PROGRAMA DE DOCTORADO 3139 EN MEDICINA

"Nuevo abordaje en el diagnóstico del cáncer de mama y ovario hereditario mediante secuenciación masiva"

Tesis doctoral presentada por Vanesa Felipe Ponce

Directores: Dr. Javier Benítez Ortiz
Dra. Lucía Pérez Cabornero

Tutora: Dra. María Concepción López Ginés

Junio, 2023

VNIVERSITAT DE VALÈNCIA Facultat de Medicina i Odontologia

INFORME DIRECTORES/AS Y TUTOR/A PARA DEPÓSITO DE TESIS

Director (es) / Codirector (es):

1.- Apellidos y nombre Javier Benitez Ortiz N.I.F. 37255854V, Departamento/Instituto: Centro: Fundación Quaes, Valencia

2.- Apellidos y nombre: Lucía Pérez Cabornero N.I.F. 71105892N, Departamento/Instituto: Análisis clínicos/Bioquímica Clínica. Unidad de Referencia Regional de Enfermedades raras DiERCyL Centro: Complejo Asistencial Universitario de Salamanca

Tutor o tutora (si procede)

Apellidos y nombre: Concha López Ginés. N.I.F. 73751473H, Departamento/Instituto: Patología Centro: Universidad de Valencia

Directores/as y tutor/a, respectivamente, de la tesis doctoral: "Nuevo abordaje en el diagnóstico del cáncer de mama y ovario hereditario mediante secuenciación masiva"

de D/Dña. Vanesa Felipe Ponce,

estudiante del Programa de Doctorado **3139 Medicina** (RD99/2011) en Medicina de la Universitat de València, emiten informe favorable para la realización del depósito y la defensa de la tesis doctoral.

Fecha: 16/06/2023

BENITEZ ORTIZ FRANCISCO JAVIER - 37255854V

PEREZ CABORNERO LUCIA - 71105892N Fecha: 2023.06.20 09:35:37 +02'00'

MARIA CONCEPCION|LOPEZ GINES

Fdo.: Fdo.: Fdo.: Fdo.:

Director/a Director/a Director/a Tutor/a

ESCUELA DOCTORAL
UNIVERSITAT DE VALÈNCIA

Facultat de Medicina i Odontologia, *Av.Blasco Ibañez nº 15, 46010-Valencia , Tel. 963864100 - Fax: 963864173*
email: doctorado.medicina@uv.es
Web: www.uv.es/doctoralmedicina

Agradecimientos

En primer lugar, quisiera dar las gracias a mis directores de tesis Javier Benítez y Lucía Pérez, por haber tenido la idea de este proyecto, haber contado conmigo para llevarlo a cabo y por todo el apoyo, ayuda y paciencia que me han mostrado durante todo este largo tiempo.

A la empresa SSGG, por dejarme utilizar todos los datos de los que he dispuesto para la elaboración de esta tesis.

A todos mis compañeros de SSGG, y en especial a mis compañeros de Genética Molecular y Médica, a los que están y los que ya no están, porque sin ellos este trabajo tampoco hubiera sido posible. Alex, por fin estoy viendo la luz, todavía no me lo creo, eres un valioso amigo. Lucía… ¡sorpresa!, quien te lo iba a decir. Mención especial para la Dra. Sonia Santillán, gracias por animarme a embarcarme en este proyecto y gracias por apoyarme siempre de todo corazón.

A mis padres, por su ayuda, cariño, paciencia, estar siempre que los necesito y por inculcarme valores como son la responsabilidad y el esfuerzo. Sin vosotros hubiese sido imposible terminar de escribir esta tesis.

A Christian, por tu apoyo incondicional, creer en mí más que yo misma, por animarme en los momentos difíciles y por estar siempre a mi lado. No sé qué haría sin ti.

A Laia y Aylén, por quitaros tanto tiempo de estar con vuestra madre, y por ser la alegría y sentido de mi vida.

Gracias a todas aquellas personas que me han regalado su conocimiento y su tiempo.

ÍNDICE

ABREVIATURAS: *6*

ÍNDICE DE FIGURAS *7*

ÍNDICE DE TABLAS *9*

INTRODUCCIÓN *11*

1. Epidemiología del cáncer de mama y ovario. 11

1.1 Factores de riesgo del cáncer de mama: *12*

2. Síndrome de cáncer de mama y ovario hereditario 13

3. Genes de susceptibilidad con alta penetrancia 17

3.1 Estructura y función de los genes BRCA1 y BRCA2 ..... *17*

3.2 Variantes patogénicas en los genes BRCA1 y BRCA2 ..... *19*

3.3 Variantes de significado incierto en los genes BRCA1 y BRCA2. ..... *20*

3.4 Herramientas disponibles para la clasificación de variantes *21*

3.5 Estimaciones del riesgo de cáncer asociado a BRCA1 y BRCA2 ..... *25*

3.6 Implicaciones terapéuticas *26*

3.7 Consejo genético en el cáncer de mama y ovario hereditario *28*

3.8 Criterios de selección para el estudio genético de los genes BRCA1 y BRCA2. ..... *32*

4. Diagnóstico genético del cáncer de mama y ovario hereditario: de las técnicas convencionales a la secuenciación masiva 35

OBJETIVOS: *39*

MATERIALES Y MÉTODOS: *40*

5. Familias seleccionadas 40

6. Extracción de ácidos nucleicos 43

6.1 Extracción de ADN a partir de sangre periférica *43*

6.2 Cuantificación del ADN *43*

7. Rastreo de variantes en los genes *BRCA1* y *BRCA2* mediante secuenciación Sanger. 44

7.1 Diseño de oligos *44*

7.2 PCR *44*

7.3 Secuenciación Sanger *45*

7.4 Análisis de secuencias *45*

8. Rastreo de grandes deleciones/duplicaciones en los genes *BRCA1* y *BRCA2* mediante MLPA 46

9. Rastreo de variantes en los genes *BRCA1* y *BRCA2* mediante secuenciación masiva. 48

9.1 Captura de las regiones diana de los genes BRCA1 y BRCA2 y evaluación de fragmentos. *48*

9.2. Pirosecuenciación con 454 GS Junior de Roche. *49*

9.3 Secuenciación por terminación cíclica reversible con MiSeq de Ilumina. *50*

9.4 Pipelines de análisis de secuenciación masiva *50*

9.5 Evaluación de la cobertura del análisis. *52*

9.6 Identificación de inserciones/deleciones y Copy Number Variants (CNVs). *53*

10. Clasificación de variantes 54

RESULTADOS *57*

12. Objetivo 1.1: Valoración de la nueva estrategia de análisis: Cribado de variantes patogénicas en los genes *BRCA1* y *BRCA2* empleando la plataforma de secuenciación masiva *454 GS Junior* 57

12.1 Selección de muestras control *57*

12.2 Concordancia genotípica *59*

12.3 Evaluación de deleciones e inserciones (indels) mediante análisis de fragmentos. *61*

12.4 Precisión *63*

12.5 Análisis del software AVA (Roche) versus pipeline propio. *64*

12.6 Cobertura *65*

12.7 Reproducibilidad *66*

13. Objetivo 1.2: Valoración de la nueva estrategia de análisis: Cribado de variantes patogénicas en los genes *BRCA1* y *BRCA2* empleando la plataforma *MiSeq*. 68

13.1 Selección de muestras control *68*

13.2 Concordancia genotípica *70*

13.3 Análisis del software MiSeqReporter (Illumina) versus pipeline propio *71*

13.4 Cobertura y reproducibilidad *74*

13.5 Evaluación de cambios en el número de copias (CNVs) *74*

14. Objetivo 2.1: Flujo de trabajo propuesto 77

15. Objetivo 2.2: Comparativa de resultados por técnica de detección: secuenciación Sanger *versus* secuenciación masiva 79

16. Objetivo 3: Caracterización del espectro de variantes de los genes *BRCA1* y *BRCA2* en nuestra cohorte de 455 casos de pacientes con sospecha clínicas de cáncer de mama y ovario hereditario. 81

16.1 Variantes patogénicas/probablemente patogénicas detectadas en los genes *BRCA1 y BRCA2*. *82*

16.2 Cambios en el número de copias *91*

16.3 Variantes patogénicas recurrentes *93*

16.4 Variantes de nueva descripción *94*

16.5 Variantes detectadas en base a su localización *97*

16.6 Variantes de significado incierto *98*

17. Objetivo 4: Reclasificación de variantes. Estudio retrospectivo. 105

18. Correlación genotipo-fenotipo. 108

19. Correlación genotipo-tipo de cáncer del caso índice 109

19.1 Cáncer de mama *110*

19.2 Cáncer de mama bilateral *111*

19.3 Cáncer de mama en varones *111*

19.4 Cáncer de ovario *111*

19.5 Cáncer de mama y ovario *112*

19.6 Cáncer de mama triple negativo *112*

DISCUSIÓN *113*

20. Aplicación de un flujograma de trabajo basado en la secuenciación masiva al estudio de variantes en los genes *BRCA1* y *BRCA2*. 116

21. Caracterización del espectro de variantes en los genes *BRCA1* y *BRCA2*. 121

22. Conclusiones de la discusión.....128
23. Perspectiva futura.....128
CONCLUSIONES: *130*
BIBLIOGRAFÍA.....*131*
ANEXOS.....*147*
Anexo I: Secuencias de los oligos usados para la secuenciación Sanger.....147
Anexo II: Concentraciones de reactivos y condiciones de PCR para generar amplicones para los genes *BRCA1* y *BRCA2*. 149
Anexo III: Reactivos y condiciones de PCR para la secuenciación: 152
Anexo IV: Campos contenidos en los ficheros bioinformáticos, así como la clasificación de las variantes en diferentes categorías dependiendo de la frecuencia.....153
Anexo V: Descripción de los criterios de clasificación de variantes según el grado de evidencia de patogenicidad y de benignidad, así como el algoritmo que emplea la combinación de ambos para clasificar las variantes. 154
Anexo VI: Descripción de las variantes detectadas en las 12 muestras utilizadas para la validación en los genes *BRCA1* y *BRCA2*. 157
Anexo VII: Variantes reclasificadas a probablemente benignas o benignas detectadas en nuestra serie de pacientes en los genes *BRCA1* y *BRCA2*. 163

ABREVIATURAS:

ACMG*: American College of Medical Genetics*

AMP: Association for Molecular Pathology

AVA: *Amplicon Variant Analyzer*

BRCA1: Breast Cancer Susceptibility Gene 1

BRCA2: Breast Cancer Susceptibility Gene 2

BRCA1/2: genes *BRCA1* y *BRCA2*

CMOH: Cáncer de mama y ovario hereditario

CNV: *Copy Number Variation*

Depth: Profundidad. Es el número total de lecturas para esa posición.

GWAS: *Genome-Wide Association Studies*

HGVS: Human Genome Variation Society

IARC: *International Agency for Research on Cancer*

Indel: pequeñas inserciones o deleciones

MAF: *Minor Allele Frequency*

MLPA: *Multiplex Ligation-dependent probe Amplification*

NGS: *Next Generation Sequencing*

OCCR: *Ovarian cancer cluster región*

OR: *Odds Ratio*

PARP: Poli-ADP-ribosa polimerasa

PARPi: Inhibidores de Poli-ADP-ribosa polimerasa

PCR: Reacción en cadena de la polimerasa

RH: Recombinación homóloga

RR: riesgo relativo

SSGG: Sistemas Genómicos

SNP: *Single Nucleotide Polymorphism*

SNV: *Single Nucleotide Variant*

Var/Depth: Este parámetro indica el cociente entre el número de lecturas que reportan la variante y el número total de lecturas para esa posición, por lo que oscila entre valores de 0 y 1.

VB: Variante benigna

VP: Variante patogénica

VPB: Variante probablemente benigna

VPP: Variante probablemente patogénica

VSI: Variante de significado incierto

ÍNDICE DE FIGURAS

Figura 1. Modelo oligogénico del riesgo de cáncer de mama. 14

Figura 2. Esquema de los dominios funcionales de las proteínas BRCA1 y BRCA2. 19

Figura 3. Representación esquemática de la reparación del ADN mediada por PARP y la recombinación homóloga. 27

Figura 4. Esquema de los pasos seguidos por los profesionales de salud ante un posible caso de CMOH. 29

Figura 5. Esquema cronológico de la metodología empleada en la identificación de variantes en los genes BRCA1 y BRCA2 de los pacientes de nuestro estudio. 42

Figura 6. Flujograma empleado en la identificación de variantes en los genes BRCA1/2 mediante la metodología Sanger. 42

Figura 7. Esquema cronológico de la metodología empleada para la extracción de ADN de los pacientes de nuestro estudio..... 44

Figura 8. Esquema de las distintas etapas que componen la técnica de MLPA.. 47

Figura 9. Tecnología del kit BRCA MASTR. 48

Figura 10. Perfiles de las 5 reacciones multiplex del kit BRCA MASTR. 49

Figura 11. Esquema resumen de las variantes clasificadas por el software AVA. 51

Figura 12. Ejemplo de un alineamiento de lectura con el software AVA de una deleción. 51

Figura 13. Lectura media calculada para cada PCR y ordenadas según su posición en el gen para una carrera……..……..54

Figura 14. Visualización de los resultados distribuidos por valores medios de profundidad de lectura por amplicón ordenado por exones y genes en la misma carrera que la anterior……..…..55

Figura 15. Algoritmo empleado para la clasificación clínica de las variantes genéticas 56

Figura 16. Diagrama de Venn del análisis de las muestras control mediante el software AVA. 59

Figura 17. Confirmación mediante secuenciación Sanger del cambio puntual rs4987046, descrito en la línea HapMap NA12144 no estaba presente, es un falso positivo de esta línea celular. 60

Figura 18. Confirmación mediante secuenciación Sanger de que el cambio puntual rs1801406, no descrito en la línea HapMap NA12144 es un falso negativo de esta línea celular. 61

Figura 19. Análisis de fragmentos de las cinco muestras control, que presentaban indels como VP. 62

Figura 20. Análisis de fragmentos de una indel en zona de homopolímeros. 62

Figura 21. Análisis de fragmentos con el programa Peak Scanner..... 63

Figura 22. Diagrama de Venn del análisis de la línea HapMap mediante el software AVA (figura de la izquierda) y el pipeline propio (figura de la derecha)..... 65

Figura 23. Gráfico de la cobertura obtenida en la línea HapMap. 66

Figura 24. Gráfico de la cobertura obtenida en las muestras de validación..... 66

Figura 25. Perfiles de cobertura de la misma muestra secuenciada en dos carreras distintas. 67

Figura 26. Diagrama de Venn del análisis de las muestras control mediante el software MiSeqReporter. 71

Figura 27. Diagrama de Venn del análisis de las muestras control con el pipeline propio. ... 72

Figura 28. Representación de la cobertura por muestra ordenada por exones. 74

Figura 30. Detección de deleciones grandes a partir de los datos de profundidad de lectura por amplicón normalizados para una carrera. 76

Figura 31. Normalización del depth por fragmento amplificado y ordenado por genes de las muestras. 77

Figura 32. Flujo de trabajo implementado en el laboratorio de diagnóstico molecular..... 78

Figura 33. Tasa de diagnóstico 82

Figura 34. Análisis de fragmentos y análisis mediante MLPA. 93

*Figura 35. Árbol familiar de la familia 63 con una VPP no descrita c.4106_4107delinsGCT (p.Ala1369Glyfs*5) en el gen BRCA1. ..... 96*

*Figura 36. Detección del cambio complejo c.4106_4107delinsGCT (p.Ala1369Glyfs*5) en el gen BRCA1 mediante diferentes metodologías. 96*

Figura 37. VP/VPP detectadas en el gen BRCA1. 98

Figura 38. VP/VPP detectadas en el gen BRCA2. 98

Figura 39. Reclasificación de variantes. 105

Figura 40. Diagrama de barras de los tipos de cáncer de los casos índice, en función del resultado del test genético. 109

Figura 41. Representación en el gen BRCA1 de los tipos de cáncer de los casos índice portadores de VP en este gen. *124*

Figura 42. Representación en el gen BRCA2 de los tipos de cáncer de los casos índice portadores de VP en este gen. *125*

Figura 43. Reclasificación de VSI. *127*

ÍNDICE DE TABLAS

Tabla 1: Riesgos de cáncer en portadoras de VP en genes de alto, moderado y bajo riesgo. *15*

Tabla 2. Términos utilizados actualmente para describir las categorías de clasificación de variantes genéticas en cinco niveles y su recomendación clínica *24*

Tabla 3. Riesgo de distintos tipos de cánceres en población general y portadores de VP *25*

Tabla 4. PARPi aprobados por la FDA y la EMA para cáncer de mama y cáncer de ovario *27*

Tabla 5. Medidas preventivas *31*

Tabla 6. Comparativa de las distintas plataformas de secuenciación utilizadas en la realización de este trabajo. *37*

Tabla 7. Criterios de selección de pacientes homogenizados y tipos de cáncer asociados *41*

Tabla 8. Variantes en las muestras control detectadas previamente mediante Sanger y línea *HapMap* *57*

Tabla 9. Concordancia genotípica de una misma muestra secuenciada en dos carreras distintas *67*

Tabla 10. Variantes testadas para la validación de la secuenciación usando la plataforma *MiSeq* en 12 muestras *68*

Tabla 11. Variantes únicamente identificadas usando *MiSeqReporter* en las 12 muestras ... *72*

Tabla 13. Diferencias entre los softwares utilizados durante la validación de *MiSeq*. *73*

Tabla 14. Grandes deleciones testadas por MLPA e identificadas mediante análisis del *detph* *75*

Tabla 15. Porcentajes de VP según metodología utilizada *79*

Tabla 16. Comparativa de costes de las tres plataformas utilizadas en este trabajo *79*

Tabla 17. Variantes detectadas en nuestra cohorte. *81*

Tabla 18. VP y VPP detectadas y características de las mismas. *83*

Tabla 19. Cambios en el número de copias. *91*

Tabla 20. Cambios en el número de copias detectados en los genes BRCA1/2. ..... *91*

Tabla 21. VP recurrentes en nuestro estudio. *93*

Tabla 22. Técnica de detección de variantes de nueva descripción. *94*

Tabla 23: Variantes VSI detectadas en el gen *BRCA1* y características de las mismas. *99*

Tabla 24. Variantes VSI detectadas en el gen *BRCA2* y características de las mismas. *102*

Tabla 25. Variante reclasificada de VSI a VP detectadas en nuestra serie de pacientes en los genes *BRCA1/2* *107*

Tabla 26. Relación fenotipo-genotipo según los criterios clínicos *108*

Tabla 27. Porcentajes de VP y VPP identificadas en nuestra serie de pacientes en los genes *BRCA1/2*, según el fenotipo del caso índice. *109*

Tabla 28. Porcentajes de VP en pacientes con cáncer de mama *110*
Tabla 29. Técnicas de secuenciación empleadas.....*119*
Tabla 30. Estudios predictivos en las familias *125*

INTRODUCCIÓN

1. Epidemiología del cáncer de mama y ovario.

El cáncer constituye la segunda causa de muerte en España (25,2% de los fallecimientos), por detrás de las enfermedades del sistema circulatorio (26,4%). Sin embargo, su mortalidad en España ha experimentado un descenso de un 1,3% en el año 2020. Esta tendencia refleja las mejoras en la supervivencia de los pacientes con tumores debido a las actividades preventivas, las campañas de diagnóstico precoz, y los avances terapéuticos. En mujeres, el tumor responsable del mayor número de muertes en España fue el de mama, con 6.528 fallecimientos[1].

La incidencia del cáncer en **España** ha sido estimada en 279.260 casos para el año 2023, lo que supone una estabilización con respecto al año anterior. Al igual que se espera un incremento en la incidencia del cáncer a nivel mundial, en España se estima que en el año 2040 la incidencia alcance los 341.000 casos. Sin embargo, la realidad puede ser ligeramente diferente, ya que esta estimación no incluye todavía el posible efecto de la pandemia de la COVID-19. Estos datos han sido extraídos de "Las cifras del cáncer en España 2023" editado por la Sociedad Española de Oncología Médica (SEOM), con datos de la Red Española de Registros de Cáncer (REDECAN), *Global Cancer Observaroty* (CGO), Agencia Internacional de Investigación del Cáncer (IARC) de la Organización Mundial de la Salud (OMS), del Instituto Nacional de Estadística (INE) y de GLOBOCAN-2020[1]. Los cánceres más frecuentemente diagnosticados en España en 2023, independientemente del sexo, son los de colon y recto (42.721 nuevos casos), seguidos del cáncer de mama (35.001), pulmón (31.282), próstata (29.002) y vejiga urinaria (21.694).

El número absoluto de cánceres diagnosticados en España continúa en aumento desde hace décadas relacionado probablemente con el aumento poblacional, el envejecimiento de la población (la edad es un factor de riesgo fundamental en el desarrollo del cáncer), la exposición a factores de riesgo (como el tabaco, el alcohol, la contaminación, la obesidad, el sedentarismo entre otros muchos) y, en algunos tipos de cáncer como el colorrectal, de mama, cérvix o próstata, por el aumento en la detección precoz. De hecho, si los programas de detección precoz son adecuados, deberían conducir a un aumento del número de casos, pero también una reducción de la mortalidad.

Según el último informe de estadísticas globales de cáncer en 2020, (GLOBOCAN 2020), se estima que se diagnostican aproximadamente 2.261.419 nuevos casos de cáncer de mama en todo el mundo y se producen unas 684.996 muertes al año a nivel mundial[2]. El cáncer de mama, por tanto, es un problema de salud pública presente en todo el mundo. También afecta al hombre, aunque representa <1% del cáncer de mama.

Si bien el cáncer de mama presenta una supervivencia relativa de un 95% en el primer año del diagnóstico, es el tumor más prevalente en mujeres, y el riesgo de padecer cáncer de

mama a lo largo de la vida es aproximadamente de un 12% (una de cada ocho mujeres), en nuestro país. Por el contrario, el cáncer de ovario presenta una incidencia mucho menor (3.584 nuevos casos para el año 2023), sin embargo, presenta una supervivencia de un 70% para el primer año, y este valor decrece hasta el 38% a los 5 años (estos valores son similares a la media europea). A día de hoy, el cáncer de ovario sigue siendo un tumor con pronóstico desfavorable y no se ha detectado en las últimas décadas un incremento de la supervivencia[1], debido principalmente a que se diagnostica en estadios avanzados.

Todas estas cifras anteriores, reflejan la magnitud de la incidencia del cáncer de mama y cáncer de ovario en las sociedades de todo el mundo y la urgencia de medidas preventivas y de tratamiento. Gracias a los avances tecnológicos en ciencias biomédicas y salud, es posible una detección temprana de la enfermedad, y es posible iniciar el tratamiento en estadios iniciales, evitando la progresión de la enfermedad a un estado metastático. Por lo que el diagnóstico precoz, es la mejor herramienta para luchar contra esta enfermedad[3].

1.1 Factores de riesgo del cáncer de mama:

Los estudios epidemiológicos han identificado múltiples factores de riesgo tanto genéticos como hormonales o medioambientales. En cuanto al cáncer de mama en mujeres premenopáusicas destacan: la edad, la densidad mamaria, los antecedentes personales previos y sobre todo los antecedentes familiares[3]. Por ejemplo, estudios epidemiológicos muestran que el riesgo se duplica en parientes de primer grado de mujeres con cáncer de mama, sin embargo para el cáncer de ovario se triplican con familiares de primer grado afectas de esta enfermedad[4].

Entre los factores de riesgo se encuentran[3,5]:

- **Edad:** La incidencia y la mortalidad por cáncer de mama aumentan proporcionalmente con la edad. La enfermedad alcanza su máximo alrededor de los 60 años, con un aumento de la curva a partir de los cuarenta. La edad media de fallecimiento por cáncer de mama se encuentra alrededor de los 68 años.
- **Raza:** más frecuente en mujeres de raza blanca.
- **Menarquia temprana / menopausia tardía/ nuliparidad**: Los factores reproductivos que aumentan la exposición a estrógenos endógenos y por tanto aumenta el riesgo de cáncer de mama serían: una edad temprana en la menarquia y una edad avanzada en la menopausia o la nuliparidad (ausencia de embarazos).
- **Estilo de vida:** Alcohol, obesidad, tabaco.
- **Historia familiar:** El riesgo aumenta si existen familiares que hayan padecido la enfermedad, y este es mayor aún, si es un familiar de primer grado (madre, hermana o hija).
- **Antecedentes personales:** Una mujer que ha tenido cáncer de mama, tiene más riesgo de padecer otro cáncer en la mama contralateral.

- **Enfermedades previas de la mama:** Algunas enfermedades de la mama como la hiperplasia atípica o el carcinoma lobulillar *in situ* pueden aumentar el riesgo.
- **Terapia hormonal sustitutiva prolongada:** Durante el tratamiento hormonal sustitutivo, existe un mayor riesgo de carcinoma de mama, dependiendo de los tipos de preparados. Este riesgo, vuelve al riesgo habitual una vez que cesa el tratamiento.

2. Síndrome de cáncer de mama y ovario hereditario

El cáncer de mama y de ovario, al igual que otros tipos de cánceres son el resultado de la combinación de factores medioambientales y factores genéticos. Aproximadamente cerca del 10-30% de los cáncer de mama y cáncer de ovario presentan una agregación familiar, aunque solamente un 5-10% de los casos se estiman que son hereditarios, y se asocian a la presencia de una variante deletérea germinal en genes de predisposición a cáncer[6,7].

Uno de los objetivos principales en la lucha contra el cáncer, es la identificación de los genes implicados en el desarrollo tumoral. Las proteínas codificadas por estos genes, normalmente regulan procesos tales como la proliferación, diferenciación y muerte celular, aunque también se han asociado a la tumorogénesis las que intervienen en la reparación del ADN.

Hoy en día es sabido que la mayor parte de la susceptibilidad al cáncer sigue un modelo poligénico[8,9], según este modelo, variantes genéticas con distinto grado de penetrancia y de prevalencia en la población, darían lugar a un fenotipo común, que es el riesgo de cáncer de mama. Un riesgo elevado, se puede adquirir por herencia de una única variante con una alta penetrancia, de unas pocas variantes con penetrancia moderada o de un conjunto amplio de variantes con penetrancia baja. De esta forma, a medida que disminuye la penetrancia de cada variante se requerirá un mayor número de variantes de menor riesgo presentes en un mismo individuo para lograr el mismo nivel de riesgo individual. Por tanto, variantes de alta penetrancia estarán presentes en familias que se ajustan de una manera óptima a un patrón de herencia mendeliana de un único factor, mientras que, según disminuye la penetrancia de las variantes, los patrones de herencia familiar se alejarán de la herencia mendeliana.

La penetrancia de una variante genética que causa enfermedad, en este caso, cáncer es la proporción de portadores de esa variante que desarrollan la enfermedad; mientras que el riesgo relativo (RR) es la medida del riesgo de desarrollar una enfermedad en comparación con el riesgo de la población general. Una variante genética generalmente se define con alta penetrancia cuando el RR para el portador es ≥4, con penetrancia moderada cuando el RR está entre 2 y 4, y con baja penetrancia cuando el RR está entre 1 y 2[10,11].

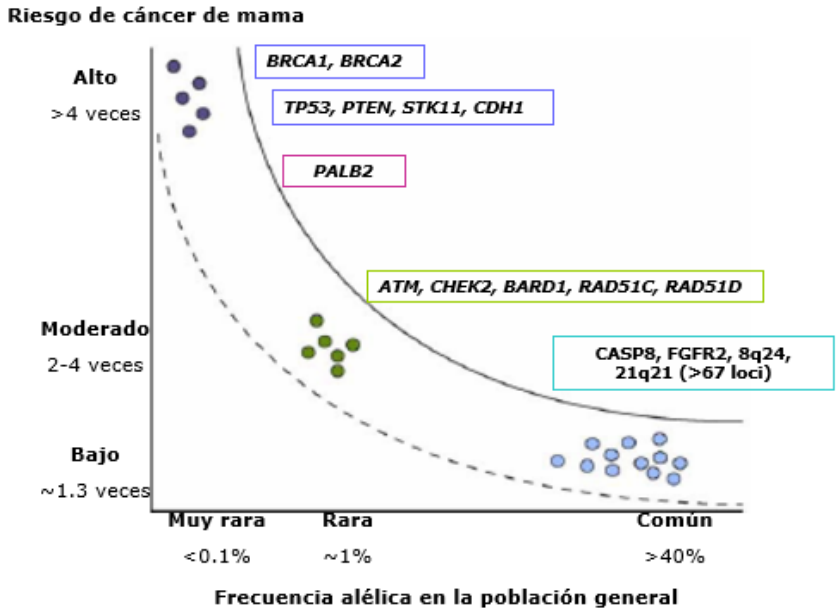


Figura 1. Modelo oligogénico del riesgo de cáncer de mama. **Este diagrama ilustra la relación entre el riesgo de cáncer de mama y la frecuencia alélica en la población general de los genes implicados en el cáncer de mama. Figura modificado de Foulkes et al**[12]

Clasificación de los genes según su riesgo relativo y prevalencia[13]:

Genes de alto riesgo: Se considera alto riesgo, cuando el riesgo relativo a padecer la enfermedad es superior a 4 veces el riesgo poblacional. Dentro de este grupo, alrededor del 15-20% de los pacientes se pueden atribuir a la presencia de una variante patogénica (VP) en línea germinal en dos genes de alta susceptibilidad para cáncer de mama: *BRCA1/2*[14]*.* En cambio, para el cáncer de ovario se estima que explicarían un porcentaje bastante más alto, entre un 30-40% de los casos[14]. Se destacan también VP en otros genes en los que el cáncer de mama tiene una prevalencia menor, y el RR es mayor de 4 y se estima que causan alrededor del 1-2% de los cáncer de mama familiares: *TP53, PTEN, STK11* y *CDH1*[15,16]*.* Estos genes no están implicados en el síndrome de cáncer de mama y ovario hereditario (CMOH), sino en otros síndromes que incluyen otros signos clínicos bien definidos[17]*.* En la actualidad, hay que destacar que el gen *PALB2* implicado en la vía de la anemia de Fanconi se considera un gen de alto riesgo[18,19].

Genes de moderado riesgo: Se considera moderado riesgo, cuando el riesgo relativo a padecer la enfermedad está entre 2 y 4 veces el riesgo poblacional. Incluyen variantes infrecuentes en *ATM*, *CHEK2,* entre otros. Estas variantes infrecuentes en estos genes suponen un 5% del componente genético del riesgo y junto con los genes *BRCA1/2*, comprenden menos del 30% del riesgo familiar de cáncer de mama. Se han descrito genes implicados en la vía de anemia de Fanconi, tales como *BRIP1 y RAD51C*. Los genes *RAD51D y RAD51C* presentan asociación confirmada con un aumento de riesgo de cáncer de ovario[20,21], y explicarían hasta un 5-10% de los casos junto con el gen *BRIP1*[14,22]. Otros genes cuyas alteraciones aumentan

el riesgo de cáncer de ovario se asocian al síndrome de Lynch (*MLH1, MSH2/EPCAM, MSH6* y *PMS2*), explicarían un 1% de los casos aproximadamente[14] y existen datos discrepantes entre la predisposición al cáncer de mama[23].

Las estimaciones de riesgos asociados a la presencia de VP en estos genes no eran precisas, hasta ahora, por lo que este hecho dificultaba el diagnóstico y el consejo genético[11,24,25]. Como consecuencia, no existía un consenso sobre qué genes incluir en los paneles de diagnóstico genético para CMOH, ya que se encuentran discrepancias entre diversos estudios, como presentamos en la tabla 1. Un ejemplo es el gen *CHEK2*, considerado un gen de moderado riesgo para cáncer de mama, que según el meta-análisis de 48 estudios en los cuales se han utilizado paneles multi-genes para el estudio de cáncer de mama y cáncer de ovario[25] estaría por debajo de un RR de 2 respecto al riesgo poblacional, aunque con amplios intervalos de confianza.

Tabla 1: Riesgos de cáncer en portadoras de VP en genes de alto, moderado y bajo riesgo.

GEN	OR CM Suszynska2019[25]	OR CM Dorling2021[11]	OR CM Couch2017[24]	OR CM Hu[26]	OR CO Suszynska2019[25]
ATM	**2,417** (2,157-2,708)	**2,10** (1,71–2,57)	**2,78** (2,22-3,62)	1,82 (1,46–2,27)	1,977 (1,330-2,939)
BARD1	**2,365** (1,857-3,013)	**2,09** (1,35–3,23)	**2,16** (1,31-3,63)	1,37 (0,87–2,16)	1,407 (0,685-2,891)
BRCA1	**8,860** (7,852-9,998)	**10,57** (8,02–13,93	ND	**7,62** (5,33–11,27)	**35,260** (29,560-42,050)
BRCA2	**5,450** (4,936-6,019)	**5,85** (4,85–7,06	ND	**5,23** (4,09–6,77)	**11,10** (9,865-14,390)
BRIP1	1,389 (1,169-1,651)	1,11 (0,80–1,53)	1,63 (1,11-2,41)	ND	**4,878** (3,729-6,380)
CDH1	**2,665** (1,678-4,233)	0,86 (0,37–1,98)	ND	**2,50** (1,01–7,07)	ND
CHEK2	0,962 (0,899-1,029)	**2,54** (2,21–2,91)	**2,26** (1,89-2,72)	**2,47** (2,02–3,05)	0,427 (0,287-0,634)
EPCAM	0,094 (0,035-0,254)	0,73 (0,36–1,49)	ND	ND	ND
MLH1	0,677 (0,464-0,990)	0,58 (0,19–1,77)	1,15 (0,30-4,19)	ND	1,436 (0,528-3,906)
MRE11A	0,894 (0,639-1,251)	0,88 (0,59–1,32)	0,86 (0,46-1,57)	ND	1,082 (0,344-3,400)
MSH2	1,633 (1,140-2,340)	1,06 (0,47–2,36)	**2,46** (0,81-6,93)	ND	**3,976** (1,818-8,695)
MSH6	1,777 (1,403-2,250)	1,96 (1,15–3,33)	1,93 (1,16-3,27)	ND	**4,077** (2,427-6,848)
NBN	1,182 (0,942-1,482)	0,90 (0,67–1,20)	1,13 (0,73-1,75)	ND	**2,166** (1,346-3,488)
NF1	1,036 (0,740-1,451)	1,76 (0,96–3,21)	0,94 (0,55-1,62)	1,93 (0,91–4,31)	ND
PALB2	**4,793** (4,128-5,566)	**5,02** (3,73–6,76)	**7,46** (5,12-11,19)	**3,83** (2,68–5,63)	**2,134** (1,420-3,207)
PMS2	0,974 (0,789-1,202)	1,16 (0,73–1,85)	0,82 (0,44-1,47)	ND	0,707 (0,292-1,716)
PTEN	**5,250** (3,058-9,014)	**2,25** (0,85–6,00)	ND	ND	**5,473** (1,257-23,820)
RAD50	0,509 (0,404-0,641)	1,08 (0,83–1,40)	0,77 (0,52-1,61)	ND	0,420 (0,157-1,125)
RAD51C	1,127	1,93	0,78	1,20	**4,241**

	(0,883-1,439)	(1,20–3,11)	(0,47-1,37)	(0,75–1,93)	(2,562-7,022)
RAD51D	1,253 (0,896-1,753)	1,80 (1,11–2,93)	**3,07** (1,21-7,88)	1,72 (0,88–3,51)	**7,276** (4,028-13,140)
STK11	1,220 (0,357-4,170)	1,60 (0,48–5,28)	ND	ND	ND
TP53	**4,129** (3,087-5,523)	**3,06** (0,63–14,91)	ND	ND	**5,046** (2,407-10,580)

ND: no determinado; CM: cáncer de mama; CO: cáncer de ovario

Sin embargo, en el trabajo de Dorling y colaboradores, evalúan genes de susceptibilidad para el cáncer de mama haciendo un estudio a 60.000 pacientes y 53.000 controles, lo que permite el cálculo de las estimaciones de riesgo mucho más precisas que las obtenidas en los anteriores trabajos[11]. En él concluyen que existe fuerte evidencia para las variantes de pérdida de función en los genes de alto riesgo *BRCA1, BRCA2* y *PALB2* y para los genes de riesgo moderado *ATM, BARD1, CHEK2, RAD51C* y *RAD51D.* Estos resultados son concordantes con los obtenidos en el trabajo de Lee et al[27], excepto para los genes *RAD51C* y *RAD51D*, donde sí que se ha encontrado una asociación de riesgo moderado para el cáncer de mama.

Genes de bajo riesgo: Comprende genes de susceptibilidad cuyas variantes son frecuentes en la población y confieren cada una de ellas un RR entre 1 y 2 para cáncer de mama o cáncer de ovario. Un riesgo inferior a 2 no se considera relevante para el diagnóstico y el consejo genético[28]. Se han identificado mayoritariamente mediante estudios pangenómicos de asociación (GWAS: *genome-wide association study*) de cientos de miles de SNPs en extensas series de miles de casos con cáncer de mama y controles. Se han descrito algunos centenares de este tipo de alelos. Se estima que estas variantes en su conjunto podrían explicar hasta un 40% del exceso de riesgo familiar para cáncer de mama[14]. De momento estas variantes no están siendo utilizadas para el diagnóstico genético, aunque se están abriendo vías de investigación en la actualidad para incluirlas en los modelos de riesgo con el fin de poder aplicarlas en la práctica clínica[29].

Síndrome de cáncer de mama y ovario hereditario

El **Síndrome de CMOH**, está causado por VP en línea germinal en los genes *BRCA1/2*, se caracterizan por presentar un patrón de herencia autosómico dominante, con una edad temprana de diagnóstico y un incremento del riesgo a desarrollar: cáncer de mama, cáncer de ovario, cáncer de próstata y cáncer de páncreas. La frecuencia poblacional de las VP se estima en 1/400-1/800 y el riesgo de cáncer en portadoras es superior a 10 veces el poblacional. *BRCA1/2* son los genes de alta penetrancia asociados a una mayor proporción de casos de CMOH[30].

El cáncer hereditario suele aparecer antes de la edad en la que es frecuente la aparición de la forma esporádica de ese mismo tipo de cáncer. En España, la media de edad de aparición del cáncer de mama es de 57-58 años. Por el contrario, los cánceres de mama asociados a VP en los genes *BRCA1/2* tienen una edad de aparición de 45 años[31].

A las mujeres portadoras de VP en los genes *BRCA1/2,* se les puede **evaluar el riesgo**, y pueden disponer de **estrategias preventivas**, tales como la cirugía y/o la quimio prevención profiláctica con terapias anti-estrógeno. Por otra parte, el consejo genético es importante para el asesoramiento preventivo respecto a segundos tumores malignos, pero, además, el genotipo *BRCA1/2*, puede ayudar a la **planificación del tratamiento**. Estudios preclínicos y clínicos sugieren que los cánceres de mama y ovario con VP o VPP en los genes *BRCA1/2*, son más sensibles a terapias que alteran el ADN, como son el cisplatino e inhibidores del PARP (PARPi), debido a la capacidad para inhibir vías de reparación de ADN[32].

