive breast cancer undergoing neoadjuvant chemotherapy. *Breast cancer research and treatment* 2015;153:435-43.
380. Beiner ME, Zhang WW, Zhang S, Gallinger S, Sun P, Narod SA. Mutations of the MYH gene do not substantially contribute to the risk of breast cancer. *Breast cancer research and treatment* 2009;114:575-8.
381. Vogt S, Jones N, Christian D, Engel C, Nielsen M, Kaufmann A, Steinke V, Vasen HF, Propping P, Sampson JR, Hes FJ, Aretz S. Expanded extracolonic tumor spectrum in MUTHYH-associated polyposis. *Gastroenterology* 2009;137:1976-85 e1-10.
382. Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Salguero I, Sherborne A, Chubb D, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nature genetics* 2013;45:136-44.
383. Valle L, Hernandez-Illan E, Bellido F, Aiza G, Castillejo A, Castillejo MI, Navarro M, Segui N, Vargas G, Guarinos C, Juarez M, Sanjuan X, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Human molecular genetics* 2014;23:3506-12.

384. Hansen MF, Johansen J, Bjernevoll I, Sylvander AE, Steinsbekk KS, Saetrom P, Sandvik AK, Drablos F, Sjurson W. A novel POLE mutation associated with cancers of colon, pancreas, ovaries and small intestine. *Familial cancer* 2015;14:437-48.
385. Rohlin A, Zagoras T, Nilsson S, Lundstam U, Wahlstrom J, Hulten L, Martinsson T, Karlsson GB, Nordling M. A mutation in POLE predisposing to a multi-tumour phenotype. *International journal of oncology* 2015;45:77-81.
386. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, Bojesen SE, Nordestgaard BG, Axelsson CK, Arias JI, Milne RL, Ribas G, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS genetics* 2008;4:e1000054.
387. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;39:870-4.
388. Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, Vermold B, Engel C, Meindl A, Arnold N, Hofmann W, Sutter C, et al. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *American journal of human genetics* 2008;82:937-48.
389. Evans DG, Ingham SL, Buchan I, Woodward ER, Byers H, Howell A, Maher ER, Newman WG, Lalloo F. Increased rate of phenocopies in all age groups in BRCA1/BRCA2 mutation kindred, but increased prospective breast cancer risk is confined to BRCA2 mutation carriers. *Cancer Epidemiol Biomarkers Prev*;22:2269-76.
390. Jin J, Cai Y, Yao T, Gottschalk AJ, Florens L, Swanson SK, Gutierrez JL, Coleman MK, Workman JL, Mushegian A, Washburn MP, Conaway RC, et al. A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *The Journal of biological chemistry* 2005;280:41207-12.
391. Wen H, Li Y, Malek SN, Kim YC, Xu J, Chen P, Xiao F, Huang X, Zhou X, Xuan Z, Mankala S, Hou G, et al. New fusion transcripts identified in normal karyotype acute myeloid leukemia. *PloS one* 2012;7:e51203.
392. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;10:57-63.
393. Hillier LW, Graves TA, Fulton RS, Fulton LA, Pepin KH, Minx P, Wagner-McPherson C, Layman D, Wylie K, Sekhon M, Becker MC, Fewell GA, et al. Generation and annotation of the DNA sequences of human chromosomes 2 and 4. *Nature* 2005;434:724-31.
394. Kazi JU, Kabir NN, Ronnstrand L. Brain-Expressed X-linked (BEX) proteins in human cancers. *Biochimica et biophysica acta* 2015;1856:226-33.
395. Kikuchi A, Yamamoto H. Tumor formation due to abnormalities in the beta-catenin-independent pathway of Wnt signaling. *Cancer science* 2008;99:202-8.
396. Lee YN, Gao Y, Wang HY. Differential mediation of the Wnt canonical pathway by mammalian Dishevelleds-1, -2, and -3. *Cellular signalling* 2008;20:443-52.
397. Dai W, Teodoridis JM, Zeller C, Graham J, Hersey J, Flanagan JM, Stronach E, Millan DW, Siddiqui N, Paul J, Brown R. Systematic CpG islands methylation profiling of genes in the wnt pathway in epithelial ovarian cancer identifies biomarkers of progression-free survival. *Clin Cancer Res* 2011;17:4052-62.
398. Mulligan AM, Couch FJ, Barrowdale D, Domchek SM, Eccles D, Nevanlinna H, Ramus SJ, Robson M, Sherman M, Spurdle AB, Wappenschmidt B, Lee A, et al. Common breast cancer susceptibility alleles are associated with tumour subtypes in BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2. *Breast Cancer Res* 2011;13:R110.
399. Greenman C, Stephens P, Smith R, Dalgleish GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, et al. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446:153-8.
400. Dutt A, Salvesen HB, Chen TH, Ramos AH, Onofrio RC, Hatton C, Nicoletti R, Winckler W, Grewal R, Hanna M, Wyhs N, Ziaugra L, et al. Drug-sensitive FGFR2 mutations in endometrial carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 2008;105:8713-7.
401. Devi S, Markandeya Y, Maddodi N, Dhingra A, Vardi N, Balijepalli RC, Setaluri V. Metabotropic glutamate receptor 6 signaling enhances TRPM1 calcium channel function and increases melanin content in human melanocytes. *Pigment cell & melanoma research* 2013;26:348-56.
402. Ozturk OG, Gun FD, Polat G. Killer cell immunoglobulin-like receptor genes in patients with breast cancer. *Medical oncology (Northwood, London, England)* 2012;29:511-5.
403. Arnheim L, Dillner J, Sanjeevi CB. A population-based cohort study of KIR genes and genotypes in relation to cervical intraepithelial neoplasia. *Tissue antigens* 2005;65:252-9.
404. Shi YK, Yu YP, Tseng GC, Luo JH. Inhibition of prostate cancer growth and metastasis using small interference RNA specific for minichromosome complex maintenance component 7. *Cancer gene therapy* 2010;17:694-9.
405. Tsuji T, Ficarro SB, Jiang W. Essential role of phosphorylation of MCM2 by Cdc7/Dbf4 in the initiation of DNA replication in mammalian cells. *Molecular biology of the cell* 2006;17:4459-72.
406. Orr SJ, Gaymes T, Ladon D, Chronis C, Czepulkowski B, Wang R, Mufti GJ, Marcotte EM, Thomas NS. Reducing MCM levels in human primary T cells during the G(0)-->G(1) transition causes genomic instability during the first cell cycle. *Oncogene* 2010;29:3803-14.
407. Cortez D, Glick G, Elledge SJ. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101:10078-83.

408. Venesio T, Molatore S, Cattaneo F, Arrigoni A, Risio M, Ranzani GN. High frequency of MYH gene mutations in a subset of patients with familial adenomatous polyposis. *Gastroenterology* 2004;126:1681-5.
409. Wrzeszczynski KO, Varadan V, Byrnes J, Lum E, Kamalakaran S, Levine DA, Dimitrova N, Zhang MQ, Lucito R. Identification of tumor suppressors and oncogenes from genomic and epigenetic features in ovarian cancer. *PLoS one* 2011;6:e28503.
410. Hampton OA, Den Hollander P, Miller CA, Delgado DA, Li J, Coarfa C, Harris RA, Richards S, Scherer SE, Muzny DM, Gibbs RA, Lee AV, et al. A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. *Genome research* 2009;19:167-77.
411. Schulte I, Batty EM, Pole JC, Blood KA, Mo S, Cooke SL, Ng C, Howe KL, Chin SF, Brenton JD, Caldas C, Howarth KD, et al. Structural analysis of the genome of breast cancer cell line ZR-75-30 identifies twelve expressed fusion genes. *BMC genomics* 2012;13:719.
412. Yoon JK, Kim do H, Koo JS. Implications of differences in expression of sarcosine metabolism-related proteins according to the molecular subtype of breast cancer. *Journal of translational medicine* 2014;12:149.
413. Liu ZY, Wang JY, Liu HH, Ma XM, Wang CL, Zhang XP, Tao YQ, Lu YC, Liao JC, Hu GH. Retinoblastoma protein-interacting zinc-finger gene 1 (RIZ1) dysregulation in human malignant meningiomas. *Oncogene* 2013;32:1216-22.
414. Shimura H, Mori N, Wang YH, Okada M, Motoji T. Aberrant methylation and decreased expression of the RIZ1 gene are frequent in adult acute lymphoblastic leukemia of T-cell phenotype. *Leukemia & lymphoma* 2012;53:1599-609.
415. Tokumaru Y, Nomoto S, Jeronimo C, Henrique R, Harden S, Trink B, Sidransky D. Biallelic inactivation of the RIZ1 gene in human gastric cancer. *Oncogene* 2003;22:6954-8.
416. He L, Yu JX, Liu L, Buysse IM, Wang MS, Yang QC, Nakagawara A, Brodeur GM, Shi YE, Huang S. RIZ1, but not the alternative RIZ2 product of the same gene, is underexpressed in breast cancer, and forced RIZ1 expression causes G2-M cell cycle arrest and/or apoptosis. *Cancer research* 1998;58:4238-44.
417. Longtine MS, DeMarini DJ, Valencik ML, Al-Awar OS, Fares H, De Virgilio C, Pringle JR. The septins: roles in cytokinesis and other processes. *Current opinion in cell biology* 1996;8:106-19.
418. Russell SE, McIlhatton MA, Burrows JF, Donaghy PG, Chanduloy S, Petty EM, Kalikin LM, Church SW, McIlroy S, Harkin DP, Keilty GW, Cranston AN, et al. Isolation and mapping of a human septin gene to a region on chromosome 17q, commonly deleted in sporadic epithelial ovarian tumors. *Cancer research* 2000;60:4729-34.
419. Kalikin LM, Sims HL, Petty EM. Genomic and expression analyses of alternatively spliced transcripts of the MLL septin-like fusion gene (MSF) that map to a 17q25 region of loss in breast and ovarian tumors. *Genomics* 2000;63:165-72.
420. Ferrer CM, Reginato MJ. Sticking to sugars at the metastatic site: sialyltransferase ST6GalNac2 acts as a breast cancer metastasis suppressor. *Cancer discovery* 2014;4:275-7.
421. Murugaesu N, Irvani M, van Weverwijk A, Ivetic A, Johnson DA, Antonopoulos A, Fearn A, Jamal-Hanjani M, Sims D, Fenwick K, Mitsopoulos C, Gao Q, et al. An in vivo functional screen identifies ST6GalNac2 sialyltransferase as a breast cancer metastasis suppressor. *Cancer discovery* 2014;4:304-17.
422. Lorschach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* 2003;17:637-41.
423. Wu MZ, Chen SF, Nieh S, Benner C, Ger LP, Jan CI, Ma L, Chen CH, Hishida T, Chang HT, Lin YS, Montserrat N, et al. Hypoxia Drives Breast Tumor Malignancy through a TET-TNFalpha-p38-MAPK Signaling Axis. *Cancer research* 2015;75:3912-24.
424. Yang L, Yu SJ, Hong Q, Yang Y, Shao ZM. Reduced Expression of TET1, TET2, TET3 and TDG mRNAs Are Associated with Poor Prognosis of Patients with Early Breast Cancer. *PLoS one* 2015;10:e0133896.
425. Fu HL, Ma Y, Lu LG, Hou P, Li BJ, Jin WL, Cui DX. TET1 exerts its tumor suppressor function by interacting with p53-EZH2 pathway in gastric cancer. *Journal of biomedical nanotechnology* 2014;10:1217-30.
426. Neri F, Dettori D, Incarnato D, Krepelova A, Rapelli S, Maldotti M, Parlato C, Paliogiannis P, Oliviero S. TET1 is a tumour suppressor that inhibits colon cancer growth by derepressing inhibitors of the WNT pathway. *Oncogene* 2015;34:4168-76.
427. Dong XY, Sun X, Guo P, Li Q, Sasahara M, Ishii Y, Dong JT. ATBF1 inhibits estrogen receptor (ER) function by selectively competing with AIB1 for binding to the ER in ER-positive breast cancer cells. *The Journal of biological chemistry* 2010;285:32801-9.
428. Li M, Fu X, Ma G, Sun X, Dong X, Nagy T, Xing C, Li J, Dong JT. Atbf1 regulates pubertal mammary gland development likely by inhibiting the pro-proliferative function of estrogen-ER signaling. *PLoS one* 2012;7:e51283.
429. Sun X, Frierson HF, Chen C, Li C, Ran Q, Otto KB, Cantarel BL, Vessella RL, Gao AC, Petros J, Miura Y, Simons JW, et al. Frequent somatic mutations of the transcription factor ATBF1 in human prostate cancer. *Nature genetics* 2005;37:407-12.
430. Zhang Z, Yamashita H, Toyama T, Sugiura H, Ando Y, Mita K, Hamaguchi M, Kawaguchi M, Miura Y, Iwase H. ATBF1-a messenger RNA expression is correlated with better prognosis in breast cancer. *Clin Cancer Res* 2005;11:193-8.
431. Dong JT. Prevalent mutations in prostate cancer. *Journal of cellular biochemistry* 2006;97:433-47.
432. Smith JD, Hing AV, Clarke CM, Johnson NM, Perez FA, Park SS, Horst JA, Mecham B, Maves L, Nickerson DA, Cunningham ML. Exome sequencing identifies a recurrent de novo ZSWIM6 mutation associated with acromelic frontonasal dysostosis. *American journal of human genetics* 2014;95:235-40.