3. Genes de susceptibilidad con alta penetrancia

El análisis de ligamiento de familias con múltiples casos de cáncer de mama y cáncer de ovario permitió identificar en 1990 la primera región cromosómica (17q21) candidata a contener un gen de alto riesgo en cáncer de mama hereditario[33]. En 1994, mediante clonación posicional se identificó el gen causal, *BRCA1* (BReast CAncer gene 1) (OMIN 113705)[34]. En el año 1995 se identificó el gen *BRCA2* (BReast CAncer gene 2) (OMIN 600185), el segundo gen de alta penetrancia implicado en el cáncer de mama hereditario[35].

3.1 Estructura y función de los genes BRCA1 y BRCA2

Son genes supresores tumorales implicados en muchos procesos, que incluyen la recombinación y reparación del daño de la doble cadena de ADN, el control del ciclo celular y la transcripción. Las proteínas que codifican desempeñan un papel importante en el mantenimiento de la estabilidad[36].

La expresión de los genes *BRCA1/2* está relacionada con el crecimiento y la diferenciación de las células epiteliales mamarias, especialmente durante la gestación y la lactancia. Estos genes se expresan de forma ubicua, aunque se ha visto que su actividad repercute especialmente en las células de la mama[36].

El gen ***BRCA1***, se localiza en la región cromosómica 17q21, ocupando 89kb. Su secuencia está distribuida en 22 exones codificantes y dos exones que corresponden a dos sitios alternativos de origen de la transcripción. La región 5´de *BRCA1* contiene el pseudogen *ΨBRCA1* que consiste en una duplicación parcial de *BRCA1,* que contiene los exones 1a, 1b y 2 con sus correspondientes intrones[37]. El mRNA transcrito es de 7.8kb, de los cuales 5.592 nucleótidos son codificantes y se traducen en una proteína de 1.863 aminoácidos, se expresa de forma abundante en testículo y timo, así como en mama y ovario. El exón 11 codifica más del 60% de la proteína, esta región de la proteína contiene dos señales de localización nuclear y puede fosforilarse en diversos puntos por ATM y ATR (proteína de ataxia telangiectasia y proteína relacionada con Rad3), después de una lesión en el ADN (Figura 2).

La proteína BRCA1 tiene cuatro dominios principales:

1) El dominio N-terminal RING, codificado por los exones 2-7 (aminoácidos del 1 al 109). Se compone de un motivo dedo RING correspondiente a los residuos 24-64, que es un dominio altamente conservado y rico en cisteínas. Juega un papel clave en la ruta de ubiquitinación, este domino interactúa formando un heterodímero con la proteína BARD1 (BRCA1-*associated RING domain 1*).
2) El dominio P300/CBP regula la transcripción al interactuar con los dos cofactores transcripcionales P300 y CBP.
3) El dominio coiled-coil se une al dominio coiled-coil de la proteína PALB2, que a su vez se une al extremo N-terminal de BRCA2. Este complejo colabora en la unión de RAD51 al ADN durante su reparación de recombinación homóloga (RH).
4) El extremo C-terminal posee dos copias en tándem del dominio BRCT (BRC1 C-*terminal*) con capacidad de unión a serinas fosforiladas y a otras proteínas: Abraxas, BRIP1 y CtIP, que corresponden respectivamente a los complejos A, B y C. El complejo A está implicado en la reparación del ADN por recombinación homóloga, y los complejos B y C actúan en los puntos de control G1/S Y G2/M del ciclo celular, respectivamente. Las secuencias BRCT de BRCA1 pueden activar la transcripción de ciertos genes. Las variantes genéticas que afectan estos dominios interfieren en sus actividades durante la transcripción y la reparación del ADN mediada por RH y suele asociarse con un aumento de la predisposición al cáncer.

Entre las funciones de *BRCA1* se incluyen la reparación del ADN y regulación del ciclo celular, junto a *RAD51, BARD1, RAD50, MRE11, NBS1*, actividad transcripcional de otros genes, como *p21, GADD45*, remodelación de la cromatina e implicación en la inactivación del cromosoma X y ubiquitinación, con la finalidad de mantener la estabilidad genómica y, por lo tanto, en relación con la etiología del cáncer de mama y ovario[38].

En resumen, la proteína BRCA1 interactúa con muchas otras regiones de genes supresores de tumores, oncogenes, activadores y represores de la transcripción, reguladores del ciclo celular[37].

El gen ***BRCA2***, se localiza en la región cromosómica 13q12, ocupando 70kb[35]. Presenta 27 exones, de los cuales 26 codifican una proteína de 3.418 aminoácidos. El transcrito es de 12kb y está presente en células del epitelio mamario y en la placenta. El exón 11 codifica más de la mitad de la proteína. El extremo N-teminal se une a PALB2, que ayuda a la localización de RAD51 en el ADN dañado. Este exón 11 contiene ocho secuencias de repetición interna llamadas motivos BRC, los cuales se consideran el mayor dominio de interacción con RAD51. La proteína RAD51 es un componente clave en la recombinación homóloga y reparación de la rotura de la doble cadena de ADN[39]. Gracias a esta interacción, BRCA2 regula la disponibilidad y la actividad de RAD51, que recubre el ADN de una sola cadena para formar el filamento de nucleoproteína que introduce y junta el ADN dúplex homólogo para iniciar el intercambio de cadenas (Figura 2).

La estructura del dominio C-terminal de BRCA2 es la parte más conservada de la proteína, contiene una región de unión con el ADN, de unos 800 aminoácidos, en la que se disponen una estructura hélice-giro-hélice y tres módulos de unión a oligonucleótidos/oligosacáridos

(OB). Se han observado un gran número de variantes identificadas en este dominio que derivan en tumores. Contiene dos señales de localización nuclear y una segunda unión a RAD51 (en el exón 27) muy conservada en vertebrados. Presenta también dominios de activación transcripcional y forma complejos con otras proteínas, como BRAF35 y P/CAF, indicando su posible participación e la remodelación cromatínica y el control del ciclo celular en el punto G2/M.

BRCA2 participa específicamente en la reparación del ADN mediante RH y en la progresión del ciclo celular (citoquinesis y meiosis).

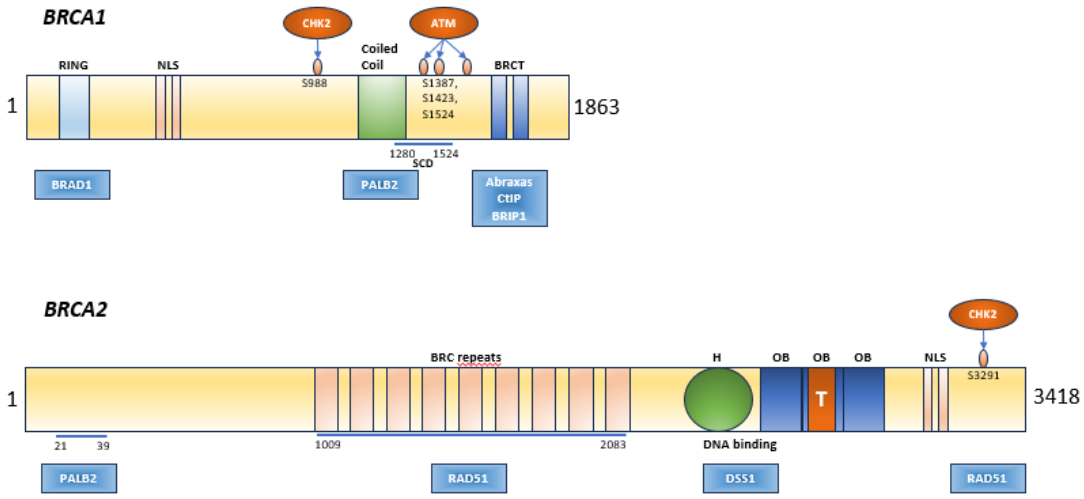


Figura 2. Esquema de los dominios funcionales de las proteínas BRCA1 y BRCA2. **En la parte superior de la proteína se denominan los dominios funcionales (RING, NLS, etc.). Los rectángulos indican el nombre de las proteínas con que se unen BRCA1 y 2 y el lugar de unión. Los óvalos señalan las quinasas que fosforilan las proteínas y el lugar de fosforilación. Los números bajo las proteínas indican el número de orden de los nucleótidos. Tomada de Roy et al. (2012)[40].**

3.2 Variantes patogénicas en los genes BRCA1 y BRCA2

Hasta la fecha se han identificado millares de VP distintas en los genes *BRCA1/2*. El número de variantes depende de la base de datos consultada. El consorcio internacional de investigadores ENIGMA (*Evidence-Based for the interpretation of Germline Mutant Allele*)[28], determina la importancia clínica de las variantes de los genes *BRCA1/2* en cáncer de mama y ovario. Además, también proporciona la opinión de expertos para tener una clasificación global y presenta sus resultados en distintas bases de datos, como son: ClinGen (Clinical Genome Resource, https://clinicalgenome.org/), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) y BRCA EXCHANGE[41] (https://brcaexchange.org/). En estos repositorios aparecen 64.359 entradas con variantes diferentes, descritas en estos genes, 2.228 se clasifican como variantes patogénicas (VP) para el gen *BRCA1* y 2.672 en el gen *BRCA2*. En la base de datos HGMD® Professional 2023.1 se contabilizaron 2.977 VP para *BRCA1* y 2.909 VP para *BRCA2*, en la fecha de revisión de estos datos (junio 2023)[42]. Mientras que en la base de datos BIC (Breast Cancer Information Core), aparecen 15.311 entradas para *BRCA1* y 14.914 entradas para *BRCA2*, aunque esta base de datos está obsoleta actualmente.

Las VP suelen consistir en pequeñas inserciones o deleciones, que provocan un cambio en el marco de lectura, dando lugar a la aparición de un codón de parada prematuro, y cuya consecuencia, es una proteína truncada. También se han descrito sustituciones nucleotídicas que provocan codones de parada prematuros. En otras ocasiones, puede alterarse el proceso de eliminación de intrones (*splicing*) y generarse un transcrito anormal, causando la pérdida completa o parcial de exones o bien, la inserción de secuencias intrónicas en el transcrito generado. También son frecuentes las sustituciones que causan el cambio de aminoácido y pueden ocurrir grandes duplicaciones o deleciones que comprenden uno o más exones del gen. Este un tipo de variantes se conocen como variantes del número de copia (CNVs o *Copy Number Variants*)

Las **CNVs**, aproximadamente representan, aproximadamente, un 10% de las VP en estos genes[43]. Existe un mayor porcentaje de grandes duplicaciones/deleciones en el gen *BRCA1* frente al gen *BRCA2* (11-13% y 2-3% respectivamente) debido a la acumulación de las secuencias *Alu*. Los elementos *Alu*, son la clase de secuencias repetidas más abundantes en el genoma, presentan una longitud de aproximadamente 300pb[44]. Aunque en el caso de *BRCA2* estas CNVs son raras, en el caso de la población portuguesa, existe una excepción, ya que presenta una inserción fundadora c.156_157ins*Alu*[45], siendo una variante altamente frecuente.

La prevalencia y tipos de VP en línea germinal en los genes *BRCA1/2* se conoce en determinadas poblaciones, la mayoría de Europa y América del Norte, a partir de numerosos estudios y gracias a la base de datos de ExAC (*Exome Aggregation Consortium*)[46]. Estos datos pueden reflejar la existencia de VP fundadoras en grupos étnicos específicos. En este sentido, el consorcio CIMBA (*Consortium of Investigatiors of Modifiers of BRCA1/BRCA2*) ha reunido datos de unas treinta mil familias de aproximadamente 1700 VP distintas en cada gen, observando diferencias en el tipo y en la frecuencia de las variantes según el área geográfica y etnia. Un ejemplo del efecto fundador, es la población judía Asquenazí, en la que existen tres VP (c.68_69delAG y c.5266dupC en *BRCA1* y c.5946delT en *BRCA2*) que explicarían el 10% de los cánceres hereditarios de esa población[47], identificándose en un 2% en población general y que comprenden un. La variante c.68_69delAG también se ha descrito en poblaciones judías no askenazis, indicando un origen más antiguo, y es una de las VP recurrente en población española (12-16%)[48].

En España, se han realizado diferentes estudios donde se han visto que hay VP frecuentes, que difieren significativamente en sus frecuencias y en su distribución geográfica[31]. En el gen *BRCA1,* cinco variantes (c.68_69delAG, c.211A>G, c.5117G>A, c.5123C>A y c.470_471delCT) representarían el 46,6% de las VP identificadas en este gen, mientras que en el gen *BRCA2*, cuatro variantes (c.2808_2811del4, c.6629_6630delAA, c.9026_9030del5 y c.9310_9311delAA), suponen el 56,6% de las VP en este gen en familias españolas[48]. En total, estas 9 variantes representan la mitad de las VP detectadas en familias españolas.

3.3 Variantes de significado incierto en los genes BRCA1 y BRCA2.

El aumento del número de test *BRCA1/2* en línea germinal ha conllevado al descubrimiento de un gran número de VSI[49]. Se requieren estudios adicionales como análisis de segregación en

la familia, evaluación de las consecuencias funcionales de esa variante para clasificarla en otra categoría superior o inferior. En la base de datos ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), hay un 37% de VSI para el gen *BRCA1* y un 45% de VSI para el gen *BRCA2*. En la base de datos BRCA Exchange (https://brcaexchange.org) presenta un 66% de VSI para el gen *BRCA1* y un 67% de VSI para el gen *BRCA2* (se han incluido las variantes sin revisar en esta base de datos, como VSI)[49]. Este tipo de variantes se identifican en un 10-20% de las pacientes que se realizan un test genético de los genes *BRCA1/2*[50].

3.4 Herramientas disponibles para la clasificación de variantes

La interpretación del posible efecto patológico de una variante, es difícil, ya que dependiendo del dominio funcional de la proteína en el que se localicen, pueden o no alterar su función. Por ello, deben examinarse diversos aspectos: como la ausencia de la variante en individuos de población sana, su cosegregación con el cáncer en la familia (solo los afectados portan la variante) y su localización en un dominio potencialmente crítico para la función de la proteína. Debe evaluarse también, si el aminoácido se ha sustituido por uno similar o muy diferente, o si se localiza en una zona de la proteína que ha permanecido muy conservada durante la evolución filogenética, lo que indicaría su importancia funcional. La demostración definitiva de su valor patológico, puede obtenerse mediante estudios funcionales, en los que se intenta generar una proteína con el cambio en cuestión, para observar si sus funciones se mantienen en un medio celular. Sin embargo, estos estudios forman parte de proyectos de investigación, no aplicables al diagnóstico de rutina[51].

Para la clasificación de variantes se recomienda seguir los criterios y guías de las principales sociedades científicas, como el Colegio Americano de Genética Médica y Genómica (ACMG), la Asociación de Patología Molecular (AMP) y la Asociación de Ciencias de Genética Clínica (ACGS)[52,53]. Estas sociedades establecen guías que evalúan los criterios de evidencia de patogenicidad. Cada criterio evaluado se categoriza, en función de la evidencia descrita como muy fuertes, fuertes, moderados o secundarios para patogenicidad, además de tener en cuenta otros criterios que apoyan la benignidad de dichas variantes. que pueden ser muy fuertes, fuertes, moderados y secundarios, así como otros criterios que apoyan la benignidad de dichas variantes. Dentro de los criterios fuertes y muy fuertes de patogenicidad están aquellas variantes que son *de novo*, que producen cambios sin sentido, que alteran el marco de lectura o que afectan a sitios de *splicing* en genes en los que la pérdida de heterocigosidad en tejido es un mecanismo conocido de la enfermedad. También hay otros criterios menos determinantes pero que también se utilizan para clasificar VP como la co-segregación de la variante con la enfermedad en los familiares afectados o que los estudios *in silico* sugieran una capacidad de la variante para alterar el *splicing*[52,53]. Finalmente, estas guías establecen un sistema de clasificación de variantes en cinco clases: VP, VPP, VSI, VPB o VB, teniendo en cuanta la combinación de estos criterios.

Por tanto, conocer los principales recursos disponibles para valorar clínicamente las variantes génicas halladas en un estudio genético, es de gran importancia para poder clasificar las variantes adecuadamente. Este punto nos permitirá saber si estas variantes han sido previamente descritas y clasificadas clínicamente o si, por el contrario, son variantes no descritas y por tanto se desconoce su implicación en la clínica de los pacientes. A continuación, se describen los principales recursos disponibles para este propósito:

Bases de datos

Las bases de datos consultadas para la clasificación de las variantes son:

1) **DbSNP**[54]: Plataforma pública estadounidense. Es una base de datos de polimorfismos de un solo nucleótido (SNP) y múltiples variaciones de pequeña escala, que incluyen inserciones/deleciones, microsatélites y variantes no polimórficas). Está desarrollado por el centro National Center for Biotechnology Information (NCBI) en colaboración con el *National Human Genome Research Institute* (NGHRI). Obtenemos información sobre el posible efecto fisiopatológico de la variante y de su frecuencia en población control (MAF).
2) ***Human Gene Mutation Database* 2023.1 (HGMD)**[42]: Es una base de datos con gran información sobre enfermedades hereditarias. Recoge citas bibliográficas de las variantes, analiza las propiedades biofísicas de los aminoácidos sustituidos y calcula predicciones de algunos programas *in silico*.
3) ***The breast cancer information core* (BIC)**[55]: Base de datos que se utiliza como referente en la búsqueda de información de las diferentes variantes en *BRCA1* y en *BRCA2*. Según su efecto fisiopatológico, clasifica a las variantes como clínicamente importantes, desconocido o clínicamente no importantes. Aunque en la actualidad se encuentra ya obsoleta.
4) ***Leiden Open Variation Database* (LOVD)**[56,57]: Contiene información sobre citas bibliográficas y permite conocer la clasificación de la *International Agency for Research on Cancer* (IARC)[58].
5) **Universal Mutation Database. UMD-BRCA1, UMD-BRCA2:** Son bases de datos pertenecientes a una red de 16 laboratorios de diagnóstico franceses, que tienen información actualizada sobre las variantes en los genes *BRCA1/2* identificadas en pacientes con CMOH. Actualmente ha sido actualizada y renombrada como **BRCA Share™**.
6) **Alamut® v2.5.** *Software* que integra información genética y genómica de diferentes fuentes para describir variantes y ayudar a la interpretación. Enlaza con diferentes predicciones de algunos programas *in silico*, como son: Align GVGD, SIFT, MutationTaster y predictores de *splicing* como: MaxEnt, NNSPLICE y HSF.
7) **ClinVar-NCBI** (https://www.ncbi.nlm.nih.gov/clinvar/): Base de datos pública donde tanto laboratorios privados como públicos reportan y clasifican variantes. Se realizan revisiones por paneles de expertos.
8) **BRCAExchange** (https://brcaexchange.org/): Base de datos pública que contiene la revisión e interpretación de variantes que han sido sometidas a la revisión de

expertos por parte del consorcio ENIGMA.

Programas de predicción in silico

Hay una gran variedad de herramientas *in silico* que pueden ayudar a la correcta interpretación de las variantes. Los algoritmos utilizados son diferentes y pueden incluir la importancia de la alteración tanto a nivel de nucleótido como a nivel de aminoácido. Estos programas se pueden dividir en dos grupos: los que predicen si el cambio es perjudicial para la función o estructura de la proteína resultante y los que predicen si se altera el *splicing*.

El impacto de las alteraciones *missense* depende de varios criterios. Los principales se resumen a continuación[52]:

- La localización de la proteína
    - Si la variante se encuentra en un dominio funcional de conocida importancia para la función de la proteína y en el que las variantes *missense* conocidas son todas patogénicas, existe una elevada probabilidad de que ésta sea patogénica.
    - Existen zonas donde se ha descrito una mayor frecuencia de VP. Se conocen como *hot spot* y si una variante se encuentra en esta zona, la posibilidad de que sea patogénica es elevada.
    - Que la variante se halle en el mismo codón que otra VP apoya una probable patogenicidad de la nueva variante.
- Grado de conservación: si una zona del genoma está altamente conservada a lo largo de la evolución, se presupone de gran importancia estructural o funcional, por lo que una alteración tiene altas posibilidades de ser deletérea
- Consecuencias bioquímicas del cambio de aminoácidos: si se produce una sustitución de un aminoácido por otro de características similares, la posibilidad de que este cambio sea dañino es menor que si el nuevo aminoácido tiene unas características fisicoquímicas (tamaño, polaridad, etc.) que el original

La principal limitación de estas herramientas es la dificultad de su interpretación debido a su baja especificidad, lo que conlleva que muchas variantes se predigan falsamente como patogénicas. Entre los programas están: *Align GVGD, SIFT, MutationTaster* y *Revel*.

En cuanto a los programas que predicen el efecto en el *splicing*. Estos predicen sitios de empalme probables. Entre los programas están: *MaxEnt, NNSPLICE, HSF* y *Splice AI*.

Evaluación de la frecuencia de la variante en población control

Conocer la frecuencia de una variante en la población control es de utilidad para conocer su patogenicidad. En general, una frecuencia alélica en población control superior a la esperada para una determinada enfermedad apoya la benignidad de esa variante. Esto se puede hacer mediante una búsqueda en las bases de datos poblacionales como la del Proyecto 1000 Genomas, del Consorcio *Genome Aggregation Database* (gnomAD)[52] o del Consorcio de Agregación de Exomas (ExAc). La primera tiene una gran representación de diferentes

poblaciones y la segunda, reúne y armoniza datos de secuenciación de exomas y genomas de una variedad de proyectos de secuenciación a gran escala y pone los datos a disposición de la comunidad científica.

Estudios funcionales

Los estudios funcionales son una herramienta poderosa para determinar la patogenicidad, pero no todos son efectivos para determinar el impacto sobre el gen o la funcionalidad de la proteína. La evidencia es mayor si el ensayo refleja la totalidad de la función biológica de la proteína comparada con sólo una parte de esta función. Los estudios funcionales que determinan el impacto de las variantes a nivel de ARN mensajero (ARNm) pueden ser altamente informativos a la hora de evaluar los efectos en la estabilidad del ARNm, su procesamiento y traducción entre otros[52].

Estudio de co-segregación

Un estudio de cosegregación, consiste en analizar en otros miembros de la familia (afectos y no afectos) la variante identificada. Si la variante está presente en familiares afectados y ausente en los familiares no afectados, se podría afirmar que la variante co-segrega con la enfermedad. Sin embargo, se debe tener precaución a la hora de usar la co-segregación para clasificar clínicamente una variante, ya que ésta evidencia una relación entre el *locus* y la enfermedad, pero no confirma la patogenicidad de la variante en sí. Además, se deben tener en cuenta la presencia de posibles fenocopias o de que la co-segregación se deba a un desequilibrio de ligamiento con la verdadera VP. Actualmente, también se debe profundizar en la relación biológica entre familiares, ya que existe la posibilidad de no paternidad biológica (donación de esperma y ovocitos o adopción, principalmente)[52].

Clasificación de variantes desde el punto de vista clínico

Con el objetivo de facilitar la interpretación de las variantes de las variantes halladas en un análisis genético, así como para mejorar el asesoramiento genético, Lindor y colaboradores[59] y posteriormente tanto el consorcio internacional ENIMGA[28] así como ACMG/AMP[52] recomiendan el siguiente modelo de clasificación de variantes, que se muestra en la tabla 2.

Tabla 2. Términos utilizados actualmente para describir las categorías de clasificación de variantes genéticas en cinco niveles y su recomendación clínica. Modificada de Spurdle[28]

Clase numérica	Acrónimo	Descripción	Probabilidad de patogenicidad	Descripción de la relevancia clínica en variantes de alta penetrancia
5	P	Patogénica	>0,999	Toma de decisiones clínicas. Ofrecer pruebas de riesgo en familiares para la variante
4	PP	Probablemente patogénica	0,9-0,999	Toma de decisiones clínicas. Ofrecer pruebas de riesgo en familiares para la variante
3	VSI	Incierta	0,1-0,899	Las recomendaciones clínicas deben basarse en la historia personal y familiar y otros factores de riesgo, no por la

				presencia de esta variante. No utilizar como pruebas de predicción en los familiares en riesgo.
2	PB	Probablemente benigna	0,001-0,099	Esta variante no es la causa de la patología. Las recomendaciones clínicas deben basarse en la historia personal y familiar y otros factores de riesgo, no por la presencia de esta variante
1	B	Benigna	<0,001	Esta variante no es la causa de la patología. Las recomendaciones clínicas deben basarse en la historia personal y familiar y otros factores de riesgo, no por la presencia de esta variante

3.5 Estimaciones del riesgo de cáncer asociado a BRCA1 y BRCA2

El riesgo de desarrollar cáncer no es el mismo para todas las portadoras de VP en los genes *BRCA1* y en el gen *BRCA2*, ya que puede estar influido por la heterogeneidad alélica, genes modificadores, cofactores ambientales y hormonales[45]. Las estimaciones de riesgo varían según la población estudiada incluso en diferentes estudios[60].

El riesgo acumulado a lo largo de la vida, alrededor de los 80 años, para desarrollar cáncer de mama en portadoras de VP en los genes *BRCA1/2* está en torno a un 70% (72% para el gen *BRCA1* y de un 69% para el gen *BRCA2*). El riesgo para desarrollar cáncer de ovario es de un 44% para portadores de VP en el gen *BRCA1* y un 17% en el caso de VP en el gen *BRCA2*[60].

El riesgo de este tipo de cánceres en pacientes con VP en los genes *BRCA1* o *BRCA2*, se muestran en la tabla 3[43].

Tabla 3. Riesgo de distintos tipos de cánceres en población general y portadores de VP

Tipo de cáncer	**Riesgo de la población general**	**Riesgo con una VP**	
		BRCA1	***BRCA2***
Mama	12%	46%-87%	38-84%
Segundo cáncer primario	2% en 5 años	21,1% en 10 años 83% a los 70 años	10,8% en 10 años 62% a los 70 años
Ovario	1%-2%	39%-63%	16,5%-27%
Mama en varón	0,1%	1,2%	Más de 8,9%

Próstata	6%	8,6% a los 65 años	15% a los 65 años 20%
Páncreas	0,50%	1%-3%	2%-7%
Melanoma	1,6%		Riesgo elevado

En el año 2008 se llevó a cabo un estudio de estos riesgos en **población española** sobre 319 familias y la estimación del riesgo acumulado para cáncer de mama con historia familiar, a la edad de 70 años fue del 52% (95% CI, 26-69%) para portadoras de VP en *BRCA1* y 47% (95% CI, 29-60%) para *BRCA2*. Las estimaciones correspondientes para cáncer de ovario fueron 22% (95% CI, 0-40%) y 18% (95% CI, 0-35%)[61].

Por otro lado, la edad de diagnóstico y el tipo de cáncer pueden manifestarse de forma muy heterogénea entre portadoras, incluso en una misma familia. Estas diferencias son debidas a efectos modificadores inherentes al tipo y localización de las VP y a variantes en otros genes, así como a factores hormonales y ambientales[62].

Actualmente se está investigando sobre este tema para la identificación de estos factores a través de consorcios internacionales para tener estimaciones más precisas de los riesgos. En este sentido existen dos consorcios internacionales, el IBCCS (International *BRCA1/2* Carrier Cohort Study) y CIMBA creados para la identificación de factores exógenos y genéticos modificadores de riesgo, respectivamente tanto para pacientes afectas testadas como para sus familiares portadores de estas variates[63,64]. Uno de los resultados más relevantes publicados por CIMBA es la confirmación de que un polimorfismo en el gen *RAD51* (c.135G>C) incrementa el riesgo de desarrollar cáncer de mama en portadoras de VP en *BRCA2*, especialmente cuando aparece en homocigosis[65]. Actualmente el consorcio se centra en el estudio de los loci de susceptibilidad del cáncer de mama de elevada frecuencia y bajo riesgo identificados a través de GWAS que modifican el riesgo de cáncer en portadoras de VP en los genes *BRCA1/2*[66].

3.6 Implicaciones terapéuticas

Actualmente, la determinación del estado de los genes *BRCA1/2* es fundamental para decidir un tratamiento con PARPi. Su empleo de los PARPi han cambiado totalmente el abordaje del cáncer de mama y ovario, permitiendo actualmente haya un solapamiento entre el manejo preventivo y terapéutico del estudio genético de los genes *BRCA1/2*[32]. Por ejemplo, en pacientes con cáncer de ovario con una VP, el tratamiento con quimioterapia y mantenimiento con PARPi ha hecho que se reduzca el ritmo de progresión y muerte hasta un 70%[67]. Por tanto, determinar si el cáncer de ovario en estas pacientes es hereditario o esporádico es fundamental para tomar medidas terapéuticas y preventivas.

Los PARPi son una clase de terapia oral dirigida cuyo efecto depende de la incapacidad de un tumor para reparar el daño en su ADN. Las proteínas PARP1 y PARP2 intervienen en los mecanismos de reparación del ADN, reparando la rotura de cadena simple mediante la escisión

de bases (REB), se une directamente al punto de daño del ADN, atrae y ayuda a los efectores que participan en la REB. La inhibición de PARP impediría la reparación del daño producido en la cadena de ADN, esto colapsaría la horquilla de replicación y produciría una rotura de ADN de doble cadena. Las células que tienen un déficit de recombinación homóloga (DHR), es decir, en el sistema de reparación de roturas de doble cadena de DNA, hecho que está presente en los tumores que se desarrollan en los portadores de VP en los genes *BRCA1/2*, no pueden repararlo y da lugar a la muerte celular[68,69]. Tendríamos por tanto una diana accionable para el tratamiento de las pacientes con cáncer de ovario portadoras de una VP en los genes *BRCA1/2*.

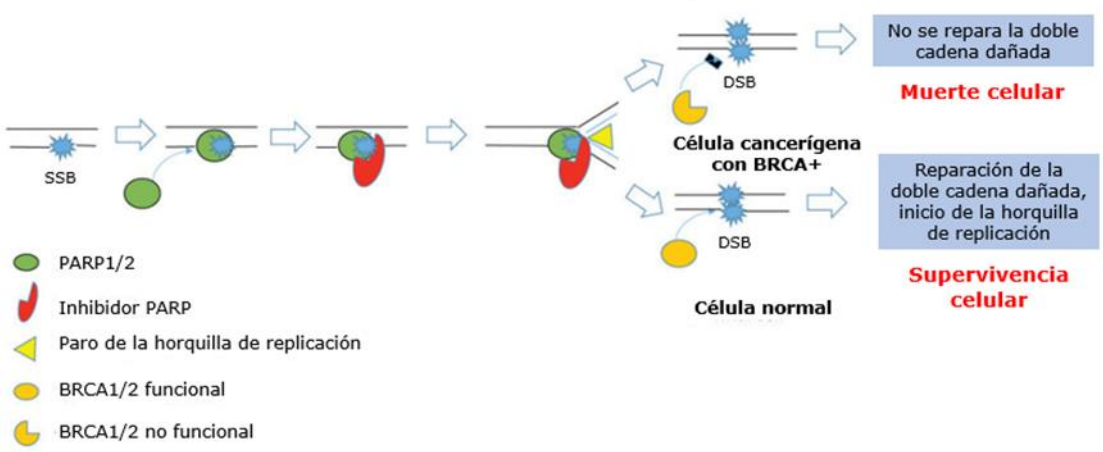


Figura 3. Representación esquemática de la reparación del ADN mediada por PARP y la recombinación homóloga. Las roturas de ADN de cadena simple (SSB) conllevan el reclutamiento del PARP y su activación. La inhibición del PARP y su retención, ocasionan roturas del ADN de cadena simple permanentes y una parada de la horquilla de replicación. Esto conlleva a roturas de la doble cadena (DSB) que requieren de la recombinación homologa (HR) para su reparación. En las células cancerígenas con VP en los genes *BRCA1/2* (BRCA+), la pérdida de heterocigosidad conduce a un BRCA no funcional y la HR no puede ocurrir. Esto da como resultado la muerte celular. Las células normales tienen su HR funcional. Como resultado tenemos Figura modificada del trabajo de Turk et al (2018)[69].

Desde el descubrimiento de los PARPi en 2005, ha habido un rápido desarrollo clínico de cinco agentes diferentes (olaparib, rucaparib, niraparib, veliparib, talazoparib), lo que lleva a seis indicaciones aprobadas por la Administración de Drogas y Alimentos de los Estados Unidos (FDA) y la EMA (Agencia europea del medicamento) para cáncer de mama y ovario[32,70] (Tabla 4).

Tabla 4. PARPi aprobados por la FDA y la EMA para cáncer de mama y cáncer de ovario

	Cáncer de ovario	**Cáncer de mama**
Mantenimiento (después de primera línea)	<u>BRCA+ a nivel germinal o somático:</u> • Olaparib (2018)	-

Terapia de primera línea y mantenimiento	Cualquier estado BRCA, enfermedad platino-sensible: • Olaparib (2017) • Rucaparib (2018) • Niraparib (2017)	-
Tratamiento recurrente	BRCA+ a nivel germinal o somático: • Olaparib (2014) • Rucaparib (2016) BRCA+ o HRD positiva: • Niraparib (2019)	BRCA+ a nivel germinal, CM metastásico o locamente avanzado triple negativo o HR+ HER2 negativo • Olaparib (2018) Resistente a tratamiento hormonal: • Talazoparib (2018)

Abreviaturas: BRCA+: VP/VPP en los genes *BRCA1/BRCA2;* HRD: Deficiencia en la recombinación homóloga.

En 2014, Olaparib se convirtió en el primer PARPi aprobado por la FDA para su uso como tratamiento del cáncer de ovario recurrente, seguido de Rucaparib en 2016 y Niraparib en 2019.

Olaparib, un inhibidor de PARP muy potente, ha sido recientemente aprobado para el tratamiento del cáncer de ovario por la FDA y por la EMA, en pacientes adultos con cáncer de ovario seroso de alto grado, cáncer de trompa de Falopio o peritoneal en recaída que presenten VP en los genes *BRCA1* o *BRCA2* y que responden a quimioterapia basada en platino. Actualmente se están realizando ensayos clínicos con varios PARPi para evaluar las toxicidades, las eficacias y el beneficio de los fármacos como monoterapias o combinados con radiación u otros agentes quimioterapéuticos, en ovario, mama, próstata, rectal, pulmón entre otros tipos de cáncer[71].

Por tanto, las VP en los genes *BRCA1/2* representan el primer **biomarcador predictivo** en cáncer de ovario y actualmente se recomienda la incorporación rutinaria del test de estos genes *BRCA1/2*, en pacientes con cáncer de ovario recién diagnosticado.

Se han detectado VP en los genes *BRCA1/2* en aproximadamente un 19,3% de los tumores de ovario serosos de alto grado, tanto germinales (13,3%) como somáticas (5,9%)[72]. El análisis de los genes *BRCA1/2* en tumores de ovario tiene un interés clínico creciente ya que las pacientes con VP tanto germinales como somáticas pueden beneficiarse del tratamiento con PARPi.

Actualmente también se están empleando PARPi en pacientes con VP germinales en los genes *BRCA1/2* con cáncer de mama metastático HER2-negativo, con cáncer de páncreas metastásico sensible a platino y cáncer de próstata resistente a la castración (en este último también con VP a nivel somático)[73].

3.7 Consejo genético en el cáncer de mama y ovario hereditario

El proceso de Consejo Genético incluye como elementos fundamentales, la recopilación de información clínica personal y familiar, la evaluación y el asesoramiento del riesgo de desarrollar una enfermedad, la comunicación de información, y el ofrecimiento de un soporte continuado (Ley de Investigación Biomédica 14/2007). En esta ley se establece además que cualquier análisis genético diagnóstico se debe realizar por prescripción médica, y antes de llevarlo a cabo, los pacientes o personas a estudiar deben recibir un asesoramiento genético pretest que incluya la información adecuada para el síndrome de sospecha, una explicación en qué consiste el estudio genético, sus limitaciones, posibles resultados esperados y su significado. Tras recibir toda esta información, deben firmar un consentimiento informado, para poder proceder con el estudio. La entrega de resultados al paciente se debe de realizar en el contexto de un asesoramiento genético post test.

Se consideran individuos candidatos a recibir asesoramiento genético en cáncer, a aquellas personas que: son atípicamente jóvenes, su diagnóstico es muy inusual, presenta neoplasias múltiples, neoplasias asociadas a defectos congénitos o existen múltiples miembros de la familia afectos de la misma neoplasia o asociadas, o bien para la identificación de biomarcadores predictivos de respuesta al tratamiento, como las terapias dirigidas con PARPi. Las técnicas de diagnóstico genético aplicadas en cáncer hereditario se focalizan en el análisis de alteraciones genéticas en línea germinal (SNV, como CNV) en aquellos genes responsables de los síndromes de sospecha[74].

Un paciente puede ser visitado por un médico de atención primaria o por un médico especialista que valoran el caso y lo remiten a una unidad de consejo genético, que está formado por un equipo multidisciplinar que valora toda la información relevante del paciente (Figura 4).

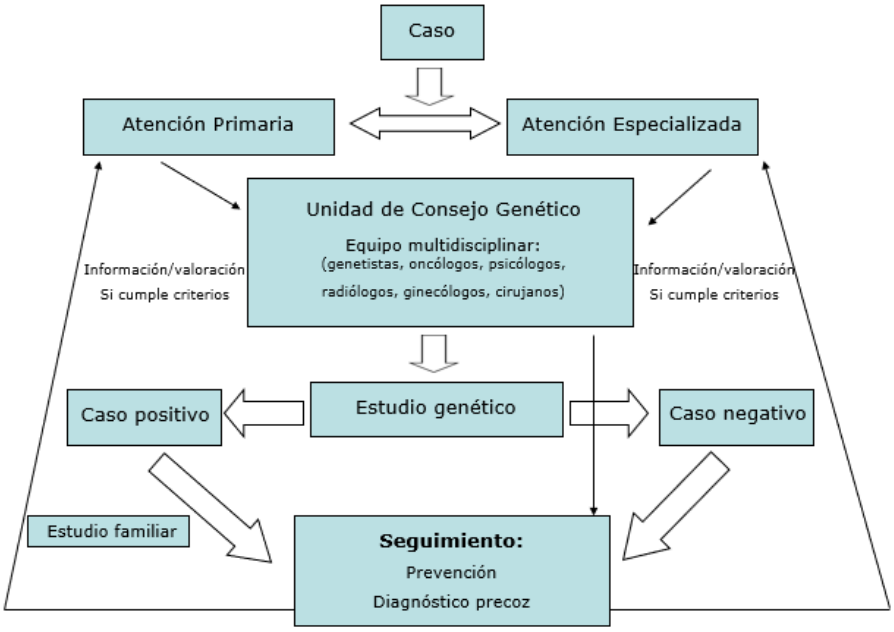


Figura 4. Esquema de los pasos seguidos por los profesionales de salud ante un posible caso de CMOH.