433. 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. 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. *Journal of medical genetics* 2008;45:438-46.
434. Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornero L, Esteban-Cardenosa E, Lastra E, Pagani F, Miner C, Velasco EA. 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.
435. Scully R, Livingston DM. In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* 2000;408:429-32.
436. Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch ED, Rowell SE, King MC. Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat Genet* 1994;8:399-404.
437. Phelan CM, Dapic V, Tice B, Favis R, Kwan E, Barany F, Manoukian S, Radice P, van der Luijt RB, van Nesselrooij BP, Chenevix-Trench G, kConFab, et al. Classification of BRCA1 missense variants of unknown clinical significance. *Journal of medical genetics* 2005;42:138-46.
438. Slaoui M, Razine R, Ibrahim A, Attaleb M, Mzibri ME, Amrani M. Breast cancer in Morocco: a literature review. *Asian Pac J Cancer Prev* 2014;15:1067-74.
439. De Haas H. Morocco's Migration Experience: a Transitional Perspective. *International Migration* 2007;45(4):39-70.
440. Bouwman P, van der Gulden H, van der Heijden I, Drost R, Klijn CN, Prasetyanti P, Pieterse M, Wientjens E, Seibler J, Hogervorst FB, Jonkers J. A high-throughput functional complementation assay for classification of BRCA1 missense variants. *Cancer discovery* 2013;3:1142-55.
441. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y)* 2013;339:819-23.
442. Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, Smogorzewska A. Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet* 2011;43:142-6.
443. Stoepker C, Hain K, Schuster B, Hilhorst-Hofstee Y, Rooimans MA, Steltenpool J, Oostra AB, Eirich K, Korthof ET, Nieuwint AW, Jaspers NG, Bettecken T, et al. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet* 2011;43:138-41.
444. Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillon J, Ramirez MJ, Pujol R, Casado JA, Banos R, et al. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *American journal of human genetics* 2013;92:800-6.
445. Fekairi S, Scaglione S, Chahwan C, Taylor ER, Tissier A, Coulon S, Dong MQ, Ruse C, Yates JR, 3rd, Russell P, Fuchs RP, McGowan CH, et al. Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* 2009;138:78-89.
446. Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capella G, Brunet J, et al. Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. *BMC cancer* 2011;12:84.
447. Osorio A, Bogliolo M, Fernandez V, Barroso A, de la Hoya M, Caldes T, Lasa A, Ramon y Cajal T, Santamarina M, Vega A, Quiles F, Lazaro C, et al. Evaluation of rare variants in the new fanconi anemia gene ERCC4 (FANCP) as familial breast/ovarian cancer susceptibility alleles. *Human mutation* 2013;34:1615-8.
448. Landwehr R, Bogdanova NV, Antonenkova N, Meyer A, Bremer M, Park-Simon TW, Hillemanns P, Karstens JH, Schindler D, Dork T. Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer. *Breast cancer research and treatment* 2011;130:1021-8.
449. Roberts TM, Kobor MS, Bastin-Shanower SA, Li M, Horte SA, Gin JW, Emili A, Rine J, Brill SJ, Brown GW. Slx4 regulates DNA damage checkpoint-dependent phosphorylation of the BRCT domain protein Rtt107/Esc4. *Molecular biology of the cell* 2006;17:539-48.
450. Gonzalez-Perez A, Lopez-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *American journal of human genetics* 2011;88:440-9.
451. Crossan GP, van der Weyden L, Rosado IV, Langevin F, Gaillard PH, McIntyre RE, Gallagher F, Kettunen MI, Lewis DY, Brindle K, Arends MJ, Adams DJ, et al. Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nature genetics* 2011;43:147-52.
452. Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ, Harper JW. Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* 2009;138:63-77.
453. Betts MJ RR. Amino acid properties and consequences of substitutions. *Bioinformatics of Geneticists* 2003;289-316.
454. Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP. A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 2008;105:10762-7.
455. Shah S, Kim Y, Ostrovskaya I, Murali R, Schrader KA, Lach FP, Sarrel K, Rau-Murthy R, Hansen N, Zhang L, Kirchoff T, Stadler Z, et al. Assessment of Mutations in Hereditary Breast Cancers. *PloS one* 2013;8:e66961.
456. Bakker JL, van Mil SE, Crossan G, Sabbaghian N, De Leeneer K, Poppe B, Adank M, Gille H, Verheul H, Meijers-Heijboer H, de Winter JP, Claes K, et al. Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases. *Human mutation* 2012;34:70-3.

457. Catucci I, Colombo M, Verderio P, Bernard L, Ficarazzi F, Mariette F, Barile M, Peissel B, Cattaneo E, Manoukian S, Radice P, Peterlongo P. Sequencing analysis of SLX4/FANCP gene in Italian familial breast cancer cases. *PLoS one* 2012;7:e31038.
458. 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. Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families. *Eur J Hum Genet* 2012;21:883-6.
459. Kurian AW, Ford JM. Multigene Panel Testing in Oncology Practice: How Should We Respond? *JAMA oncology* 2015;1:277-8.
460. Spinella JF, Healy J, Saillour V, Richer C, Cassart P, Ouimet M, Sinnett D. Whole-exome sequencing of a rare case of familial childhood acute lymphoblastic leukemia reveals putative predisposing mutations in Fanconi anemia genes. *BMC cancer* 2015;15:539.
461. Cavanagh H, Rogers KM. The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. *Hereditary cancer in clinical practice*;13:16.
462. Phelan CM, Iqbal J, Lynch HT, Lubinski J, Gronwald J, Moller P, Ghadirian P, Foulkes WD, Armel S, Eisen A, Neuhausen SL, Senter L, et al. Incidence of colorectal cancer in BRCA1 and BRCA2 mutation carriers: results from a follow-up study. *British journal of cancer*;110:530-4.
463. King TA, Li W, Brogi E, Yee CJ, Gemignani ML, Olvera N, Levine DA, Norton L, Robson ME, Offit K, Borgen PI, Boyd J. Heterogenic loss of the wild-type BRCA allele in human breast tumorigenesis. *Annals of surgical oncology* 2007;14:2510-8.
464. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. Genetic susceptibility to breast cancer. *Molecular oncology* 2010;4:174-91.
465. Tischkowitz M, Xia B. PALB2/FANCN: recombining cancer and Fanconi anemia. *Cancer research* 2010;70:7353-9.
466. Gregg SQ, Robinson AR, Niedernhofer LJ. Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. *DNA repair* 2011;10:781-91.
467. Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, et al. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 1996;86:811-22.
468. Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, van Leeuwen W, Theil AF, Vermeulen W, et al. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotrophic axis. *Nature* 2006;444:1038-43.
469. Ding DP, He XF, Zhang Y. Lack of association between XPG Asp1104His and XPF Arg415Gln polymorphism and breast cancer risk: a meta-analysis of case-control studies. *Breast cancer research and treatment* 2011;129:203-9.
470. Jorgensen TJ, Visvanathan K, Ruczinski I, Thuita L, Hoffman S, Helzlsouer KJ. Breast cancer risk is not associated with polymorphic forms of xeroderma pigmentosum genes in a cohort of women from Washington County, Maryland. *Breast cancer research and treatment* 2007;101:65-71.
471. Lee SA, Lee KM, Park WY, Kim B, Nam J, Yoo KY, Noh DY, Ahn SH, Hirvonen A, Kang D. Obesity and genetic polymorphism of ERCC2 and ERCC4 as modifiers of risk of breast cancer. *Experimental & molecular medicine* 2005;37:86-90.
472. Romanowicz-Makowska H, Smolarz B, Kulig A. [Polymorphisms in XRCC1 and ERCC4/XPF DNA repair genes and associations with breast cancer risk in women]. *Pol Merkuri Lekarski* 2007;22:200-3.
473. Smith TR, Levine EA, Perrier ND, Miller MS, Freimanis RI, Lohman K, Case LD, Xu J, Mohrenweiser HW, Hu JJ. DNA-repair genetic polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2003;12:1200-4.
474. Yang Z, Fang X, Pei X, Li H. Polymorphisms in the ERCC1 and XPF genes and risk of breast cancer in a Chinese population. *Genetic testing and molecular biomarkers* 2013;17:700-6.
475. de Laat WL, Sijbers AM, Odijk H, Jaspers NG, Hoeijmakers JH. Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Nucleic acids research* 1998;26:4146-52.
476. Kohlhase S, Bogdanova NV, Schurmann P, Bermisheva M, Khusnutdinova E, Antonenkova N, Park-Simon TW, Hillemanns P, Meyer A, Christiansen H, Schindler D, Dork T. Mutation analysis of the ERCC4/FANCP gene in hereditary breast cancer. *PLoS one* 2014;9:e85334.
477. Oldenburg RA, Kroeze-Jansema K, Kraan J, Morreau H, Klijn JG, Hoogerbrugge N, Ligtenberg MJ, van Asperen CJ, Vasen HF, Meijers C, Meijers-Heijboer H, de Bock TH, et al. The CHEK2*1100delC variant acts as a breast cancer risk modifier in non-BRCA1/BRCA2 multiple-case families. *Cancer research* 2003;63:8153-7.
478. Bergman A, Karlsson P, Berggren J, Martinsson T, Björck K, Nilsson S, Wahlstrom J, Wallgren A, Nordling M. Genome-wide linkage scan for breast cancer susceptibility loci in Swedish hereditary non-BRCA1/2 families: suggestive linkage to 10q23.32-q25.3. *Genes, chromosomes & cancer* 2007;46:302-9.
479. Huusko P, Juo SH, Gillanders E, Sarantaus L, Kainu T, Vahteristo P, Allinen M, Jones M, Rapakko K, Eerola H, Markey C, Vehmanen P, et al. Genome-wide scanning for linkage in Finnish breast cancer families. *Eur J Hum Genet* 2004;12:98-104.
480. Park DJ, Lesueur F, Nguyen-Dumont T, Pertesi M, Odefrey F, Hammet F, Neuhausen SL, John EM, Andrusis IL, Terry MB, Daly M, Buys S, et al. Rare mutations in XRCC2 increase the risk of breast cancer. *American journal of human genetics* 2012;90:734-9.

481. Wen H, Kim YC, Snyder C, Xiao F, Fleissner EA, Becirovic D, Luo J, Downs B, Sherman S, Cowan KH, Lynch HT, Wang SM. Family-specific, novel, deleterious germline variants provide a rich resource to identify genetic predispositions for BRCAx familial breast cancer. *BMC cancer* 2014;14:470.
482. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology* 2012;30:434-9.
483. Didraga MA, van Beers EH, Joosse SA, Brandwijk KI, Oldenburg RA, Wessels LF, Hogervorst FB, Ligtenberg MJ, Hoogerbrugge N, Verhoef S, Devilee P, Nederlof PM. A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients. *Breast cancer research and treatment*;130:425-36.
484. Larsen MJ, Thomassen M, Tan Q, Laenkholm AV, Bak M, Sorensen KP, Andersen MK, Kruse TA, Gerdes AM. RNA profiling reveals familial aggregation of molecular subtypes in non-BRCA1/2 breast cancer families. *BMC medical genomics* 2014;7:9.
485. He M, Hu X, Chen L, Cao AY, Yu KD, Shi TY, Kuang XY, Shi WB, Ling H, Li S, Qiao F, Yao L, et al. A recessive variant of XRCC4 predisposes to non- BRCA1/2 breast cancer in chinese women and impairs the DNA damage response via dysregulated nuclear localization. *Oncotarget* 2014;5:12218-32.
486. Yan J, Takahashi T, Ohura T, Adachi H, Takahashi I, Ogawa E, Okuda H, Kobayashi H, Hitomi T, Liu W, Harada KH, Koizumi A. Combined linkage analysis and exome sequencing identifies novel genes for familial goiter. *Journal of human genetics* 2012;58:366-77.
487. Norton N, Li D, Rampersaud E, Morales A, Martin ER, Zuchner S, Guo S, Gonzalez M, Hedges DJ, Robertson PD, Krumm N, Nickerson DA, et al. Exome sequencing and genome-wide linkage analysis in 17 families illustrate the complex contribution of TTN truncating variants to dilated cardiomyopathy. *Circulation* 2013;6:144-53.
488. Louis-Dit-Picard H, Barc J, Trujillano D, Miserey-Lenkei S, Bouatia-Naji N, Pylypenko O, Beaurain G, Bonnefond A, Sand O, Simian C, Vidal-Petiot E, Soukaseum C, et al. KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nature genetics* 2103;44:456-60, S1-3.
489. Osborn MJ, Gabriel R, Webber BR, DeFeo AP, McElroy AN, Jarjour J, Starker CG, Wagner JE, Joung JK, Voytas DF, von Kalle C, Schmidt M, et al. Fanconi anemia gene editing by the CRISPR/Cas9 system. *Human gene therapy* 2015;26:114-26.
490. Weber J, Ollinger R, Friedrich M, Ehmer U, Barenboim M, Steiger K, Heid I, Mueller S, Maresch R, Engleitner T, Gross N, Geumann U, et al. CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proceedings of the National Academy of Sciences of the United States of America* 2015;112:13982-7.
491. Dixon-Salazar TJ, Silhavy JL, Udpa N, Schroth J, Bielas S, Schaffer AE, Olvera J, Bafna V, Zaki MS, Abdel-Salam GH, Mansour LA, Selim L, et al. Exome sequencing can improve diagnosis and alter patient management. *Science translational medicine* 2012;4:138ra78.
492. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, McDonald MT, Meisler MH, Goldstein DB. Clinical application of exome sequencing in undiagnosed genetic conditions. *Journal of medical genetics* 2012;49:353-61.

ANEXOS

Artículo ANEXO: "Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes."

Menéndez M, Castellsagué J, Mirete M, Pros E, Feliubadaló L, Osorio A, Calaf M, Tornero E, Del Valle J, Fernández-Rodríguez J, **Quiles F**, Salinas M, Velasco A, Teulé A, Brunet J, Blanco I, Capellá G, Lázaro C.

Breast Cancer Research and Treatment. 2011 Jul 7.

Impact Factor (2012): 4.859

Contribución del doctorando: Al inicio de su tesis, el doctorando estuvo ayudando en este proyecto, a modo de entrenamiento, para aprender la metodología utilizada en nuestro laboratorio en el análisis de VSD a nivel de RNA. Para ello, estuvo trabajando bajo la supervisión de la Dra. Mireia Menendez en el análisis de algunas de las variantes estudiadas. También estuvo aprendiendo el uso de los programas de predicción *in silico* de *splicing* así como de los programas de análisis de secuencias.

Assessing the RNA effect of 26 DNA variants in the *BRCA1* and *BRCA2* genes

Mireia Menéndez · Joan Castellsagué · Marc Mirete · Eva Pros · Lúdia Feliubadaló · Ana Osorio · Mónica Calaf · Eva Tornero · Jesús del Valle · Juana Fernández-Rodríguez · Francisco Quiles · Mónica Salinas · Angela Velasco · Alex Teulé · Joan Brunet · Ignacio Blanco · Gabriel Capellá · Conxi Lázaro

Received: 7 April 2011 / Accepted: 24 June 2011 / Published online: 7 July 2011
© Springer Science+Business Media, LLC. 2011

Abstract Comprehensive genetic testing of the breast cancer susceptibility genes *BRCA1* and *BRCA2* identified approximately 16% of variants of unknown significance (VUS), a significant proportion of which could affect the correct splicing of the genes. Our aim is to establish a workflow for classifying VUS in these complex genes, the first stage of which is splicing analysis. We used a combined approach consisting of five in silico splicing prediction programs and RT-PCR analysis for a set of 26 variants not previously studied at the mRNA level and six variants that had already been studied, four of which were used as positive controls as they were found to affect the splicing of these genes and the other two were used as negative controls. We identified a splicing defect in 8 of the 26 newly studied variants and ruled out splicing alteration in the remaining 18 variants. The results for the four

positive and the two negative control variants were consistent with results presented in the literature. Our results strongly suggest that the combination of RNA analysis and in silico programs is an important step towards the classification of VUS. The results revealed a very high correlation between experimental data and in silico programs when using tools for predicting acceptor/donor sites but a lower correlation in the case of tools for identifying ESE elements.

Keywords *BRCA* · HBOCS · VUS · Splicing mutation · In silico prediction

Introduction

Breast cancer is the leading cause of cancer mortality and morbidity among women. It is currently estimated that 5–10% of all breast cancers are hereditary and attributable to mutations in highly penetrant susceptibility genes [1]. Germline mutations in the human breast cancer susceptibility genes *BRCA1* (MIM#113705) and *BRCA2* (MIM#600185) are responsible for most hereditary breast and ovarian cancers [2]. Extensive analyses of *BRCA1* and *BRCA2* have identified a large number of different disease-causing mutations, polymorphisms and variants of unknown significance (VUS) [3, 4]. More than 3,000 different sequence variants are currently listed in the Breast Cancer Information Core database (BIC: <http://research.nhgri.nih.gov/bic/>).

Most of the reported disease-associated alleles of *BRCA1* and *BRCA2* have been attributed to frameshift, nonsense or splice site alterations leading to truncated *BRCA1* or *BRCA2* proteins [5–7]. However, a large number of variants routinely encountered in clinical and research laboratories cannot be readily distinguished as

The authors Mireia Menéndez and Joan Castellsagué contributed equally to the study.

Electronic supplementary material The online version of this article (doi:10.1007/s10549-011-1661-5) contains supplementary material, which is available to authorized users.

M. Menéndez · J. Castellsagué · M. Mirete · E. Pros · L. Feliubadaló · M. Calaf · E. Tornero · J. del Valle · J. Fernández-Rodríguez · F. Quiles · M. Salinas · A. Velasco · A. Teulé · J. Brunet · I. Blanco · G. Capellá · C. Lázaro (✉)
Hereditary Cancer Program, Genetic Diagnosis Unit, Laboratori de Recerca Translacional, Institut Català d'Oncologia, Hospital Duran i Reynals—Bellvitge Biomedical Research Institute (ICO-IDIBELL), L'Hospitalet de Llobregat, Barcelona and Hospital Josep Trueta, IDIBGI, Gran Via 199-203, Girona 08908, Spain
e-mail: clazaro@iconcologia.net

A. Osorio
Human Cancer Genetics, Spanish National Cancer Centre (CNIO), Madrid, Spain

either disease-associated mutations or benign polymorphisms. These variants include missense variants leading to amino acid changes, variants located in intronic regions and silent mutations [8–12]; they are usually referred to as VUS and pose a real problem in genetic counseling. An increasing number of recent studies have demonstrated the effect of several variants on the correct splicing of the *BRCA1* and *BRCA2* genes [13–19].

The aim of this study is to determine the putative splicing effect of 26 DNA variants identified in *BRCA1* and *BRCA2* genetic testing genes in Spanish hereditary breast–ovarian cancer syndrome (HBOCS) families. Some of the studied variants have previously been classified as pathogenic mutations and some are thought to be VUS. We used a combination of in silico prediction tools and analyses of mRNA transcripts from lymphocyte cell cultures derived from carriers of these DNA variants to ascertain their effect on the mRNA splicing process and, if possible, their pathogenicity.

Materials and methods

Patients

Since its creation in 1999, the Hereditary Cancer Program at ICO (Institut Català d'Oncologia) has identified a set of high-risk families with suspected HBOCS. Following the Catalan Consensus *OncoGuidelines* (http://www.genecat.net/salut/depsan/pdf/og0601esgrr_din4.pdf) on genetic testing for this syndrome, the patients were tested for mutations in the *BRCA1* and *BRCA2* genes (unpublished data), after receiving appropriate genetic counseling and providing written informed consent (see more details in Supplemental Data). For this study, we selected a total of 26 DNA variants (Table 1, Fig. 1) located in these two genes and investigated their effect on the mRNA splicing process, using an additional four DNA variants as positive controls and two as negative controls. The variants were selected on the basis of RNA availability for splicing studies; some have previously been classified as pathogenic mutations, whereas others are classified as VUS or described for the first time in this study (see complete description in Table 1).

Mutation nomenclature

Human Genome Variation Society (HGVS) approved guidelines (<http://www.hgvs.org/mutnomen/>) were used for *BRCA1* and *BRCA2* nomenclature, although Breast Cancer Information Core (BIC; <http://research.nhgri.nih.gov/projects/bic/>) nomenclature is also shown. Reference sequences are detailed in the Supplemental Data. For mutation nomenclature (except for BIC naming, Table 1),

nucleotide 1 is the A of the ATG-translation initiation codon, in accordance with approved guidelines.

Cell cultures, RNA isolation and RT-PCR

Human lymphocytes from fresh and frozen samples were used. To prevent potential degradation of unstable transcripts by nonsense-mediated mRNA decay (NMD), cells were treated with puromycin. To improve the detection and interpretation of splicing aberrations, transcripts from carriers were compared with transcripts from control lymphocyte cultures. Total RNA was isolated using TRIzol reagent. Technical details are provided in the Supplemental Data.

Splicing prediction programs

Mutant and normal sequences were analyzed with several bioinformatic tools to identify potential splicing mutations. Disruption/creation of splice sites was evaluated with three programs (Table 2): Splice Site Prediction by Neural Network (SSPNN; http://www.fruitfly.org/seq_tools/splice.html), NetGene2 Server (<http://www.cbs.dtu.dk/services/NetGene2/>) and SoftBerry (<http://linux1.softberry.com/berry.phtml?topic=fssplice&group=programs&subgroup=gfind>). In the case of intronic mutations, predictions were not considered informative when the wild-type splice donor or acceptor site was not recognized (nr) by the program. For some exonic variants, scores at both the donor and the acceptor sites are shown (Table 2).

When the DNA variant was located in the coding region, two programs were used to determine whether an alteration of a putative exonic splicing enhancer (ESE) element could be observed: ESEfinder (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) and RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>).