El proceso de asesoramiento genético en cáncer se divide en varias etapas:

- **Sospecha clínica de un síndrome de cáncer hereditario**: se elabora el árbol genealógico detallado incluyendo al menos tres generaciones consecutivas, motivos por los que el paciente acude a la visita y se evalúa las razones que tiene para solicitar las pruebas: se valora el riesgo de cáncer; se orienta sobre la prevención primaria; e intenta que se adopten hábitos de vida saludables y promoción de la salud.
- **Valoración de la indicación del estudio genético:** se determina si existe la sospecha clínica de un síndrome determinado, del cual se conocen los genes asociados y cuyo resultado puede interpretarse a nivel clínico. Se informa sobre las implicaciones médicas personales y familiares de la identificación de una susceptibilidad genética.
- **Asesoramiento sobre los beneficios y limitaciones de cada tipo de estudio genético:** Se asesora sobre las implicaciones acerca de la toma de decisiones médicas de detección precoz, prevención y tratamiento, así como las limitaciones relacionadas con la sensibilidad limitada de algunas técnicas diagnósticas o la posibilidad de identificar VSI.

Realización del estudio genético, previo consentimiento informado en los casos en que se considere indicado. Se recomienda iniciar el estudio genético en una familia a partir de la persona que tenga más posibilidades de ser portadora de la VP (caso índice o probando). El test genético de predisposición al cáncer, es el análisis que nos informa si un individuo ha heredado una alteración genética que aumente el riesgo para ciertos tipos de cánceres. Los resultados del test genético tienen repercusiones tanto en el individuo como en sus familiares. El objetivo de las pruebas genéticas es determinar la etiología de sospecha de una alteración genética, estos resultados influyen en el manejo médico, o proporcionan un diagnóstico que permite guiar el consejo genético, por esta razón, las variantes genéticas deben clasificarse con precisión y consecuentemente[75].

- **Interpretación e información del resultado del estudio genético** cuando se haya realizado.

Podemos definir tres tipos de resultados genéticos:

1. **Resultado positivo:** Identificación de una alteración deletérea responsable del cuadro clínico: se ha encontrado una VP en *BRCA1/2* en el paciente que se ha sometido a la prueba genética de ambos genes. En el asesoramiento post test deben discutirse:
    a. Implicaciones personales del hallazgo: Esto significa que una persona con un resultado positivo tiene un mayor riesgo de cáncer de mama, y/o cáncer de ovario, así como otros tipos de cáncer. Para ayudar a llevar este tipo de riesgo se pueden incluir una serie de recomendaciones como un aumento de las pruebas de cribado, protocolos de quimio prevención o cirugía profiláctica.
    b. Implicaciones familiares: Los familiares de primer grado de un paciente positivo para una VP *BRCA1/2* tienen un 50% de probabilidades de tener esa misma VP.
2. **Resultado negativo:** Ausencia de detección de alteración deletérea: Se deben establecer una pauta de seguimiento clínico basada en los datos clínicos personales y familiares.

3. **Resultado incierto:** Se ha encontrado una alteración en los genes *BRCA1/2*, pero no está claro lo que significa este resultado para el riesgo de cáncer. Ocurre cuando se identifica en el paciente una VSI. Este tipo de resultado podrá ser reclasificado en un futuro si se obtienen más conocimientos sobre este tipo de variantes[76]. Pueden solicitarse muestras de otros familiares tanto sanos como afectos, para ver si cosegrega con la enfermedad, o bien realizar estudios funcionales de pérdida de función de la proteína. A las personas con una VSI se le dan recomendaciones según los antecedentes personales y familiares de cáncer. Esto se discute en el consejo genético pre test.

Tanto para resultados negativos, como de significado incierto, se puede valorar la posibilidad de ampliar el estudio genético a otros genes implicados en las patologías personales o familiares[74].

- **Plan de seguimiento médico** en función del resultado del estudio genético o de la historia personal y familiar si no se consideró oportuno realizar el test.

Medidas preventivas para portadores de VP en los genes *BRCA1/2*[73]*:*

Tabla 5. Medidas preventivas

En mujeres	**CM**	Vigilancia	Mamografías anuales a partir de los 30-35 años y resonancia mamaria anual a partir de los 25 años
		Cirugía profiláctica	Mastectomía y salpingo-ovariectomía
	CO	Vigilancia	Ecografía transvaginal y marcador tumoral Ca 125 cada 6-12 meses a partir de los 30 años
		Cirugía profiláctica	Salpingooforectomía bilateral
En varones	**CPro**	Vigilancia	Cribado anual a partir de los 40-45 años
Ambos sexos	**Melanoma**	Vigilancia	Cribado dermatológico
	CP y CC	Vigilancia	Cribado en función de la historia familiar

CC: cáncer de colon; CM: cáncer de mama; CO: cáncer de ovario; CP: cáncer de páncreas; CPro: cáncer de próstata.

Una de las características que entorpece el establecimiento de una probabilidad pre test es la alta frecuencia de cáncer de mama esporádico. En el caso del CMOH es frecuente la existencia

de **fenocopias**, es decir, la presencia de un cáncer de mama no relacionado con el síndrome de CMOH en una familia con CMOH.

3.8 Criterios de selección para el estudio genético de los genes BRCA1 y BRCA2.

Aunque hasta hace poco tiempo no existía un consenso a nivel internacional, los criterios para la selección de familias para el estudio de los genes *BRCA1/2* a nivel germinal, que se adoptan por los diferentes grupos de trabajo, suelen ser bastante similares y por lo general, todos ellos tienen en cuenta: el número de casos de cáncer de mama o cáncer de ovario en la familia, la edad de aparición, la presencia de cáncer de mama bilateral o cáncer de mama en varón. La mayoría de ellos comportan una probabilidad de detección de VP superior al 10%.

A nivel nacional, los criterios consenso de selección utilizados por el Grupo de Trabajo de la Sociedad Española de Oncología Médica (SEOM2019) son los siguientes[73]:

a) **Independientemente de la historia familiar si:**

- Paciente con cáncer de mama y cáncer de ovario sincrónico o metacrónico
- Cáncer de mama menor ≤ 40 años
- Cáncer de mama bilateral <50 años el primer diagnóstico
- Cáncer de mama triple negativo ≤60 años
- Cáncer de ovario epitelial no mucinoso
- Ascendencia con VP fundadora
- Detección de una mutación a nivel somático con una frecuencia alélica >30%.
- Pacientes con cáncer de mama metastásico HER2 negativo como candidatas a terapia con PARPi

b) **Familias con dos miembros afectos con cáncer de mama y/o cáncer de ovario** y al menos una de las siguientes características:

- Varón con cáncer de mama
- Cáncer de mama <50 años y cáncer de próstata o cáncer de páncreas <60 años
- Cáncer de mama y ovario
- Dos casos de cáncer de mama ambos casos diagnosticados <50 años
- Uno de ellos es bilateral y otro cáncer de mama diagnosticado <60 años

c) **Familias con tres o más individuos con cáncer de mama** (al menos uno premenopaúsico) y/o cáncer de ovario y/o cáncer de próstata o cáncer de páncreas en la misma rama de la familia

No obstante, existen **limitaciones** reconocidas de estos criterios, y su interpretación debe incluir la posibilidad de que la prueba se realizó sobre un individuo con cáncer esporádico, dentro de una familia de alto riesgo. Además dos estudios centrados en poblaciones no seleccionadas de pacientes con cáncer de mama, hallaron que el 30-50% de las mujeres con una VP en los genes *BRCA1/2* no tenían antecedentes familiares importantes o conocidos de cáncer de mama o de ovario[77,78].

A finales del año 2020 se publicó una guía consenso internacional cuya finalidad es actualizar y estandarizar el estudio genético de los *BRCA1/2* tanto a nivel somático como geminal.[32] Estas guías incluyen un abordaje multidisciplinar, realizada por un grupo de expertos, con especialistas de diferentes ámbitos: genetistas, oncólogos médicos y quirúrgicos, patólogos, especialistas en ética y representantes de pacientes. Este grupo siguió una metodología para el desarrollo de unas directrices específicas, para ello se basaron en varios puntos:

- Evaluación de la probabilidad de encontrar una VP en los genes *BRCA1/2* realizando una revisión de la literatura, los modelos de riesgos y citas de los expertos.
- Valor terapéutico para portadoras de VP en los genes *BRCA1/2* para la terapia con los PARPi en cáncer de mama precoz y avanzado.

Las recomendaciones sugeridas tras este trabajo abordan, por tanto, el enfoque preventivo y terapéutico y se resumen en:

1) Preventivas:
    a. Pacientes con historia personal o familiar con alta probabilidad de encontrar VP en *BRCA1/2* (Grado A).
    b. Pacientes con probabilidad media de encontrar VP en *BRCA1/2* (Grado B).
    c. Independientemente de la probabilidad de encontrar una VP en los *BRCA1/2*, el paciente debe decidir libremente, previo asesoramiento genético.
    d. No está recomendado el estudio rutinario de los genes *BRCA1/2* en pacientes con baja probabilidad de encontrar VP en *BRCA1/2* (Grado C), sin valor terapéutico en la práctica clínica.
    e. Se recomienda el asesoramiento genético antes y después de realizar un test predictivo de los *BRCA1/2*.
2) Tratamiento personalizado para el cáncer de mama:
    a. Para pacientes con cáncer de mama recién diagnosticado y que cumplen los criterios de alta probabilidad de encontrar una VP en línea germinal debe considerarse realizar el test de *BRCA1/2* si el estado de portador afecta al tratamiento del cáncer de mama.
    b. Pacientes con cáncer de mama metastásico que requieren quimioterapia se recomienda realizar la prueba a nivel germinal, ya que se recomienda la

terapia de platino frente a la de taxanos, en pacientes sin tratamiento previo de platino.

c. En el cáncer de mama metastásico HER2 negativo, el test germinal se recomienda porque debe ofrecerse el tratamiento con olaparib o talazoparib de primera línea a tercera línea de quimioterapia para mujeres con VP.

3) PARPi:

a. El test *BRCA1/2* debe ofrecerse con fines terapéuticos con PARPi a pacientes con cáncer de mama metastásico HER2 negativo, cáncer de próstata resistente a la castración, cáncer de páncreas metastásico sensible a platino y cáncer de ovario epitelial de alto grado estadio III/IV de FIGO, recién diagnosticado o recurrente, este proceso debe ser rápido una vez obtenida la información genética específica.

b. Se recomienda el test *BRCA1/2* a nivel germinal para terapia dirigida con PARPi, independientemente del criterio alto o moderado de probabilidad de encontrar VP en los *BRCA1/2*, ya que entre un 10 a un 75% de los pacientes con cáncer de mama, cáncer de ovario, cáncer de próstata o cáncer de páncreas no cumple con estos criterios.

c. El cáncer de ovario epitelial cumple los criterios de alta probabilidad de encontrar VP en los *BRCA1/2*, con finalidades de reducción de riesgos y valor terapéutico. Por lo tanto, el test germinal debe ser ofrecido a cualquier mujer con carcinoma de ovario epitelial no-mucinoso en el momento del diagnóstico. Debe recomendarse el estudio de los *BRCA1/2* en tumor, en caso de no encontrar una VP a nivel germinal.

d. Los oncólogos y cirujanos debidamente formados que están involucrados en la atención del cáncer, podrían dar la información inicial de asesoramiento, en coordinación con un equipo multidisciplinario, incluidos los genetistas.

4) Test en tumor:

a. Cuando la prueba en tumores tiene finalidad diagnóstica y terapéutica: el paciente debe ser consciente de los riesgos genéticos implícitos, incluidos los familiares, así como problemas de prevención que pueden surgir de factores genéticos en las pruebas tumorales, porque la mayoría de los hallazgos de mutaciones en *BRCA1/2* tumorales reflejan una predisposición de la línea germinal. Por lo tanto, esta prueba genética requiere información y un consentimiento informado antes de realizarse. En caso de resultar con valor terapéutico, esta información debe ser proporcionada por profesionales, como oncólogos con experiencia en el diagnóstico genético y con el manejo de enfermedades hereditarias, como el cáncer de mama, trabajando en conjunto con un el consejo genético.

b. Para realizar una interpretación adecuada de los resultados de secuenciación ADN tumoral, debe tenerse en cuenta el % de células tumorales, la profundidad de lectura adecuadas para detectar CNVs y la fracción alélica mutada. Las técnicas y *pipelines* adecuados que permitan la detección de SNV y CNV, como la secuenciación masiva.

c. Se debe ofrecer el test a nivel germinal a cualquier paciente en el cual se haya identificado una mutación patogénica o probablemente patogénica de los *BRCA1/2* a nivel tumoral.
d. Deben desarrollarse programas educativos para aumentar la sensibilización y formación de los profesionales sanitarios en oncología, en particular oncólogos, cirujanos, especialistas y representantes ético legales del paciente, para mejorar sus habilidades comunicativas tanto para la realización de los test *BRCA1/2* con fines terapéuticos como para la atención personalizada según los resultados genéticos.

4. Diagnóstico genético del cáncer de mama y ovario hereditario: de las técnicas convencionales a la secuenciación masiva.

La gran mayoría de las alteraciones genéticas responsables del CMOH corresponden a variantes puntuales en los genes *BRCA1/2,* consisten en cambios de una sola base o deleciones/inserciones de pocas bases que producen truncamiento de proteínas, interrupción del procesamiento del ARN mensajero o sustituciones de aminoácidos que tienen un impacto significativo en la función de la proteína. Estas variantes son detectadas mediante la **secuenciación Sanger**, que es el método *gold standard* para el análisis genético de variantes puntuales. Sin embargo, el gran tamaño de estos genes, la laboriosidad del método de secuenciación y el alto coste, hacen complicada su aplicación cuando el número de muestras a procesar es elevado.

Este método fue desarrollado por Fred Sanger en 1975 y se basa en la copia del proceso natural de replicación del ADN de nuestras células. El método original fue mejorándose y automatizándose hasta convertirse en la tecnología de referencia para el diagnóstico de alteraciones genéticas en el genoma humano[79,80].

Pasaron casi 30 años desde la publicación del método de Sanger, hasta que apareciera una nueva tecnología de secuenciación de ácidos nucleicos. La principal característica de estas nuevas metodologías, es la posibilidad de secuenciación masiva de forma paralela, esto significa, son capaces de secuenciar millones de fragmentos de ADN de forma paralela a un precio mucho más barato por base que la secuenciación Sanger[81].

Hace unos años, debido al alto coste de secuenciación por base que se obtenía con la tecnología Sanger (tabla 6), en algunos laboratorios, se solía realizar un cribado de variantes inicial mediante DHPLC (Denaturing High- Performance Liquid Chromatography) o HRMCA High-Resolution Melting Curve Analysis) para detectar VP y posteriormente secuenciarlas mediante secuenciación Sanger[82,83]. Estas técnicas tradicionales eran muy laboriosas y poco eficientes, y muchas veces se tenía que realizar posteriormente la secuenciación del gen completo.

En el año 2005, comienza la carrera de avances para lograr el objetivo de secuenciar un genoma completo por un precio de 1000 dólares. El primer método de secuenciación masiva fue el método de **pirosecuenciación** después, surgen nuevas alternativas de secuenciación,

que utilizan el mismo principio de didesoxinucleotidos terminadores en sus protocolos, aunque con mejoras innovadoras[84]. Estos sistemas presentan una gran capacidad de generación de datos, basada en la producción de varios millones de secuencias de un tamaño definido (entre 35 y 400pb) en varios días de reacción (de 1 a 12 días), esto permite que cada una de las posiciones estudiadas sean secuenciadas decenas o centenas de veces. Estos nuevos métodos de secuenciación masiva no ofrecen lecturas tan largas como el método clásico de Sanger, aunque en pocos años se ha mejorado la longitud de las secuencias obtenidas y en algunos casos alcanzan longitudes comparables a Sanger, pero en cambio, son capaces de generar cientos de miles de reacciones en paralelo (de alto rendimiento) gracias a la inmovilización de las reacciones en una superficie sólida. De esta manera, la cantidad de reactivos necesario se minimiza al máximo y se abarata el coste por base leída[85,86]. La generación de secuencias de tamaño relativamente pequeño, dificulta el proceso de alineamiento, a diferencia de lo que sucede con la secuenciación Sanger. Para solucionar estos problemas, es necesario generar un número de lecturas elevado en cada una de las posiciones genómicas, punto clave en los análisis con finalidad diagnóstica, como los que nos ocupan. La ausencia de una buena cobertura genera problemas para la detección de variantes, por lo que disminuye la fiabilidad de los resultados. Este problema, se agrava, ante la presencia de VP del tipo deleciones o inserciones[87].

Para poder detectar CNVs, que consisten en deleciones de uno o varios exones o bien, las duplicaciones, menos frecuentes, se requieren ensayos alternativos, ya que, con la técnica de PCR, no es posible detectarlos. Se emplean técnicas adicionales como el **MLPA** (*Multiplex Ligation-dependent probe Amplification*) para poder identificarlos. Recientemente se ha demostrado, que mediante la tecnología de la secuenciación masiva es posible detectar estas CNVs, sin necesidad de utilizar una técnica alternativa en todas las muestras a ensayar, únicamente se utilizarían para confirmación de casos positivos[88].

Durante los últimos años se han ido desarrollando distintas **técnicas de captura de ADN basadas en la hibridación de sondas** o en **sistemas de PCR multiplex**, que permiten capturar zonas concretas del genoma, que van desde pocas kilobases hasta la captura del exoma completo, y que combinados con la secuenciación masiva hacen posible realizar un cribado de variantes en todos los genes implicados en el cáncer hereditario, proporcionando una cantidad de información muy extensa sobre la genética de un individuo[89].

Por ello, la tecnología NGS está suponiendo una revolución en el ámbito asistencial tanto en el diagnóstico de patologías hereditarias, como en el estudio de alteraciones moleculares de tipo somático, gracias a su gran sensibilidad y especificidad analíticas y de costes, ya que son asequibles y los tiempos de respuesta son muy cortos. Además, la llegada de los mini-secuenciadores con tecnología NGS nos ofrecieron nuevas oportunidades para el diagnóstico de rutina, a través de la secuenciación de paneles de genes de pequeño tamaño, de una forma mucho más rentable, y en ocasiones mucho más eficiente, que las tecnologías tradicionales de diagnóstico[90]. Su menor coste de adquisición, en comparación con las plataformas de NGS de mayor tamaño, han favorecido la introducción de estas novedosas tecnologías de secuenciación en pequeños y medianos laboratorios[91].

Las nuevas plataformas de NGS van aumentando sus capacidades técnicas ofreciendo un nivel de resolución mayor y minimizando los errores. Paralelamente se está produciendo una evolución rápida de las herramientas bioinformáticas necesarias para poder obtener el máximo aprovechamiento de toda la cantidad de información que se genera con estas plataformas. Así, como consecuencia, obtenemos unos niveles de sensibilidad, especificidad y exactitud cada vez mayores en el análisis, y esto nos permite un elevado rendimiento en la detección y caracterización de variantes puntuales y también CNVs y alteraciones estructurales como inversiones y translocaciones[74].

Estas nuevas tecnologías dan cobertura de forma eficiente a los estudios genéticos de tipo diagnóstico, es decir, a la búsqueda de variantes causales en los casos índice de familias con sospecha de CMOH. En cambio, para los estudios predictivos de riesgo en familiares, cuando ya se conoce la variante causal en la familia, no se realizan por NGS sino mediante secuenciación Sanger, que el coste es más eficiente para estos casos. Los estudios predictivos de riesgo, suponen la mayoría de los estudios de diagnóstico que se realizan en los laboratorios de diagnóstico[92].

La implementación de las nuevas tecnologías en la práctica clínica requiere de una evaluación de las mismas y una adecuada planificación. La evaluación tecnología hace referencia a los procesos de validación analíticos y clínicos de estos nuevos procedimientos, como a su utilidad clínica. El análisis de un panel de genes responsables de uno o varios síndromes de sospecha se debe de realizar en el contexto asistencial. En la actualidad, la implementación en la práctica clínica de estos sistemas de NGS requieren una confirmación de los resultados mediante la secuenciación Sanger o alternativamente, se requiere una adecuada validación analítica y clínica por los laboratorios clínicos acreditados, hasta que estas nuevas plataformas y sistemas de análisis puedan acreditarse para diagnóstico por las entidades reguladoras[87].

En este estudio se describe el desarrollo de un protocolo de trabajo orientado al diagnóstico genético del CMOH, secuenciando los genes *BRCA1/2* a partir de datos procedentes de tres plataformas de secuenciación: secuenciación Sanger ABI 3730 Applied Biosystems, y dos de secuenciación masiva; 454 *GS Junior* (*Roche Diagnostics, Switzerland*) y *MiSeq (Illumina, USA)*. Así como la búsqueda de un nuevo algoritmo diagnóstico y su validación. En la tabla 6 se describen las principales diferencias entre las tres técnicas empleadas, las mayores diferencias se observan en las longitudes de lectura, tiempos de procesamiento y costes. Mediante la plataforma de *Roche Diagnostics* obtenemos secuencias de 400pb de longitud de lectura, 1 millón de lectura por carrera y un tiempo de 10 horas, el error más frecuente son las variantes tipo *indels*. Con la plataforma de *Illumina* obtenemos una longitud de lectura de 150pb, 10 millones de lectura por carrera y un tiempo de procesado de 26 horas, el coste es menor que con la plataforma de *Roche Diagnostics* y finalmente la plataforma de secuenciación Sanger, con la cual obtenemos una mayor longitud de lectura 800pb, en menor tiempo, pero con un alto coste y pocas lecturas.

Tabla 6. Comparativa de las distintas plataformas de secuenciación utilizadas en la realización de este trabajo. Modificado de Liu et al y Escalante et al[93,94]

	Plataforma		
Características	***454 GS Junior***	***MiSeq***	***SANGER 3730xl***
Longitud de lectura máxima (pb) /procesamiento (Gb/carrera)	400 pb/0,5	2x 150 pb/2	800 bp/ 0,00006
Millones de lecturas/carrera	1 millón	10 millones	0,000096
Tiempo carrera	10 horas	26 horas	3 horas
Tipo de librerías/secuenciación	PCR en emulsión /pirosecuenciación	PCR puente/sintesis	PCR/sintesis
Tipo de error/ratio error	*Indels*/1%	Sustituciones/0,80%	Sustituciones/0,1%
Coste por Gb	1000 €	445 €	5.300 €
Coste aparato aproximadamente	95.500 €	110.000 €	84.000 €

Tp: tiempo; pb: pares de bases; Gb: Gigabytes

Las principales ventajas de la tecnología de la NGS que han revolucionado el campo del diagnóstico genético han sido la **rapidez**, el **bajo coste** y la **precisión**.

A lo largo de 10 años (2005-2014) nuestro centro ha estado trabajando en la identificación y estudio de familias con cáncer de mama y/o cáncer de ovario sospechosas de ser portadoras de una VP en los genes *BRCA1* o *BRCA2*. El estudio se realizó con diferentes tipos de técnicas desde las más manuales a las más automatizadas. Este hecho nos ha permitido ir valorando la evolución de los distintos aspectos técnicos y genéticos que han girado en torno al diagnóstico de estas familias. Una consecuencia ha sido valorarlos y establecer lo que hoy constituye una rutina de trabajo, que, sin dejar de incorporar la automatización, la bioinformática y las técnicas más avanzadas todavía requiere de una buena parte de interpretación y valoración personal.

En esta tesis se recoge esta evolución y su resultado final.

OBJETIVOS:

1) Validación y optimización de una técnica de análisis de variantes en los genes *BRCA1/2*, basada en la amplificación de fragmentos, y empleando distintas plataformas de secuenciación masiva, que permita reducir el tiempo de estudio en cada paciente, respecto a la secuenciación Sanger.
   1.1) Establecer una estrategia de análisis empleando la plataforma de pirosecuenciación *454 GS Junior*
   1.2) Establecer una estrategia de análisis empleando la plataforma de secuenciación por terminación cíclica reversible *MiSeq*

2) Diseño de una estrategia de cribado de variantes altamente sensible y específica, más eficiente en cuanto a coste y tiempo de diagnóstico que permita su aplicación a la práctica clínica de forma rutinaria.
   2.1) Flujo de trabajo propuesto
   2.2) Determinar las limitaciones y ventajas de cada técnica

3) Caracterizar el espectro de variantes de los genes *BRCA1/2* en una cohorte de pacientes con sospecha clínica de síndrome de cáncer de mama y ovario hereditario.

4) Reevaluación y reclasificación de VSI, de forma retrospectiva, en base al conocimiento actual.

MATERIALES Y MÉTODOS:

5. Familias seleccionadas

En este trabajo se ha realizado un estudio retrospectivo, que abarca desde el mes de marzo del año 2005 hasta marzo de 2014. En este periodo, se recogieron muestras de sangre periférica en EDTA y ADN de los probandos pertenecientes a 455 familias con sospecha clínica de CMOH, procedentes de diferentes hospitales españoles, después de obtener el consentimiento informado. De ellas, 435 eran mujeres y 20 varones (14 con cáncer de mama, 1 cáncer de páncreas, 1 cáncer de colon y 4 individuos asintomáticos).

Las muestras proceden de distintos hospitales pertenecientes a diferentes comunidades autónomas. La selección de los casos ha sido realizada siguiendo los criterios clínicos empleados por cada hospital. Para unificar los criterios de selección de los pacientes, se han empleado unos criterios laxos, que permitan englobar a todas las muestras. De este modo, los criterios empleados para la clasificación de las familias han sido los siguientes:

a) Un solo caso familiar

- Paciente con cáncer de mama y cáncer de ovario sincrónico o metacrónico
- Cáncer de mama ≤ a 50 años
- Cáncer de mama bilateral
- Cáncer de ovario
- Cáncer de mama triple negativo ≤ a 50 años

b) Familias con dos miembros afectos con cáncer de mama y/o cáncer de ovario y al menos una de las siguientes características:

- Varón con cáncer de mama
- Uno de ellos es un cáncer de ovario, cáncer tubárico o primario peritoneal
- Al menos uno de ellos diagnosticados < de los 50 años
- Uno de ellos es bilateral y el otro diagnosticado <50 años

c) Familias con tres o más individuos con cáncer de mama y/o cáncer de ovario en la misma rama de la familia independientemente de la edad de diagnóstico

Cuando no se disponía de información clínica de los casos, se clasificaron como no referido (NR). En la tabla 7 se recoge la distribución de los casos índice según los criterios anteriormente descritos para homogeneizar las muestras de nuestro estudio.

Tabla 7. Criterios de selección de pacientes homogenizados y tipos de cáncer asociados		
Sexo	**N.º de casos**	**Edad media de diagnóstico**
Mujer	435	43 años
Varón	20	58 años
Criterios	**N.º de casos**	**Edad media de diagnóstico**
1: Un caso único	101	38 años
2: Dos familiares	114	42 años
3: Tres casos	109	46 años
No cumple criterios (NC)	12	60 años
No referida la historia familiar (NR)	119	-
Tipos de cáncer	**N.º de casos**	**Edad media de diagnóstico**
CM	357	43 años
CM y CO	13	58 años
CM bilateral	21	44,9 años
CO	23	41,16 años
CM varón	14	60,3 años
CP	1	-
CC	1	47 años
CM triple negativo	9	42,85 años
Asintomático	16	-

CC: cáncer de colon; CM: cáncer de mama; CO: cáncer de ovario; CP: cáncer de páncreas; -: no se dispone de la edad de diagnóstico

La edad media de diagnóstico de los pacientes estudiados es 43 años, rango entre 19 y 70 años. Este parámetro ha sido estimado en 230 casos de los que se disponía de la edad de diagnóstico.

De los 455 probandos estudiados, 163 casos han sido estudiados mediante secuenciación Sanger y 292 casos han sido estudiados mediante secuenciación masiva, de ellos 201 casos se estudiaron mediante pirosecuenciación y 91 casos se estudiaron mediante secuenciación por síntesis utilizando la plataforma *MiSeq*. En 308 casos (67,69%) se estudiaron CNVs de los genes *BRCA1/2*, de estos, en 175 casos se realizó el estudio de los genes *BRCA1/2* por MLPA (*Multiplex Ligation-dependent probe Amplification*), representando un 38,46% de los casos y

en 133 casos se estudiaron mediante análisis de fragmentos, representando un 29,23% de los casos (Figura 5).

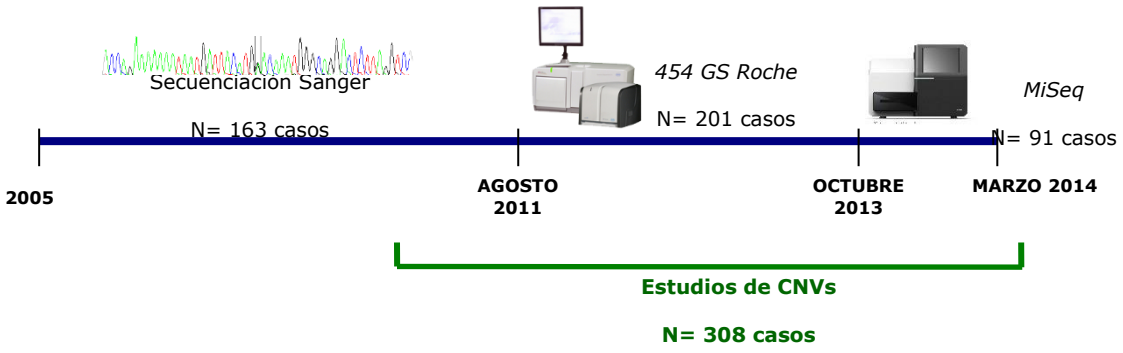


Figura 5. Esquema cronológico de la metodología empleada en la identificación de variantes en los genes BRCA1/2 de los pacientes de nuestro estudio. N= número de casos estudiados.

El flujograma empleado para los casos de secuenciación Sanger se recoge en la figura 6. Las muestras de ADN se secuencian mediante secuenciación Sanger, de manera secuencial, primero se secuencia el gen *BRCA1* y después el gen *BRCA2* (o viceversa, según petición de facultativo especialista), si el resultado es negativo, en algunos casos, se estudiaron CNVs mediante MLPA.

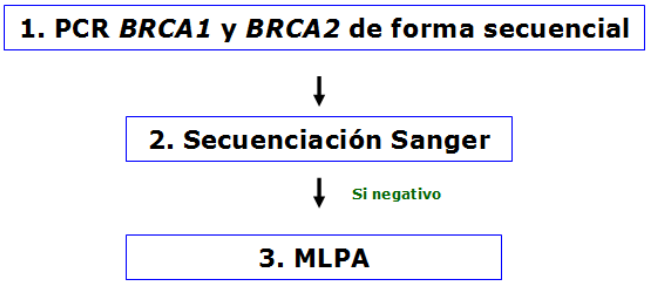


Figura 6. Flujograma empleado en la identificación de variantes en los genes BRCA1/2 mediante la metodología Sanger.

Los casos estudiados mediante secuenciación Sanger, son los más antiguos del estudio, siendo recopilados desde el año 2005 hasta agosto del 2011. En los primeros años de estudio, la metodología se centraba en la identificación de variantes puntuales o de pequeñas deleciones o inserciones que fueran patogénicas o probablemente patogénicas y tan solo en 4 estudios se estudiaron grandes deleciones o duplicaciones mediante MLPA.

El flujograma empleado para los casos estudiados mediante secuenciación masiva se encuentra descrito en la validación de la técnica de secuenciación masiva, en el apartado de los resultados, figura 32.

En todos los casos estudiados mediante secuenciación masiva, se ha realizado el estudio de ambos genes, y se ha completado el estudio con el análisis de CNVs empleando la metodología MLPA en 171 casos, y realizando un análisis de fragmentos en 133 casos.

6. Extracción de ácidos nucleicos

6.1 Extracción de ADN a partir de sangre periférica

Para la obtención del ADN genómico de las muestras de sangre periférica se realizaron dos tipos de extracciones de ácidos nucleicos, una manual y otra automatizada (figura 7).

Se extrajo el ADN genómico de todos los casos índices a partir de 200 o 400 µl de sangre periférica, extraída en su centro de origen, a partir de en un tubo con anticoagulante EDTA.

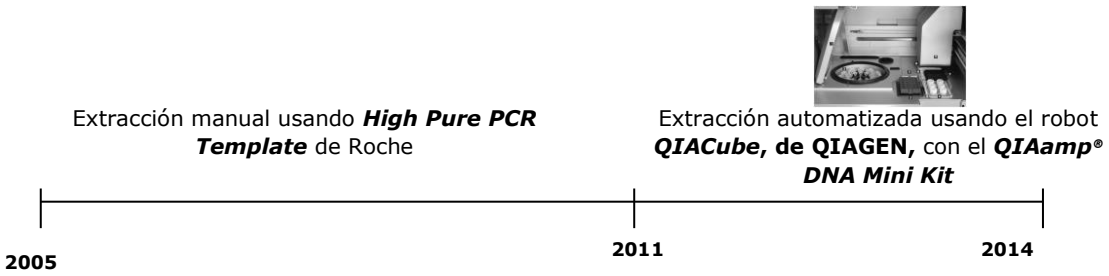


Figura 7. Esquema cronológico de la metodología empleada para la extracción de ADN de los pacientes de nuestro estudio.

Extracción de ADN mediante con el kit High Pure PCR Template (Roche)

Se realizó la extracción manual de ácidos nucleicos empleando el kit *High Pure PCR Template*, siguiendo las instrucciones del fabricante. Esta técnica consta de una fase de lisis, tras la que se procede al aislamiento del ADN, después se realiza una fase de purificación y finalmente se eluye el ADN en agua. Este tipo de extracción se llevó a cabo para todas las muestras analizadas mediante secuenciación Sanger, desde el año 2005 hasta el año 2011.

Extracción de ADN mediante el kit QIAamp DNA Mini Kit (QIAgen) y automatización mediante QIACube.

Se realizó la extracción automatizada de ácidos nucleicos mediante el kit *QIAamp DNA Mini Kit*, utilizando el robot *QIACube* (QIAgen), siguiendo las instrucciones del fabricante. Esta extracción se llevó a cabo para todas las muestras desde el año 2011 hasta la actualidad.

El protocolo de purificación puede automatizarse completamente mediante *QIACube*, siguiendo los mismos pasos que el procedimiento manual (lisis, unión, lavado y elución).

6.2 Cuantificación del ADN

La calidad y concentración de las muestras de ADN se determinó con el espectrofotómetro *Nanodrop* (Thermo Scientific), que mide la absorbancia de las muestras a las longitudes de onda de 230nm, 260nm y 280nm. La absorbancia del ADN es máxima a 260nm. La absorbancia a 230nm (longitud de onda de mínima absorbancia del ADN) está influenciada por la cantidad de sales presentes en la muestra y a 280nm por la cantidad de proteínas. Por ello,

consideramos valores óptimos para realizar un estudio de MLPA o de secuenciación masiva, una relación de absorbancia de 260/280 y 260/230 entre 1.8 y 2, y una concentración de ADN adecuada de unos 50 ng/µl. Muestras con valores inferiores a los señalados fueron precipitadas mediante etanol.

7. Rastreo de variantes en los genes *BRCA1* y *BRCA2* mediante secuenciación Sanger.

7.1 Diseño de oligos

Para el diseño de oligos se utilizó el programa bioinformático Primer3 (*https://primer3.ut.ee/*) que se encuentra para su libre uso en distintos servidores *web*. Se diseñaron los oligos específicos para la amplificación de los exones y regiones intrónicas adyacentes (aproximadamente +/- 20 nucleótidos) de los genes *BRCA1* (32 amplicones) y *BRCA2* (50 amplicones). Los 82 fragmentos obtenidos tienen un tamaño de entre 250-450pb a excepción de un fragmento que alcanza los 976pb. Para los exones de gran tamaño los oligos diseñados permitan la amplificación de fragmentos solapantes en aproximadamente 30pb para cubrir adecuadamente todas las regiones codificantes. El transcrito empleado para cada gen ha sido para *BRCA1*: NM_007294.3 y para *BRCA2*: NM_000059.3. Las secuencias de los oligos se diseñaron basándonos en las secuencias nucleotídicas de los genes a estudiar (*BRCA1/2*, disponibles en la red (*https://genome.ucsc.edu/).*

En el anexo I se detallan las secuencias de los 82 pares de *oligos* utilizadas para la amplificación de los genes *BRCA1/2*.

7.2 PCR

La PCR (***P****olymerase* ***C****hain* ***R****eaction*) o reacción en cadena de la polimerasa es una técnica de biología molecular que permite la amplificación de un fragmento de ADN utilizando la propiedad natural de las ADN polimerasas para replicar hebras de ADN.

Las tablas con los reactivos de la PCR y las condiciones de ciclado de PCR y tamaño del fragmento que genera cada par de oligos tanto para los genes *BRCA1/2*, se encuentran descritas en el anexo II.

En general, el volumen de la PCR con el que trabajamos es 50 µl.

Evaluación del resultado de la PCR

Se realiza una electroforesis en gel de agarosa para comprobar la presencia de ADN después de la realización de una PCR, así como la separación por tamaños de las moléculas de ADN obtenidas. El gel se obtiene disolviendo la agarosa en tampón TBE y se tiñe con bromuro de etidio. La concentración de agarosa empleada fue del 2%. La migración de los fragmentos de ADN varía dependiendo del tamaño, conformación molecular, concentración de la agarosa en el gel, voltaje, dirección del campo eléctrico, presencia de colorantes añadidos que se intercalan en el ADN y de la composición del tampón en el gel.

A 5 µL del producto de PCR se le añaden 1,5 µL de tampón de carga, azul de bromofenol. Nos servimos de un marcador de peso molecular, de 100pb, para así poder inferir el tamaño de los fragmentos.

Purificación de productos de PCR

Para secuenciar productos de PCR es necesario purificarlos previamente, para eliminar los dNTPs y oligos no consumidos en la amplificación y, dímeros de oligos generados en la PCR; todos ellos interfieren en la reacción de secuenciación. Para purificar los productos de PCR empleamos la placa de filtración de MACHEREY-NAGEL: NucleoFast 96 PCR Plate (Macharey-Nagel), que no requiere lavados, ni pasos de centrifugación y permite la purificación simultánea de 96 muestras con una manipulación mínima de las mismas.

7.3 Secuenciación Sanger

El kit de secuenciación empleado está basado en la química de los *BigDyes* y es suministrado por Applied Biosystems: *ABI PRISM BigDye Terminator v3.1 Ready Reaction Cycle Sequencing Kit*.

Los reactivos y las condiciones empleadas para la secuenciación de cada fragmento se detallan en la tabla 1 y 2 del anexo III.

Purificación de los productos de secuenciación

Se ha usado el método de la filtración mediante membrana en la que oligos, dímeros, nucleótidos y sales son filtrados a través de la misma bajo presión de vacío, quedando los productos de la secuenciación retenidos en la membrana, pudiéndose recuperar de modo sencillo mediante elución. El formato es en placa lo que permite la purificación simultánea de hasta 96 muestras, la placa empleada es *Multiscreen Vacuum Manifold* (*Merck Millipore*).