Results

Description of selected DNA variants

A total of 32 DNA variants detected in our set of hereditary breast and ovarian cancer patients were selected for this study; 16 variants were located in *BRCA1* and 16 in *BRCA2* (Table 1; Fig. 1). Twenty of these variants had previously been reported in the BIC database or in the literature, whereas 12 are reported for the first time in this study. Prior to our RNA analysis, 12 could be classified as pathogenic mutations, whereas the rest were classified as VUS. Four variants for which aberrant splicing had already been described [20, 21] were included as positive controls (*BRCA1*: c.211A>G, c.212+1G>A and c.213-12A>G;

Table 1 Description of 32 DNA variants in *BRCA1* and *BRCA2* genes and their effect on splicing

Gene	Variant (BIC)	Variant (HGVS)	BIC information	Protein prediction if splicing not altered	Effect on splicing	Effect on splicing (HGVS)
<i>BRCA1</i>	<i>IVS2-14C>T^a</i>	<i>c.81-14C>T</i>	Reported 52, VUS	–	No effect on splicing	<i>r.=</i>
	<i>IVS2-11delIT</i>	<i>c.81-11delIT</i>	Reported 6, VUS	–	No effect on splicing	<i>r.=</i>
	<i>226C>A</i>	<i>c.107C>A</i>	Not reported	<i>p.Ser36Tyr</i>	No effect on splicing	<i>r. 107c>a</i>
	<i>243delA</i>	<i>c.124delA</i>	Reported 1, PAT	<i>p.Ile42TyrfsX8</i>	No effect on splicing	<i>r. 124del</i>
	<i>330A>G^a</i>	<i>c.211A>G</i>	Reported 35, PAT	<i>p.Arg71Gly</i>	Deletion of 22 bp at the end of exon 5	<i>r. 191_212del; p.Cys64X</i>
	<i>IVS5+1G>A^a</i>	<i>c.212+1G>A</i>	Reported 6, PAT	–	Deletion of 22 bp at the end of exon 5	<i>r. 191_212del; p.Cys64X</i>
	<i>IVS5-12A>G^a</i>	<i>c.213-12A>G</i>	Reported 24, PAT	–	Insertion of 11 bp of intro 5 at the beginning of exon 6	<i>r. 212_213ins213-11_213-1;p.Arg71SerfsX21</i>
	<i>367T>G</i>	<i>c.248T>G</i>	No reported	<i>p.Val83Gly</i>	No effect on splicing	<i>r. 248t>g</i>
	<i>IVS6-24delAAT</i>	<i>c.302-24_302-22delAAT</i>	Not reported	–	No effect on splicing	<i>r.=</i>
	<i>IVS13+1G>T</i>	<i>c.4357+1G>T</i>	Reported 2, PAT	–	Skipping of exon 13	<i>r.4186_4357del; p.Arg1397TyrfsX2</i>
<i>BRCA2</i>	<i>IVS13+1delG</i>	<i>c.4357+1delG</i>	Not reported	–	Skipping of exon 13 and deletion of G	<i>r.[4186_4357del, 4357del]; p.[Arg1397TyrfsX2, Ala1453GlnfsX3]</i>
	<i>IVS14+1G>T</i>	<i>c.4484+1G>T</i>	Not reported	–	Skipping of exon 14	<i>r. 4358_4484del; p.Ala1453GlyfsX10</i>
	<i>4885G>A</i>	<i>c.4766G>A</i>	Reported 2, VUS	<i>p.Arg1589His</i>	No effect on splicing	<i>r. 4766 g>a</i>
	<i>IVS17+3A>G</i>	<i>c.5074+3A>G</i>	Reported 1, VUS	–	Skipping of exon 17 and insertion of 153 bp of intron 17	<i>r.[4987_5074del, 5074_5075ins5074+1_5074+153; 5074+3a>g]; p.[Val1665SerfsX8, p.Asp1692GlyfsX15]</i>
	<i>IVS20+6G>C</i>	<i>c.5277+6G>C</i>	Not reported	–	No effect on splicing	<i>r.=</i>
	<i>IVS21+15G>C</i>	<i>c.5332+15G>C</i>	Reported 1, VUS	–	No effect on splicing	<i>r.=</i>
	<i>IVS6+14C>T</i>	<i>c.516+14C>T</i>	Not reported	–	No effect on splicing	<i>r.=</i>
	<i>IVS6-19C>T^a</i>	<i>c.517-19C>T</i>	Reported 15, VUS	–	No effect on splicing	<i>r.=</i>
	<i>IVS 6-1G>A</i>	<i>c.517-1G>A</i>	Reported 1, PAT	–	Skipping of exon 7	<i>r.517_631del; p.Gly173SerfsX19</i>
	<i>1991A>G^a</i>	<i>c.1763A>G</i>	Not reported	<i>p.Asn588Ser</i>	Deletion of 147 bp at the end of exon 10	<i>r.1763_1909del; p.Asn588_Gly637delinsSer</i>
	<i>IVS 11-20T>A</i>	<i>c.6842-20T>A</i>	Reported 5, VUS	–	No effect on splicing	<i>r.=</i>
	<i>7675A>G</i>	<i>c.7447A>G</i>	Reported 2, VUS	<i>p.Ser2483Gly</i>	No effect on splicing	<i>r. 7447a>g</i>
<i>7708C>T</i>	<i>c.7480C>T</i>	Reported 10, PAT	<i>p.Arg2494X</i>	No effect on splicing	<i>r. 7480c>u</i>	
<i>7861G>A</i>	<i>c.7633G>A</i>	Reported 3, VUS	<i>p.Val2545Ile</i>	No effect on splicing	<i>r. 7633 g>a</i>	
<i>IVS18-1G>A</i>	<i>c.8332-1G>A</i>	Not reported	–	Deletion of 14 bp at the beginning of exon 19	<i>r.8332_8345del; p.Ile2778TyrfsX15</i>	

Table 1 continued

Gene	Variant (BIC)	Variant (HGVS)	BIC information	Protein prediction if splicing not altered	Effect on splicing	Effect on splicing (HGVS)
	IVS 21+11A>C	c.8754+11A>C	Not reported	–	No effect on splicing	r. =
	IVS 22-15T>G	c.8954-15T>G	Not reported	–	Insertion of 14 bp of intron 22 at the beginning of exon 23	r.8953_8954ins8954-14_8954-1; p.Val2985AlafsX8
	IVS22-5A>G	c.8954-5A>G	Not reported	–	Insertion of 4 bp of intron 22 at the beginning of exon 23	r.8953_8954ins8954-4_8954-1; p.Val2985AspfsX34
	9193delA	c.8965delA	Not reported	p.Ile2989PhefsX12	No effect on splicing	r. 9193del
	9254delATCAT	c.9026_9030del5	Reported 10, PAT	p.Tyr3009SerfsX7	No effect on splicing	r. 9026_9030del
	934C>T	c.9116C>T	Reported 4, VUS	p.Pro3039Leu	No effect on splicing	r. 9116C>u
	9485T>C	c.9257-16T>C	Reported 6, VUS	–	No effect on splicing	r. =

PAT pathogenic variant, BIC Breast Cancer Information Core, HGVS Human Genetic Variation Society

^a Mutations previously studied at mRNA level by other groups are in italics. To facilitate published data comparison, both the BIC and the HGVS have been used

BRCA2: c.1763A>G) and two variants with no effect on splicing were included as negative controls (*BRCA1*: c.81-14C>T; *BRCA2*: c.517-19C>T) [15, 18]. For the rest of the variants ($n = 26$), RNA data were not available and their effect at the transcriptional level was unknown. Of the 12 pathogenic mutations, 6 were located in the invariant GT and AG intronic dinucleotides and 6 were located outside this region, in either intronic or coding regions. We used a combination of prediction tools and experimental RNA analysis to determine the splicing effect of each variant under study and to evaluate the usefulness of splicing prediction programs and the validity of the predictions for identifying disease-associated mutations.

Assessing splicing alterations by RT-PCR analysis

To assess the impact of each mutation on splicing, total RNA was isolated from 32 independent patients and several controls (cDNA primers provided in Supplemental Data Table 1). About 12 of the 32 variants led to aberrantly expressed transcripts: 7 in *BRCA1* and 5 in *BRCA2*. The remaining 20 variants had no apparent effect on mRNA splicing (Table 1). Of the 12 variants that led to aberrant transcripts, 6 affected the highly conserved donor (GT) or acceptor (AG) dinucleotide splice sites (*BRCA1*: c.212+1G>A, c.4357+1G>T, c.4357+1delG, c.4484+1G>T and *BRCA2*: c.517-1G>A and c.8332-1G>A), four were located in intronic sequences outside the invariant dinucleotide sites (*BRCA1*: c.5074+3A>G and c.213-12A>G, and *BRCA2*: c.8954-15T>G and c.8954-5A>G) and two were exonic variants that created a cryptic donor site. Although mutations located in the invariant dinucleotide sites were already classified as pathogenic mutations because of their position, the RNA effect was only known for the mutation c.212+1G>A [19]; the descriptions of the other mutations are newly reported in this study. Aberrant RNA splicing had already been described for the *BRCA1* mutations c.211A>G and c.213-12A>G, and for the *BRCA2* variant c.1763A>G, which were used as internal controls in our study [19–21].

mRNA sequence characterization of variants shown to affect splicing

All aberrant transcripts described for the first time in this study are described briefly in this section and represented in the corresponding figures. As controls, we analyzed four variants that were already known to alter correct splicing: three (*BRCA1*: c.211A>G, c.212+1G>A and c.213-12A>G) are reported in the BIC database as pathogenic mutations [19–21] and the fourth (*BRCA2* c.1763A>G) has been described as a VUS [19]. Two variants with no effect on splicing were included as negative controls

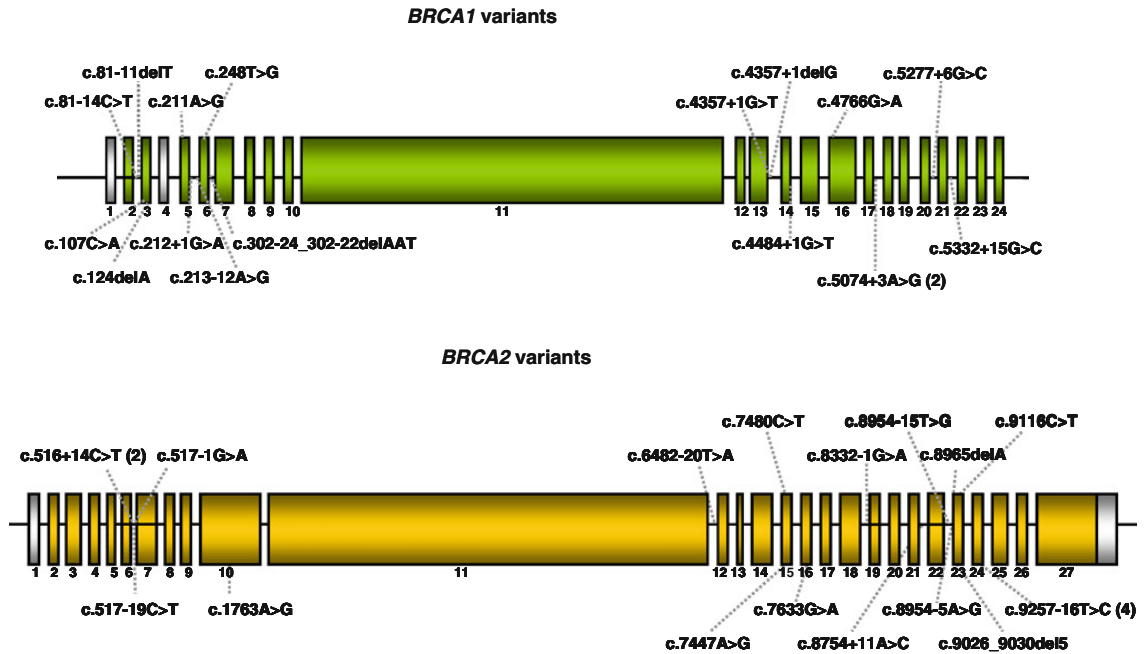


Fig. 1 Schematic representation of the *BRCA1* and *BRCA2* genes showing the 32 DNA variants considered in this study. The number of probands in our set of families for each DNA variant is shown in *parenthesis* when there is more than one

(*BRCA1*: c.81-14C>T; *BRCA2*: c.517-19C>T) [15, 18]. RNA analysis of the six variants is consistent with data reported in the literature.

The *BRCA1* c.4357+1G>T mutation is reported as a pathogenic mutation in the BIC database and has been described elsewhere [22], but its effect on the mRNA splicing process has not been assessed. RT-PCR analysis of RNA showed that exon 13 (172 bp) was deleted (r.4186_4357del). This mutant transcript was not observed in controls. The alteration is predicted to lead to a truncated protein (p.Arg1397TyrfsX2) (Fig. 2a). In addition, the cDNA assay identified an alternatively spliced *BRCA1* transcript in the mutation carrier and in all of the controls studied; this alternative transcript consists of the deletion of the first three bases of exon 14 (r.4358_4360del), leading to an in-frame deletion of the amino acid alanine at codon 1,453 (p.Ala1453del) (data not shown).

BRCA1 c.4357+1delG had not previously been reported. RT-PCR analysis showed a complex skipping pattern including the generation of two abnormal transcripts. The first consists of a skipping of exon 13, as observed for c.4357+1G>T, although in the case of c.4357+1delG this aberrant band is observed in a lower proportion; this frameshift deletion is predicted to generate a premature stop codon (p.Arg1397TyrfsX2). The second aberrant transcript consists of the deletion of a guanine at the last

base of exon 13 (r.4357del) because of the use of a cryptic donor splice site created by the DNA mutation; this mutated transcript also produces a truncated protein (p.Ala1453GlnfsX3) (Fig. 2b). The alternative transcript r.4358_4360del was also observed in this sample.

The *BRCA1* c.4484+1G>T mutation is not reported in the BIC database. This mutation arises in the invariant guanine in the consensus sequence of the 5' donor splice site of exon 14. cDNA analysis showed that this mutation abolishes the natural 5' splice site of exon 14, causing complete deletion of this exon (r.4358_4484del). The alteration is predicted to lead to a truncated protein (p.Ala1453GlyfsX10). This mutant transcript was not observed in controls (Fig. 3a).

The *BRCA1* c.5074+3A>G variant was identified in a woman with bilateral breast cancer; her mother, who was diagnosed with breast cancer at age 53 and with colorectal cancer at age 68, carried the same variant, which is reported once in the BIC database as a VUS. RNA analysis of this variant identified two aberrant transcripts that were not observed in controls: one shows a skipping of exon 17 (r.4987_5074del), the other shows an insertion of the first 153 bp of intron 17 (r.5074_5075ins5074+1_5074+153). These aberrant transcripts are predicted to lead to the truncated proteins p.Val1665SerfsX8 and p.Asp1692GlyfsX15, respectively (Fig. 3b). Therefore, the pathogenicity

Table 2 In silico analysis of 32 DNA variants in *BRCA1* and *BRCA2* genes

Gene	Variant (BIC)	Variant (HGVS)	SS	SSPNN		SoftBerry		NetGene2		ESEfinder	RESCUE-ESE	Effect on splicing
				Wild-type	Variant	Wild-type	Variant	Wild-type	Variant			
<i>BRCA1</i>	IVS2-14C>T	c.81-14C>T	A	0.52	0.58	nr	4.45	0.17	0.26	na	na	No effect on splicing
	IVS2-11delIT	c.81-11delIT	A	0.52	0.68	nr	nr	0.17	0.28	na	na	No effect on splicing
	226C>A	c.107C>A	A	0.52	0.58	nr	4.45	0.17	0.26	Multiple sites disrupted; 1 new site	No change	No effect on splicing
243delA	c.124delA	c.124delA	D	0.89	0.89	14.36	14.36	0.54	0.59	Multiple sites disrupted & new sites	No change	No effect on splicing
			A	0.52	0.58	nr	4.45	0.17	0.38			
330A>G	c.211A>G	c.211A>G	D	0.89	0.89	14.36	14.36	0.54	0.82	1 new sites	1 new site	Deletion of 22 bp at the end of exon 5
			D	0.92	0.43	9.74	nr	0.38	nr			
IVS5+1G>A	c.212+1G>A	c.212+1G>A	D	0.92	nr	9.74	nr	0.38	nr	na	na	Deletion of 22 bp at the end of exon 5
IVS5-12A>G	c.213-12A>G	c.213-12A>G	A	nr	nr	nr	nr/7.12	0.23	0.20	na	na	Insertion of 11 bp of intro 5 at the beginning of exon 6
367T>G	c.248T>G	c.248T>G	A	nr	nr	nr	nr	0.23	0.23/0.07	No change	No change	No effect on splicing
			D	nr	nr	10.44	10.44	0.41	0.45			
IVS6-24delAAT	c.302-24_302-22delAAT	c.302-24_302-22delAAT	A	0.99	0.99	12.28	12.68	0.52	0.64	na	na	No effect on splicing
			D	0.99	nr	10.86	nr	0.92	nr			
IVS13+1G>T	c.4357+1G>T	c.4357+1G>T	D	0.99	nr	10.86	nr	0.92	nr	na	na	Skipping of exon 13
			D	0.99	nr/0.70	10.86	nr	0.92	nr/0.55	na	na	Skipping of exon 13 and deletion of G
IVS13+1delG	c.4357+1delG	c.4357+1delG	D	0.66	0.66	nr	nr	0.96	0.95	1 new site	NC	No effect on splicing
			D	0.92	nr	9.18	nr	nr	nr		na	na
IVS14+1G>T	c.4484+1G>T	c.4484+1G>T	D	1	nr	14.92	nr	0.99	nr	na	na	Skipping of exon 14
4885G>A	c.4766G>A	c.4766G>A	D	0.66	0.66	nr	nr	0.96	0.95	1 new site	NC	No effect on splicing
IVS17+3A>G	c.5074+3A>G	c.5074+3A>G	D	0.92	nr	9.18	nr	nr	nr	na	na	Skipping of exon 17 and insertion of 153 bp of intron 17
IVS20+6G>C	c.5277+6G>C	c.5277+6G>C	D	0.93	0.93	13.24	12.26	0.92	0.90	na	na	No effect on splicing
IVS21+15G>C	c.5332+15G>C	c.5332+15G>C	D	1	1	14.64	14.64	1	0.99	na	na	No effect on splicing

Table 2 continued

Gene	Variant (BIC)	Variant (HGVS)	SS		SSPNN		SoftBerry		NetGene2		ESEfinder	RESCUE-ESE	Effect on splicing
			Wild-type	Variant	Wild-type	Variant	Wild-type	Variant	Wild-type	Variant			
BRCA2	IVS6+14C>T	c.516+14C>T	D	0.98	0.98	12.96	12.96	0.62	0.60	na	na	No effect on splicing	
	IVS6-19C>T	c.517-19C>T	A	0.98	0.98	14.55	14.05	0.78	0.77	na	na	No effect on splicing	
	IVS 6-1G>A	c.517-1G>A	A	0.98	nr/0.98	14.55	nr/10.43	0.78	nr/0.56	na	na	Skipping of exon 7	
	1991A>G	c.1763A>G	D	0.95	0.95/0.99	9.74	9.74/14.36	0.71	0.71/0.76	No change	2 new sites	Deletion of 147 bp at the end of exon 10	
	IVS 11-20T>A	c.6842-20T>A	A	0.92	0.88	7.92	7.42	0.43	0.31	na	na	No effect on splicing	
	7675A>G	c.7447A>G	A	0.90	0.93/0.72	nr	nr	0.14	0.16/0.94	I new site	I new site	No effect on splicing	
	7708C>T	c.7480C>T	A	0.90	0.90	nr	nr	0.14	nr	NC	No change	No effect on splicing	
	7861G>A	c.7633G>A	A	0.69	0.99	12.12	12.12	0.46	0.46	na	na	No effect on splicing	
	IVS18-1G>A	c.8332-1G>A	A	0.91	nr	4.80	nr/5.97	0.17	nr/0.14	I new site	No change	No effect on splicing	
	IVS 21+11A>C	c.8754+11A>C	D	0.98	0.98	10.72	10.72	1	1	na	na	Deletion of 14 bp at the beginning of exon 19	
	IVS 22-15T>G	c.8954-15T>G	A	0.53	nr/0.95	8.20	4.83/8.82	nr	nr/0.31	na	na	No effect on splicing	
	IVS22-5A>G	c.8954-5A>G	A	0.53	nr/0.89	8.20	nr/10.22	nr	nr/0.14	na	na	Insertion of 4 bp of intron 22 at the beginning of exon 23	
9193delA	c.8965delA	A	0.53	0.45	8.2/5.8	8.2/6.3	nr	0.14	I site disrupted	No change	No effect on splicing		
9254delATCAT	c.9026_9030del15	A	0.53	0.53	8.20	8.20	nr	nr	Multiple sites disrupted	No change	No effect on splicing		
9344C>T	c.9116C>T	D	0.57	0.57	nr	nr	nr	nr	I site broken; 2 sites increase	NC	No effect on splicing		
9485T>C	c.9257-16T>C	A	0.98	0.96	14.28	14.15	0.82	0.82	na	na	No effect on splicing		

SS splice site, D donor splice site, A acceptor splice site, nr wild-type splice site—not recognized, na not applied

Splice prediction programs (SSPNN, SoftBerry and NetGene2) predict wild-type scores from acceptor/donor sites. Sometimes a new cryptic splice site was predicted and indicated as “score wild-type splice site/score new cryptic splice site”. ESEfinder and RESCUE-ESE predict alterations of ESE elements