Electroforesis capilar

En nuestro laboratorio, el análisis de las reacciones de secuenciación se realizó mediante un secuenciador 3730xl DNA Analyzer (*Applied Biosystems*, USA). Se trata de un sistema de electroforesis de 96 capilares que realiza los procesos de rellenado de capilares, inyección de las muestras, separación de los fragmentos y detección de la fluorescencia de forma automatizada. La modalidad de secuenciación que aplicamos habitualmente es la denominada Long read, la cual implica el uso de capilares de 50 cm y tiempos de carrera de dos horas aproximadamente. El voltaje aplicado en la electroforesis son 8.500 V a una temperatura de 60ºC y con polímero POP-7TM.

7.4 Análisis de secuencias

Las secuencias se analizaron utilizando el programa *SeqScape 2.5.0 (Applied Biosystems, USA),* en el cual hemos introducido previamente la secuencia consenso creando el RDG (*Reference Data Group*) de los genes *BRCA1* (NM_007294.3) y *BRCA2 (*NM_000059.3). Esta secuencia de referencia, ha sido descargada de la página *web UCSC Genome Browser* (URL*: http://genome.ucsc.edu/*). El análisis de los datos para la detección de variantes fue realizado

mediante la comparación de la secuencia de cada una de las muestras respecto de la secuencia de referencia.

8. Rastreo de grandes deleciones/duplicaciones en los genes *BRCA1* y *BRCA2* mediante MLPA.

El análisis de CNVs en los genes *BRCA1/2* se realizó mediante la técnica de amplificación por PCR multiplex dependiente de ligación, MLPA (*Multiplex Ligation-dependent probe Amplification*). Este método fue descrito por Schouten en 2002[95] y consta de 5 etapas: desnaturalización, hibridación, ligación, amplificación y electroforesis capilar.

El principio del MLPA se basa en la hibridación de dos oligonucleótidos o sondas adyacentes al ADN de estudio, seguida por una unión (ligación) de ambas sondas y posterior amplificación por PCR de los productos ligados. Como se observa en la figura adjunta (Figura 8), la sonda de la izquierda, presenta una secuencia de unión específica para cada fragmento a estudio y una secuencia terminal, que es común a todas las sondas y que posteriormente servirá de región de unión para los oligos en la PCR. La sonda de la derecha, es más larga y está compuesta por las dos regiones presentes en la sonda de la izquierda, más una secuencia espaciadora. Esta secuencia espaciadora, diferente para cada fragmento, es la que determina el tamaño final de la sonda y, por lo tanto, de cada uno de los productos amplificados por PCR, lo que permite diseñar múltiples fragmentos diferentes que, tras migrar en un secuenciador, se separaran según su tamaño.

El gen *BRCA1* se analizó con el kit P002-C2 (*MRC-Holland, Amsterdan, Holanda*) y el gen *BRCA2* con el kit P045-B3 (*MRC-Holland, Amsterdan, Holanda*) siguiendo en ambos casos las instrucciones del fabricante. En el caso del gen *BRCA1*, la mezcla contiene 35 sondas diferentes y entre ellas, 9 fragmentos control. En el caso del gen *BRCA2*, se incluyen tres sondas para el gen *CHEK2*. Una de ella, solo amplifica en el caso de que esté presente la VP c.1100delC. La mezcla contiene 43 sondas diferentes y entre ellas, 8 fragmentos control. Una de las limitaciones de esta técnica es que algunos cambios puntuales en la secuencia de ADN a analizar, pueden interferir en la hibridación de alguna sonda y se pueden generar falsos positivos.

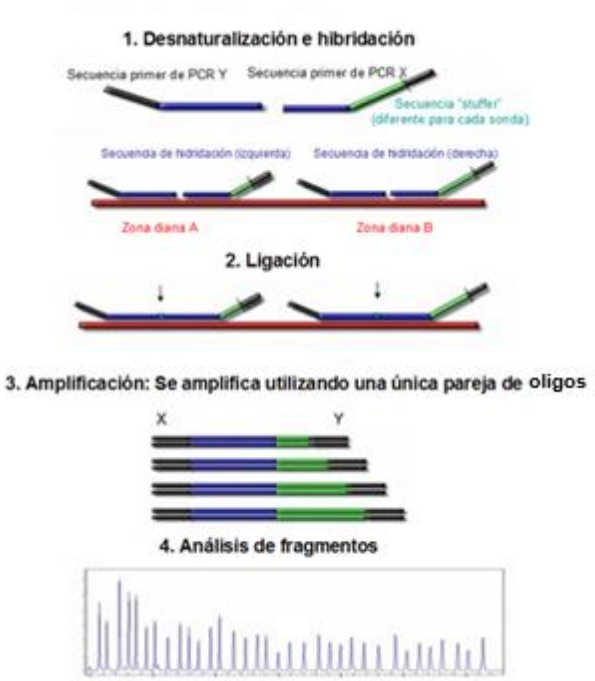


Figura 8. Esquema de las distintas etapas que componen la técnica de MLPA. (Figura modificada de MRC HOLLAND (http://www.mlpa.com/).

En resumen, 125 ng de ADN genómico fueron desnaturalizados a 98ºC e hibridados durante 18h a 60ºC con la mezcla de sondas (*probe mix*). Las muestras fueron entonces tratadas con ligasa-65 a 54ºC durante 15 minutos. Las reacciones de PCR subsecuentes fueron realizadas usando un cebador marcado con 6-carboxyfluoresceína (FAM). En cada estudio se incluyeron tres controles y un control negativo. Los productos de PCR se resolvieron usando un secuenciador *3730xl DNA Analyzer*, mediante electroforesis capilar y el resultado fue analizado con el *software Genemapper (Applied Biosystems*) para análisis de fragmentos y *software Coffalyser* (*MRC-Holland*), para realizar el análisis comparativo entre las muestras. Los límites relativos normales se establecieron entre 0,75 y 1,30, Se consideró que existía una deleción cuando las sondas que mostraban un valor por debajo de 0,75 y una duplicación, cuando las sondas mostraban un valor por encima de 1,30 respecto a los valores de los controles normales.

En caso de resultados con posibles deleciones o duplicaciones, el ensayo se repitió, esta vez utilizando el kit de sondas confirmatorias, que contienen sondas que hibridan en distintas secuencias de cada exón, en ensayo para el gen *BRCA1* con el kit P087-C1 (*MRC-Holland*) y para el gen *BRCA2* con el kit P077-A2 (*MRC-Holland*). En el caso del gen *BRCA1*, la mezcla contiene 36 sondas diferentes y entre ellas, 10 fragmentos control. En el caso del gen *BRCA2* se emplean sondas para todos los exones, excepto para el exón 22. La mezcla contiene 48 sondas diferentes y entre ellas, 10 fragmentos control.

9. Rastreo de variantes en los genes *BRCA1* y *BRCA2* mediante secuenciación masiva.

9.1 Captura de las regiones diana de los genes BRCA1 y BRCA2 y evaluación de fragmentos.

Las amplificaciones se realizaron utilizando el *kit BRCA MASTR Assay (Multiplicom, Niel, Belgium).* Se trata de un sistema de captura, que está diseñado para amplificar por PCR todas las regiones codificantes de los genes *BRCA1/2* utilizando cinco mezclas de oligos multiplexados. Para la validación y primeras tandas de muestras se empleó la v2.0 y la v2.1 para el resto de las muestras. La diferencia entre el kit 2.0 y 2.1, es el número de amplicones para la amplificación de los genes *BRCA1/2*, son 94 y 93 respectivamente.

A los amplicones resultantes, se les agrega, mediante una segunda PCR, un identificador molecular (MID) y un adaptador específico para cada aparato de secuenciación masiva, después se combinan y se secuencian. Hasta octubre de 2013 se utilizó un 454 *GS Junior* (*Roche Diagnostics, Switzerland*) y a partir de noviembre de 2013 un secuenciador *MiSeq (Illumina, USA)* siguiendo las instrucciones del fabricante.

Los amplicones debidamente etiquetados, se mezclan y cada librería amplificada se purifica posteriormente. A continuación, las librerías amplificadas y marcadas de diferentes muestras, se pueden agrupar y se procesan mediante PCR en emulsión o bien mediante amplificación en puente seguida de la secuenciación. Las lecturas de las secuencias resultantes son analizadas para identificar variantes en las posiciones al comparar con la secuencia de referencia de los genes *BRCA1/2* (Figura 9).

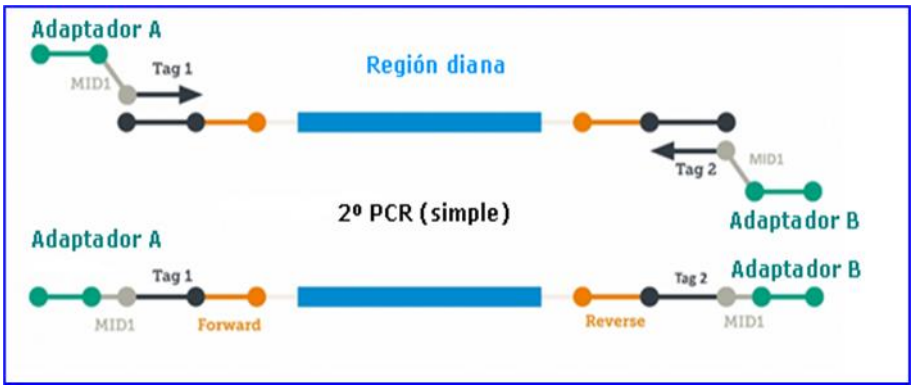


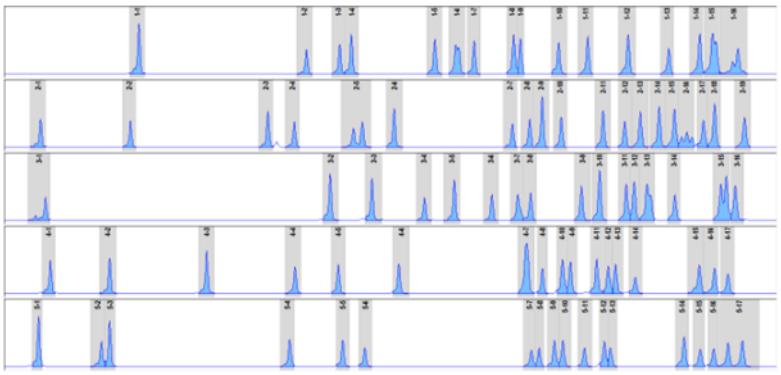
Figura 9. Tecnología del kit BRCA MASTR. **En la figura de arriba, se representa la PCR multiplex, específica de las regiones a amplificar de los genes *BRCA1/2*. En la gráfica de abajo, se lleva a cabo una PCR universal, para el marcaje MID y los adaptadores específicos del secuenciador masivo a emplear.**

Paralelamente, se realizó un análisis de fragmentos mediante electroforesis capilar con un doble objetivo:

* confirmar la amplificación uniforme de los productos de PCR.

* analizar el tamaño de cada uno de los fragmentos incluidos en cada PCR multiplex, para detectar *indels* y para evaluar regiones de homopolímeros de 6pb o superiores.

Los amplificados se corren en un sistema electroforesis capilar de alta resolución, mediante un secuenciador *3730xl DNA Analyzer*, (*Applied Biosystems, USA)* y se analizan empleando el *software GeneMapper (Applied Biosystems, USA*) o bien el *Peak Scanner* (*Applied Biosystems, USA*). En la figura 10, se muestra los perfiles que observamos en el análisis de fragmentos.



Bin	Amplicon	Bin	Amplicon	Bin	Amplicon	Bin	Amplicon	Bin	Amplicon
1-1	BRCA2_ex09_01	2-1	BRCA1_ex20_01	3-1	BRCA2_ex06_01	4-1	BRCA1_ex03_01	5-1	BRCA1_ex22_01
1-2	BRCA1_ex23_01	2-2	BRCA1_ex19_01	3-1	BRCA1_ex20_01	4-2	BRCA2_ex11_13	5-1	BRCA1_ex20_01
1-3	BRCA1_ex11_12	2-3	BRCA1_ex16_02	3-2	BRCA1_ex14_01	4-3	BRCA1_ex17_01	5-2	BRCA1_ex09_01
1-4	BRCA2_ex11_07	2-4	BRCA2_ex25_02	3-3	BRCA2_ex19_01	4-4	BRCA1_ex13_01	5-3	BRCA2_ex11_14
1-5	BRCA2_ex13_01	2-5	BRCA2_ex02_01	3-3	BRCA1_ex02_01	4-5	BRCA1_ex16_01	5-4	BRCA2_ex21_01
1-6	BRCA1_ex11_07	2-5	BRCA2_ex11_09	3-4	BRCA2_ex10_04	4-6	BRCA2_ex11_11	5-5	BRCA1_ex10_01
1-6	BRCA2_ex11_19	2-6	BRCA1_ex21_01	3-5	BRCA1_ex11_05	4-7	BRCA1_ex11_06	5-6	BRCA2_ex12_01
1-7	BRCA2_ex15_01	2-7	BRCA2_ex16_01	3-6	BRCA1_ex06_01	4-7	BRCA2_ex14_02	5-7	BRCA2_ex22_01
1-8	BRCA2_ex11_01	2-8	BRCA2_ex08_01	3-7	BRCA1_ex11_08	4-8	BRCA2_ex10_03	5-8	BRCA1_ex11_10
1-9	BRCA2_ex11_10	2-9	BRCA1_ex11_01	3-8	BRCA1_ex12_01	4-9	BRCA1_ex07_01	5-9	BRCA2_ex03_02
1-10	BRCA2_ex24_01	2-10	BRCA1_ex08_01	3-9	BRCA2_ex11_12	4-10	BRCA1_ex05_01	5-10	BRCA2_ex11_02
1-11	BRCA2_ex14_01	2-11	BRCA1_ex11_14	3-10	BRCA2_ex11_05	4-11	BRCA2_ex05_01	5-11	BRCA2_ex11_06
1-12	BRCA2_ex11_15	2-12	BRCA1_ex15_01	3-11	BRCA1_ex18_01	4-12	BRCA2_ex27_03	5-12	BRCA2_ex11_17
1-13	BRCA2_ex11_21	2-13	BRCA2_ex27_02	3-12	BRCA1_ex11_13	4-13	BRCA1_ex11_03	5-13	BRCA1_ex11_02
1-14	BRCA1_ex11_04	2-14	BRCA2_ex11_18	3-13	BRCA2_ex11_08	4-14	BRCA2_ex25_01	5-14	BRCA2_ex27_01
1-15	BRCA2_ex18_01	2-15	BRCA2_ex11_04	3-13	BRCA2_ex11_16	4-15	BRCA1_ex11_09	5-15	BRCA1_ex24_01
1-15	BRCA2_ex10_01	2-16	BRCA2_ex10_05	3-14	BRCA2_ex23_01	4-16	BRCA2_ex11_03	5-16	BRCA2_ex04_01
1-16	BRCA2_ex03_01	2-17	BRCA2_ex10_02	3-15	BRCA2_ex18_02	4-17	BRCA1_ex11_11	5-17	BRCA2_ex26_01
		2-18	BRCA2_ex11_20	3-15	BRCA2_ex07_01			5-17	BRCA2_ex11_22
		2-19	BRCA2_ex17_01	3-16	BRCA2_ex20_01				

Figura 10. Perfiles de las 5 reacciones multiplex del kit BRCA MASTR. **Se observan 5 perfiles diferentes, correspondientes a una muestra control.**

9.2. Pirosecuenciación con 454 GS Junior de Roche.

En nuestro laboratorio se puso a punto la tecnología de la pirosecuenciación mediante la plataforma *454 GS Junior (Roche Diagnostics, Switzerland*)) y se utilizó desde agosto de 2011 hasta octubre de 2013. La pirosecuenciación es una tecnología que permite determinar el orden de una secuencia de ADN mediante luminiscencia. Se basa en el principio de "secuenciación por síntesis" en el que se detecta la incorporación de nucleótidos al ADN por acción de la ADN polimerasa. La particularidad de este método es que detecta la liberación de pirofosfato cuando los nucleótidos son incorporados. A partir del pirofosfato se generará luz

por medio de diversas reacciones enzimáticas en cadena. Nuestro flujo de trabajo, fue primero validado en 6 muestras de pacientes previamente secuenciados por Sanger y la línea *Hapmap* NA12144, para evaluar la concordancia genotípica. Después de la validación se realizaron 201 estudios mediante esta tecnología. Para obtener óptimos de cobertura, profundidad de lectura superior a 50x, solo se incluían de 6-8 muestras por carrera.

9.3 Secuenciación por terminación cíclica reversible con MiSeq de Ilumina.

En nuestro laboratorio se puso a punto la tecnología de secuenciación por síntesis mediante terminación cíclica reversible con la plataforma *MiSeq* (*Illumina*) y se utiliza desde noviembre de 2013 hasta la actualidad. Durante la secuenciación por síntesis, los reactivos y los nucleótidos marcados con fluorescencia, se añaden a la cadena de ADN. Cada una de las bases emite una intensidad de longitud de onda única que la hará identificable. Nuestro algoritmo de trabajo, fue primero validado con 12 muestras de pacientes previamente secuenciados por pirosecuenciación y confirmación por secuenciación Sanger de las variantes con relevancia clínica. Después de la validación, se llevaron a cabo 91 estudios mediante esta tecnología hasta marzo de 2014. Las muestras secuenciadas por carrera varían de dos a quince muestras, para obtener una media de profundidad de lectura de 1000x.

9.4 Pipelines de análisis de secuenciación masiva

Software de análisis de Amplicon Variant Analyzer (AVA)[96]

El análisis de datos obtenidos mediante el pirosecuenciador *454 GS Junior* se ha realizado mediante el *software* integrado en el equipo de secuenciación, es decir, el analizador de variantes *AVA (Roche).*

El programa *AVA* alinea las lecturas e identifica las diferencias entre la lectura y la secuencia de referencia de *BRCA1/2* (hg19). Las variaciones se muestran gráficamente en un histograma, donde se indica las posiciones de la variante, y también, en forma de texto, con un alineamiento múltiple con códigos de color, destacando las regiones y las bases de diferencia a partir de la secuencia de referencia. Posteriormente, se emplean un conjunto de filtros, definidos por el usuario, para procesar todas las variantes detectadas. Los criterios del filtrado empleados fueron:

- Cobertura mínima de cada región >50x
- Frecuencia de las variantes ≥20%
- Detección de la variante tanto en la dirección directa como en la reversa

Esta selección permite reducir al mínimo el número de falsos positivos. Finalmente, se obtiene una tabla resumen, dónde se representa la frecuencia de las variantes identificadas para cada

muestra, que previamente han pasado el filtrado definido. Las variantes detectadas de relevancia clínica son confirmadas mediante Secuenciación Sanger.

Este *software* permite la visualización de varias muestras simultáneamente, lo que nos permite de una manera rápida, identificar una variante en relación al resto de muestras de la carrera (Figura 11 y 12).

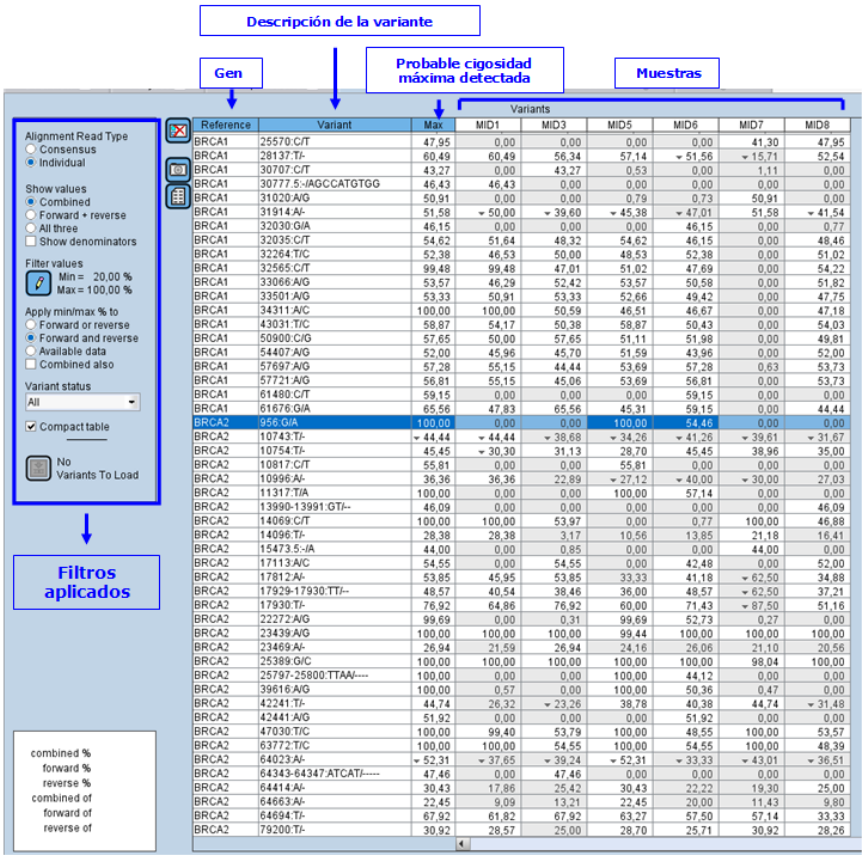


Figura 11. Esquema resumen de las variantes clasificadas por el software AVA. En la parte de arriba, a la derecha, tenemos las muestras a analizar marcadas con los MID, cada muestra en una columna. En la parte de la izquierda, se encuentran las variantes identificadas en las muestras y en los genes *BRCA1/2*, cada variante está en una fila.

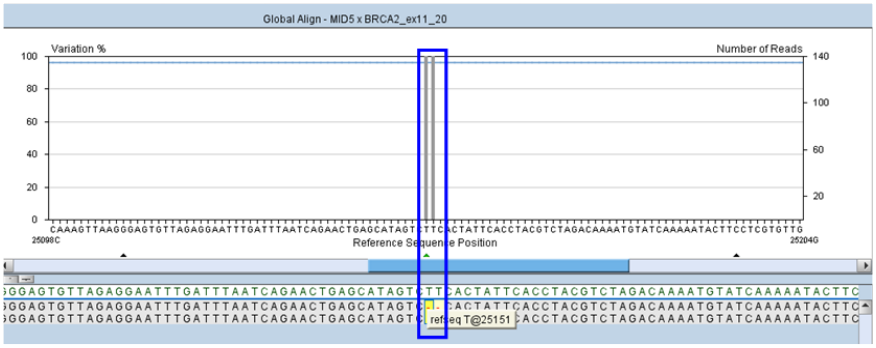


Figura 12. Ejemplo de un alineamiento de lectura con el software AVA de una deleción.

Software MiSeqReporter (Illumina)

El análisis de datos obtenidos por la plataforma *MiSeq* se realizó mediante el *software MiseqReporter (Illumina).* Se trata de un *software* de análisis de datos, que procesa las llamadas de variantes generadas durante la secuenciación de la carrera, en tiempo real en el secuenciador, y produce información sobre el alineamiento, variantes estructurales y el ensamblaje de secuencias de los genes *BRCA1/2*.

Pipeline desarrollado por SSGG.

SSGG desarrolló y puso a punto un *pipeline* para el filtrado y selección de variantes propio, a partir de los datos del fichero que generan las plataformas *454 GS Roche y MiSeq*. Este desarrollo fue llevado a cabo por el grupo de bioinformáticos de SSGG. Este *pipeline* resultante se divide en:

- **La identificación de variantes y pequeñas indels**, fue llevada a cabo mediante la utilización de diferentes *"callers"* (programa para la llamada de variantes), buscando un equilibrio entre la especificidad y la sensibilidad. Se emplearon las herramientas *Samtools v.0.1.9* y *GATK v.1.0.5777*.
- **La anotación de variantes** fue realizada empleando una versión evolucionada del *script* "*Variant Effect Predictor*" de *Ensembl* frente a la versión 62 de esta base de datos[97]. Además de la información obtenida de *Ensembl*, se incorporó en el análisis de estas muestras, la base de datos de *Human Gene Mutation Database Professional*, base de datos de variantes de importancia biológica y su relación con la enfermedad[98].

Los campos contenidos en los ficheros, así como la clasificación de las variantes en diferentes categorías dependiendo de la frecuencia se detallan en el anexo IV.

9.5 Evaluación de la cobertura del análisis.

Empleamos el *kit Multiplicom*, el cual, utiliza un sistema de amplificación mediante PCR en multiplex, por lo tanto, se hace imprescindible verificar que todos los amplicones se encuentran representados, y que, además, la profundidad de lectura obtenida en cada uno de ellos, alcanza unos niveles mínimos para la correcta identificación de variantes.

Para visualizar la cobertura de cada amplicón, se generan dos gráficas, en ellas se representan los valores medios de profundidad de lectura por amplicón, para las muestras dentro de la misma carrera, ordenando estos valores, bien por la posición genómica de cada una de las PCR individuales en el gen, o bien, por grupos de PCRs incluidos dentro de la misma reacción en multiplex (Figura 13).

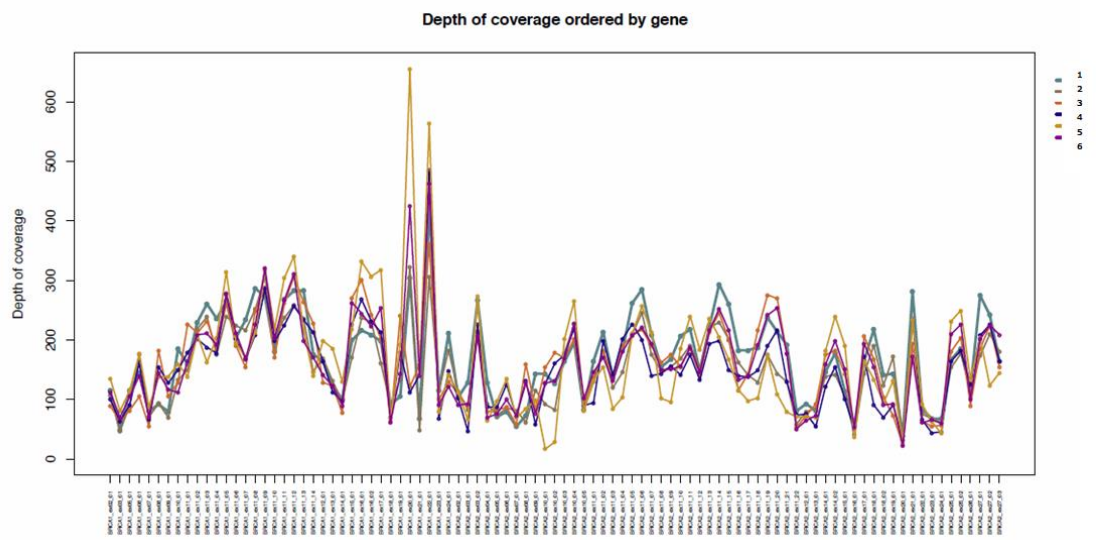


Figura 13. Lectura media calculada para cada PCR y ordenadas según su posición en el gen para una carrera.

9.6 Identificación de inserciones/deleciones y Copy Number Variants (CNVs).

Posteriormente, con la experiencia del análisis de muestras con *454 GS Junior* y los análisis de fragmentos, se implementó una estrategia para la detección de CNVs a partir de los resultados obtenidos para el estudio de la cobertura. A partir de los valores medios de profundidad de lectura por amplicón para cada una de las muestras pertenecientes a la misma carrera, con el fin de confirmar una presencia mínima de lectura a lo largo de todo el panel y una distribución de lecturas "homogénea" entre muestras.

En cada carrera, se normalizaron los valores de la profundidad media de lectura por amplicón, dividiendo cada uno de ellos por el número de lecturas totales y posteriormente, por el valor medio para todas las muestras en ese amplicón[99]. Los datos de cobertura de dos muestras de ensayo se compararon después de la normalización con el promedio de tres muestras de referencia con número de copias normal. El cálculo de normalización y el procedimiento de comparación de los datos de recuento se realizaron por separado para cada uno de las cuatro PCR BRCA1/2 multiplex. En resumen, para determinar el cociente de dosificación (DQ) de un amplicón en un individuo dado, se determinó el recuento de lectura absoluto total por multiplex específico como la suma de los recuentos de lectura para todos los amplicones de ese múltiplex específico[99].

Las potenciales deleciones detectadas tras la observación de los gráficos generados, son contrastadas con los datos obtenidos en el laboratorio mediante la visualización de los perfiles de fragmentos de las PCRs y mediante la técnica de MLPA (Figura 14).

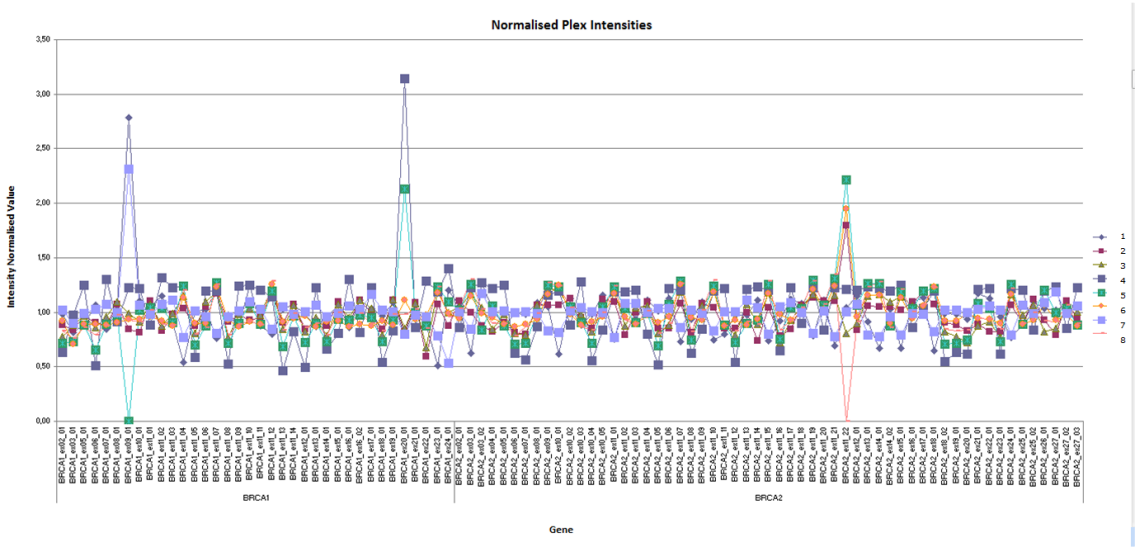


Figura 14. Visualización de los resultados distribuidos por valores medios de profundidad de lectura por amplicón ordenado por exones y genes en la misma carrera que la anterior.

10. Clasificación de variantes

Una vez realizado el análisis genético, se debe llevar a cabo la clasificación de las variantes halladas en él. Dicha clasificación se puede realizar desde el punto de vista molecular, en función de cómo afecta el cambio a la estructura del gen o de la proteína que traduce éste; o bien desde el punto de vista clínico, según la implicación que dicha variante tenga en el desarrollo de una determinada enfermedad. La nomenclatura de las distintas variantes localizadas sigue la normativa propuesta por la *Human Genome Variation Society* (HGVS)[100].

En este trabajo, la clasificación de variantes se ha realizado integrando la información disponible para cada una de ellas, teniendo en cuenta:

- Tipo de cambio
- Clasificación en la literatura y en las bases de datos consultadas (dbSNP, HGMD, LOVD, BRCAExchange, ClinVar, UMD y BIC (éstas dos últimas solo empleadas en la primera clasificación)
- *Minor Allele Frequency* (MAF)
- Valor de los predictores *in silico* indicados en la tabla de resultados 18, 23 y 24 para las variantes *missense* (*SIFT, PolyPhen, Align GVGD*) y de *splicing* (MaxEnt, NNSPLICE, HSF)

Adicionalmente, se han considerado otros indicadores de patogenicidad proporcionados por el programa **Alamut®** para variantes de tipo *missense* (*phyloP score, amino acid conservation, Grantham distance, protein domain*).

El efecto ha sido atribuido atendiendo a los siguientes criterios en la clasificación inicial de las variantes y el algoritmo diagnóstico de decisión de la figura 15.

Patogénica*:

- Variantes descritas: variantes de tipo *frameshift, missense, nonsense*, y *splicing region*, descritas como tal en HGMD y ampliamente referidas en la literatura.

Probablemente patogénica:

- Variantes descritas: variantes de tipo *frameshift, missense, nonsense*, y *splicing region*, previamente descritas como patogénicas en HGMD y en la que la mayoría de la bibliografía consultada, pero no toda, refiere su probable carácter patogénico. La clasificación de la variante como deletérea por parte de la mayoría de los predictores sirven para apoyar su clasificación en esta categoría.
- Variantes nuevas: variantes de tipo *frameshift, missense, nonsense*, y *splicing region*, en la que la tendencia de los predictores evaluados apunta a la probable patogenicidad del cambio.

***Significado Incierto*:**

- Variantes descritas: variante de tipo *frameshift, missense, nonsense*, y *splicing region* en las que existe una elevada discrepancia en la bibliografía consultada o tanto en la bibliografía consultada, como en los predictores.
- Variantes nuevas: variantes de tipo *frameshift, missense, nonsense*, y *splicing region* en las que existe una elevada discrepancia entre los valores de los predictores evaluados.

Probablemente benigna

- Variantes descritas*:* variantes de tipo *missense, nonsense*, y *splicing region*, en la que la mayoría de la bibliografía consultada, pero no toda, refiere su probable carácter benigno. La clasificación de la variante como no deletéreo por parte de la mayoría de los predictores, así como un valor de MAF cercano al 1% sirven para apoyar su clasificación en esta categoría.
- Variantes nuevas: variantes de tipo *missense* y *splicing region*, en la que la tendencia de los predictores evaluados apunta a un probable efecto neutral del cambio, y además co-ocurre con una VP.

***Benigna**:**

- Variantes descritas: variantes de tipo *missense, nonsense*, y *splicing region*, previamente descritas como tal en la literatura. La clasificación de la variante como no deletérea por parte de la mayoría de los predictores, así como un valor de MAF cercano al 1% sirven para apoyar su clasificación en esta categoría.

*****En ningún caso se ha atribuido un efecto patogénico o benigno a una variante nueva.

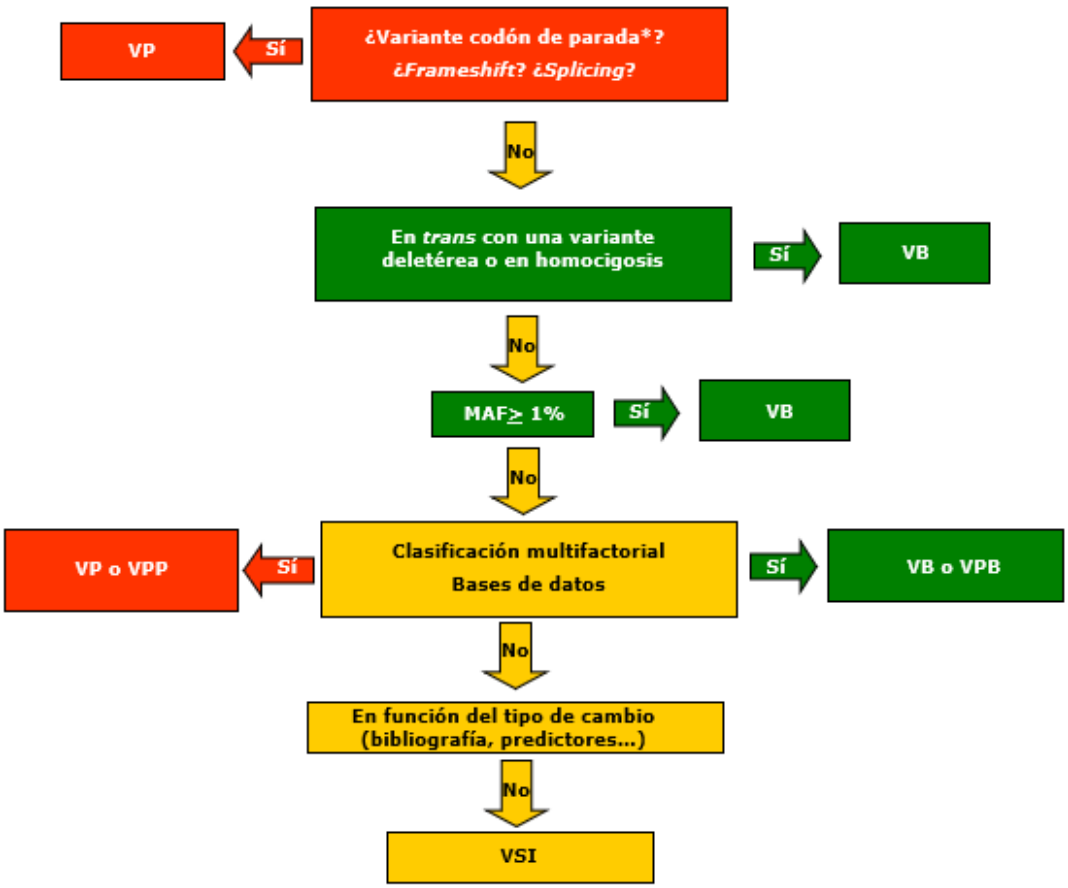


***Figura 15. Algoritmo empleado para la clasificación clínica de las variantes genéticas* (Adaptado del curso de la SEOM de Miguel de la Hoya). *Excepto en el gen *BRCA2*, p.Lys3326*.**

Se realizó la validación mediante secuenciación Sanger de los cambios detectados con relevancia clínica y las VSI.

Para la **reclasificación** de las variantes se utilizaron los criterios de la guía clínica del Colegio Americano de Genética Médica y Genómica (ACMG), de la guía de ACSG (Association for Clinical Genomic Science) y las recomendaciones de ENIGMA[28,52,53]. Se volvió a realizar una búsqueda de las variantes en las bases de datos.

Los criterios de clasificación de variantes según el grado de evidencia de patogenicidad y de benignidad, así como el algoritmo que emplea la combinación de ambos para clasificar las variantes se detallan en el anexo V.

RESULTADOS

12. Objetivo 1.1: Valoración de la nueva estrategia de análisis: Cribado de variantes patogénicas en los genes *BRCA1* y *BRCA2* empleando la plataforma de secuenciación masiva *454 GS Junior*

Con el objetivo de evaluar una nueva estrategia de análisis de los genes *BRCA1/2*, basada en la secuenciación masiva, se estableció un protocolo para evaluar esta técnica, comparando los resultados frente a los obtenidos mediante secuenciación Sanger, el método *gold standard* establecido hasta la fecha. Para ello se seleccionó un set de muestras con variantes previamente conocidas y estudiadas mediante secuenciación Sanger, y se secuenciaron mediante NGS con el fin de determinar la concordancia genotípica, cobertura de las mismas y estudiar eficacia de distintos *softwares* de análisis de datos NGS; *AVA* y un *pipeline* propio.

12.1 Selección de muestras control

La selección de muestras incluidas para validar el panel con *454 GS Junior* incluía:

- Seis muestras control de pacientes con CMOH, secuenciadas mediante secuenciación Sanger. Este set de muestras incluía 37 variantes previamente descritas, entre ellas hay 6 VP (5 *indels* y 1 *missense*) y estaban representadas tanto variantes puntuales como indels de distinto tamaño.
- Una línea celular *HapMap* NA12144, genotipada mediante *arrays* de genotipado masivo. Esta muestra presentaba 9 variantes conocidas, 8 variantes eran SNV y una deleción de 4pb.