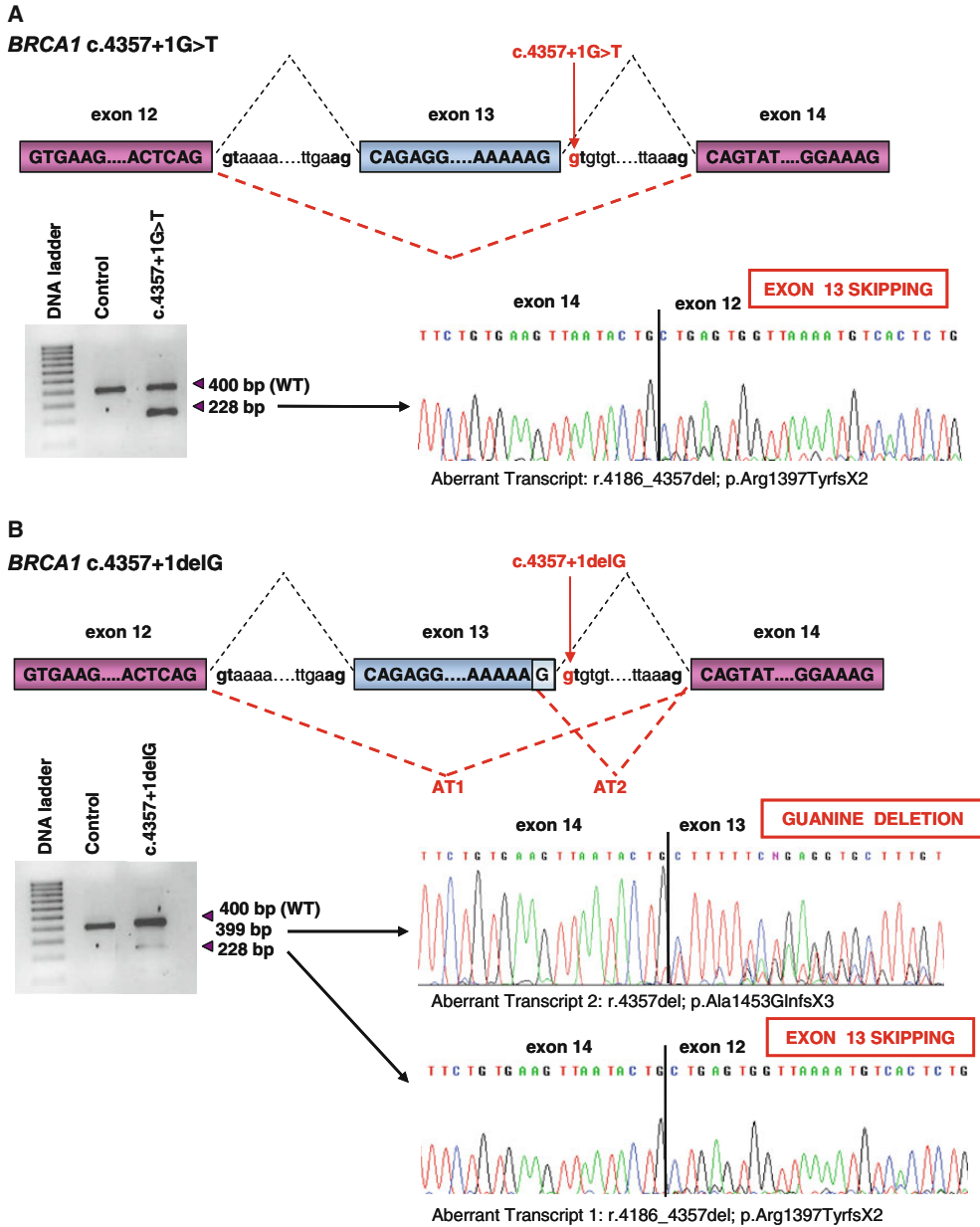


Fig. 2 a Scheme of the aberrant splicing caused by the *BRCA1* c.4357+1G>T variant, which showed the abolition of the wild-type donor site of exon 13. The cDNA amplification showed a double band, one corresponding to the wild-type transcript (400 bp) and the other corresponding to a transcript lacking exon 13 (288 bp). Sequencing of the aberrant cDNA showed complete skipping of

exon 13. **b** *BRCA1* 4357+1delG abolished the wild-type donor site of exon 13 and activated a cryptic donor splice site. The *upper band* contained the wild-type transcript and a transcript showing deletion of the last base of exon 13; the *lower band* contained a transcript lacking exon 13. The *upper sequence* showed overlapping after exon 14; the *lower sequence* showed skipping of exon 13

of this mutation may derive from the splicing anomalies detected and is supported by the co-segregation with breast cancer observed in this family.

The *BRCA2* c.517-1G>A mutation is reported once in the BIC database as a pathogenic mutation, but its effect on mRNA splicing has not been described. We performed RT-

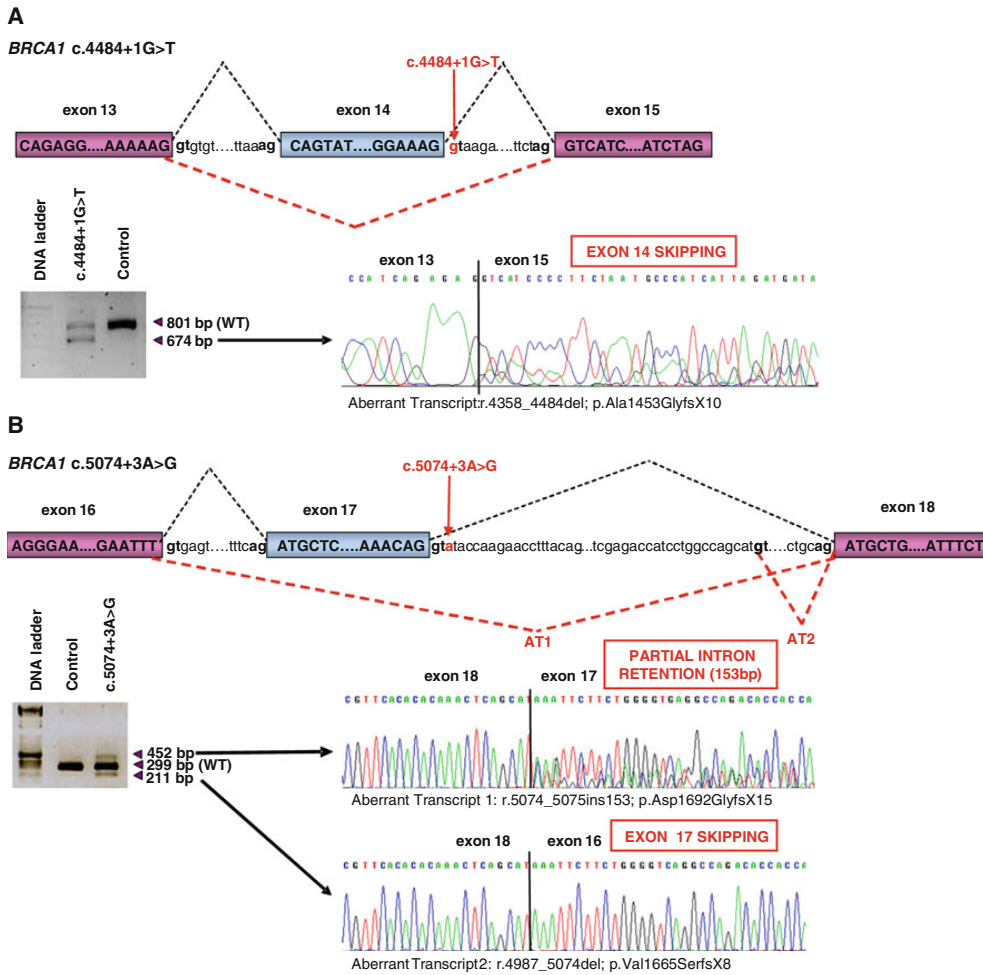


Fig. 3 **a** Scheme of the aberrant splicing caused by the *BRCA1* c.4484+1G>T variant, which showed the abolition of the wild-type donor site of exon 14. The cDNA amplification showed a double band, one corresponding to the wild-type transcript (801 bp) and the other corresponding to exon 14 skipping (674 bp). The sequence showed overlapping of the two sequences starting at the end of exon 13. **b** *BRCA1* c.5074+3A>G abolished the wild-type donor site of

exon 17 and activated a cryptic donor splice site. The RT-PCR analysis showed a triple band, corresponding to the wild-type transcript (299 bp), a transcript lacking exon 17 (211 bp) and a transcript showing a 153-bp insertion from intron 17 (452 bp), respectively. The *upper sequence* showed overlapping from exon 18; the *lower sequence* showed only exon 17 skipping

PCR to amplify the region from exon 6 to 10 and showed that this variant leads to a skipping of exon 7 (r.517_631del) (Fig. 4a). This deletion is predicted to lead to a truncated protein (p.Gly173SerfsX19) and is not observed in controls. In addition, using another pair of primers, we identified an alternatively spliced *BRCA2* transcript that is present in the mutation carrier and in all of the controls studied. This alternative transcript is generated by the activation of a cryptic donor splice site that leads to the skipping of most of exon 6 and all of exon 7

(r.478_631del). This transcript is predicted to lead to a truncated protein (p.Val160SerfsX19) (data not shown).

BRCA2 c.1763A>G is a missense variant (p.Asn588Ser) located in exon 10 that was identified in a woman diagnosed with breast cancer at age 31. It is not reported in the BIC database but was described elsewhere while this study was ongoing [19]. This variant creates a new cryptic donor site, which leads to an alternative transcript consisting of an in-frame deletion of 147 nucleotides (r.1763_1909del) that, in turn, produces an in-frame deletion of 49 amino

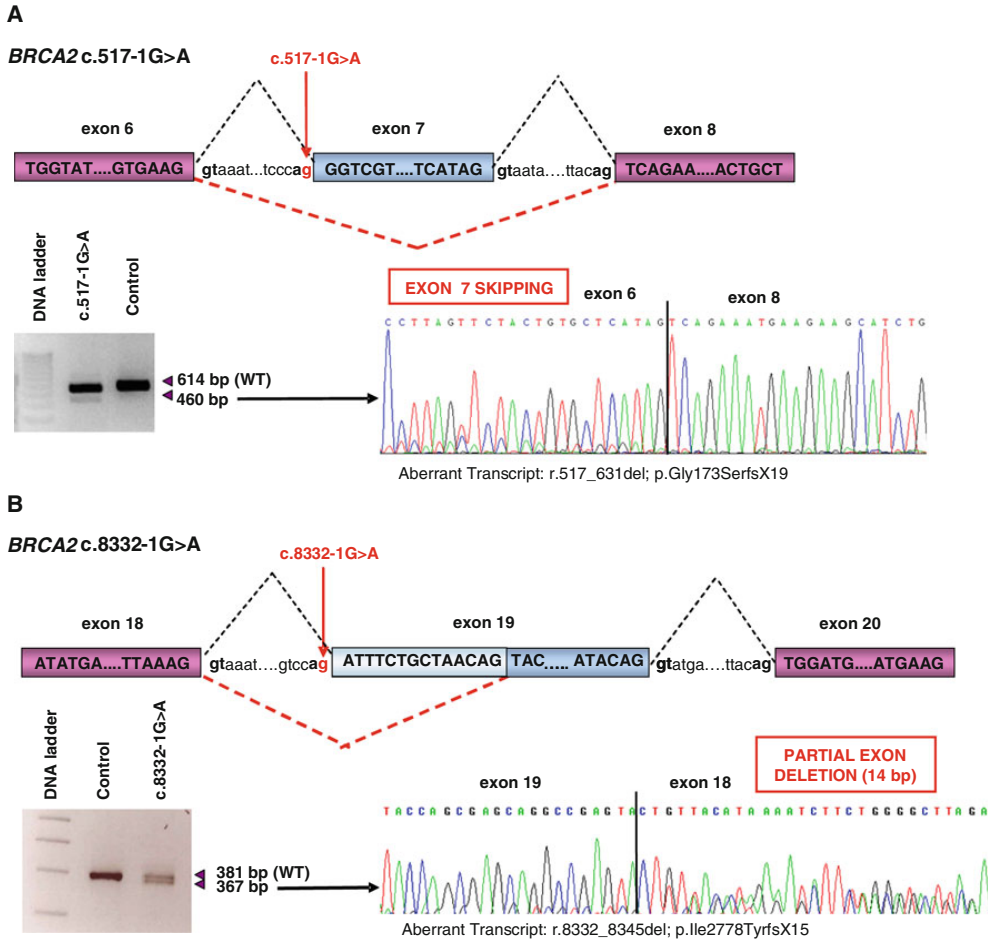


Fig. 4 a Scheme of the aberrant splicing caused by the *BRCA2* c.517-1G>A variant, which showed the abolition of the wild-type donor site of exon 19 and activation of a cryptic acceptor splice site. The RT-PCR analysis showed a double band, one corresponding to the wild-type transcript (614 bp) and the other showing a deletion at the beginning of exon 19 (460 bp). The sequence showed overlapping after exon 19

1G>A variant showed abolition of the wild-type donor site of exon 19 and activation of a cryptic acceptor splice site. The RT-PCR analysis showed a double band, one corresponding to the wild-type transcript (614 bp) and the other showing a deletion at the beginning of exon 19 (460 bp). The sequence showed overlapping after exon 19

acids with the insertion of a serine (p.Asn588_Gly637de-linsSer) (data not shown). This mutation retains its classification as a VUS because it is an in-frame deletion and does not affect any recognized *BRCA2* functional domain. Moreover, in the family considered here, this mutation does not appear to segregate with the disease, since the mother of the mutation carrier (also diagnosed with breast cancer at age 40) does not harbor the variant.