La descripción del conjunto de variantes por muestra se encuentra en la tabla 8 (en negrita se resaltan las VP).

Tabla 8. Variantes en las muestras control detectadas previamente mediante Sanger y línea *HapMap*

Muestras	Nomenclatura HGVS	ID	Tipo
1	***BRCA1*: c.815_824dup (p.Thr276Alafs*14)**	*rs387906563*	Duplicación de 10pb
2	*BRCA1*: c.2311T>C (p.Leu771=)	*rs16940*	Cambio puntual
	BRCA1: c.3548A>G (p.Lys1183Arg)	*rs16942*	Cambio puntual
	BRCA1: c.4308T>C (p.Ser1436=)	*rs1060915*	Cambio puntual
	***BRCA2*: c.9026_9030del (p.Tyr3009Serfs*7)**	*rs80359741*	Deleción de 5 pb
	BRCA2: c.1114A>C (p.Asn372His)	*rs144848*	Cambio puntual
3	***BRCA1*: c.68_69del (p.Glu23Valfs*17)**	*rs386833395*	Deleción de 2pb

4	***BRCA1*: c.211A>G (p.Arg71Gly)**	*rs80357382*	Cambio puntual
	BRCA1: c.3548A>G (p.Lys1183Arg)	*rs16942*	Cambio puntual
	BRCA1: c.3113A>G (p.Glu1038Gly)	*rs16941*	Cambio puntual
	BRCA1: c.2311T>C (p.Leu771=)	*rs16940*	Cambio puntual
	BRCA1:c.4308T>C (p.Ser1436=)	*rs1060915*	Cambio puntual
	BRCA1: c.2077G>A (p.Asp693Asn)	*rs4986850*	Cambio puntual
	BRCA1: c.2082C>T (p.Ser694=)	*rs1799949*	Cambio puntual
	BRCA2: c.3396A>G (p.Lys1132=)	*rs1801406*	Cambio puntual
	BRCA2: c.-26G>A	*rs1799943*	Cambio puntual
	BRCA2: c.7796A>G (p.Glu2599Gly)	*NR*	Cambio puntual
	BRCA2: c.7806-14T>C	*rs9534262*	Cambio puntual
5	*BRCA2*: c.681+56C>T	*rs2126042*	Cambio puntual
	***BRCA2*: c.715dup (p.Ser239Lysfs*6)**	NR	Duplicación de 1pb
6	*BRCA1*: c.2311T>C (p.Leu771=)	*rs16940*	Cambio puntual
	BRCA1: c.3548A>G (p.Lys1183Arg)	*rs16942*	Cambio puntual
	BRCA1: c.2082C>T (p.Ser694=)	*rs1799949*	Cambio puntual
	BRCA1: c.4308T>C (p.Ser1436=)	*rs1060915*	Cambio puntual
	BRCA1: c.4676-14C>T	NR	Cambio puntual
	BRCA2: c.7008-63C>G	NR	Cambio puntual
	***BRCA2*: c.658_659del (p.Val220Ilefs*4)**	*rs80359604*	Deleción de 2pb
	BRCA2: c.1151C>T (p.Ser384Phe)	*rs41293475*	Cambio puntual
***HapMap* NA12144**	*BRCA2*: c.-26G>A	*rs1799943*	Cambio puntual
	BRCA2: c.631+183T>A	*rs3752451*	Cambio puntual
	BRCA2: c.6841+78_6841+81del	*rs138193280*	Deleción de 4pb
	BRCA2: c.7242A>G (p.Ser2414=)	*rs1799955*	Cambio puntual
	BRCA2: c.125A>G (p.Tyr42Cys)	*rs4987046*	Cambio puntual
	BRCA2: c.7806-14T>C	*rs9534262*	Cambio puntual
	BRCA2: c.8755-66T>C	*rs4942486*	Cambio puntual
	BRCA2: c.*105A>C	*rs15869*	Cambio puntual
	BRCA1: c.442-34C>T	*rs799923*	Cambio puntual

NR: No referenciado

12.2 Concordancia genotípica

Las regiones diana de las 6 muestras control y la *HapMap* fueron capturadas empleando la tecnología del kit *BRCA MASTR v2.0,* basada en PCRs multiplex. En la carrera de validación se obtuvieron más de 50 lecturas por fragmento. Los valores de los filtros aplicados en relación a la profundidad de lectura fueron de un mínimo de 20% hasta un máximo del 100%, estos filtros fueros aplicados a las lecturas *forwards* y *reverse*.

Con estos parámetros se detectaron 36 de las 37 variantes conocidas, 30 SNV y 6 *indels*. Sin embargo, se detectó la presencia de un falso negativo (rs4987046) y 70 falsos positivos, entre ellos, 69 *indels* y el cambio puntual *rs1801406* en homocigosis en la línea *HapMap*.

En la figura 16, se muestran el número de variantes detectadas en las 6 muestras control y la línea *HapMap* testadas para la validación de la técnica, utilizando el *software* de análisis *AVA* propio del secuenciador *454 GS Junior*.

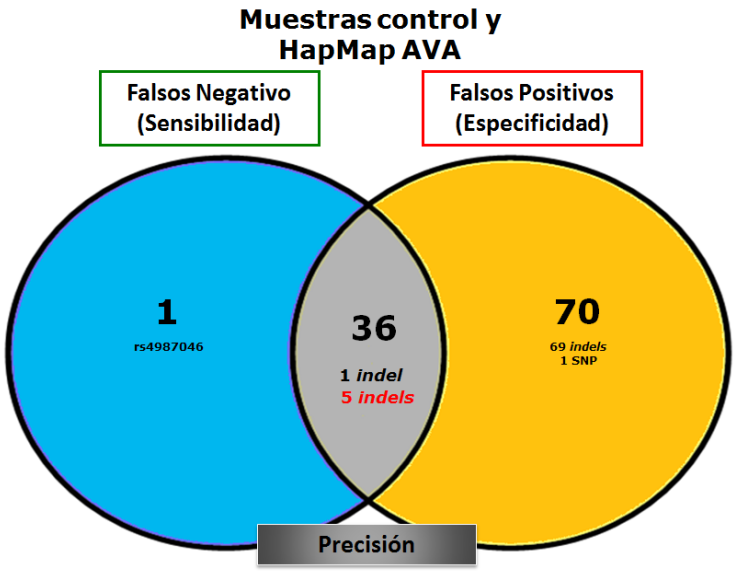


Figura 16. Diagrama de Venn del análisis de las muestras control mediante el software AVA. En rojo, se encuentran marcados los *indels* patogénicos de las muestras control.

12.2.1 Evaluación de los falsos negativos

Se detectó la presencia de un único SNV *rs4987046*, no detectado mediante secuenciación masiva (falso negativo) en la línea NA12144 *HapMap*. Para aclarar este hecho, se llevó a cabo la confirmación de la presencia de esta variante en la línea celular, mediante secuenciación Sanger. Esta secuenciación reveló que el SNV no estaba presente en esta línea celular, por lo tanto, con estos datos, podemos afirmar que no obtuvimos ningún falso negativo tipo SNV mediante la utilización de esta técnica, y que este *rs4987046,* se trata de un falso positivo de la línea *HapMap* NA12144 (Figura 17).

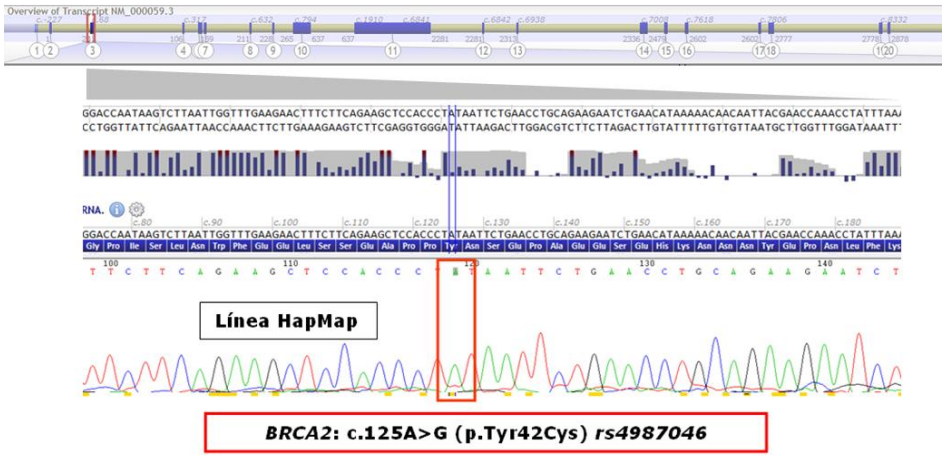


Figura 17. Confirmación mediante secuenciación Sanger del cambio puntual rs4987046, descrito en la línea HapMap* NA12144 *no estaba presente, es un falso positivo de esta línea celular. Se muestra el cromatograma de la posición c.125A>G (p.Tyr42Cys) en el gen *BRCA2 rs4987046* en la línea *HapMap* empleada para el ensayo, observamos una adenina en homocigosis, en esa posición.

12.2.2 Evaluación de falsos positivos

En el estudio de validación se identificaron 70 falsos positivos, entre ellos 69 *indels* y el cambio puntual *rs1801406* en homocigosis en la línea *HapMap*.

a) Cambio puntual

Para aclarar la presencia del SNP c.3396A>G (p.Lys1132Lys), *rs1801406,* en homocigosis, detectado mediante secuenciación masiva y no detectado mediante *arrays* de genotipado en la línea celular NA12144, se llevó a cabo la secuenciación mediante Sanger de este SNP en la línea *HapMap* NA12144. Esta secuenciación reveló, que el SNP estaba presente en esta línea celular en homocigosis, tal y como habíamos detectado mediante la secuenciación masiva empleando *454 GS Junior*, por lo tanto, con estos datos, podemos afirmar que no obtuvimos ningún falso positivo tipo SNV (Figura 18).

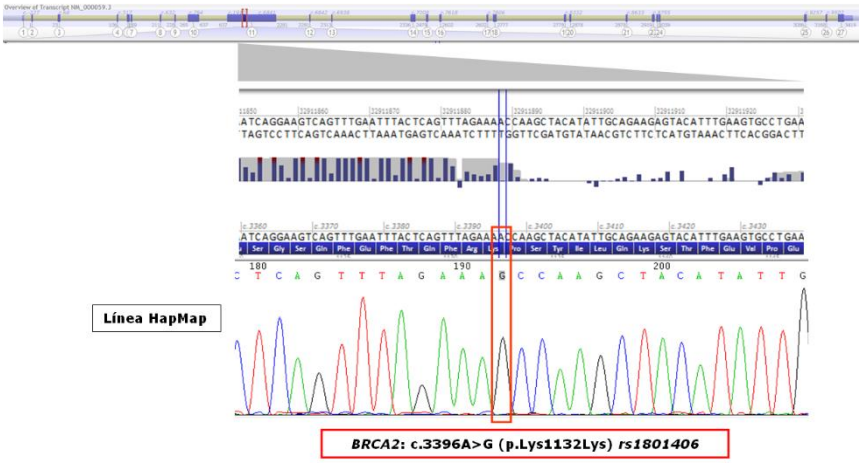


Figura 18. Confirmación mediante secuenciación Sanger de que el cambio puntual rs1801406, no descrito en la línea HapMap NA12144 es un falso negativo de esta línea celular.

b) Pequeñas deleciones/inserciones

Mediante la pirosecuenciación y el analizador de datos *AVA,* se detectaron un total de 69 *indels* falsos positivos, en el set de muestras control utilizadas para la validación. Este es un problema inherente a la técnica de pirosecuenciación, la detección de variantes en las zonas de homopolímero produce falsos positivos del tipo *indels*.

Para intentar solucionar el problema inherente a la técnica de pirosecuenciación, como es la detección de variantes en zona de homopolímero, y así reducir la gran cantidad de *indels* falsos positivos, previamente identificados en las muestras de validación, se llevó a cabo, un nuevo protocolo de trabajo basado en el **análisis de los fragmentos** generados mediante PCR-Multiplex. Este procedimiento se describe en el siguiente apartado.

12.3 Evaluación de deleciones e inserciones (indels) mediante análisis de fragmentos.

La evaluación de *indels* mediante análisis de fragmentos se realizó a partir de las 5 PCR-Multiplex utilizadas para la amplificación de los genes *BRCA1/2*. Un examen visual de los datos electroforéticos de estas PCR-Multiplex de las muestras control (genotipo conocido) sugiere que las muestras con la misma secuencia tienen patrones electroforéticos similares, mientras aquellos fragmentos con pequeñas deleciones/inserciones de nucleótidos dan un patrón claramente alterado. En este patrón, generalmente, se pueden distinguir dos picos, uno de ellos correspondiente al alelo normal y el otro, al alelo con la deleción o inserción, separados por el número de bases de las que conste la alteración. Adicionalmente, en la intensidad del pico primario (un alelo normal) se veía reducido a la mitad respecto del patrón normal (dos alelos normales).

Se realizó la evaluación de los patrones electroforéticos, distinguiendo entre patrones normales y alterados. Inicialmente se incluyó un individuo sin VP, como control de referencia de análisis.

Así determinamos los patrones que corresponden al fragmento normal (sin ninguna variante de tipo *indel*) y evaluamos fragmentos alterados. En estos fragmentos que presentaban patrones anómalos, es en los que pudimos detectar las 5 inserciones/deleciones reales que habían sido previamente detectadas en las muestras de validación (ver tabla 8). En cuanto al tamaño de estas *indels*, este análisis de fragmentos detectó inserciones que implicaban a una sola base (*BRCA2* c.715dup) hasta inserciones de 10bp (*BRCA1*: c.815_824dup), (Figura 19).

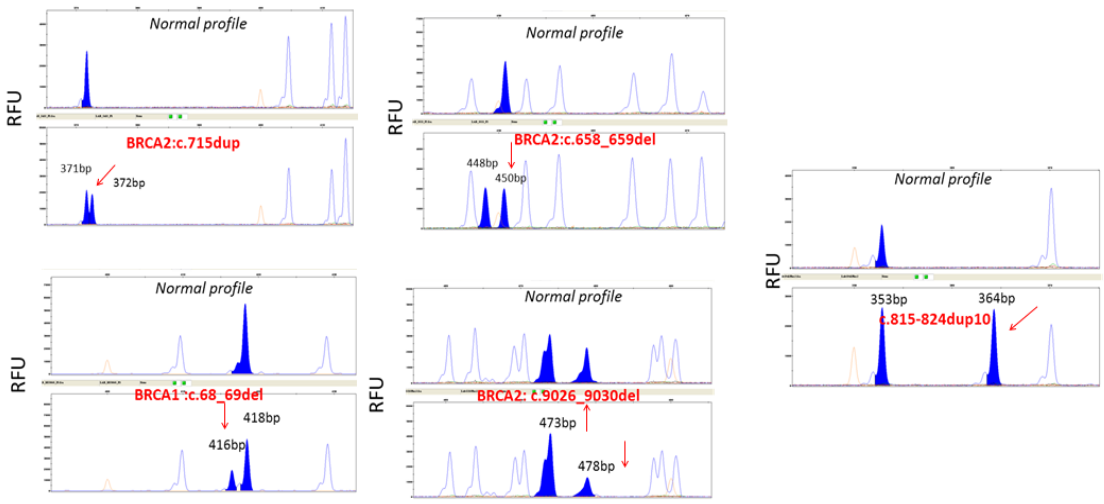


Figura 19. Análisis de fragmentos de las cinco muestras control, que presentaban indels como VP. Se observaron patrones distintos en estas muestras, frente a los patrones observados en la muestra usada como control (no presenta *indels*). En las gráficas de arriba observamos los perfiles observados en la muestra control (sin *indels*) y las muestras de abajo presentan las *indels*. RFU: unidades de fluorescencia relativa.

Por tanto, mediante esta metodología pudimos descartar los 69 *indels* falsos positivos, anotados por el *software* de análisis e identificar aquellos fragmentos donde se localizaba una *indel* real.

Sorprendentemene, mediante el análisis de fragmentos pudimos identificar la detección de *indels* reales localizadas en zonas de homopolímero, lo ilustra la detección de la variante c.956dupA (p.Asn319Lysfs*8) en el gen *BRCA2*, en una muestra de validación adicional, ilustrada en la figura 20. Esta muestra presenta una duplicación de una adenina en un tracto poliA (constituido por cinco adeninas consecutivas). Esta variante no fue detectada mediante la anotación del *pipeline* propio, ni mediante el *software* de análisis *AVA* Roche. Solo fuimos capaces de detectarla gracias al empleo de los análisis de fragmentos (Figura 20).

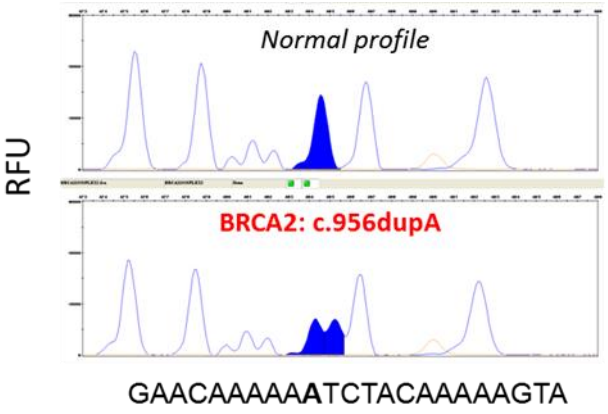


Figura 20. Análisis de fragmentos de una indel en zona de homopolímeros. Se muestra la detección mediante los perfiles de intensidad, donde se observan dos picos (dos alelos que se diferencian en una base). En la imagen superior, observamos una muestra control con un perfil esperado para el este plex.

En cuanto a los **inconvenientes** derivados del empleo de los perfiles de intensidad, debemos decir, que el análisis es laborioso, se aumenta el tiempo que hay que dedicar al análisis de la muestra, y hay que tener cuidado con las *indels* polimórficas de las zonas intrónicas, ya que, pueden inducir confusión. Un ejemplo que representa este último hecho es la deleción polimórfica en el intrón 11 del gen *BRCA2,* el *rs138193280*, de modo que en las muestras que lo portan, se identifica un patrón electroforético anómalo (Figura 21).

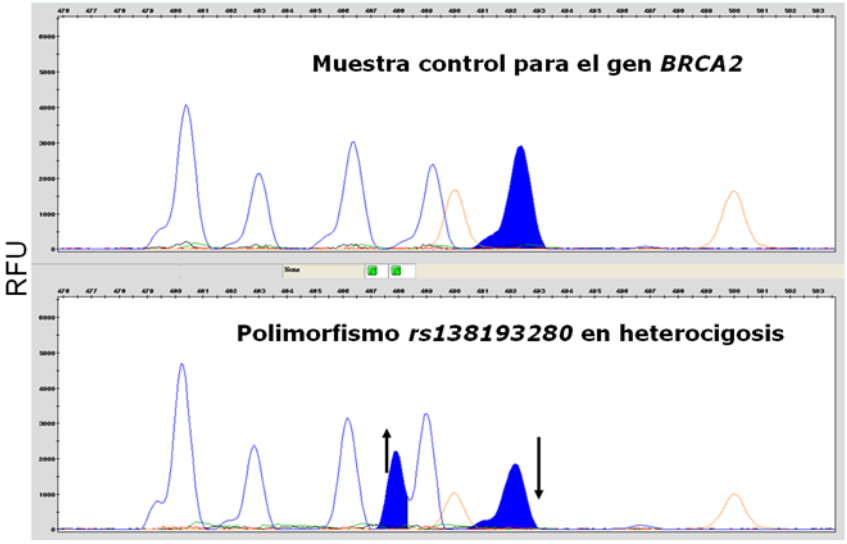


Figura 21. Análisis de fragmentos con el programa Peak Scanner *de una muestra control (zona de arriba) y una muestra en análisis con el polimorfismo* rs138193280 *en heterocigosis (abajo).*

El análisis de fragmentos estaba orientado inicialmente, a confirmar la amplificación uniforme de todos los fragmentos de la PCR multiplex, pero como hemos visto adicionalmente nos permitió:

- Detectar las pequeñas deleciones/inserciones en una muestra antes de la secuenciación.
- Utilizar este método como herramienta para descartar los falsos positivos (*indels*) que genera la pirosecuenciación.

12.4 Precisión

Todas las variantes previamente conocidas en los genes *BRCA1/2* incluidas en las muestras de validación se detectaron con éxito. Por tanto, mediante al uso de la plataforma *454 GS Junior* se detectaron el 100% de las VP asociadas a cáncer.

Además, como hemos comentado anteriormente, se detectó un falso negativo y hasta 70 falsos positivos, la mayoría de los cuales eran variantes tipo *indels,* debido a los errores de esta tecnología en la zona de homopolímeros. El falso positivo y el falso negativo de tipo SNV detectado durante esta fase de validación resultaron ser errores de genotipado de la línea

celular empleada, dado que, el falso positivo no fue confirmado mediante secuenciación Sanger, en cambio el falso negativo sí que fue confirmado mediante esta tecnología.

Por otro lado, mediante la implementación del análisis de fragmentos para la detección de inserciones y deleciones en las muestras de validación, logramos establecer un proceso para discriminar los *indels* verdaderos, de los artefactos introducidos por el proceso de pirosecuenciación. Por ello, tras la aplicación de la combinación del análisis de fragmentos y la aclaración de los falsos positivos y negativos detectados en la línea celular *HapMap*, obtuvimos una precisión del 100%.

12.5 Análisis del software AVA (Roche) versus pipeline propio.

Se llevó a cabo un proceso de comparación de los datos de la línea *HapMap* analizada, mediante el empleo del *software* comercial *AVA* de *Roche* y el empleo del *pipeline* propio desarrollado en SSGG.

El *software* comercial *AVA*, detectó 8 variantes previamente descritas, un cambio puntual falso negativo, (el rs4987046, descrito anteriormente como falso positivo de la línea celular) y la presencia de 15 falsos positivos, entre ellos, 14 *indels* y un cambio puntual.

El *pipeline* de desarrollo propio detectó 8 variantes previamente descritas, un cambio puntual falso negativo, (el *rs4987046,* descrito anteriormente como falso positivo de la línea celular), y la presencia de 70 falsos positivos, entre ellos 63 *indels* y siete cambios puntuales (SNVs) adicionales. Estos cambios puntuales son detectados únicamente por el *pipeline* propio, esto es debido a que el *software AVA* elimina la zona de adaptadores del kit de *Multiplicom*, y con el *pipeline* propio vemos más zona intrónica. Por un lado, 5 de los 7 SNVs identificadas, estaban localizadas en regiones intrónicas profundas y zonas UTR: c.*201A>G, *rs77863879, rs206080, rs76584943* y c.8632+174G>T. Por este motivo, nuestro *pipeline* identifica un mayor número de SNVs. Por otro lado, los otros dos cambios puntuales detectados que no están localizados en estas zonas son: *rs80357050* (vemos el alelo de referencia) y el cambio detectado empleando *AVA rs1801406*, (confirmado posteriormente mediante secuenciación Sanger). Los parámetros empleados para el análisis con el *pipeline* propio, fueron un mínimo de $50x^{101}$ para cobertura y una frecuencia alélica de al menos un 20% para la detección de variantes en heterocigosis.

Lo más destacable que se observó, fue la gran diferencia de falsos positivos detectados con el *pipeline* propio frente al *software AVA* (Figura 22).

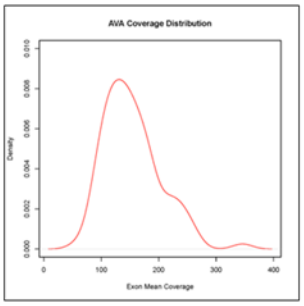
Figura 22. Diagrama de Venn del análisis de la línea HapMap *mediante el software AVA (figura de la izquierda) y el pipeline propio (figura de la derecha).*

12.6 Cobertura

La cobertura, es un parámetro que indica el número de veces que cada base del gen está presente en las lecturas de secuenciación producidas. Es un factor determinante para evaluar la fiabilidad del nucleótido asignado a esa posición en el gen. Para garantizar la detección de variantes se debe comprobar que todos los fragmentos alcanzan un mínimo de 50x, valor mínimo establecido para la identificación de variantes[101,102].

Para la **línea *HapMap* NA12144**, en general, observamos una buena cobertura para ambos genes, aproximadamente de 150x por exón. El mínimo de cobertura se observó en el exón 20 del gen *BRCA1*, y fue superior a 50x (Figura 23).

1) *Software* AVA: cobertura media de la línea NA12144.



2) *Pipeline* propio: cobertura media de la línea NA12144.

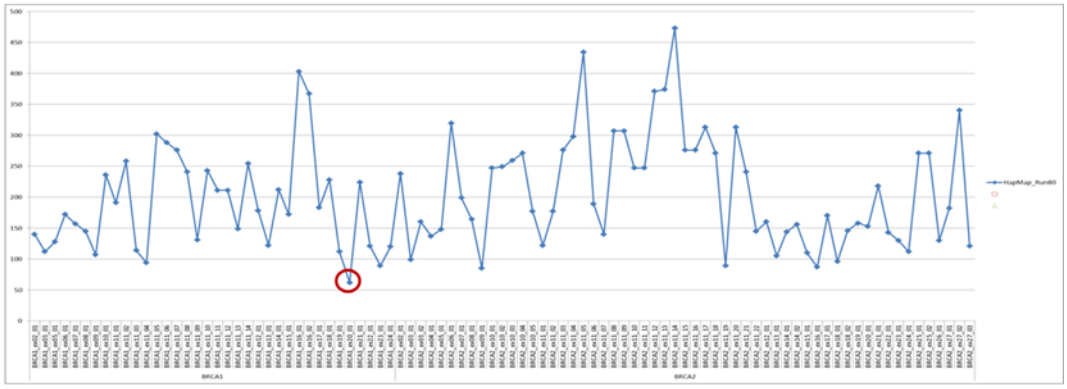


Figura 23. Gráfico de la cobertura obtenida en la línea *HapMap*. *En el gráfico de arriba, se muestra en color rojo, la cobertura global observada para la muestra con el software AVA, se muestra una cobertura homogénea. En el gráfico de abajo, la línea azul muestra la cobertura (número de lecturas) de cada uno de los amplicones de los genes BRCA1/2 obtenida mediante el pipeline propio, el mínimo de cobertura se observa en el exón 20 del gen BRCA1, señalizado con un círculo rojo.*

Los valores de profundidad de lectura observados en las **6 muestras de validación** fueron **homogéneos.** La profundidad media de lectura por amplicón en las muestras control fue de 163x (Figura 24). Se obtuvieron más de 50 lecturas por fragmento en las muestras control.

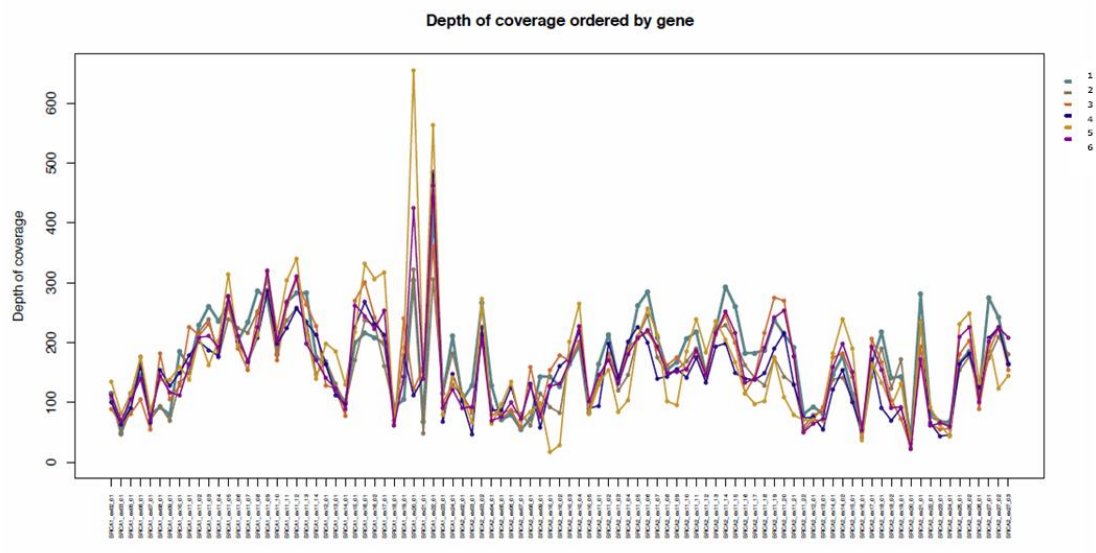


Figura 24. Gráfico de la cobertura obtenida en las muestras de validación. *Profundidad de lectura media calculada para cada PCR y ordenadas según su posición en el gen para cada muestra de validación.*

12.7 Reproducibilidad

Para comprobar la reproducibilidad del método, se realizó el análisis de la misma muestra en dos carreras distintas. Se compararon las coberturas observadas y la concordancia genotípica, mediante el *pipeline* propio (Figura 25).

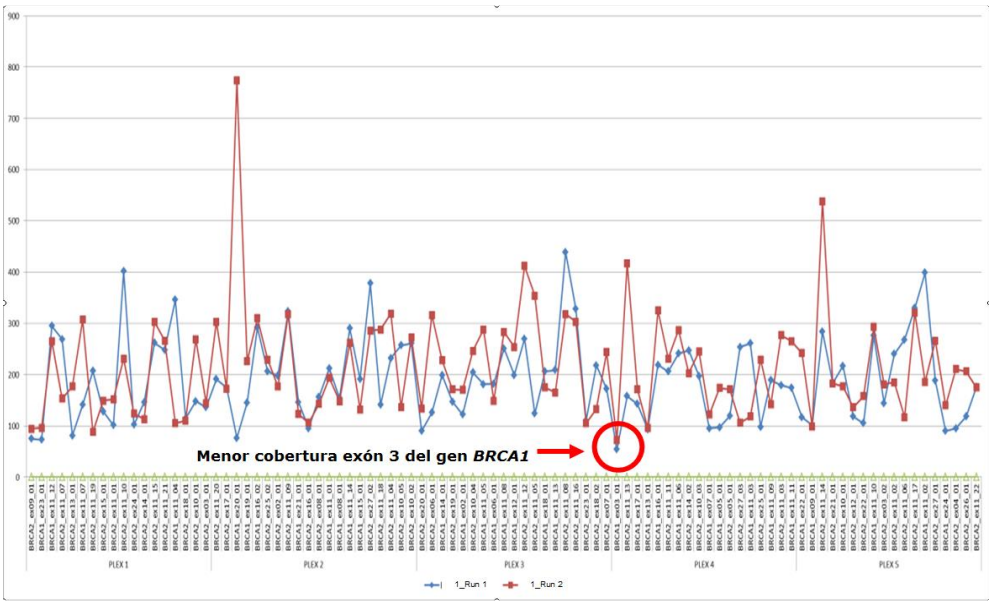


Figura 25. Perfiles de cobertura de la misma muestra secuenciada en dos carreras distintas. Los datos de la primera carrera se muestran en azul y los de la segunda carrera en rojo. Se resalta en rojo el exón amplificado con menos cobertura.

En cuanto a la cobertura, se observa como el exón 3 del gen *BRCA1*, es el exón menos cubierto, debido a este motivo, se decidió cubrir este exón por secuenciación Sanger, cuando la cobertura no fuera la adecuada.

En cuanto a la concordancia genotípica se observó una gran reproducibilidad del método, con una concordancia del 100% para variantes conocidas y se identificó la presencia de diferentes falsos positivos *indels* en las zonas de homopolímero, inherentes a la pirosecuenciación (Tabla 9).

Tabla 9. Concordancia genotípica de una misma muestra secuenciada en dos carreras distintas

Muestra 1					
	Carrera 1	**Carrera 2**		**Carrera 1**	**Carrera 2**
BRCA1	*rs376540*	*rs376540*	***BRCA2***	*rs21226042*	*rs21226042*
	indel1	*No detectado*		*rs543304*	*rs543304*
	indel2	*No detectado*		*rs206075*	*rs206075*
	indel3	indel3		*rs206076*	*rs206076*
	rs1799949	*rs1799949*		p.Glu2599Gly	p.Glu2599Gly
	rs16940	*rs16940*		*rs9534262*	*rs9534262*
	rs799917	*rs799917*		*rs4942486*	*rs4942486*
	rs16941	*rs16941*		*rs15869*	*rs15869*
	rs16942	*rs16942*		indel1	indel1
	rs799916	*rs799916*		indel2	indel2

	rs1090915	*rs1090915*		indel3	indel3
	rs8176212	*rs8176212*		indel4	*No detectado*
	rs1799966	*rs1799966*		indel5	*No detectado*
	rs8176233	*rs8176233*		indel6	*No detectado*
	rs8176234	*rs8176234*		indel7	indel7
	rs30922994	*rs30922994*		indel8	indel8
				indel9	indel9
				indel10	indel10
				indel11	indel11
				indel12	indel12

13. Objetivo 1.2: Valoración de la nueva estrategia de análisis: Cribado de variantes patogénicas en los genes *BRCA1* y *BRCA2* empleando la plataforma *MiSeq*.

En este objetivo propuesto, nos planteamos valorar una nueva estrategia de análisis de los genes *BRCA1/2* basada en la plataforma *MiSeq* de *Illumina*, que permitiese solventar los inconvenientes de la pirosecuenciación.

Para llevar a cabo la validación, se secuenciaron 12 muestras control utilizando la plataforma *MiSeq*. Estas muestras habían sido previamente secuenciadas por pirosecuenciación en la plataforma *454 GS Junior*. Las variantes de relevancia clínica, fueron confirmadas mediante secuenciación Sanger. Se determinó la concordancia genotípica de estas muestras mediante ambas metodologías (pirosecuenciación y secuenciación por terminador reversible), cobertura de las muestras y comparación de dos *softwares* de análisis de datos NGS: el *software* *MiSeqReporter* (*Illumina*) y el *pipeline* propio desarrollado en SSGG.

13.1 Selección de muestras control

Se incluyeron doce muestras control que presentan 211 variantes. Entre las VP, confirmadas mediante secuenciación Sanger se incluían 4 inserciones/duplicaciones, 2 deleciones, 1 *missense*, y 1 deleción-inserción compleja, además de 4 grandes CNVs confirmadas mediante MLPA (ver Apartado 13.5 de resultados). En la tabla 10, se muestran las variantes identificadas en las 12 muestras testadas para la validación de la técnica.

Tabla 10. Variantes testadas para la validación de la secuenciación usando la plataforma *MiSeq* en 12 muestras

Gen	ID dbSNP	Variantes HGVS	Nº de muestras
BRCA1	*rs3765640*	c.-19-115T>C	8

BRCA1	*rs80357642*	**c.65_66insTT (p.Leu22Phefs*2)**	1
BRCA1	*rs397509310*	**c.70_73dupTGTC (p.Pro25Leufs*17)**	1
BRCA1	*rs799923*	c.442-34C>T	6
BRCA1	*rs8176144*	c.548-58del	8
BRCA1	*rs1799965*	c.591C>T (p.Cys197=)	1
BRCA1	*rs1799950*	c.1067A>G (p.Gln356Arg)	2
BRCA1	*rs4986850*	c.2077G>A (p.Asp693Asn)	3
BRCA1	*rs1799949*	c.2082C>T (p.Ser694=)	8
BRCA1	*rs16940*	c.2311T>C (p.Leu771=)	8
BRCA1	*rs799917*	c.2612C>T (p.Pro871Leu)	8
BRCA1	*rs16941*	c.3113A>G (p.Glu1038Gly)	8
BRCA1	*rs4986852*	c.3119G>A (p.Ser1040Asn)	1
BRCA1	-	c.3356C>G (p.Thr1119Ser)	1
BRCA1	*rs2227945*	c.3418A>G (p.Ser1140Gly)	1
BRCA1	*rs16942*	c.3548A>G (p.Lys1183Arg)	8
BRCA1	*rs28897689*	c.4039A>G (p.Arg1347Gly)	2
BRCA1	*rs799916*	c.4097-141A>C	8
BRCA1	-	**c.4106delCAinsGCT (p.Ala1369Glyfs*25)**	1
BRCA1	*rs1060915*	c.4308T>C (p.Ser1436=)	8
BRCA1	*rs8176212*	c.4485-63C>G	8
BRCA1	*rs1799966*	c.4837A>G (p.Ser1613Gly)	8
BRCA1	*rs8176233*	c.4987-92A>G	8
BRCA1	*rs8176234*	c.4987-68A>G	8
BRCA1	*rs28897696*	**c.5123C>A (p.Ala1708Glu)**	1
BRCA1	*rs3092994*	c.5152+66G>A	8
BRCA2	*rs1799943*	c.-26G>A	6
BRCA2	*rs367871824*	c.631+25C>T	1
BRCA2	*rs3752451*	c.631+183T>A	6
BRCA2	*rs2126042*	c.681+56C>T	4
BRCA2	**-**	c.767C>T (p.Thr256Ile)	1
BRCA2	*rs756186207*	**c.956dupA (p.Asn319Lysfs*8)**	1
BRCA2	*rs144848*	c.1114A>C (p.Asn372His)	7
BRCA2	*rs28897711*	c.1938C>T (p.Ser646=)	1
BRCA2	*rs1801406*	c.3396A>G (p.Lys1132=)	6

BRCA2	*rs543304*	c.3807T>C (p.Val1269Val)	5
BRCA2	*rs28897724*	c.4068G>A (p.Leu1356=)	2
BRCA2	-	**c.4263dupT (p.Glu1422*)**	1
BRCA2	*rs80359788*	c.4584C>T (p.Ser1528=)	1
BRCA2	*rs80358755*	c.5312G>A (p.Gly1771Asp)	1
BRCA2	*rs276174862*	**c.5720_5723del (p.Ser1907*)**	1
BRCA2	*rs138193280*	c.6841+80_6841+83del	5
BRCA2	*rs81002811*	c.6842-20T>A	1
BRCA2	*rs1799955*	c.7242A>G (p.Ser2414=)	3
BRCA2	*rs9534262*	c.7806-14T>C	9
BRCA2	*rs4942486*	c.8755-66T>C	9
BRCA2	*rs80359756*	**c.9310_9311delAA (p.Lys3104Valfs*6)**	1
BRCA2	*rs15869*	c.*105A>C	6

En el Anexo VI, se encuentran descritas de manera individualizada las variantes detectadas en cada una de las 12 muestras utilizadas para la validación en los genes *BRCA1* y *BRCA2*.

13.2 Concordancia genotípica

Tanto con el *MiSeqReporter* (*Illumina*), como con el *pipeline* propio detectamos un total de 211 variantes previamente identificadas, entre ellas las 7 *indels* patogénicas, por lo tanto, obtenemos un **100% de concordancia**. No hubo falsos negativos, en cambio, sí detectamos falsos positivos (descritos en la siguiente sección).