The *BRCA2* c.8332-1G>A mutation is not reported in the BIC database. This mutation arises in the invariant dinucleotide in the consensus sequence of the 3' acceptor splice site of exon 19. cDNA analysis of the mutation

showed that this mutation abolishes the natural 3' splice site of exon 19 and leads to the activation of a cryptic splice site 14 bp downstream. Consequently, the mutated transcript shows a 14-bp deletion at the beginning of exon 19 (r.8332_8345del). This change is predicted to generate a truncated protein (p.Ile2778Tyrf sX15) (Fig. 4b) and is not observed in normal controls.

The *BRCA2* c.8954-15T>G variant is not reported in the BIC database. It creates a new cryptic donor site that leads to an alternative transcript with an insertion of the last 14 nucleotides of intron 22 (r.8953_8954ins8954-14_8954-1), generating a truncated protein (p.Val2985Alaf sX8). This

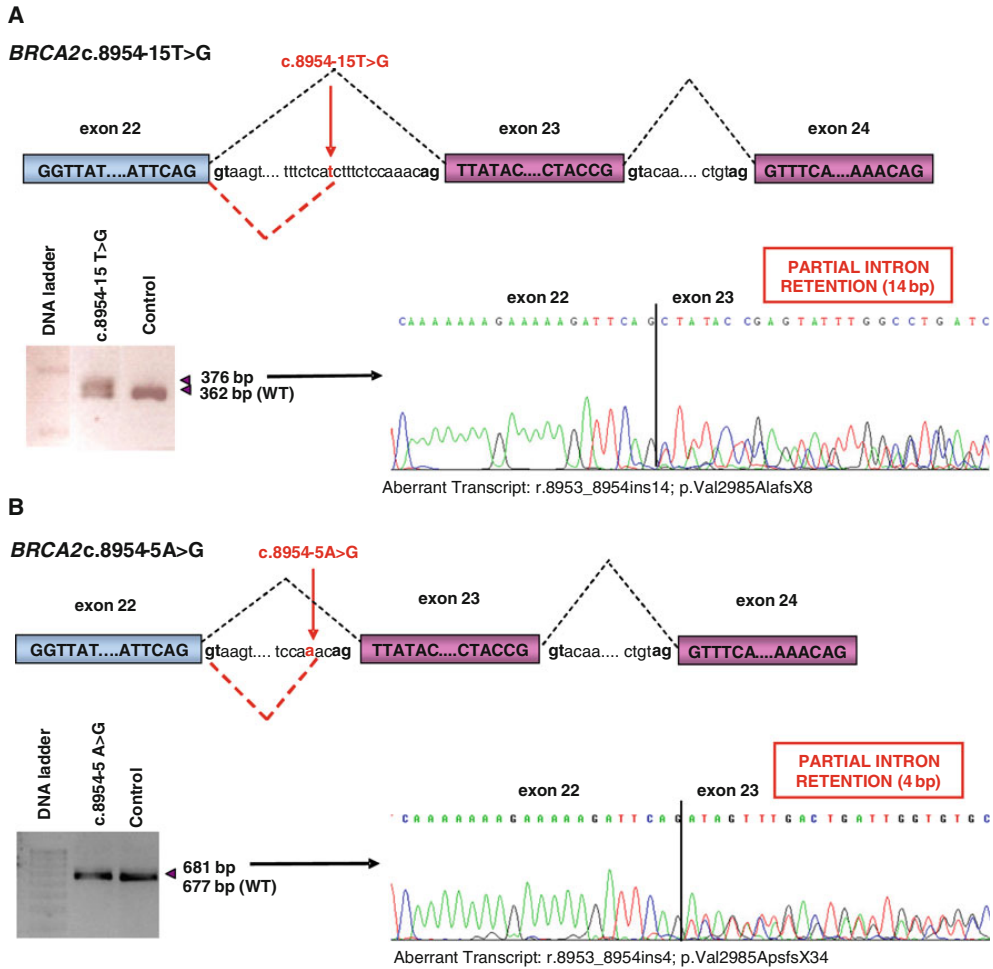


Fig. 5 a Scheme of the aberrant splicing caused by the *BRCA2* c.8954-15T>G variant, showing the creation of a new acceptor site at the end of intron 22. The RT-PCR analysis showed a double band, one corresponding to the wild-type transcript (362 bp) and the other showing a 14-bp intron retention (376 bp). The cDNA sequence showed these overlapping after exon 22. b The *BRCA2* c.8954-5A>G

variant created a new acceptor site at the end of intron 22. Although only one band was detected in the agarose gel, the sequence of this band revealed the presence of two transcripts, overlapping after exon 22, corresponding to the wild-type transcript (677 bp) and the transcript showing a 4-bp intronic retention (681 bp), respectively

aberrant transcript was not observed in controls (Fig. 5a). Segregation analysis showed that the two sisters with breast cancer (diagnosed at 29 and 40 years old, respectively) are carriers of this variant, which was inherited from their unaffected father whose history is non-informative of HBOCS.

The *BRCA2* c.8954-5A>G variant is not reported in the BIC database. It creates a novel cryptic donor site that leads to an alternative transcript with an insertion of the last four nucleotides of intron 22 (r.8953_8954ins8954-4_8954-1), generating a truncated protein (p.Val2985AspfsX34)

(Fig. 5b). This aberrant pattern is not observed in control samples. The variant was observed in a young female patient with metachronic bilateral cancer diagnosed at 36 and 42 years who has no relative affected by tumors associated with HBOCS.

The 20 remaining *BRCA1* and *BRCA2* variants showed the same RT-PCR agarose gel pattern as normal controls. Sequencing of the amplified RT-PCR products did not reveal any differences to wild-type sequences with the specific primer designed for each variant, which suggests that these mutations do not alter normal splicing (data not shown).

Splicing prediction programs

For the 32 mutations for which detailed molecular analysis of RNA was performed, we used three bioinformatic programs to predict possible changes in the efficiency of splice donor and acceptor sites and two programs specifically designed to identify exonic enhancer elements that could potentially alter correct splicing (Table 2). The first three programs (SSPNN, NatGene2 and SoftBerry) correctly predicted whether an alternative RNA splicing effect would be observed in the vast majority of cases (Table 2). We also used two prediction programs, ESEfinder and RESCUE-ESE, to ascertain the putative effect of our set of 12 exonic VUS on exonic splicing enhancers (Table 2).

Discussion

In genetic counseling, it is crucial to determine the pathogenicity of every DNA change identified through mutation screening protocols. Although it is widely accepted that mutations that generate putative truncated proteins (i.e., nonsense or frameshift mutations) or large deletions are disease-causing mutations, it is more difficult to determine the biological consequences of missense mutations, silent mutations and mutations in intronic regions [23]. Several approaches can help to overcome this difficulty: co-segregation of the VUS in the affected family, analysis of the evolutionary conservation of the observed DNA change, analysis of the frequency of the variant in a control population, analysis of the change in the protein structure and functional studies. A technically easy approach is to study the effect of a variant on mRNA as it provides a mechanism to study RNA splicing, RNA processing and RNA steady-state levels that could be affected by transcription or degradation. Our study of 32 *BRCA1* and *BRCA2* variants using experimental RNA analysis and in silico prediction programs revealed aberrant mRNA splicing in 12 different DNA variants, 6 of which had previously been described and were used as internal controls. In 6 cases, the DNA change was observed in an intronic region of the canonical dinucleotide splice sites and had already been classified as pathogenic; these variants were included in the study to determine their specific effect on splicing. In four cases, the mutations were intronic but located outside the invariable dinucleotide sites (*BRCA1*: c.5074+3A>G and c.213-12A>G and *BRCA2*: c.8954-15T>G and c.8954-5A>G) and in two cases, the variants were exonic. *BRCA1* c.211A>G is a known pathogenic mutation occurring before the last nucleotide of an exon and had previously been characterized at the mRNA level. *BRCA2* c.1763A>G was reported to alter splicing while our study was in the experimental stage [19],

but it is still considered to be VUS. Twenty of the mutations studied did not show any aberrant pattern under our study design, which suggests that they would be unlikely to affect the splicing of the affected region. By using a dual strategy combining experimental analysis of transcripts and in silico prediction tools, we were able to draw several conclusions and identify strategies for performing this type of study, discussed in greater detail in the following paragraphs.

In experimental RNA analysis, it is important to be able to distinguish between aberrant transcripts caused by a DNA mutation and aberrant alternative transcripts that are also found in normal controls. In our study, we identified r.4358_4360del (*BRCA1*) and r.478_631del (*BRCA2*) as alternative transcripts also found in controls. It is also important to bear in mind that a different proportion of transcripts is observed for some variants. Although quantitative analysis was beyond the scope of our study, it could be useful to include it in the classification algorithm in some cases. There are two clear examples in our set of variants: in the first, two different pathogenic mutations described as causing the same aberrant transcript in *BRCA1* (c.212+1G>A and c.211A>G) produced different proportions of mutated and wild-type transcripts (see Supplemental Data Fig. 1); in the second, although the *BRCA1* variants c.4357+1G>T and c.4357+1delG affect the same wild-type donor splice site, c.4357+1G>T only produces one aberrant transcript showing skipping of exon 13, whereas c.4357+1delG produces a second aberrant transcript that shows loss of the last nucleotide of exon 13 (Fig. 2a, b).

The in silico prediction programs used in this study to identify donor and acceptor sites produced largely consistent results, with only slight differences observed. For example, the programs correctly distinguished between *BRCA1* c.212+1G>A and c.211A>G changes, in agreement with our experimental data. Predictive differences were observed for the *BRCA1* variants c.4357+1G>T and c.4357+1delG, the latter of which was shown to create a novel donor site. The effects observed experimentally for the *BRCA2* variants c.8954-15T>G and c.8954-5A>G were described accurately by all three programs. However, in silico predictions for the *BRCA1* c.5074+3A>G and *BRCA2* c.517-1G>A variants were less reliable. In the case of the *BRCA1* c.5074+3A>G variant, the aberrant transcript with skipping of exon 17 was predicted by all of the programs, but the other aberrantly spliced transcript identified, which showed a 153-bp insertion from intron 17, was not predicted by any of them. The *BRCA2* c.517-1G>A mutation abolishes the consensus acceptor site and leads to exon 7 skipping; however, the splicing programs predicted the creation of a new cryptic acceptor site that our RNA analysis did not confirm. *BRCA2* c.1763A>G is a missense variant that creates a cryptic donor site predicted by all three programs, although the canonical donor site scores

were not reduced. For *BRCA2* c.8332-1G>A, all programs predicted destruction of the wild-type acceptor site. A novel acceptor site, located 14 bp downstream of the wild-type acceptor, was used and was predicted by all of the programs except SSPNN. Interestingly, the *BRCA2* c.9116C>T mutation was located in the base before the last nucleotide of exon 23, and SSPNN predicted that the C>T change completely abolishes the score at the 5' splice site, thus suggesting skipping of exon 23 as a possible mRNA effect. However, this prediction was not confirmed experimentally by amplifying the cDNA region containing exons 22–24 (data not shown). The two programs used to identify ESE elements (ESEfinder and RESCUE-ESE) produced more complex results that were difficult to interpret because they systematically recognize a set of different motifs that can act as putative regulators. Our experimental results appear to indicate that none of the splicing alterations identified involve an ESE but that they are all related to activation or disruption of donor and acceptor sites.

From our study, we conclude that the analysis of splicing alterations in all *BRCA1* and *BRCA2* variants would be useful not only for classifying VUS but also for determining the real effect of each variant on the mRNA splicing process. Our analysis identified aberrant splicing in 12 of the 32 variants studied and strongly suggests that all but one (*BRCA2* c.1763A>G) of these alterations are true pathogenic mutations, predisposing carriers to the development of HBOCS. The analysis facilitates genetic counseling, risk assessment and preventive surveillance strategies for family members of these probands. Our data showed that although in silico programs are useful, some discrepancies with experimental results can be expected. Therefore, when possible, it is preferable to combine predictive models for splicing mutations with RT-PCR to ascertain the real splicing effect of the variant under study.