Se obtuvo una media de 20 variantes/muestra, **menos variantes por muestra que con la pirosecuenciación** (20 *versus* 26), debido principalmente a que con el empleo de la plataforma *MiSeq* **se eliminan falsos positivos que se observan en las zonas de homopolímeros** introducidos por la pirosecuenciación (salvo algunas excepciones).

Un claro ejemplo de la mejora en la detección de *indels* en zonas de homopolímero, por la secuenciación de *MiSeq* frente a la pirosecuenciación, lo ilustra, la detección de la variante c.956dupA (p.Asn319Lysfs*8) en el gen *BRCA2*, en la muestra de validación número 10. Esta variante se encuentra en una zona de homopolímero, presenta una duplicación en un tracto poliA (presenta seis adeninas, en lugar de 5), y se ha identificado adecuadamente tanto por el *pipeline* propio, como *MiSeqReporter*, en cambio, mediante la pirosecuenciación solo fuimos capaces de su detección gracias al empleo de los análisis de fragmentos.

Aunque el empleo de esta nueva técnica de secuenciación redujo el número de variantes en zona de homopolímero, debido principalmente al cambio de tecnología, y ya no era indispensable el uso del análisis de fragmentos de las PCR multiplex, se siguió utilizando estos

perfiles, con las mismas ventajas (confirmación de la amplificación uniforme de todos los fragmentos de la PCR multiplex, detección de pequeñas *indels* previo a la secuenciación) e inconvenientes (análisis más laborioso). Además, gracias a ellos se intuían de forma gráfica, los posibles cambios en el número de copias presentes en las muestras, que posteriormente eran confirmados con la técnica del MLPA.

El *Var/Depth* (más bajo detectado para una variante tipo SNV en heterocigosis fue de **0,29.** Por este motivo, se decidió utilizar una frecuencia mínima de 0,20 para la detección de una variante en heterocigosis.

13.3 Análisis del software MiSeqReporter (Illumina) versus pipeline propio

Se obtuvieron los siguientes resultados con ambos *softwares*:

- Con ***MiSeqReporter* (*Illumina*)** se observan las 211 variantes previamente identificadas, entre ellas, las 7 *indels* patogénicas previamente identificadas, por lo tanto, obtenemos un 100% de concordancia. No identificamos ningún falso negativo, en cambio, sí que obtenemos falsos positivos, en concreto, 50 SNV. El porcentaje de falsos positivos obtenidos con este software es alto. No detectamos CNVs (Figura 26)

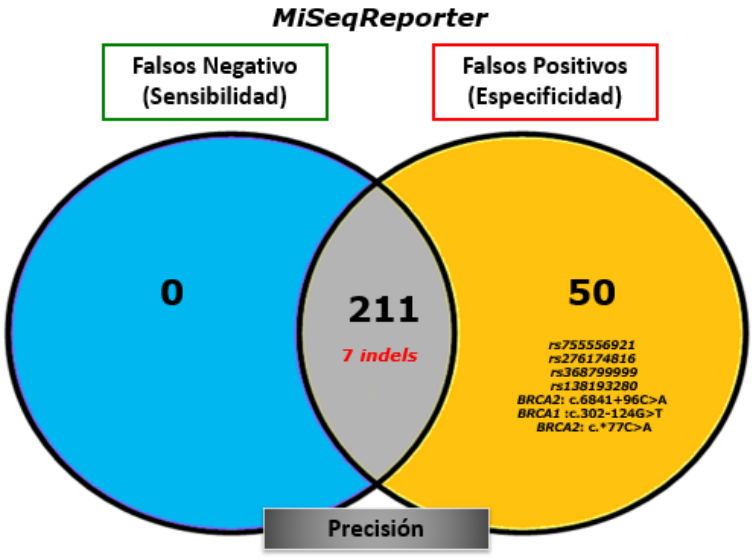


Figura 26. Diagrama de Venn del análisis de las muestras control mediante el software MiSeqReporter.

- Con el ***pipeline* propio**, también obtuvimos un 100% de concordancia. No identificamos ningún falso negativo, en cambio, sí que obtuvimos 26 falsos positivos, todos ellos variantes del tipo SNV. El porcentaje de falsos positivos obtenidos con este *software* es más bajo que el obtenido con *MiSeqReporter*. No detectamos CNVs (Figura 27).

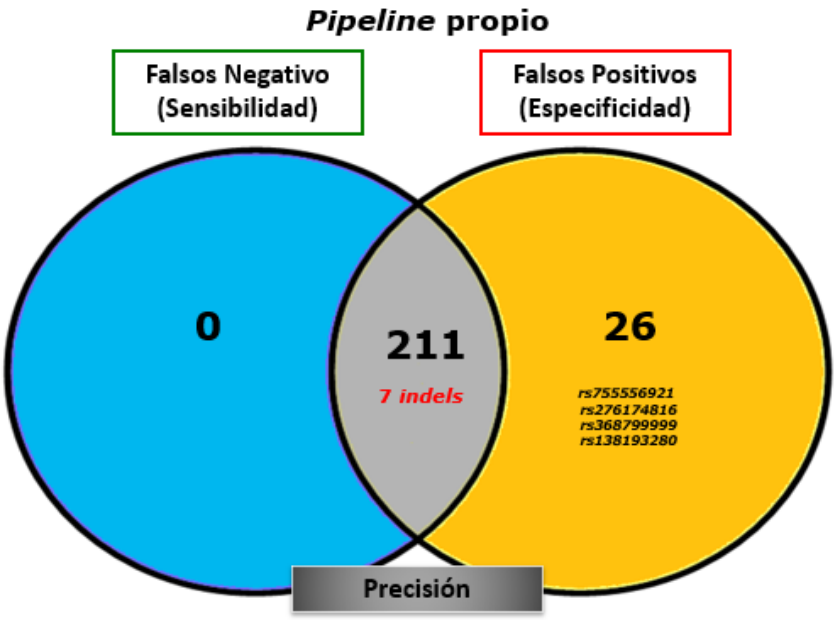


Figura 27. Diagrama de Venn del análisis de las muestras control con el pipeline propio.

Con el ***software* de *MiSeqReporter* (*Illumina*)** hemos detectado en todas las muestras incluidas en la validación, los siguientes c**ambios no detectados previamente mediante *454 GS Junior* ni mediante el *pipeline* propio** (Tabla 11):

Tabla 11. Variantes únicamente identificadas usando *MiSeqReporter* en las 12 muestras

Gen	ID dbSNP	Variantes HGVS	Nº de muestras
BRCA2	-	c.6841+96C>A	10
BRCA2	-	c.*77C>A	3
BRCA1	-	c.302-124G>T	11

- Las variantes en el gen *BRCA2*: c.6841+96C>A y *BRCA1*: c.302-124G>T. Asumimos que ambas variantes son falsos positivos en la zona final del oligonucleótido, debido a que se detectan en todas las muestras y no tiene un rs asociado con frecuencias alélicas descritas.
- La variante en el gen *BRCA2*: c.*77C>A, solo es detectada por *MiSeqReporter* (*Illumina*) en tres muestras control. Este cambio no había sido detectado previamente mediante la pirosecuenciación, tampoco es detectado con el *pipeline* propio, por lo tanto, asumimos que un falso positivo.

***MiSeqReporter* (*Illumina*) y el *pipeline* propio han detectado en todas las muestras,** con una frecuencia similar, los siguientes cambios no detectados previamente por *454 GS Junior* (Tabla 12):

Tabla 12. Variantes únicamente identificadas usando *MiSeqReporter* y *pipeline* propio en las 12 muestras			
Gen	**ID dbSNP**	**Variantes HGVS**	**Nº de muestras**
BRCA2	*rs755556921*	c.68-7del	12
BRCA2	*rs276174816*	c.1909+22del	12
BRCA2	*rs368799999*	c.516+54T>C	1
BRCA2	*rs138193280*	c.6841+80_6841+83del	1

- La variante en *BRCA2*: c.68-7del (zona de poliT). Se asumió que era un falso positivo en la zona de homopolímero, de la tecnología *MiSeq*, debido a que se detectó en todas las muestras con una frecuencia similar y aunque está descrito en dbSNP[54] (*rs755556921*) no se han descrito frecuencias en población control.
- La variante *en BRCA2*: c.1909+22del (zona de poliT). Se asumió que es un falso positivo en la zona de homopolímero, de la tecnología *MiSeq*, debido a que se detecta en todas las muestras con una frecuencia similar y aunque está descrito en dbSNP (*rs276174816*) no se han descrito frecuencias en población control.

Estos dos errores en la zona de homopolímero (politiminas), una en el intrón 2 aceptor (10 timinas) y la otra en el intrón 10 dador (11 timinas), pueden fácilmente confirmarse mediante un análisis visual del archivo bam en el *software* IGV o bien en el *software* Alamut.

- Las variantes *BRCA2*: c.516+54T>C (*rs368799999*) y c.6841+80_6841+83del (*rs138193280*), fueron detectadas en una única muestra. Estos dos cambios se encuentran en zonas intrónicas profundas, de modo que con la tecnología de *454 GS Junior* no los identificamos ya que las zonas de adaptadores de los extremos se eliminan mediante esta tecnología; en cambio, utilizando la tecnología de *MiSeq*, estos adaptadores no son eliminados y por tanto, somos capaces de detectar estas variantes que se encuentran localizadas en la zona de hibridación de los oligos.

En cuanto a las anotaciones de las variantes con ambos *softwares*, *MiSeqReporter* (*Illumina*) y el *pipeline* propio, observamos que la anotación de las inserciones fue mejor mediante el empleo del *pipeline* propio, ya que detectaba este tipo de variante como una variante *frameshift*, a diferencia de la anotación de *MiSeqReporter* (*Illumina*) que la mostraba como dos cambios puntuales. La anotación de cambios complejos (*indels*) es similar en ambos sistemas.

En la tabla 13, se resumen de las diferencias observadas entre los dos *softwares*:

Tabla 13. Diferencias entre los *softwares* utilizados durante la validación de *MiSeq*.		
	MiSeqReporter	***Pipeline* propio**

Detección de todas las VP	Sí	Sí
Falsos positivos	Más falsos positivos	Menos falsos positivos
Anotaciones	Anota las inserciones como dos cambios puntuales	Mejor anotación en cuanto a las inserciones
Visualización de todas las muestras de la carrera	No	Sí

Como **conclusión**, desde un punto de vista analítico, **nos decantamos por el empleo del *pipeline* propio para realizar el análisis de los *BRCA1/2*.**

13.4 Cobertura y reproducibilidad

Se observó que el porcentaje de lecturas útiles medio, se situó en torno al ~95% y la reproducibilidad media entre muestras atendiendo a los valores de cobertura fue del 0,92.

Se observa una alta cobertura para todas ella, con una media de aproximadamente 900x (Figura 28).

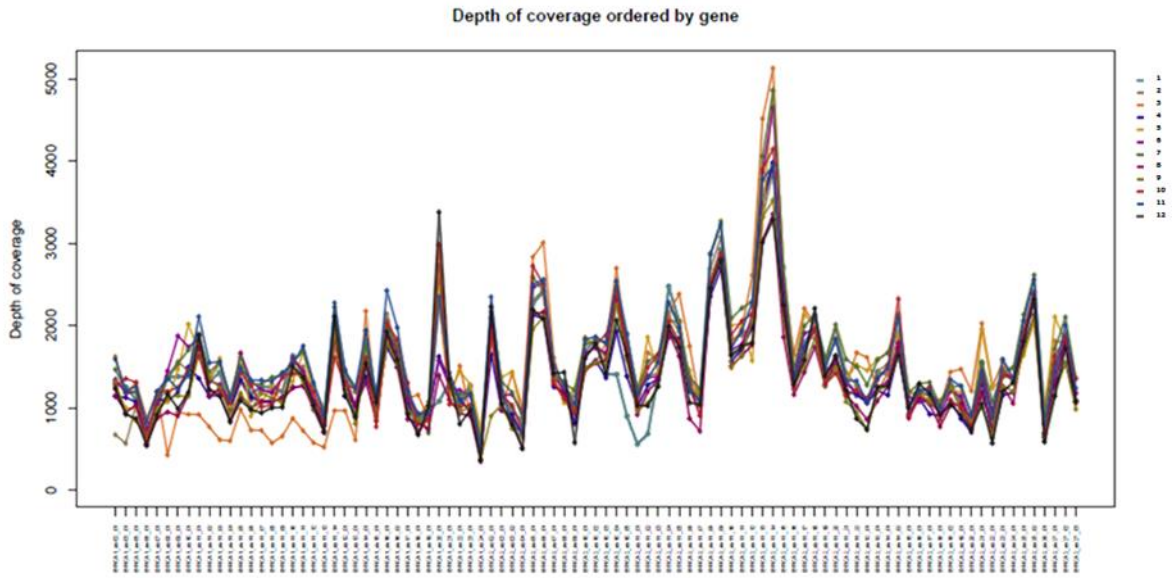


Figura 28. Representación de la cobertura por muestra ordenada por exones. Se muestra la profundidad ordenada por exones para las 12 muestras de validación. Observamos una menor cobertura en el exón 24 del gen BRCA1.

13.5 Evaluación de cambios en el número de copias (CNVs)

La implementación del uso de la profundidad de lectura (*depth*) tanto con el uso de la plataforma *454 GS Junior*, como con el uso de la plataforma *MiSeq*, se realizó principalmente para:

- Confirmar las deleciones reales y descartar los falsos positivos generados por la pirosecuenciación.
- Detectar las posibles deleciones/duplicaciones que impliquen a uno o varios exones de un gen.

Los valores de profundidad media por amplicón sin normalizar, que nos servían inicialmente para estudiar la cobertura global, nos hicieron intuir que, en algunas carreras, había ligeros descensos en la profundidad de lectura media por amplicón, que no tenían relación con una disminución en la eficiencia en alguna de las PCR multiplex. Aplicando una serie de pasos de normalización, fuimos capaces de disminuir el ruido entre muestras, destacando de forma más clara, las variaciones en el número de copias presentes en las muestras control. Así, la presencia de CNVs, se podían intuir de forma gráfica, mediante el análisis normalizado del *depth*.

El análisis del *depth* se realiza mediante el análisis gráfico del *depth* por fragmentos y gen generado por bioinformática. En carreras posteriores, se utilizaron muestras previamente ya estudiadas por la técnica de MLPA (M*ultiplex Ligation-dependent Probe Amplification*), y se intentó encontrar una estrategia para la detección de las variaciones en el número de copias (CNVs), así, se incluyeron cuatro muestras con CNVs. En la tabla 14, se describen las VP detectadas.

Tabla 14. Grandes deleciones testadas por MLPA e identificadas mediante el análisis del *detph*

Muestras	Deleciones detectadas
1	**Deleción de los exones 1 al 3 del gen *BRCA1***
2	**Deleción de los exones 8 al 13 del gen *BRCA1***
3	**Deleción del exón 2 del gen *BRCA2***
4	**Deleción del exón 11 del gen *BRCA2***

La misma estrategia descrita para la detección de CNVs basada en el análisis del *depth* fue evaluada para *MiSeq*, empleando las mismas muestras como controles de grandes deleciones.

En las siguientes figuras, se representa la normalización entre todas las muestras del *depth* por fragmento amplificado, ordenado por genes en dos carreras distintas de *454 GS Junior* (Figuras 29 y 30). Como podemos observar, existen carreras con mucha disparidad tras

normalizar los datos de cobertura. Sin embargo, una disminución del *depth* de cobertura en un fragmento parece ser indicativa de la presencia de una posible deleción.

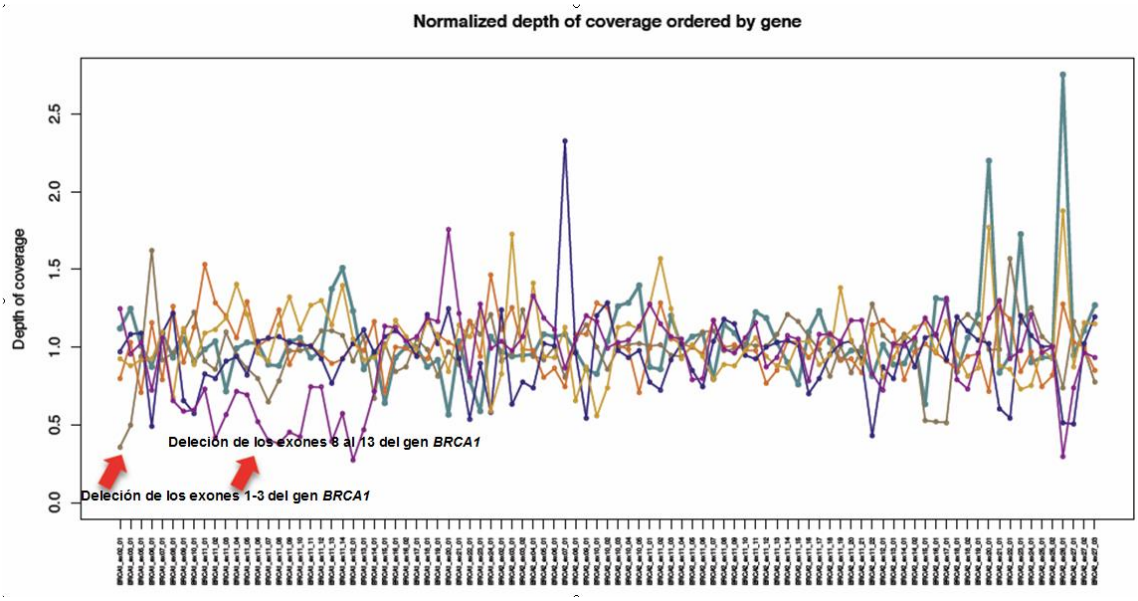


Figura 29. Detección de deleciones grandes a partir de los datos de profundidad de lectura por amplicón normalizados para una carrera. *Cada línea de color representa una muestra distinta. Se aprecia la deleción de los exones 1-3 del gen BRCA1 (muestra verde) y la deleción del final de los exones 8-13 para el gen BRCA1 (muestra morada). Se aprecia un alto ruido de fondo.*

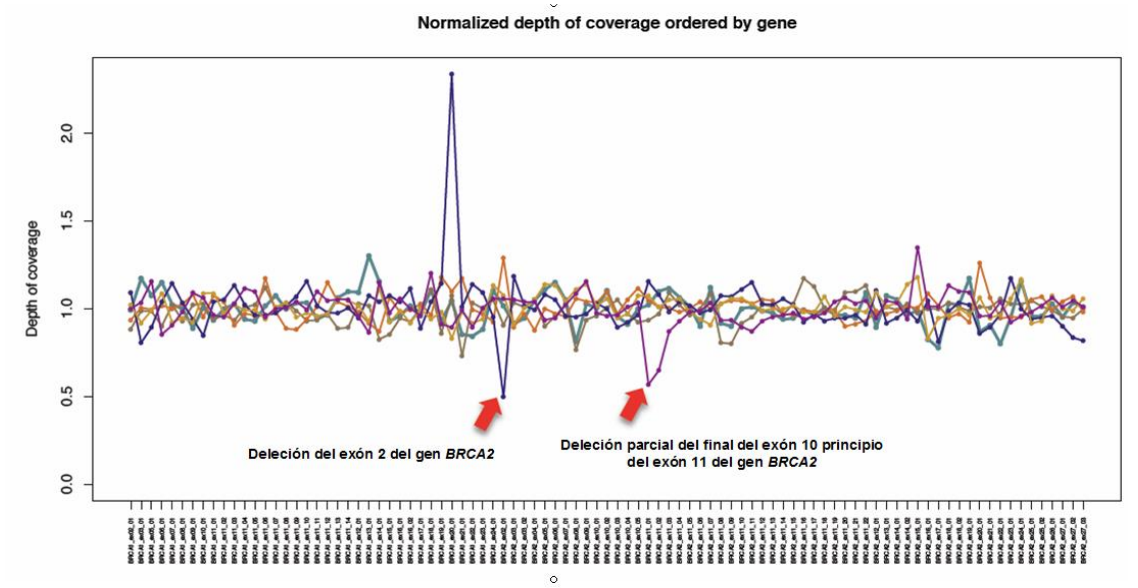


Figura 30. Detección de deleciones grandes a partir de los datos de profundidad de lectura por amplicón normalizados para una carrera. Cada línea de color representa una muestra distinta. Se aprecia la deleción del exón 2 del gen *BRCA2* (muestra azul) y la deleción del final del exón 10 y principio del exón 11 para el gen *BRCA2* (muestra morada).

En los gráficos observamos un elevado ruido de fondo en el análisis de CNVs, donde todos los puntos señalados en rojo serían sospechosos de ser una deleción/duplicación por lo que habría que validarlo por una tecnología complementaria (MLPA), (Figura 30).

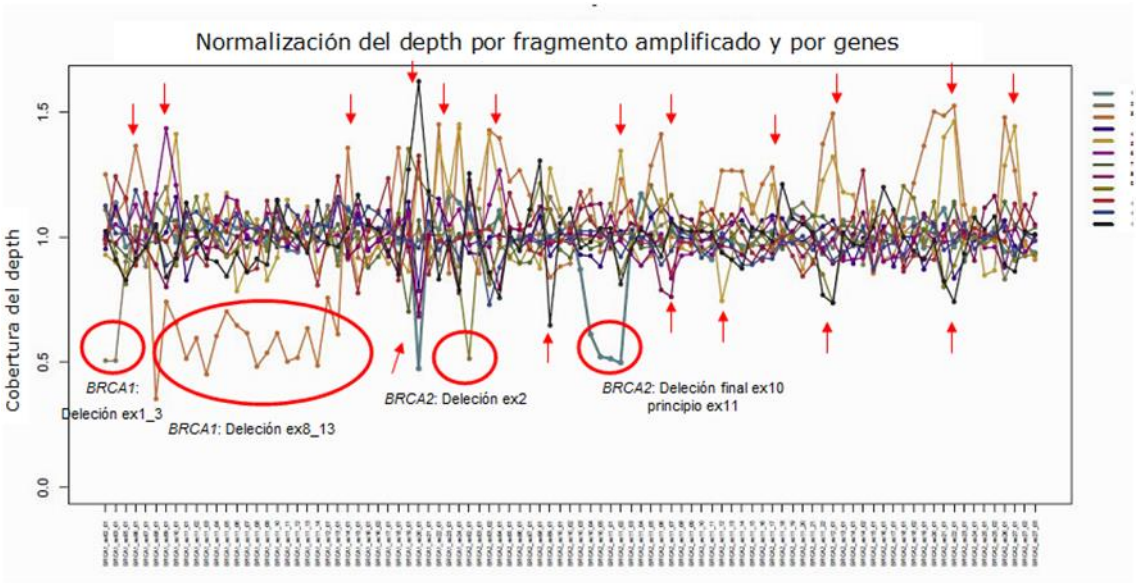


Figura 31. Normalización del depth por fragmento amplificado y ordenado por genes de las muestras. Cada línea de color representa una muestra distinta. Mediante las flechas de color rojo indicamos los falsos positivos debidos al ruido de fondo. Con los círculos rojo indicamos los CNVs de las muestras control.

14. Objetivo 2.1: Flujo de trabajo propuesto

Los datos de validación de las dos plataformas mediante secuenciación masiva nos han permitido elaborar un protocolo para tratar de obtener la máxima fiabilidad y sensibilidad en el estudio de variantes en los pacientes con sospecha de CMOH.

Cada tecnología tiene sus propias características y parten de un mismo sistema de captura de las regiones diana de los genes *BRCA1/2*, PCR multiplex mediante el sistema de *Multiplicom*.

En el caso de la **plataforma *454 GS Junior*** el flujo de trabajo, consistía en la combinación de resultados entre el *software AVA* junto con los resultados obtenidos a partir del análisis de fragmentos de las PCRs multiplex. Con esta estrategia, como hemos comentado, se hizo posible una identificación eficiente de los *indels* y se solventó así, las limitaciones de la pirosecuenciación en estas zonas de homopolímero. En estas regiones se genera una alta tasa de error, siendo más acusada cuanto mayor sea la longitud del homopolímero[102].

La estrategia de análisis de los datos de la **plataforma *MiSeq*** fue similar a la seleccionada para la plataforma *454 GS Junior* basada en la captura por amplicones.

Teniendo en cuenta los resultados obtenidos en la validación de cada plataforma de secuenciación, el flujo de trabajo elegido para el estudio de los genes *BRCA1/2* por secuenciación masiva fue el siguiente (Figura 32):

- Preparación de la librería de ADN, mediante PCR multiplex (purificación y cuantificación).
- Análisis de los fragmentos de la PCR multiplex con dos objetivos:

- confirmación de la amplificación uniforme de todos los fragmentos de la PCR multiplex,
- detección de pequeñas *indels* previo a la secuenciación.

- Amplificación clonal
  - **Plataforma *454 GS Junior*** mediante una *em*PCR (PCR en emulsión en agua y aceite de las cadenas simples seleccionadas), inmovilización de bolas enriquecidas.
  - **Plataforma *MiSeq***. Mediante PCR puente.
- Evaluación de las variantes detectadas.
  - **Plataforma *454 GS Junior***. Mediante *software* comercial *AVA* (*Roche*) para la identificación de variantes y el *pipeline* propio para la anotación de las mismas.
  - **Plataforma *MiSeq***. Mediante el *pipeline* propio.
- Confirmación de las VP/VPP mediante secuenciación Sanger.
- Análisis del *depth* para la evaluación de CNVs y confirmación mediante MLPA.

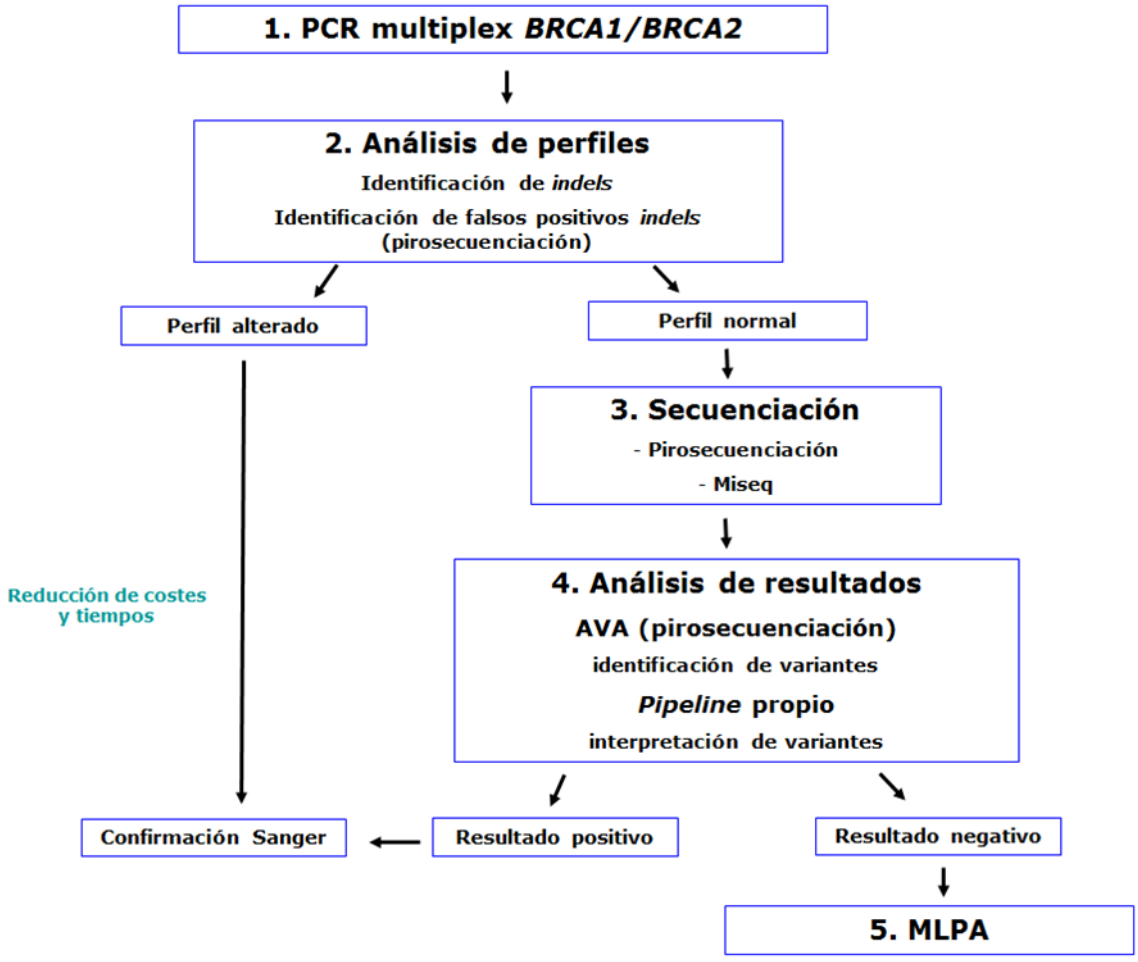


Figura 32. Flujo de trabajo implementado en el laboratorio de diagnóstico molecular.

15. Objetivo 2.2: Comparativa de resultados por metodología: secuenciación Sanger *versus* secuenciación masiva.

La aplicación del flujo de trabajo previamente descrito, nos permitió analizar los genes *BRCA1/2* en pacientes con sospecha clínica de CMOH. Los resultados obtenidos en cuanto a la detección de variantes, comparando agrupando los datos con respecto a la tecnología de secuenciación *gold standard*, Sanger, se muestran en la siguiente tabla (Tabla 15).

Tabla 15. Porcentajes de VP según metodología utilizada

	SANGER				***MASIVA***			
Fenotipo familiar	**Tamaño muestral (N)**	***BRCA1* (%)**	***BRCA2* (%)**	***BRCA1/2* (%)**	**Tamaño muestral (N)**	***BRCA1* (%)**	***BRCA2* (%)**	***BRCA1/2* (%)**
1: Un caso CM ≤50 años CM bilateral CM+CO CO	27	1 (3,7%)	1 (3,7%)	**2 (7,4%)**	74	5 (6,75%)	5 (6,75%)	**10 (13,51%)**
2: Dos casos CM varón CO CM ≤50 años CM bilateral + CM ≤50 años	56	0	4 (7,14%)	**4 (7,14%)**	58	5 (8,62%)	4 (6,89%)	**9 (15,51%)**
3: Tres casos o más a cualquier edad	23	7 (30,43%)	3 (13,04%)	**10 (43,47%)**	86	7 (8,13%)	10 (11,62%)	**17 (19,76%)**
NR: No referido	56	7 (12,5%)	3 (5,35%)	**10 (17,85%)**	63	2 (3,17%)	2 (3,17%)	**4 (6,34%)**
NC: No cumple criterios	1	0	0	**0**	11	0	0	**0**
Totales	163	**15 (9,2%)**	**11 (6,74%)**	**26 (15,95%)**	292	**19 (6,5%)**	**21 (7,19%)**	**40 (13,69%)**

CM: cáncer de mama; CO: cáncer de ovario; CMO: cáncer de mama y ovario; N= nº de casos estudiados

Con respecto a **los resultados de la serie de datos Sanger**, tenemos un total de 163 casos estudiados mediante esta tecnología. La edad media de diagnóstico de estos casos fue de 43 años. En esta serie obtenemos un 15,95% de pacientes diagnosticados genéticamente. El tiempo promedio de entrega de resultados en un estudio secuencial de secuenciación mediante el método Sanger es de 12 semanas.

Hemos estudiado 292 casos mediante **secuenciación masiva**. La edad media de diagnóstico de estos casos fue de 44 años. En esta serie obtenemos un 13,69% de pacientes diagnosticados genéticamente. El tiempo promedio de entrega de resultados en estudios de secuenciación masiva y MLPA/análisis de fragmentos es de 5 semanas.

La comparativa de costes de las tres plataformas empleadas en este trabajo, se resume en la tabla 16, modificada de Minucci *et al*[103].

Tabla 16. Comparativa de costes de las tres plataformas utilizadas en este trabajo

	Plataforma	

Parámetros	**SANGER 3730xl (€)**	***454 GS JUNIOR* (€)**	***MiSeq* (€)**	**Observaciones**
Captura de las regiones diana	82 PCRs simples (CDS+intrones +/- 20 nucleótidos) 300	*Multiplicom* 5 PCR multiplex 140	*Multiplicom* 5 PCR multiplex 140	Para Sanger son 82 reacciones de PCR para cubrir ambos genes
Análisis de fragmentos	-	5	5 (opcional)	
Purificación librerías NGS/ Productos PCR y post BigDie	50	3	3	Purificamos 164 reacciones *forward* y *reverse* con BigDie para Sanger
Reactivos secuenciación	500	70	48	
Material fungible	10	5	5	
Almacenamiento de datos	2	2	2	
Coste en personal	400	100	80	Las diferencias en costes de personal son debidas a que en NGS se pueden procesar hasta 12 muestras (*MiSeq*) y 8 muestras (*454 GS Junior*), en Sanger solamente de una en una
Total por muestra	1264	325	283	Los costes para NGS han sido calculados con un análisis negativo en fragmentos; a estos gastos habría que sumar las posibles 1-2 PCR de confirmación de variantes con relevancia clínica mediante Sanger

Observamos el menor coste cuando empleamos técnicas de secuenciación masiva, si las muestras resultaran positivas para el análisis de fragmentos, tendríamos un ahorro considerable en el coste de ambas. Para todas ellas, los posibles cambios en el número de copias se testarían mediante MLPA.

Si tenemos en cuenta los resultados obtenidos por secuenciación masiva, podemos observar una serie de ventajas con respecto a la secuenciación Sanger, teniendo en cuentas las diferencias existentes en cuanto a la plataforma de secuenciación utilizada. De este modo:

1. Observamos una **reducción en el número de variantes/muestra**. Con el secuenciador *454 GS Junior* se obtenía una media de aproximadamente unas 26 variantes/muestras mientras que con el secuenciador *MiSeq* se obtiene una media de 20 variantes/muestra. Esto se debe a que no se observan errores en la zona de homopolímeros con *MiSeq*, siendo una mejora con respecto a la tecnología anterior.

2. Observamos, que **el *pipeline* propio detecta todos los cambios puntuales y pequeñas deleciones/inserciones** de las muestras control. Este *pipeline* resulta más completo que el *software MiSeqReporter* (*Illumina*) por la anotación más detallada de variantes y por permitir la evaluación de todas las muestras de la carrera en paralelo. Se implantó para la rutina de trabajo del laboratorio, el uso del *pipeline* propio.

3. Las **CNVs se pueden detectar mediante el *depth* por amplicón**, pero dado el elevado ruido de fondo que se observa, es necesario completar la evaluación mediante una tecnología complementaria (observamos lo mismo con la tecnología anterior).

4. Se pueden **procesar mayor número de muestras en la misma carrera**, con *454 GS Junior* 6-8 muestras, con *MiSeq* se obtienen buenos datos con 12-14 muestras en una misma carrera.

16. Objetivo 3: Caracterización del espectro de variantes de los genes *BRCA1* y *BRCA2* en nuestra cohorte de 455 casos de pacientes con sospecha clínicas de cáncer de mama y ovario hereditario.

Tras el análisis de los genes *BRCA1* y *BRCA2* de los 455 casos índice con sospecha clínica de CMOH, se detectaron un total de 68 variantes distintas (64 cambios puntuales y 4 cambios en el número de copias). Se han obviado las VPB y VB. Las variantes detectadas, se han distribuido según la clasificación acorde a su efecto clínico, tabla 17.

Tabla 17. Variantes detectadas en nuestra cohorte

Genes	VP y VPP	VSI	Totales
BRCA1	23	7	**30 (44,11%)**
BRCA2	29	9	**38 (55,88%)**
Totales	**52 (76,47%)**	**16 (23.52%)**	**68 (100%)**

Hemos encontrado un porcentaje de variantes con relevancia clínica similar en ambos genes, aunque ligeramente superior en el gen *BRCA2* en ambos tipos de variantes.

En nuestra cohorte, hemos diagnosticado genéticamente un **15,38%** casos índices (70/455), setenta casos estudiados presentaban una VP o VPP, un 3,95% (18/455) de los casos índice presentaban VSI y un 80,65% presentaban VPB o VB (Figura 33).

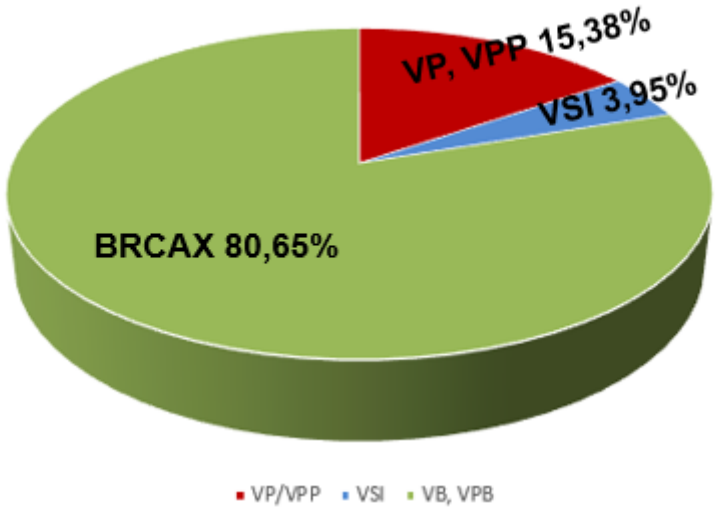


Figura 33. Tasa de diagnóstico. Porcentaje de casos índices con variantes clínicamente relevantes, en los genes *BRCA1/2*. En rojo: casos índices con VP, azul: casos índices con VSI y verde: casos índices con variantes sin relevancia clínica.

En los siguientes apartados, se detallan en profundidad las variantes identificadas en los casos índices estudiados.

16.1 Variantes patogénicas/probablemente patogénicas detectadas en los genes BRCA1 y BRCA2.

Se han clasificado como VP/VPP un total de **52 variantes distintas en un total de 455 familias**, lo que supone un 69,33% del total de variantes clínicamente relevantes detectadas. De las 52 variantes detectadas, 23 fueron detectadas en el gen *BRCA1* y 29 en *BRCA*2. De todas ellas 6 son de nueva descripción (deleción parcial del exón 11, c.1792del (p.Thr598Hisfs*16), c.4263dup (p.Glu1422*), c.4677del (p.Phe1559Leufs*9), c.6034del (p.Ser2012Profs*28) en *BRCA2* y c.4106_4107delinsGCT (p.Ala1369Glyfs*5) en el gen *BRCA1*) y se han detectado por primera vez en este estudio.

En base al **tipo de variante**, la mayoría de las VP o VPP detectadas en este estudio han sido de tipo *frameshift* (32/52), seguido de las variantes *nonsense* (9/52), variantes en zonas de *splicing* (5/52), CNVs (4/52) y por último las variantes *missense* (2/52). En la tabla 18, se realiza una descripción de las VP y VPP y las características de los casos índice.

Tabla 18. VP y VPP detectadas y características de las mismas.

BRCA1

Exón	Cambio nucleotídico[#]	Tipo	Nº de casos índice	Nº familia (característica)	Predictores *in silico*	DbSNP/ MAF	ClinVar	HGMD	LOVD[3]	BRCA EXCHANGE	Criterios ACMG	Ref
2	c.64_65dup (p.Leu22Phefs*2)	F	1	44 (CM 26 años)	-	rs80357642 Sin frecuencias	236275*** Patogénica	NR	BRCA1_002937 Patogénica	Patogénica	PVS1, PM2, PP5	-
2	c.68_69del (p.Glu23Valfs*17)	F	1	34 (CM 47 años, 4 estudios predictivos, 2 portadores)	-	rs80357914 delCT=0,0227% (57/251050, GnomAD_exome)	17662*** Patogénica	CD951600 DM	BRCA1_001114 Patogénica	Patogénica	PVS1, PM2, PP5	[104]
3	c.83_84del (p.Leu28Argfs*12)[Δ]	F	1	143 (CM)	MaxEnt: -0,9% NNSPLICE: -36,3% HSF: 0,0%	rs80357728 Sin frecuencias	55733*** Patogénica	CD086714 DM	BRCA1_001903 Patogénica	Patogénica	PVS1, PM2, PP5	[105]
3	c.124del (p.Ile42Tyrfs*8)	F	1	185 (CM)	-	rs80357943 Sin frecuencias	54173*** Patogénica	CD972043 DM	BRCA1_001895 Patogénica	Patogénica	PVS1, PM2, PP5	[106]
5	c.211A>G (p.Arg71Gly)	Sp	4	10 (CM) 15 (2 casos, CM) 21 (CM) 111 (CM)	Align GVGD: C65 SIFT: Deleterious MutationTaster: disease causing PolyPhen-2: possibly damaging MaxEnt: -50,5% NNSPLICE: -53,2% HSF: -6,2%	rs80357382 C=0,0004% (1/249744, GnomAD_exome)	17693*** Patogénica	CM994455 DM	BRCA1_000059 Patogénica	Patogénica	PM1, PM2, PM5, PP3	[107]

Tabla 18. VP y VPP detectadas y características de las mismas.												
I5	c.213-2A>C	Sp	2	49 (CM bilateral 38 y 40 años) 102 (CO 33 años)	MaxEnt: -100,0% NNSPLICE: 0,0% HSF: -100,0%	rs397508940 Sin frecuencias	54473* Patogénica	CS109761 DM	BRCA1_001163 Patogénica	No revisado	PVS1, PM2, PP5	[108]
11b	c.1121_1123delinsT (p.Thr374Ilefs*3)	F	1	37 (CM 54 años, 2 estudios predictivos, uno portador)	-	rs273897652 Sin frecuencias	54137*** Patogénica	CX032787 DM	BRCA1_002712 Patogénica	Patogénica	PVS1, PM2, PP5	[104]
11b	c.1674del (p.Gly559Valfs*13)	F	1	40 (CM 56 años)	-	rs80357600 Sin frecuencias	37425*** Patogénica	CD015140 DM	BRCA1_001809 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11b	c.1965C>A (p.Tyr655*)	N	1	2 (CM 41 años)	-	rs886039987 Sin frecuencias	266207*** Patogénica	CM128923 DM	BRCA1_002295 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11b	c.2901_2902dup (p.Pro968Leufs*33)	F	4	7 (CM) 14 (CM) 19 (CM) 33 (CO)	-	rs398122670 Sin frecuencias	91602*** Patogénica	CI134880 DM	BRCA1_000739 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11b	c.3331_3334del (p.Gln1111Asnfs*5)	F	2	3 (CMB y CO) 13 (CM 61 años)	-	rs80357701 delGCTT=0,0008% (1/125568, TOPMED)	37523*** Patogénica	CD961838 DM	BRCA1_001471 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11b	c.3689T>G (p.Leu1230*)	N	1	20 (CM)	-	rs80357162 Sin frecuencias	37540*** Patogénica	CM134878 DM	BRCA1_000889 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11b	c.3770_3771del (p.Glu1257Glyfs*9)	F	2	149 (CM 30 años) 152 (CM) 2 estudios predictivos, un portador	-	rs80357579 delCT=0,0008% (2/251226, GnomAD_exome) delCT=0,0024% (3/125568, TOPMED)	37546*** Patogénica	CD961840 DM	BRCA1_001550 Patogénica	Patogénica	PVS1, PM2, PP5	[107]

Tabla 18. VP y VPP detectadas y características de las mismas.												
						delCT=0,0008% (1/121396, ExAC)						
11b	c.4041_4042del (p.Gly1348Asnfs*7)[Δ]	F	1	134 (CO 35 años, 1 estudio predictivo)	-	rs80357727 Sin frecuencias	55082*** Patogénica	CD022129 DM	BRCA1_000784 Patogénica	Patogénica	PVS1, PM2, PP5	[109]
11b	**c.4106_4107delinsGCT (p.Ala1369Glyfs*5)**^	F	1	63 (Paciente de 32 años, dos CM en la familia, uno antes de los 50 años)	MaxEnt: +0,0% NNSPLICE: +3.9% HSF: +0,0%	NR	NR	NR	NR	NR	PVS1, PM2	-
14	c.4456del (p.Ser1486Valfs*19)	F	1	125 (CO) 2 estudios predictivo, uno portador	-	rs397509179 Sin frecuencias	55204*** Patogénica	CD042855 DM	BRCA1_003489 Patogénica	Patogénica	PVS1, PM2, PP5	[110]
18	c.5117G>A (p.Gly1706Glu)	M	1	25 (CM)	Align GVGD: C65 SIFT: Deleterious MutationTaster: disease causing PolyPhen-2: possibly damaging	rs80356860 G=0,0048% (12/251342, GnomAD_exome) G=0,0056% (7/125568, TOPMED) G=0,0041% (5/120634, ExAC)	37638*** Patogénica	CM000502 DM	BRCA1_000392 VPP	Patogénica	PS3, PM1, PM2, PM5, PP3	[107,111]
18	c.5123C>A (p.Ala1708Glu)	Sp	6	55 (CM triple- 47 años) 69 (CM 37 años) 82 (CO) 146 (CO) 179 (CM) 186 (CM)	Align GVGD: C65 SIFT: Deleterious MutationTaster: disease causing PolyPhen-2: probably damaging	rs28897696 A=0,0080% (10/125568, TOPMED) A=0,003% (1/31382, GnomAD)	55407*** Patogénica	CM950153 DM	BRCA1_000393 Patogénica	Patogénica	PS3, PM2, PP3	[111]
21	c.5324T>G (p.Met1775Arg)	M	1	56 (CM triple- 34 años)	Align GVGD: C45	rs41293463	17694***	CM950157	BRCA1_000453	Patogénica	PP3, PP5, PM2, PM5	[111]

Tabla 18. VP y VPP detectadas y características de las mismas.												
					SIFT: Deleterious MutationTaster: disease causing PolyPhen-2: probably damaging	C=0,0012% (3/251394, GnomAD_exome) C=0,0024% (3/125568, TOPMED)	Patogénica	DM	Patogénica			
23	c.5419del (p.Ile1807Leufs*27)[Δ]	F	1	173 (CM)	MaxEnt: 0,0% NNSPLICE: -19.5% HSF: 0,0%	rs80357934 Sin frecuencias	55571*** Patogénica	CD994462 DM	BRCA1_003177 Patogénica	Patogénica	PVS1, PM2, PP5	[112]
BRCA2												
Exón	**Cambio nucleotídico[#]**	**Tipo**	**Nº de casos índice**	**Nº Caso índice (característica)**	**Predictores *in silico***	**dbSNP MAF**	**ClinVar**	**HGMD**	**LOVD**	**BRCA EXCHANGE**	**Criterios ACMG**	**Ref**
I9	c.793+1G>A	Sp	1	133 (CM 41 años, estudio predictivo y portador)	MaxEnt: -100,0% NNSPLICE: -100,0% HSF: -100,0%	rs81002846 A=0,003% (1/31208, GnomAD)	52437** Patogénica/ Probablemente patogénica	NR	BRCA2_001256 Patogénica	No revisado	PVS1, PM2, PP5	-
10	c.1597del (p.Thr533Leufs*25)	F	1	1 (CM 67 años)	-	rs80359292 Sin frecuencias	51152*** Patogénica	CD020556 DM	BRCA2_003046 Patogénica	Patogénica	PVS1, PM2, PP5	[104]
10	**c.1792del (p.Thr598Hisfs*16)[Δ]**	F	1	131 (CM bilateral)	-	rs886040389 Sin frecuencias	266660*** Patogénica	NR	BRCA2_004272 Patogénica	Patogénica	PVS1, PM2, PP5	[113]
10	c.1813dup (p.Ile605Asnfs*11)	F	1	26 (CM 37 años)	-	rs80359306 dupA=0,0026% (3/115908, ExAC)	37762*** Patogénica	CI972557 DM	BRCA2_001802 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11	c.2095C>T (p.Gln699*)	N	1	27 (CM 44 años)	-	rs878853559	236835***	CM134881	BRCA2_004332	Patogénica	PVS1, PM2, PP5	[107]

Tabla 18. VP y VPP detectadas y características de las mismas.												
						Sin frecuencias	Patogénica	DM	Patogénica			
11	c.2808_2811del (p.Ala938Profs*21)	F	2	148 (CM) 168 (CM varón)	-	rs80359351 delACAA=0,0008% (2/250852, GnomAD_exome) delACAA=0,0040% (5/125568, TOPMED) delACAA=0,0017% (2/120744, ExAC)	9322*** Patogénica	CD951625 DM	BRCA2_001211 Patogénica	Patogénica	PVS1, PM2, PP5	[114]
11	c.2830A>T (p.Lys944*)	N	1	16 (CM)	-	rs80358533 T=0,0008% (2/250612, GnomAD_exome) T=0,0024% (3/125568, TOPMED)	51355*** Patogénica	CM970180 DM	BRCA2_001214 Patogénica	Patogénica	PVS1, PM2, PP5	[115]
11	c.2957dup (p.Asn986Lysfs*2)	F	1	166 (CM)	-	rs80359365 Sin frecuencias	51378*** Patogénica	CI063656 DM	BRCA2_001395 Patogénica	Patogénica	PVS1, PM2, PP5	[116]
11	c.3680_3681del (p.Leu1227Glnfs*5)	F	1	114 (CM)	-	rs80359395 delTG=0,0008% (1/125568, TOPMED) delTG=0,003% (1/31392, GnomAD)	51504*** Patogénica	CD011120 DM	BRCA2_002089 Patogénica	Patogénica	PVS1, PM2, PP5	[117]
11	c.4030_4035delinsC (p.Asn1344Hisfs*6)	F	2	6 (CM) 8 (CM)	-	rs886040509 Sin frecuencias	266788*** Patogénica	CX134883 DM	BRCA2_004585 Patogénica	Patogénica	PVS1, PM2, PP5	[107]
11	**c.4263dup (p.Glu1422*)[Δ]**	N	1	91 (CM Y CO)	-	rs1555283664 Sin frecuencias	419134*** Patogénica	NR	NR	Patogénica	PVS1, PM2, PP5	-

Tabla 18. VP y VPP detectadas y características de las mismas.												
11	**c.4677del (p.Phe1559Leufs*9)$^{\Delta}$**	F	1	60 (CM 61 años y 65 años, melanoma 67 años, cáncer trompa de Falopio)	-	rs1064794708 Sin frecuencias	420798*** Patogénica	NR	NR	Patogénica	PVS1, PM2, PP5	-
11	c.4894_4895del (p.Ser1632Tyrfs*6)	F	1	35 (CM Multicentrico 40 años)	-	rs397507748 Sin frecuencias	51735*** Patogénica	CD095770 DM	BRCA2_004685 Patogénica	Patogénica	PVS1, PM2, PP5	[118]
11	c.5350_5351del (p.Asn1784Hisfs*2)	F	1	99 (CO y dos estudios predictivos, portadores los dos)	-	rs80359507 delAA=0,0024% (3/125568, TOPMED) delAA=0,0008% (1/120630, ExAC) delAA=0,003% (1/31198, GnomAD)	37959*** Patogénica	CD972078 DM	BRCA2_001465 Patogénica	Patogénica	PVS1, PM2, PP5	[109]
11	c.5621_5624del (p.Ile1874Argfs*34)	F	1	118 (CM Bilateral 50 y 60 años)	-	rs80359526 delTTAA=0,0004% (1/247614, GnomAD_exome)	37980*** Patogénica	CD972080 DM	BRCA2_003286 Patogénica	Patogénica	PVS1, PM2, PP5	[119]
11	c.5720_5723del (p.Ser1907*)	N	2	58 (CM 42 años, 8 estudios predictivos, 2 portadores) 83 (CM, 1 estudio predictivo)	-	rs80359530 delCT=0,0004% (1/250888, GnomAD_exome) delCT=0,0008% (1/125568, TOPMED)	254563*** Patogénica	CD020559 DM	BRCA2_003301 Patogénica	Patogénica	PVS1, PM2, PP5	[120]
11	c.5980C>T (p.Gln1994*)	N	1	187 (CM 24 años, tres estudios predictivos)	-	rs80358831 Sin frecuencias	38011*** Patogénica	CM023027 DM	BRCA2_000150 Patogénica	Patogénica	PVS1, PM2, PP5	[121]
11	c.6024dup (p.Gln2009Alafs*9)	F	1	105 (CM a los 33 años)	-	rs80359554 dupG=0,0004% (1/250690, GnomAD_exome)	38015*** Patogénica	CI057920 DM	BRCA2_000691 Patogénica	Patogénica	PVS1, PM2, PP5	[122]

Tabla 18. VP y VPP detectadas y características de las mismas.												
11	**c.6034del (p.Ser2012Profs*28)$^{\Delta}$**	F	1	64 (CM a los 34, 1 estudio predictivo portador)	-	rs397507823 delTT=0,0004% (1/250824, GnomAD_exome) delTT=0,0008% (1/120502, ExAC)	427240*** Patogénica	NR	NR	Patogénica	PVS1, PM2, PP5	[123]
11	c.6037A>T (p.Lys2013*)	N	1	96 (CM)	-	rs80358840 G=0,0004% (1/250808, GnomAD_exome) G=0,0008% (1/125568, TOPMED)	38016*** Patogénica	CM021512 DM	BRCA2_002849 Patogénica	Patogénica	PVS1, PM2, PP5	[124]
11	c.6209_6212del (p.Glu2070Valfs*10)	F	1	167 (CM 37 años)	-	rs276174866 Sin frecuencias	52032*** Patogénica	CD129038 DM	BRCA2_002256 Patogénica	Patogénica	PVS1, PM2, PP5	[125]
11	c.6275_6276del (p.Leu2092Profs*7)	F	2	48 (CM y CO 50 y 62 años) 54 (CM)	-	rs11571658 delTT=0,0037% (9/243532, GnomAD_exome) delTT=0,0016% (2/125568, TOPMED) delTT=0,0017% (2/119988, ExAC)	9318*** Patogénica	CD951627 DM	BRCA2_000156 Patogénica	Patogénica	PVS1, PM2, PP5	[109]
11	c.6405_6409del (p.Asn2135Lysfs*3)	F	1	29 (CM 50 años) Estudio predictivo a 3 familiares	-	rs80359584 delCTTAA=0,0004% (1/240630, GnomAD_exome) delCTTAA=0,0008% (1/118050, ExAC)	38043*** Patogénica	CD972084 DM	BRCA2_001907 Patogénica	Patogénica	PVS1, PM2, PP5	[126]
11	c.6486_6489del (p.Lys2162Asnfs*5)	F	1	41 (CM 38 años, dos familiares de primer	-	rs80359598	38048***	CD982506	BRCA2_000162	Patogénica	PVS1, PM2, PP5	[120]

Tabla 18. VP y VPP detectadas y características de las mismas.												
				grado con CM, 1 estudio predictivo)		delACAA=0,0009% (2/232332, GnomAD_exome) delACAA=0,0016% (2/125568, TOPMED) delACAA=0,0017% (2/118868, ExAC)	Patogénica	DM	Patogénica			
17	c.7863T>A (p.Tyr2621*)	N	1	65 (Paciente de 35 años, tres familiares 1 grado con CM, 1 estudio predictivo)	-	rs276174896 Sin frecuencias	52428*** Patogénica	CM086257 DM	BRCA2_005259 Patogénica	Patogénica	PVS1, PM2, PP5	[127]
18	c.8023A>G (p.Ile2675Val)	M	1	22 (CM)	Align GVGD: C25 SIFT: Deleterious MutationTaster: disease causing PolyPhen-2: Probably damaging	rs397507954 G=0,0004% (1/251076, GnomAD_exome)	52475*** Patogénica	CS083875 DM	BRCA2_000251 Patogénica	Patogénica	PM2, PP3, PP5	[108]
25	c.9026_9030del (p.Tyr3009Serfs*7) VP fundadora del nordeste de España	F	1	31 (CM a los 48 años)	-	rs80359741 delATCAT=0,0004% (1/250686, GnomAD_exome)	38204*** Patogénica	CD961862 DM	BRCA2_000893 Patogénica	Patogénica	PVS1, PM2, PP5	[114]
25	c.9310_9311del (p.Lys3104Valfs*6) VP fundadora en Castilla-León	F	1	122 (CM)	-	rs80359756 Sin frecuencias	52813*** Patogénica	CD000597 DM	BRCA2_005662 Patogénica	Patogénica	PVS1, PM2, PP5	[128]

#Se ha empleado la nomenclatura recomendada por la *Human Genome Variation Society* (HGVS v.2.0). ^Las variantes no descritas, pero con un efecto sobre la proteína de cambio de fase de lectura o codón de parada prematuro se han considerado como VPP. M: *Missense*; F: *Frameshift*; Sp: *Splicing*; N: *Nonsense*; NR: No referido; CM: Cáncer de mama; CMB: Cáncer de mama bilateral; CO: Cáncer de ovario; HGMD: DM (*disease causing mutation*), DM? (variante indeterminada); IARC: clase 5: patogénica; Δ: Variantes descritas posteriormente a ser identificadas en nuestra cohorte.

16.2 Cambios en el número de copias

El estudio realizado para analizar el cambio en el número de copias en los genes *BRCA1/2*, ha sido realizado en un 67,69% (308/455) de los casos estudiados:

- En 175 casos se realizó el estudio mediante la metodología MLPA:
    - En 126 casos se realizó simultáneamente el estudio de secuenciación masiva y MLPA
    - En 49 casos se realizó el MLPA tras descartar la presencia de una variante puntual en los genes *BRCA1* o *BRCA2*
- En 133 casos se realizó el estudio mediante la metodología de análisis de fragmentos.

Los resultados se resumen en la siguiente tabla:

Tabla 19. Cambios en el número de copias

Genes	*Nº VP (%)*
BRCA1	3 (0,97%)
BRCA2	1 (0,32%)
BRCA1+BRCA2	4 (1,29%)

El análisis reveló la presencia de 4 cambios en el número de copias, 3 grandes deleciones en el gen *BRCA1* y una gran deleción en el gen *BRCA2* (Tabla 20). No se ha detectado ninguna gran duplicación genómica. Este hallazgo supone que un 1,29% de las familias testadas eran portadoras de una gran deleción. La frecuencia de casos portadores de este tipo de alteraciones genómicas supone un 0,97% (3/308) en el gen *BRCA1* y un 0,32% (1/308) en el gen *BRCA2.*

Como podemos observar, se ha detectado un mayor porcentaje cambios en el número de copias en el gen *BRCA1*.

Tabla 20. Cambios en el número de copias detectados en los genes *BRCA1/2*.

BRCA1								
Exones	**Posición genómica**	**Tipo**	**Nº de casos índice**	**Nº caso índice (característica)**	**dbSNP/ ClinVar**	**HGMD**	**ClinGen**	**Ref**
Deleción del exón 1 al 3	g.(?_43115768)_(43125304_?)del	CNV	1	47 (CM 41 años)	584117	CG056497 DM	ISCA_20827 Score haploinsuficiencia :3	[129,130]
Deleción del exón 7 al 12	g.(¿_43082459)_(43099847_?)del NM_007294.3:c.442-?_4357+?del	CNV	1	50 (CM 40 años, estudio predictivo 3 familiares, 2 portadores)	254055	CG994821 DM	ISCA_20827 Score haploinsuficiencia :3	[130]

Deleción exón 23	g.(¿_43044295)_(43045802_?)del	CNV	1	129 (CM y O)	401496	CG146555 DM	ISCA_20827 Score haploinsuficiencia :3	[131]
BRCA2								
Exones	**Posición genómica**	**Tipo**	**Nº de casos índice**	**Nº caso índice (características)**	**dbSNP/ ClinVar**	**HGMD**	**ClinGen**	**Ref**
Deleción parcial del exón 11$^{\Delta}$	g.(¿_32336265)_(32336383_?)del	CNV	1	45 (CM)	NR	NR	ISCA_33268 Score haploinsuficiencia :3	-

NR: No referido; CM: Cáncer de mama; CO: Cáncer de ovario; HGMD: DM (*disease causing mutation*).

Las deleciones encontradas en el gen *BRCA1*, se han identificado en dos pacientes con cáncer de mama clasificados en el grupo 3, según nuestra clasificación, con edad de diagnóstico de 40 y 48 años respectivamente, el tercer caso se trata de un caso índice con cáncer de mama y cáncer de ovario, desconocemos la edad de diagnóstico, es un caso clasificado como tipo 2, tiene una hermana con cáncer de mama y cáncer de ovario. Por último, hemos detectado una gran deleción en el gen *BRCA2,* desconocemos la historia familiar, lo hemos clasificado en el grupo NR, es un cáncer de mama.

a) Deleción del final del exón 10 y principio del exón 11 en el gen *BRCA2*$^{\Delta}$

Cabe resaltar un caso especial, observado por análisis de fragmentos, en los perfiles para el gen *BRCA2*. El análisis reveló, una disminución de determinados picos del final del exón 10 y principio del exón 11 del gen *BRCA2*. Por MLPA confirmamos la deleción en heterocigosis de parte del exón 11 del gen *BRCA2*, pero no del exón 10; ya que solamente observamos la deleción de la sonda 11a. Sin embargo, la sonda 11b presentaba dos copias, al igual que la sonda del exón 10. Esta última hibrida al principio del exón 10. Por tanto, probablemente la deleción comience en la parte final del exón 10 y principio del exón 11 del gen *BRCA2* y por eso no observamos la deleción del exón 10 mediante el MLPA (Figura 34).

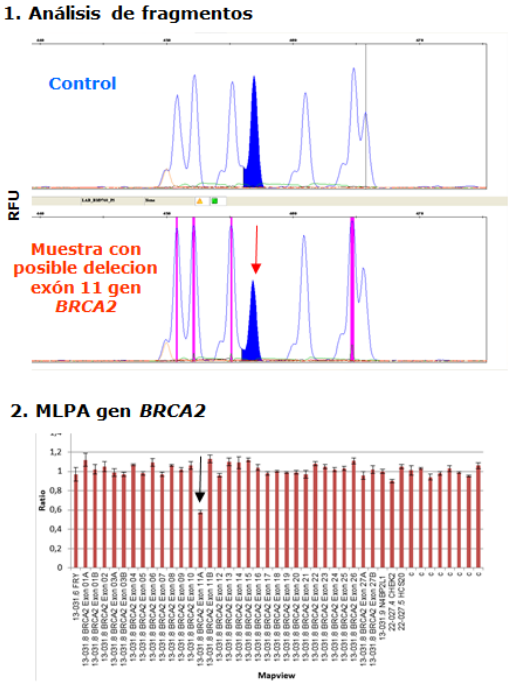


Figura 34. Análisis de fragmentos y análisis mediante MLPA. 1) Análisis de fragmentos. En la parte superior se muestra el perfil de una muestra control y en la parte inferior una disminución pronunciada de un fragmento del exón 11 del gen *BRCA2,* indicando una posible deleción. 2) Confirmación mediante MLPA de la deleción de parte del exón 11 del gen *BRCA2*.

16.3 Variantes patogénicas recurrentes

La mayoría de las familias portadoras de variantes deletéreas presentan alteraciones únicas. Sin embargo, existe un cierto número de variantes recurrentes en familias no relacionadas dentro de nuestra cohorte de pacientes, que se corresponden con VP fundadoras descritas previamente en población española, lo que sugeriría un origen común.

Veinticinco de las 70 variantes con importancia clínica detectada eran VP fundadoras en nuestra cohorte, representando un **35,71% del total de VP detectadas**. En la tabla 21 se señalan las VP más frecuentes en la población analizada.

Esto indica que sólo seis VP son suficientes para explicar el 54% (20/37) de las familias *BRCA1* positivas; mientras que cuatro dan cuenta del 24,24% (8/33) de las familias *BRCA2* positivas.

Tabla 21. VP recurrentes en nuestro estudio.

BRCA1		***BRCA2***	
Cambio nucleotídico	**N= 37 (%*BRCA1*+)**	**Cambio nucleotídico**	**N= 33 (%*BRCA2*+)**
c.5123C>A (p.Ala1708Glu)	6 (16%)	c.2808_2811del (p.Ala938Profs*21)	2 (6%)

c.211A>G (p.Arg71Gly)	4 (11%)	c.4030_4035delinsC (p.Asn1344Hisfs*6)	2 (6%)
c.2900_2901dup (p.Pro968Leufs*33)	4 (11%)	c.5720_5723del (p.Ser1907*)	2 (6%)
c.213-2A>C	2 (5%)	c.6275_6276del (p.Leu2092Profs*7)	2 (6%)
c.3331_3334del (p.Gln1111Asnfs*5)	2 (5%)		
c.3770_3771del (p.Glu1257Glyfs*9)	2 (5%)		

N: Número de casos con VP o VPP en el estudio.

Las VP recurrentes que hemos detectado en el gen *BRCA1* en nuestra cohorte de pacientes, han sido en su mayoría VP fundadoras españolas, de este modo, la variante c.5123C>A (p.Ala1708Glu), se ha detectado en 6 familias independientes, lo que supone que un 16% de la población es portadora de una variante en el gen *BRCA1*. Por otro lado, las variantes c.211A>G (p.Arg71Gly) y c.2900_2901dup (p.Pro968Leufs*33) en el gen *BRCA1*, se han detectado en 4 ocasiones cada una, lo que supone un 11% de los casos con una VP en el gen *BRCA1*.

En relación al gen *BRCA2* las variantes c.2808_2811del (p.Ala938Profs*21), c.4030_4035delinsC (p.Asn1344Hisfs*6), c.5720_5723del (p.Ser1907*) y c.6275_6276del (p.Leu2092Profs*7), se ha detectado en 2 ocasiones cada uno, suponiendo un 5% de la población portadora de una VP o VPP en el gen *BRCA2.*

Otras VP fundadoras detectadas en nuestro estudio en una única familia han sido: c.68_69del (p.Glu23Valfs*17), c.5117G>A (p.Gly1706Glu), c.9026_9030del (p.Tyr3009Serfs*7) y c.9310_9311del (p.Lys3104Valfs*6).

16.4 Variantes de nueva descripción

Durante el análisis de variantes se han encontrado 6 alteraciones que no se encuentran descritas en las bases de datos consultadas ni se han hallado publicadas anteriormente tras realizar una exhaustiva revisión bibliográfica (http://www.ncbi.nlm.nih.gov/entrez/PubMed) por lo que se las considera de nueva descripción a nivel mundial. A continuación, en la tabla 22 describimos las técnicas mediante las cuales se han detectado estas variantes.

Tabla 22. Técnica de detección de variantes de nueva descripción

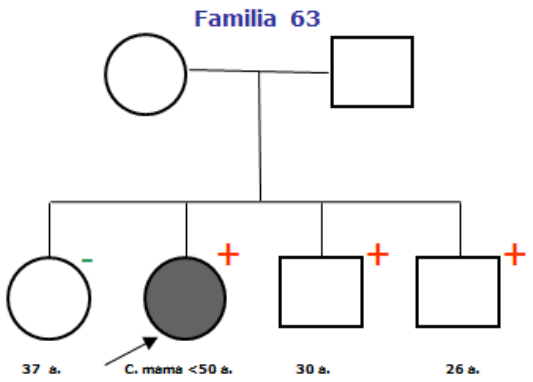
GEN	**VARIANTES**	**TÉCNICA DE DETECCIÓN**			
		Sanger	**NGS Pirosecuenciación**	***Análisis de fragmentos***	**MLPA**

BRCA2	Deleción parcial del exón 11	-	-	X	X
BRCA2	c.1792del (p.Thr598Hisfs*16)	X	X	X	-
BRCA2	c.4263dup (p.Glu1422*)	X	X	X	-
BRCA2	c.4677del (p.Phe1559Leufs*9)	X	X	-$^{\Omega}$	-
BRCA2	c.6034del (p.Ser2012Profs*28)	X	X	X	-
BRCA1	c.4106_4107delinsGCT (p.Ala1369Glyfs*5)	X$^{\Delta}$	X	X	-

- No detectada; X Detectada; $^{\Omega}$Todavía no estaba implementado el análisis de fragmentos; $^{\Delta}$Se identificó la variante por Sanger

Todas las variantes de nueva descripción fueron inserciones/deleciones y fuimos capaces de observar un patrón alterado mediante análisis de fragmentos, a excepción del c.4677del, ya que todavía no estaba implementado este paso en la rutina de laboratorio. Incluso fuimos capaces de detectar mediante esta metodología una gran deleción de parte del exón 11 en el gen *BRCA2*. Para identificar de manera adecuada las variantes complejas, es necesario la secuenciación Sanger, ya que mediante el empleo de las técnicas NGS no somos capaces de definir el cambio real presente en el probando.

En el gen *BRCA1* se ha descrito una única variante nueva que se corresponde a una variante compleja c.4106_4107delinsGCT (p.Ala1369Glyfs*5). Se detectó en una paciente de Andalucía con historia familiar de dos cánceres de mama, uno diagnosticado antes de los 50 años. Se llevaron a cabo tres estudios predictivos con dos resultados portadores de la variante en varones en edades tempranas. Figura 35.



c.4106_4107delinsGCT en el gen *BRCA1*

Figura 35. Árbol familiar de la familia 63 con una VPP no descrita c.4106_4107delinsGCT (p.Ala1369Glyfs*5) en el gen BRCA1. Se realizaron tres estudios predictivos en la familia.

En la siguiente figura se muestra la detección de esta variante compleja, fuimos capaces de detectar la variante mediante análisis de fragmentos y la secuenciación masiva, pero únicamente mediante la secuenciación Sanger logramos identificar el verdadero cambio complejo que presentaba nuestra probando.

Refer...	Variant	Max	Muestra 1	Muestra 2	Muestra 3	Muestra 4	Muestra 63	Muestra 6
BRCA1	33553:G/T	58,25	0,00	0,00	58,25	0,00	0,00	0,00
BRCA1	34311:A/C_c.4097-141A>C_SNP_rs799916	53,52	53,52	0,00	0,00	0,00	0,00	0,00
BRCA1	34460.5:-/G	48,63	0,00	0,00	0,00	0,00	48,63	0,00
BRCA1	34462:A/T	49,18	0,00	0,00	0,00	0,00	49,18	0,00

#HGNC_ID	Chro	Position	Ref_Allele	Var_Allele	Genotype	Depth	Ratio_Var/Depth
BRCA1	17	41243039	T	A	T/A	184	P_Hetero_(0.49)
BRCA1	17	41243041	-	C	-/C	184	UNC_FP_(0.20)

Figura 36. Detección del cambio complejo c.4106_4107delinsGCT (p.Ala1369Glyfs*5) en el gen BRCA1 mediante diferentes metodologías. 1) Identificación del cambio mediante análisis de fragmentos, con la flecha roja se señala la alteración del perfil (muestra morada) con respecto a una muestra control (color naranja); 2) Identificación del cambio con el *software AVA*, se observa un cambio puntual y una deleción en posiciones genómicas cercanas; 3) Identificación del cambio mediante el *pipeline* propio, se observa un cambio puntual y una deleción en posiciones genómicas cercanas; 4) Confirmación mediante secuenciación Sanger del cambio complejo, mediante esta tecnología se logra identificar el verdadero cambio que presenta la muestra.

Cinco variantes nuevas se localizan en el gen *BRCA2* y son todas de tipo *frameshift* a excepción de una gran deleción genómica descrita previamente en el apartado 4.2. Las puntuales VPP nuevas detectadas en el gen *BRCA2*, han sido:

- c.1792del (p.Thr598Hisfs*16), se detectó en un caso de Andalucía. La paciente presenta cáncer de mama bilateral. Esta variante fue detectada en nuestro estudio en mayo del año 2013, posteriormente al realizar una actualización de datos, hemos comprobado que ha sido descrita por Hernan et al en el año 2015[113].
- c.4263dup (p.Glu1422*), se detectó en otro caso de Andalucía. Se trata de una paciente de 80 años diagnosticada de cáncer de mama y cáncer de ovario. Su hija de 40 años también está diagnosticada de cáncer de mama. Esta variante fue identificada en nuestra

cohorte en el año 2012, posteriormente fue revisada por el consorcio ENIGMA en el año 2017 y fue clasificada como patogénica.

- c.4677del (p.Phe1559Leufs*9), se detectó en la comunidad Cántabra. La paciente presenta un tumor maligno de mama izquierda, con diagnóstico a los 61 años, cáncer basocelular, carcinoma ductal infiltrante de mama izquierda a los 65 años, melanoma en la espalda a los 68 años, y cáncer de trompa de Falopio izquierda. Antecedentes familiares: Madre de 85 años tratada de cáncer de mama. Hermana de 55 años tratada de cáncer de mama. Esta variante se identificó en nuestra cohorte en el año 2012, posteriormente en el año 2017 fue revisada por el consorcio ENIGMA y clasificada como VP.
- c.6034del (p.Ser2012Profs*28), se detectó en Cataluña. Se trata de una paciente diagnosticada de cáncer de mama a los 34 años. Antecedentes familiares: Dos casos en familiares de 1er grado. Tres o más casos en familiares de 2º grado. Se realizó un estudio predictivo en su hermano, y resultó ser portador de la variante. Esta variante fue detectada en nuestro estudio en mayo del año 2012, posteriormente en el año 2018 se identificó esta variante en una paciente con cáncer de mama en el trabajo de Pajares et al[123].

16.5 Variantes detectadas en base a su localización

Se han identificado VP y VPP a lo largo de los dos genes. La mayoría de las VP detectadas en ambos genes han sido identificadas en el exón 11, siendo el exón más largo en ambos genes. Hemos podido identificar que la mayoría de las variantes se encuentran descritas y correspondían a pequeñas inserciones o deleciones. Además, hemos identificado variantes nuevas en los dos genes, aunque la frecuencia sea baja.

En los siguientes gráficos se muestra la distribución de VP y VPP, identificadas en este estudio, según su localización en los genes *BRCA1/2* (Figuras 37 y 38).

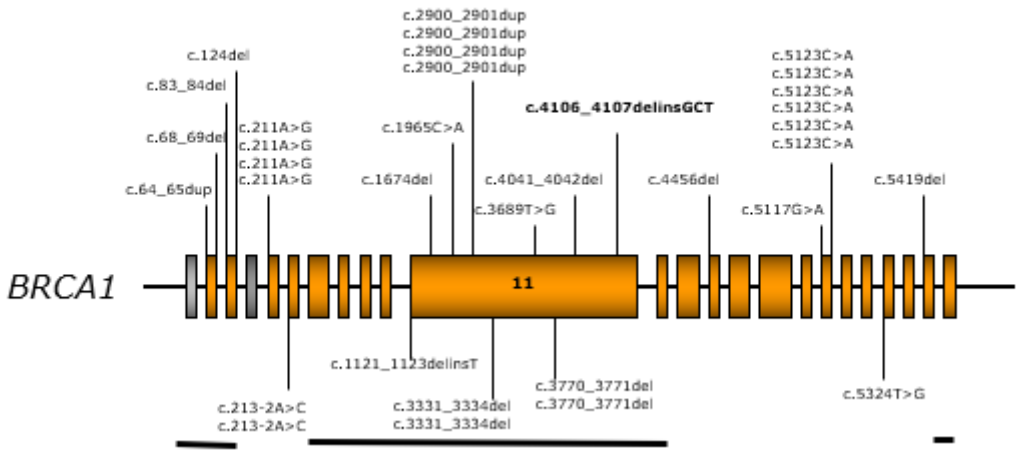


Figura 37. VP/VPP detectadas en el gen BRCA1. En negrita se resaltan las variantes de nueva descripción. Con una barra se representan las deleciones encontradas

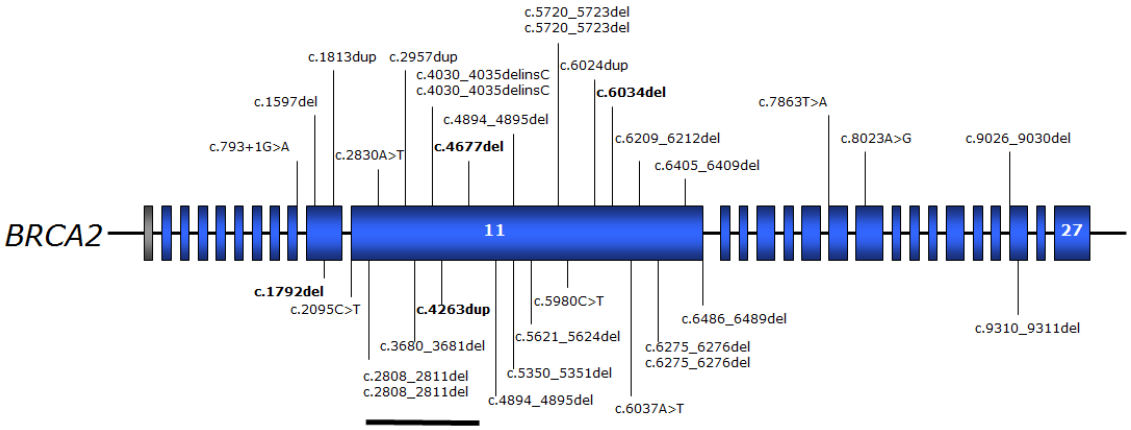


Figura 38. VP/VPP detectadas en el gen BRCA2. En negrita se resaltan las variantes de nueva descripción. Con una barra se representan las deleciones encontradas.

16.6 Variantes de significado incierto

En nuestra cohorte de estudio, se han detectado un total de 16 VSI diferentes revisadas durante este año 2023, y de ellas, 7 han correspondido al gen *BRCA1* y 9 lo han sido para el gen *BRCA2*. En total hemos detectado 16 variantes clasificadas como VSI en 18 casos índice, lo que representa un 3,95% del total de casos índice. Hay familias que presentan co-ocurrencia de VSI y VP. Por ejemplo, el caso índice 125, paciente con CO. Se detectó la VP c.4456del (p.Ser1486Valfs*19) en el gen *BRCA1,* junto con la variante nueva c.276T>G (p.Ala92Ala) en el mismo gen, clasificada como probablemente no patogénica, la variante reclasificada en el presente estudio como probablemente no patogénica c.68-7T>A en el gen *BRCA2* y la variante de nueva descripción, clasificada como VSI c.8331+38G>A en el gen *BRCA2.*

En los siguientes apartados se hace una descripción de las variantes detectadas en nuestra cohorte, descripción de las familias y predicción *in silico* de las variantes.