Acknowledgments We thank the patients described in this study for their willingness to participate and Asociación Española contra el Cáncer (AECC), which recognizes our group as one of the Grupos Estables de Investigación Oncológica 2010. Contract grant sponsor: Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute and Autonomous Government of Catalonia. Contract grant numbers: ISCIIRETIC: RD06/0020/1051, RD06/0020/1050, RD06/0020/0028, 2009SGR290, PI10/01422.

Conflict of interest None.

References

- Fackenthal JD, Cartegni L, Krainer AR, Olopade OI (2002) *BRCA2* T2722R is a deleterious allele that causes exon skipping. *Am J Hum Genet* 71:625–631
- Fackenthal JD, Olopade OI (2007) Breast cancer risk associated with *BRCA1* and *BRCA2* in diverse populations. *Nat Rev Cancer* 7:937–948
- Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J et al (1998) Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 62:676–689
- 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
- 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
- Liu HX, Cartegni L, Zhang MQ, Krainer AR (2001) A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. *Nat Genet* 27:55–58
- Mazoyer S, Puget N, Perrin-Vidoz L, Lynch HT, Serova-Sinilnikova OM, Lenoir GM (1998) A *BRCA1* nonsense mutation causes exon skipping. *Am J Hum Genet* 62:713–715
- Krawczak M, Reiss J, Cooper DN (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* 90:41–54
- Ars E, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum Mol Genet* 9:237–247
- Teraoka SN, Telatar M, Becker-Catania S, Liang T, Onengut S, Tolun A, Chessa L, Sanal O, Bernatowska E, Gatti RA et al (1999) Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am J Hum Genet* 64:1617–1631
- Zatkova A, Messiaen L, Vandenbroucke I, Wieser R, Fonatsch C, Krainer AR, Wimmer K (2004) Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum Mutat* 24:491–501
- Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, Devilee P (2009) Intronic variants in *BRCA1* and *BRCA2* that affect RNA splicing can be reliably selected by splice-site prediction programs. *Hum Mutat* 30:107–114
- Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E, Menin C, Montagna M (2003) The *BRCA2* sequence variant IVS19+1G->A leads to an aberrant transcript lacking exon 19. *Cancer Genet Cytogenet* 141:175–176
- Brose MS, Volpe P, Paul K, Stopfer JE, Colligon TA, Calzone KA, Weber BL (2004) Characterization of two novel *BRCA1* germ-line mutations involving splice donor sites. *Genet Test* 8:133–138
- Claes K, Poppe B, Machackova E, Coene I, Foretova L, De Paeppe A, Messiaen L (2003) Differentiating pathogenic mutations from polymorphic alterations in the splice sites of *BRCA1* and *BRCA2*. *Genes Chromosomes Cancer* 37:314–320
- Keaton JC, Nielsen DR, Hendrickson BC, Pyne MT, Scheuer L, Ward BE, Brothman AR, Scholl T (2003) A biochemical analysis demonstrates that the *BRCA1* intronic variant IVS10-2A->C is a mutation. *J Hum Genet* 48:399–403
- Pyne MT, Brothman AR, Ward B, Pruss D, Hendrickson BC, Scholl T (2000) The *BRCA2* genetic variant IVS7+2T->G is a mutation. *J Hum Genet* 45:351–357
- Chen X, Truong TT, Weaver J, Bove BA, Cattie K, Armstrong BA, Daly MB, Godwin AK (2006) Intronic alterations in *BRCA1* and *BRCA2*: effect on mRNA splicing fidelity and expression. *Hum Mutat* 27:427–435
- Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornera 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(6):1957–1967
20. Vega A, Campos B, Bressac-De-Paillerets B, Bond PM, Janin N, Douglas FS, 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. *Hum Mutat* 17:520–521
21. Hoffman JD, Hallam SE, Venne VL, Lyon E, Ward K (1998) Implications of a novel cryptic splice site in the BRCA1 gene. *Am J Med Genet* 80:140–144
22. Borg A, Haile RW, Malone KE, Capanu M, Diep A, Torngren T, Teraoka S, Begg CB, Thomas DC, Concannon P et al (2010) Characterization of BRCA1 and BRCA2 deleterious mutations and variants of unknown clinical significance in unilateral and bilateral breast cancer: the WECARE study. *Hum Mutat* 31:E1200–E1240
23. Baralle D, Lucassen A, Buratti E (2009) Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep* 10:810–816

INFORME DE PUBLICACIONES

Yo, Conxi Lázaro García, en calidad de directora de la tesis doctoral presentada por Francisco Quiles Vidal, titulada **“Síndrome de cáncer de mama y ovario hereditario: estudio *in vitro* de las variantes *BRCA1* y *BRCA2* de significado biológico desconocido y búsqueda de nuevos genes responsables de este síndrome.”** hago constar que el doctorando ha participado activamente en el diseño y realización experimental de los trabajos que se incluyen en esta tesis doctoral así como en el análisis de resultados, discusión y presentación de los mismos en congresos o en artículos científicos. En cuanto a los resultados publicados en los que ha participado el doctorando, a continuación se detalla el factor de impacto de la revista en la fecha en la que fue publicado. En los trabajos en coautoría, se detalla el trabajo específico realizado por el doctorando:

Artículo 1

Whiley PJ¹, de la Hoya M, Thomassen M, Becker A, Brandão R, Pedersen IS, Montagna M, Menéndez M, **Quiles F**, Gutiérrez-Enríquez S, De Leeneer K, Tenés A, Montalban G, Tserpelis D, Yoshimatsu T, Tirapo C, Raponi M, Caldes T, Blanco A, Santamariña M, Guidugli L, de Garibay GR, Wong M, Tancredi M, Fachal L, Ding YC, Kruse T, Lattimore V, Kwong A, Chan TL, Colombo M, De Vecchi G, Caligo M, Baralle D, Lázaro C, Couch F, Radice P, Southey MC, Neuhausen S, Houdayer C, Fackenthal J, Hansen TV, Vega A, Diez O, Blok R, Claes K, Wappenschmidt B, Walker L, Spurdle AB, Brown MA; ENIGMA consortium. ***Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing.*** *Clin Chem.* 2014 Feb;60(2):341-52.

Factor de Impacto (2014): 7,911

Contribución del doctorando: La participación del doctorando en este trabajo consistió en el análisis de 8 VSD en *BRCA1/2* a nivel de RNA siguiendo la metodología implantada en nuestro laboratorio para este fin. Dicha metodología consiste en extracción de RNA a partir de linfocitos, generación de cDNA mediante PCR de retro-transcripción, amplificación por PCR de la región afectada por la variante y posterior detección de transcritos alternativos mediante la combinación de electroforesis en gel de agarosa y secuenciación del producto de PCR. Los resultados obtenidos para cada variante se compararon, de una forma cualitativa, con los resultados obtenidos en una serie de 11 controles y se clasificaron según el esquema de clasificación de variantes propuestos por la IARC. El doctorando se encargó de ejecutar toda la parte experimental de nuestro centro.

Artículo 2

Quiles F, Fernández-Rodríguez J, Mosca R, Feliubadaló L, Tornero E, Brunet J, Blanco I, Capellá G, Pujana MÀ, Aloy P, Monteiro A, Lázaro C. ***Functional and structural analysis of C-terminal BRCA1 missense variants. PLoS One.*** 2013 Apr 17;8(4)

Factor de Impacto (2013): 3,534

Artículo 3

Como directora de la presente tesis doctoral hago constar que la Dra. Juana Fernandez-Rodríguez, coautora del artículo “Analysis of *SLX4/FANCP* in non-*BRCA1/2*-mutated breast cancer families” no ha utilizado estos resultados para la realización de una tesis doctoral.

Fernández-Rodríguez J*, **Quiles F***, Blanco I, Teulé A, Feliubadaló L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capellá G, Brunet J, Lázaro C, Pujana MA. ***Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. BMC Cancer.*** 2012 Mar 8;12:84.

(*) Estos autores han contribuido equitativamente en este trabajo.

Factor de Impacto (2012): 3,333

Contribución del doctorando: La participación del doctorando consistió en la secuenciación mediante Sanger del gen *SLX4* en una serie de 96 pacientes BRCAx y posterior validación de diez de las variantes identificadas en una batería de 284 controles utilizando para ello sondas Kaspar. Este trabajo fue realizado conjuntamente con la Dra. Juana Fernández-Rodríguez.

Artículo 4

Osorio A¹, Bogliolo M, Fernández V, Barroso A, de la Hoya M, Caldés T, Lasa A, Ramón y Cajal T, Santamariña M, Vega A, **Quiles F**, Lázaro C, Díez O, Fernández D, González-Sarmiento R, Durán M, Piqueras JF, Marín M, Pujol R, Surrallés J, Benítez J. ***Evaluation of rare variants in the new fanconi anemia gene ERCC4 (FANCO) as familial breast/ovarian cancer susceptibility alleles. Hum Mutat.*** 2013 Dec;34(12):1615-8.

Factor de Impacto (2013): 5,050

Contribución del doctorando: La participación del doctorando en este trabajo consistió en el análisis mutacional del gen *ERCC4* en una serie de 96 pacientes BRCA1 de nuestro centro. Las variantes identificadas fueron analizadas mediante programas de predicción *in silico* para evaluar su posible impacto en el *splicing* o en la función de la proteína.

Artículo ANEXO

Menéndez M, Castellsagué J, Mirete M, Pros E, Feliubadaló L, Osorio A, Calaf M, Tornero E, Del Valle J, Fernández-Rodríguez J, **Quiles F**, Salinas M, Velasco A, Teulé A, Brunet J, Blanco I, Capellá G, Lázaro C. **"Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes."** *Breast Cancer Research and Treatment*. 2011 Jul 7.

Impact Factor (2012): 4,859

Contribución del doctorando: Al inicio de su tesis, el doctorando estuvo ayudando en este proyecto, a modo de entrenamiento, para aprender la metodología utilizada en nuestro laboratorio en el análisis de VSD a nivel de RNA. Para ello, estuvo trabajando bajo la supervisión de la Dra. Mireia Menendez en el análisis de algunas de las variantes estudiadas. También estuvo aprendiendo el uso de los programas de predicción *in silico* de *splicing* así como de los programas de análisis de secuencias.

Resultados sometidos a publicación

Artículo 1

Francisco Quiles¹, Mireia Menéndez¹, Eva Tornero¹, Jesús del Valle¹, Àlex Teulé¹, Sarai Palanca², Angel Izquierdo¹, Carolina Gómez¹, Olga Campos¹, Raúl Santamaria³, Joan Brunet¹, Gabriel Capellá¹, Lúdia Feliubadaló¹ and Conxi Lázaro¹. **"Investigating the effect of 28 BRCA1 and BRCA2 mutations on their related transcribed mRNA"**.

Breast Cancer Research and Treatment (sometido en Octubre de 2015)

Factor de Impacto (2015): 3,940

Artículo 2

Francisco Quiles¹ (BSc), Àlex Teulé¹ (MD), Nora Martinussen Tandstad² (MD), Lúdia Feliubadaló¹ (PhD), Eva Tornero¹ (BSc), Jesús del Valle¹ (PhD), Mireia Menéndez¹ (PhD), Mónica Salinas (BSc), Vibeke Wethe Rognlien² (MD), Angela Velasco (BSc), Angel

Izquierdo¹ (PhD), Gabriel Capellá¹ (MD,MPH), Joan Brunet¹ (MD, PhD), and Conxi Lázaro¹ (PhD). ***Identification of a founder BRCA1 mutation in Moroccan population.***

Clinical Genetics (sometido a publicación en Diciembre de 2015).

Factor de Impacto (2015): 3,931

Firma del Director,

Conxi Lázaro García

En Barcelona, a 26 de Noviembre de 2015