Variantes de significado incierto detectadas en el gen BRCA1.

Se han detectado 7 variantes distintas VSI el gen *BRCA1*, dos habían sido descritas previamente como variantes de efecto desconocidos y 6 variantes fueron de nueva descripción en el momento en el cual se detectaron.

Si nos referimos al tipo de cambio en la proteína: identificamos 6 cambios *missense* y una deleción de 12pb con una inserción de un ácido aspártico, que no cambia la pauta de lectura.

En la tabla 23, se describen en detalle las variantes en el gen *BRCA1* incluyendo los predictores *in silico* empleados para su valoración.

Tabla 23: Variantes VSI detectadas en el gen *BRCA1* y características de las mismas.												
Exón	**Cambio nucleotídico#**	**Tipo**	**Nº de casos índice**	**Nº Familia (características)**	**Predictores *in silico***	**DbSNP/ MAF**	**ClinVar**	**HGMD**	**LOVD[3]**	**BRCA Exchange**	**Criterios ACMG**	**Ref**
9	**c.656A>T (p.Asp219Val)**	M	1	30 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: Benign	NR	NR	NR	NR	NR	PM2	[107]
11b	**c.2533A>G (p.Ile845Val)**	M	1	75 (CM 30 años)	Align GVGD: C0 SIFT: Tolerated MutationTast: polymorphism PolyPhen-2: Benign	rs747649874 C=0,0004% (1/251090, GnomAD_exome) C=0,0008% (1/121336, ExAC)	1792726 Conflicto en la interpretación de la variante: VSI (2), VPB (1)	NR	BRCA1_002092 VSI	NR	PM2	NR
11b	c.2662C>T (p.His888Tyr)	M	2	53 (CM) 128 (CM 64 años)	Align GVGD: C0 SIFT: Tolerated MutationTaster :polymorphism PolyPhen-2: possibly damaging	rs80357480 A=0,0020% (5/251050, GnomAD_exome) A=0,0024% (3/125568, TOPMED) A=0,0041% (5/121374, ExAC)	54632 Conflicto en la interpretación de la variante: VB (2) VPB (2), VSI (9)	NR	BRCA1_001505 VSI	NR	PM2, BP6	[132]

11b	**c.2890G>A (p.Gly964Arg)**	M	1	9 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster: disease causing PolyPhen-2: possibly damaging	rs879254027 Sin frecuencias	245974 VSI (3), VPB (1)	NR	NR	NR	PM2	NR
13	**c.4304A>G** **(p.Asp1435Gly)**	M	2	176 (CM, madre y tía con CM) 190 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster :polymorphism PolyPhen-2: Benign	rs876660809 Sin frecuencias	234039 VSI (5)	NR	NR	NR	PM2	NR
15	**c.4553A>G** **(p.Gln1518Arg)**	M	1	182 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: Benign	NR	1802916 VSI (2)	NR	NR	NR	PM2	NR
24	**c.5547_5558del (p.Glu1849_Tyr1853delinsAsp)**	In	1	39 (CM 42 años)	-	NR	NR	NR	NR	NR	PM2	NR

#Se ha empleado la nomenclatura recomendada por la Human Genome Variation Society (HGVS v.2.0)

M: *Missense*; F:*Frameshift*; Sp:*Splicing*; N: *Nonsense*;: No referido; CM: cáncer de mama; CMB: Cáncer de mama bilateral; CO: Cáncer de ovario; Δ: Variantes descritas posteriormente. VSI: Variante de significado incierto.

Variantes de significado incierto detectadas en el gen BRCA2.

De las 9 VSI distintas detectadas en el gen *BRCA2*, dos fueron de nueva descripción en el momento de su detección, identificadas en negrita.

Si nos referimos al tipo de cambio en la proteína: 7 cambios *missense*, 1 variantes intrónicos y 1 cambio *nonsense*.

En la tabla 24, se describen las variantes VSI detectadas en el gen *BRCA2* y los predictores *in silico* empleados para evaluar el efecto de las variantes a nivel de la proteína y del *splicing.*

Tabla 24. Variantes VSI detectadas en el gen *BRCA2* y características de las mismas.												
Exón/ Intrón	**Cambio nucleotídico#**	**Tipo**	**Nº casos índice**	**Nº caso índice (características)**	**Predictores *in silico***	**DbSNP/ MAF**	**ClinVar**	**BRCA Exchange**	**HGMD**	**LOVD**	**Criterios ACMG**	**Referencia bibliográfica**
10	c.1714G>A (p.Val572Ile)	M	1	5 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: Benign	rs587782713 A=0,0008% (1/125568, TOPMED) A=0,0041% (5/120850, ExAC)	142782 VSI (7), VPB (2)	NR	NR	BRCA2_002659 VSI/VPB	PM2, BP4	[107]
11	**c.2401A>G (p.Asn801Asp)**	M	1	181 (CM varón a los 51 años)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: Benign	rs765279896 G=0,0008% (2/249710, GnomAD_exome) G=0,0008% (1/125568, TOPMED) G=0,0017% (2/120582, ExAC)	422118 VSI (5), VPB (1)	NR	NR	NR	PM2, BP4	NR
11	**c.4930G>C (p.Glu1644Gln)**	M	1	139 (CM 36 años)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: probably damaging	rs1555283989 Sin frecuencias	441348 VSI (3), VPB (1)	NR	NR	NR	PM2, BP4	NR
11	c.6357C>G (p.Asn2119Lys)	M	1	81 (CM)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism PolyPhen-2: Benign	rs1064794110 G=0,0004% (1/243084, GnomAD_exome)	419795 VSI (3), VPB (1)	NR	NR	NR	PM2, BP4	NR

16	c.7759C>T (p.Leu2587Phe)	M	1	159 (CM 49 años, triple negativo	Align GVGD: C15 SIFT: Deleterious MutationTaster: disease causing	rs56335340 T=0,0044% (11/251436, GnomAD_exome) T=0,0024% (3/125568, TOPMED) T=0,0041% (5/121352, ExAC)	141335 Conflicto en la interpretación de la variante: VPB (3),VB (1), VPP (1) VSI (7)	NR	CM154277 (DM ?)	BRCA2_001204 VSI	PM2, PP3, BP6	[133,134]
18	c.8182G>C (p.Val2728Leu)	M	1	163 (CM 43 años, familia de alto riesgo)	SIFT: Tolerated PolyPhen-2: Benign	rs28897749 A=0,2437% (306/125568, TOPMED) A=0,207% (65/31406, GnomAD)	52521 Conflicto en la interpretación de la variante: VPB (1) e VSI (2)	NR	NR	BRCA2_000296 VSI y VPB	PM2	NR
I18	c.8331+38G>A	I	1	125 (CO, co-ocurre con la VP c.4456del en *BRCA1*)	SpliceAI: Benign (0)	rs753751625 A=0,0009% (2/234106, GnomAD_exome) A=0,0019% (2/105586, ExAC)	NR	NR	NR	NR	PM2, BP4	NR
27	c.10222A>T (p.Lys3408*)	N	1	238 (CM 35 años)	MutationTaster: polymorphism	rs80358402 T=0,0036% (5/140276, GnomAD_exome) T=0,0045% (12/264690, TOPMED) T=0,0008% (1/120076, ExAC)	51056 Conflicto en la interpretación de la variante: VPB (5), VB (1), VSI (1)	NR	NR	NR	PM2, BP6	NR
27	c.10250A>G (p.Tyr3417Cys)	M	1	61 (CM 40 años)	Align GVGD: C0 SIFT: Tolerated MutationTaster: polymorphism	rs730881600 Sin frecuencias	182306 Conflicto en la interpretación de la	NR	NR	BRCA2_001344 VSI	PM2, BP6	NR

					PolyPhen-2: Benign		variante: VSI (4), VPB (2)					

#Se ha empleado la nomenclatura recomendada por la Human Genome Variation Society (HGVS v.2.0)

M:*Missense*; Sp:*Splicing*; N: Nonsense; NR: No referido; CM: cáncer de mama; CMB: Cáncer de mama bilateral; CO: Cáncer de ovario

17. Objetivo 4: Reclasificación de variantes. Estudio retrospectivo.

Este estudio comenzó en el año 2005, y desde ese año hasta la actualidad, la información sobre las variantes ha ido aumentando. Por este motivo se decidió llevar a cabo **dos re-evaluación de todas las VSI, durante el año 2020 y durante el año 2023**. Se ha realizado una revisión de las bases de datos de variantes, así como las bases de datos de frecuencias, que han ido aumentando la información, debido a que se han ido secuenciando más genomas, gracias a los consorcios de 1000 genomas, Exac y el proyecto ESP[135,136]. Adicionalmente, también se han revisado los resultados de nuevos estudios funcionales de las variantes, así como la información incluida en una base de datos de variantes propia donde, realizamos una consulta de la misma (revisando si estas variantes co-ocurren con VP). Gracias a estos nuevos datos aportados, se han podido reclasificar las variantes más antiguas. Se han reclasificado un total de 64 VSI, 22 variantes en el gen *BRCA1* detectadas en 41 casos índice y 42 variantes en el gen *BRCA2* detectadas en 75 casos índice. La mayoría de las VSI han pasado a ser VPB por lo que las VSI se han reducido en un 9% de los casos índices con respecto a la clasificación original (Figura 40).

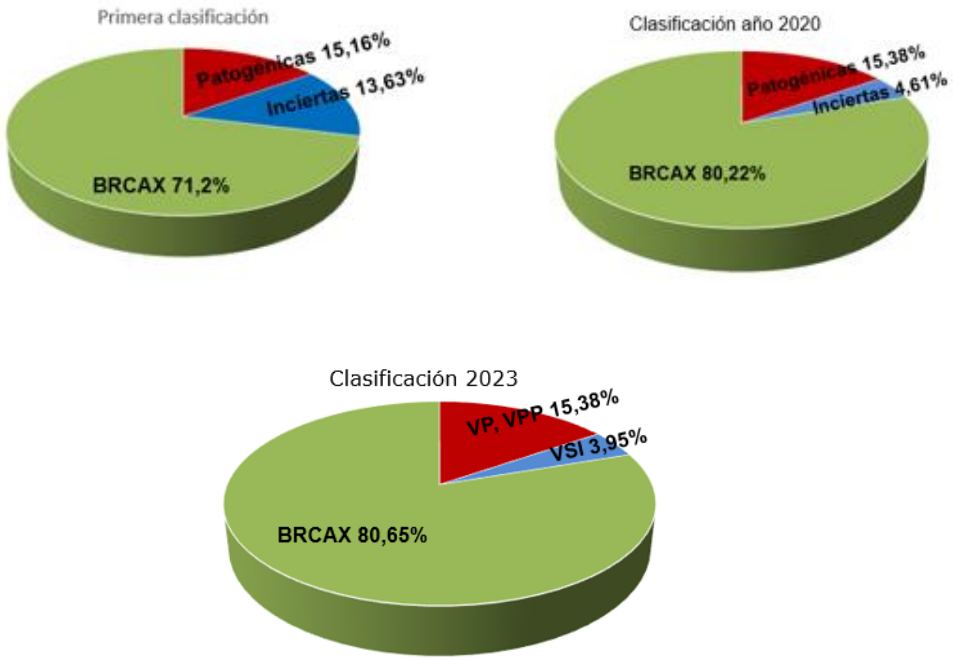


Figura 39. Reclasificación de variantes. Porcentaje de variantes antes y después de la reclasificación en el año 2020 y 2023, detectadas en nuestra serie de pacientes en los genes *BRCA1* y *BRCA2*. Se observa una disminución importante en el número de VSI, que pasan mayoritariamente a VPB.

La mayoría de las variantes reclasificadas han pasado a ser VPB (22 variantes) o VB (42 variantes), una variante ha pasado de ser VSI a ser VP.

Las variantes reclasificadas de VSI a VPB, todas ellas se caracterizan por ser variantes tipo *missense*, variantes sinónimas y variantes en zonas intrónicas. Las variantes han sido

reclasificadas en base a su frecuencia, estudios funcionales publicados, estudios de *splicing*, predictores *in-silico* y co-ocurrencia con VP en el mismo gen.

Tenemos que destacar que una **variante clasificada en primer lugar como VSI, se reclasificó como VP después de la revisión**:

Se trata de la variante **c.8023A>G (p.Ile2675Val)** en el exón 18 del gen *BRCA2*. Este cambio supone un cambio de aminoácido de isoleucina a valina en la posición 2675 de la proteína. Esta variante fue detectada en el año 2009, empleando la metodología Sanger, en este año, la información disponible sobre esta variante era escasa, ya que no se encontraba descrita en ninguna base de datos consultada y no se disponía de la frecuencia alélicas, por este motivo, la clasificación asignada fue de VSI.

Con los datos actuales, sabemos que esta posición se encuentra altamente conservada entre especies. Hay unanimidad en todas las bases de datos consultadas (LOVD, ClinVar, HGMD) clasificando esta variante como patogénica. En la base de datos HGMD (número de entrada CS083875), en el artículo de Houdayer *et al*[108] clasifican esta variante como 3S (impacto severo sobre el *splicing*), y observan una alteración del sitio de *splicing*, generándose otro sitio alternativo a 309bp *upstream* del nativo. En la base de datos BRCA Exchange fue clasificada como patogénica en abril del año 2020, Además, esta variante ha sido descrita en la base de datos dbSNP rs397507954, sin frecuencias descritas en población control. Los algoritmos de predicción *in silico* de la variante (p.Ile2675Val) asignan a este cambio un efecto probablemente deletéreo (SIFT, Align GVGD, Polyphen-2 y Mutation Taster). Con todos estos datos, hemos reclasificado esta variante como patogénica.

En el anexo V, se detallan las variantes que hemos reclasificado, describiendo las bases de datos revisadas en la actualidad y los nuevos trabajos publicados sobre esas variantes, en la tabla 25, se muestra una única variante que pasa de VSI a VP.

Tabla 25. Variante reclasificada de VSI a VP detectadas en nuestra serie de pacientes en los genes *BRCA1/2*.										
Cambio nucleotídico	**Variante informada (año)**	**dbSNP** **MAF**	**ClinVar**	**LOVD**	**HGMD**	**BRCA** **Exchange**	**Observaciones** **(Caso índice)**	**Reclasificación 2015**	**Criterios ACMG**	**Reclasificación 2020**
BRCA2: **c.8023A>G (p.Ile2675Val)**	**VSI** **(2009)**	**rs397507954** **G=0,0004% (1/251076, GnomAD_exome)**	**52475** **Patogénica**	**BRCA2_00251** **Patogénica**	**CM152116** **DM**	**Patogénica**	**Se reportó en un caso índice**	**Patogénica**	**PP5, PM2, PP3, PS3**	**Patogénica**

18. Correlación genotipo-fenotipo.

Para determinar las características de nuestra serie de pacientes respecto de las VP o VPP detectadas, hemos evaluado la relación existente de los distintos genotipos respecto al criterio de selección de las familias del estudio, el tipo de tumor detectado en el caso índice, historia familiar y la edad de diagnóstico.

Para determinar la eficacia diagnóstica de los criterios de selección empleados en este trabajo a la hora de detectar portadores de variantes (VP y VPP) se ha evaluado la relación entre el fenotipo familiar y la tasa de detección de variantes en base a cada gen (Tabla 26).

La edad media de la serie de pacientes portadores de VP es de 41 años.

Tabla 26. Relación fenotipo-genotipo según los criterios clínicos

Fenotipo familiar	**Edad media casos +**	**Tamaño muestral (N)**	***BRCA1* (%)**	***BRCA2* (%)**	***BRCA1/2* (%)**
1: Un caso CM ≤50 años CM bilateral CM+CO CO	35 años	101	6 (5,94%)	6 (5,94%)	**12 (11,88%)**
2: Dos casos CM varón CO CM ≤50 años CM bilateral + CM ≤50 años	45 años	114	6 (5,26%)	8 (7,01%)	**14 (12,28%)**
3: Tres casos o más a cualquier edad	42 años	109	16 (14,67%)	13 (11,92%)	**29 (26,6%)**
NR: No referido	-	119	9 (7,56%)	6 (5,04%)	**15 (12,6%)**
NC: No cumple criterios	-	12	0	0	**0**
Totales		**455**	**37 (8,13%)**	**33 (7,25%)**	**70 (15,38%)**

N= número de casos estudiados; CM: cáncer de mama; CO: cáncer de ovario; casos+: casos positivos

Es de destacar el **mayor porcentaje de casos positivos, cuando tenemos un mayor número de casos familiares afectos,** como observamos con el criterio de 3 o más casos diagnosticados a cualquier edad, con un **diagnóstico del 26,6%.** No se observan diferencias significativas entre los casos únicos y los casos con dos familiares afectos, ya que los

porcentajes de diagnóstico son muy similares en nuestra población. No se han observado tampoco diferencias significativas entre los dos genes testados en las distintas categorías.

Si observamos los resultados, fijándonos únicamente en los casos clasificados siguiendo los criterios 1, 2 y 3, tenemos un total de 324 casos y 55 VP, teniendo en esta serie de casos un 16,97% de pacientes genéticamente diagnosticados, mejoramos el porcentaje de diagnóstico global en casi dos puntos.

19. Correlación genotipo-tipo de cáncer del caso índice

Observando los resultados obtenidos según la prevalencia de los tipos de cáncer y según el resultado del test genético en el global de los casos, en los casos de cáncer de ovario, de cáncer de mama y ovario, y en los casos triple negativos obtenemos una mayor frecuencia de test de los genes *BRCA1/2* positivos (Figura 40).

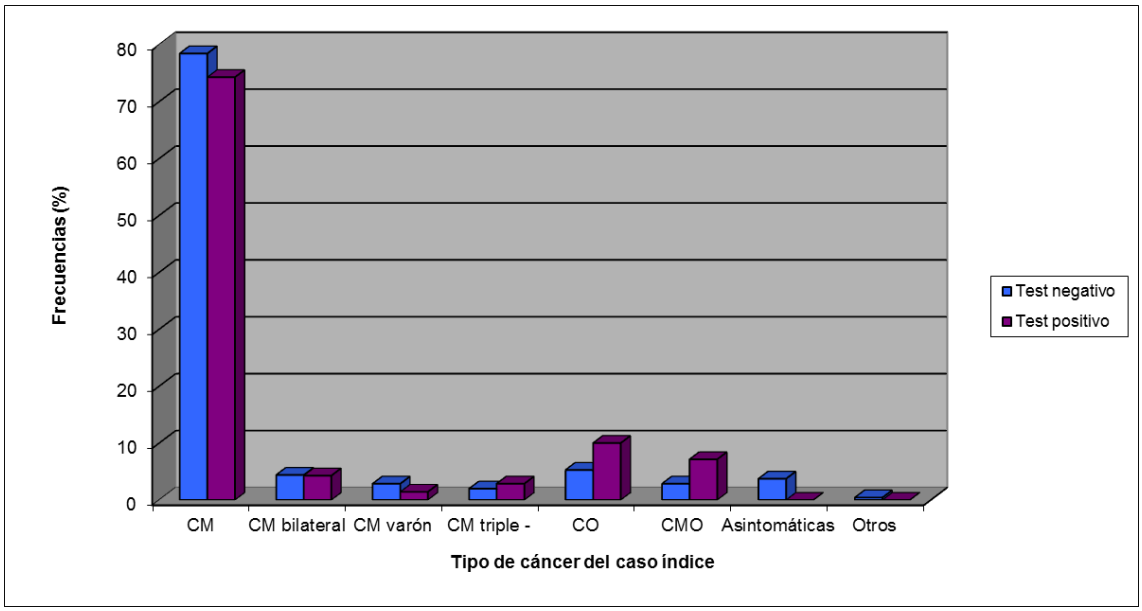


Figura 40. Diagrama de barras de los tipos de cáncer de los casos índice, en función del resultado del test genético. CM: cáncer de mama, CM triple -: cáncer de mama triple negativo, CO: cáncer de ovario, CMO: cáncer de mama y ovario, Asintomáticas: caso índice asintomático, Otros: otros tipos de cáncer

En tabla siguiente, tenemos un resumen, con la descripción de los casos atendiendo solo al tipo de cáncer del caso índice, independientemente de la historia familiar y la edad de diagnóstico.

Tabla 27. Porcentajes de VP y VPP identificadas en nuestra serie de pacientes en los genes *BRCA1/2*, según el fenotipo del caso índice.

	CM	**CM bilateral**	**CO**	**CMO**	**CM TN**	**CM varón**

	N=357	N=21	N=23	N=13	N=9	N=14
BRCA1	27 (7,56%)	1 (4,76%)	6 (26,08%)	2 (15,38%)	2 (22,22%)	0
BRCA2	25 (7%)	2 (9,52%)	1 (4,34%)	3 (23,07%)	0	1 (7,14%)
BRCA1/2	52 (14,56%)	3 (14,28%)	7 (30,43%)	5 (38,46%)	2 (22,22%)	1 (7,14%)
Edad media diagnóstico casos+	44 años	44 años	34 años	63,66 años	40,5 años	-
Edad media diagnóstico casos -	43 años	44,9 años	41,16 años	58,8 años	43,8 años	60,3 años

CM: cáncer de mama; CO: cáncer de ovario; CMO: cáncer de mama y ovario; TN: Triple negativo; N= nº de casos estudiados

19.1 Cáncer de mama

De los 357 casos de cáncer de mama, independientemente del número de familiares afectos, hemos detectado 52 VP, un **14,56%** de pacientes genéticamente diagnosticados. De estas variantes, 27 se localizaban en el gen *BRCA1*, esto supone un 7,56% de pacientes diagnosticados genéticamente y 25 en el gen *BRCA2,* lo que supone un 7% de pacientes.

Por otro lado, **si tenemos en cuenta los antecedentes familiares**, en función del número de afectos en la familia, obtenemos un **25,27%** de casos positivos en familias con más de 3 miembros familiares directos con cáncer de mama, frente a un 8,86% de casos positivos en familias con al menos de 2 casos cáncer de mama, y al 12,93% de familias con un único caso de cáncer de mama (Tabla 28).

Tabla 28. Porcentajes de VP en pacientes con cáncer de mama N=357

	NR N=116	**≥3 casos CM/CO N=91**	**<2 casos CM/CO N=79**	**CM único N=65**	**CM NC N=6**
BRCA1	9 (7,75%)	13 (14,28%)	1 (1,26%)	4 (6,15%)	0
BRCA2	6 (5,17%)	10 (10,98%)	6 (7,59%)	3 (4,61%)	0
BRCA1/2	**15 (12,93%)**	**23 (25,27%)**	**7 (8,86%)**	**7 (10,76%)**	0
Edad media diagnóstico casos +	-	43 años	44 años	33 años	-

Edad media diagnóstico casos -	-	47 años	42 años	41 años	63 años

CM: cáncer de mama; CO: cáncer de ovário; N= nº de casos estudiados; NC: no cumple critérios

Obtenemos una mayor incidencia de casos portadores de VP en familias con más de 3 miembros familiares directos con cáncer de mama, 25,27% frente a un 8,86% de casos positivos en familias con al menos de 2 casos cáncer de mama, y familias con un único caso de cáncer de mama, un 10,76%. La edad media de diagnóstico de familias con un solo caso, en casos índice portadores de VP es de 33 años, seguido de 44 años para los casos con 2 miembros de cáncer de mama en la familia y finalmente, de 43 años, en los casos de tres o más familiares afectos de cáncer de mama/cáncer de ovario.

Los casos índices con resultados negativos para *BRCA1/2* presentan una media de edad de diagnóstico superior a los casos índice con resultados positivos. Se muestra mayor diferencia en los casos únicos.

Clasificamos 6 casos de cáncer de mama únicos con una edad de diagnóstico superior a 50 años, así que no cumplían con nuestros criterios.

19.2 Cáncer de mama bilateral

En nuestra serie hemos encontrado 21 casos de cáncer de mama bilaterales, con un porcentaje diagnóstico del 14,28%, de ellos 2 presentan VP en el gen *BRCA2*, y una en el gen *BRCA1*. La edad media de diagnóstico en la serie de casos negativos para *BRCA1/2* ha sido de 44,9 años, y la edad media de casos con VP ha sido de 44 años, por lo que no se observan grandes diferencias. Se ha detectado una variante de nueva descripción c.1792del (p.Thr598Hisfs*16) en el gen *BRCA2*, en el momento del diagnóstico genético.

19.3 Cáncer de mama en varones

En nuestra serie hubo 14 varones con cáncer de mama, entre ellos uno con cáncer de mama y leucemia. Tenemos una tasa de VP de un 7,14% (1/14). La VP detectada ha sido una variante *frameshift* en el gen *BRCA2*, en un paciente cuya madre presentaba cáncer de mama. Además, hemos encontrado un paciente con una variante de nueva descripción tipo *missense* clasificada como VSI. La edad media de diagnóstico fue de 60,3 años.

19.4 Cáncer de ovario

En total hemos detectado un 30,43% (7/23) de casos genéticamente diagnosticados entre las pacientes con cáncer de ovario estudiadas, 6 VP en el gen *BRCA1* (6/23), y una VP en *BRCA2* (1/23). La edad media de los casos con VP ha sido de 34 años.

La edad media de diagnóstico del cáncer de ovario ha sido de 41,16 años. Entre los cánceres de ovario de nuestra serie, hubo un caso con cáncer de ovario y cáncer de colon, y un caso de cáncer de ovario bilateral.

19.5 Cáncer de mama y ovario

El total de casos estudiados ha sido 13, dos de ellos con cáncer de mama bilateral y cáncer de ovario, y la edad media de diagnóstico ha sido de 58,8 años. En esta serie tenemos un 38,46% de casos genéticamente diagnosticados y una edad de diagnóstico de 63,66 años. En nuestra serie de casos hemos detectado cinco VP, dos variantes en el gen *BRCA1*, (deleción del exón 24, y una variante *frameshift)*, y tres variantes en el *BRCA2*, (dos variantes son de nueva descripción y están clasificadas como VPP, debido a su efecto deletéreo sobre la proteína, y una VP *frameshift).*

19.6 Cáncer de mama triple negativo

En nuestra serie de casos hubo un total de 9 casos triples negativos, en los cuales hemos detectado dos VP en el gen *BRCA1*, representando un 22,22%. La edad media de diagnóstico ha sido de 42,85 años, y en los casos con VP ha sido de 40,5 años.

DISCUSIÓN

Ser portador de una VP en los genes *BRCA1/2* está asociado con un incremento del riesgo de desarrollar cáncer de mama, cáncer de ovario y otros tipos de cánceres[32,43], por lo tanto, una temprana identificación de portadores de VP en los genes *BRCA1/2* es importante ya que se pueden beneficiar de un adecuado consejo genético, cribado y de determinadas estrategias preventivas y terapias específicas. Se puede hacer estudios de portadores en familiares directos, y de esta manera, los portadores de la VP familiar se podrán beneficiar de las medidas de seguimiento intensivo y de aquellas encaminadas a la reducción del riesgo; y los familiares que no son portadores, se librarán de la ansiedad y el miedo a la enfermedad que suele ser común en una familia con elevada carga de tumores.

Las **principales ventajas de la tecnología de la NGS** que han revolucionado el campo del diagnóstico genético han sido la rapidez, el bajo coste y la precisión. Durante los últimos 30 años la secuenciación Sanger ha sido la técnica *gold standard* para la identificación de variantes clínicamente relevantes en el diagnóstico rutinario de los genes *BRCA1*/2. Sin embargo, el gran tamaño de estos genes, la laboriosidad del método de secuenciación y el alto coste, hacen complicada su aplicación cuando el número de muestras a procesar es elevado. En la práctica habitual, el estudio completo de los genes *BRCA1/2*, supone un alto coste económico si se realiza mediante secuenciación Sanger, esto hace que el estudio de los genes *BRCA1/2*, se realice, en determinadas ocasiones, mediante un cribado de variantes, en lugar de realizar el estudio completo ya que, por lo general, el estudio completo implica mucho tiempo y es laborioso, dificultando de esta manera, el diagnóstico clínico. Todos estos hechos, junto con la baja prevalencia de las variantes en las poblaciones de riesgo, (aproximadamente un 16%[137]), justifican el abordaje diagnóstico de estos genes mediante nuevas técnicas, como la secuenciación masiva, que reducen el trabajo y el tiempo de análisis. Estas técnicas son robustas, ofrecen un mayor rendimiento, un menor coste por nucleótido secuenciado y un menor tiempo en la obtención de resultados[138–140].

Por todos estos motivos, poco a poco la tecnología Sanger va siendo sustituida gradualmente por la tecnología NGS.

Además, en estos últimos años, empleamos el test genético, no solo para un diagnóstico precoz en familiares, sino también, para una medicina personalizada, con independencia de la historia familiar y la herencia. El análisis genético de los genes *BRCA1/2* está en aumento, debido a que la Agencia Europea del Medicamento aprobó en el año 2014 el uso de los PARPi, en el caso de pacientes con cáncer de ovario de alto grado, ante la identificación de VP en estos genes, tanto en línea germinal como en línea somática[67]. En este contexto, se necesita testar rápidamente a las pacientes para poder empezar a administrar esta terapia dirigida en estadios iniciales. En esta misma línea, el Grupo Español de Investigación en Cáncer de Ovario (GEICO) y la Sociedad Española de Oncología Médica (SEOM), han implementado unas recomendaciones para el estudio genéticos de los genes *BRCA1/2* en el cáncer de ovario. Indicando que la identificación de **variantes germinales y somáticas en los genes**

***BRCA1/2* en pacientes con cáncer de ovario tiene valor pronóstico y predictivo**, por lo que se debe de tener en cuenta a la hora de planificar el tratamiento e informar a la paciente. El tiempo recomendado para la obtención de los resultados genéticos de 4 semanas[141]. En el año 2018 se aprobó también el uso de PARPi (Olaparib) para mujeres portadoras de VP a nivel germinal en los genes *BRCA1/2* con diagnóstico de cáncer de mama metastásico HER2 negativo que han sido tratadas previamente con quimioterapia[142]. Más recientemente se ha aprobado también su empleo en cáncer de próstata metastásico resistente a la castración con VP en los genes *BRCA1/2* a nivel germinal o tumoral, así como en cáncer de páncreas metastásico a nivel germinal[32].

La secuenciación masiva está presentando en los últimos años, una bajada importante de los costes y los tiempos si la comparamos con la secuenciación Sanger.

El objetivo de esta tesis ha sido proponer el estudio completo, de manera rutinaria, de los genes *BRCA1/2*, tanto región codificante como regiones intrónicas adyacentes, basados en el empleo de la PCR multiplex (*Multiplicom*), realizada para generar las librerías de DNA de los pacientes, seguido de dos estrategias de secuenciación masiva, como han sido:

1) pirosecuenciación seguido del análisis bioinformático mediante el *software AVA*
2) terminación cíclica reversible, seguida del análisis bioinformático mediante el *software* de desarrollado propio.

Estas **nuevas tecnologías** también tenían una serie de limitaciones, como, por ejemplo, **ausencia de protocolos de procesamiento estandarizados** (esto es debido en parte, a la continua evolución de las químicas de secuenciación). Sabemos que es fundamental una buena cobertura, para que no haya problemas en la detección de variantes y fiabilidad de los resultados[143]. Actualmente ya existe consenso en cuanto a la cobertura tanto a nivel germinal como somático, ya que recientemente se publicaron unas guías internacionales[32] en donde se recomienda una cobertura horizontal de al menos un 100% de la secuencia exónica e intrónica adyacente al *splicing*; para la detección de variantes puntuales a nivel germinal se recomienda una profundidad de lectura de al menos un 30x y para la detección de grandes deleciones/duplicaciones una cobertura de al menos 200x, en cambio, para la identificación de variantes puntuales en tejido tumoral se recomienda una profundidad de lectura de al menos 300x. Se recomienda utilizar la nomenclatura para la designación de variantes de la Human Genome Variation Society (HGVS)[100], así como una categorización de variantes a nivel germinal en 5 clases, atendiendo las recomendaciones de la ACMG[52]. Actualmente se están discutiendo unas nuevas guías de clasificación de variantes para los genes *BRCA1/2*, entre miembros del consorcio internacional ENIGMA y la base de datos ClinGen basadas en las normas ACMG/AMP.

Otro punto importante, es el **manejo y el análisis de los datos obtenidos** (alineación, ensamblaje y anotaciones de variantes). La interpretación y el almacenamiento de la gran cantidad de datos que se generan en cada carrera, pueden suponer un problema para los centros hospitalarios especializados[81]. Otro inconveniente, es la ausencia de programas y

algoritmos universales de análisis de datos. Muchos investigadores utilizan programas libres, a partir de los cuales optimizar su propio proceso de análisis, para así conseguir unos resultados robustos y reproducibles[144]. Algunos de estos inconvenientes tienen una fácil solución, como el desarrollo de algoritmos bioinformáticos sencillos, potentes y eficaces para el análisis de los datos, estandarización de los protocolos experimentales, fijación de parámetros de calidad, que aseguren unos datos sólidos, reproducibles y fiables. Algunas casas comerciales, venden junto a los pequeños paneles, programas que facilitan la interpretación de los datos generados, aunque en la mayoría de los casos, se necesita el apoyo de un bioinformático para que los resultados puedan ser adecuadamente y rápidamente interpretados.

En esta tesis **se evaluaron dos métodos para la secuenciación de los genes *BRCA1/2,* pirosecuenciación (*454 GS Junior*) y terminación cíclica reversible (*MiSeq*),** mediante el uso de muestras control previamente testadas por metodología Sanger para los casos del secuenciador *454 GS Junior* y previamente testadas con *454 GS Junior* y confirmación por Sanger de las variantes de relevancia clínica, para los casos secuenciados con *MiSeq*. En ambos casos, se utilizó un ensayo PCR multiplex comercial (*Multiplicom*) ya que ofrece una mejor reproducibilidad y la configuración es más sencilla que los métodos propios.

Para el análisis de las inserciones y duplicaciones en regiones de homopolímero se ha implementado el análisis de fragmentos, es decir, a los resultados del análisis bioinformático se le acompañó el análisis manual de los perfiles de los fragmentos procedentes de las reacciones de PCR multiplex, realizados en el laboratorio. Con esto, conseguimos reducir el número de confirmaciones mediante secuenciación Sanger, ya que, la llamada de variantes *indels* producen falsos positivos[140]. Nuestros resultados nos indican que empleando este algoritmo diagnóstico alcanzamos unos valores de especificidad y sensibilidad adecuadas. Si comparamos la pirosecuenciación con otras tecnologías NGS, podemos decir que obtenemos las secuencias más largas, en cambio, como principal desventaja es la imprecisión en la determinación de los tamaños en los homopolímeros[145].

En el caso de la tecnología *MiSeq*, no necesitamos el análisis de fragmentos para la detección de deleciones e inserciones, aunque se siguió utilizando esta metodología, por el acortamiento del tiempo en la obtención de resultados positivos cuando se detectaba un patrón anómalo y por la sospecha de la presencia de CNVs.

En este estudio, se ha descrito una estrategia de trabajo optimizada y validada para maximizar el rendimiento analítico empleando NGS, y garantizando una interpretación de alta calidad, para facilitar la toma de decisiones clínicas. Se ha descrito la estrategia seguida y la validación analítica empleando NGS y MLPA/fragmentos para detectar VP en los genes *BRCA1/2*. Hemos sido capaces de detectar SNVs, *indels* y CNVs sin errores en el set de validación con muestras control, y hemos identificado variantes nuevas en el set de muestras de diagnóstico empleadas en este estudio.

20. Aplicación de un flujograma de trabajo basado en la secuenciación masiva al estudio de variantes en los genes *BRCA1* y *BRCA2*.

Se ha presentado un flujo de trabajo completo para el análisis de los genes *BRCA1/2*, basado en el uso de la estrategia de PCR multiplex (*Multiplicom*) para generar bibliotecas de ADN del paciente, seguido por pirosecuenciación, utilizando la plataforma *454 GS Junior* (Roche) y el análisis bioinformático basado en una complementación de dos *softwares* (*AVA* y un *pipeline* de desarrollo propio). Se han implementado unos filtros adecuados, para que se genere la lista de variantes, que requiere un análisis visual y una confirmación mediante secuenciación Sanger de las variantes clínicamente relevantes.

El análisis de las *indels* se realizó mediante un análisis de fragmentos de las PCR multiplex. Utilizando este flujo de trabajo logramos un 100% de sensibilidad en el diagnóstico (página 67).

En general, el ensayo permite la generación de datos robustos para los pacientes en estudio, y determinamos un número de muestras medio por carrera 6 a 8, para obtener los parámetros óptimos para el análisis. Estos resultados son similares a los encontrados en el trabajo de Feliubadaló et al (2013), en el que encuentran un tamaño muestral óptimo de 5 a 7 muestrasen su estudio[140]. Este tamaño muestral es adecuado para un laboratorio de tamaño medio, donde no es necesario esperar a un gran número de muestras para comenzar la carrera, ya que un dispositivo grande, aumentaría las listas de espera y como consecuencia, los tiempos de respuesta en el diagnóstico genético.

Entre las **ventajas del uso de esta tecnología** están el precio bajo (tabla 6 del apartado de resultados), y lecturas más largas en comparación con otras grandes plataformas NGS que ofrece ventajas en el alineamiento de posibles *indels*. En este estudio, la deleción más larga detectada ha sido de 11pb. La principal desventaja con respecto a otras tecnologías NGS, por ejemplo, el de terminación reversible, es la dificultad de establecer el número de nucleótidos idénticos correcto en zonas de homopolímero, generando una alta tasa de error en estas zonas, siendo más acusada cuanto mayor sea la longitud del homopolímero[102]. Para evitar este error, como se ha detallado anteriormente, se empleó el análisis de fragmentos.

El **análisis de fragmentos** permite, en el paso previo a la secuenciación, analizar los amplicones de todas las PCRs multiplex de los genes *BRCA1/2*, y permite evaluar la presencia de una posible *indel* cuando se observa un patrón electroforético alterado. Esto es debido a que este sistema nos permite:

- Detectar todas las *indels* presentes en la muestra, no solo las localizadas en la zona de homopolímeros, al analizar todos los fragmentos que se van a secuenciar.

- Obtener un menor tiempo para la detección de una *indel*. La amplificación de fragmentos y análisis por *Genemapper* es un paso previo a la secuenciación, por lo que, los fragmentos que se observan alterados, se pueden confirmar mediante

secuenciación Sanger o bien mediante MLPA, con anterioridad al análisis de las secuencias obtenidas mediante la pirosecuenciación.

Por lo tanto, si comparamos nuestros resultados con los resultados de la validación de los trabajos de Feliubadaló et al (2013) y Ruiz et al (2014), obtenemos resultados muy similares en cuanto a especificidad (100%) y sensibilidad (>99,99%), aunque ellos emplean para la detección de inserciones y deleciones el método BRCA HP (*Multiplicom*)[139,140].

Adicionalmente hemos detectado que existen inconvenientes derivados del empleo de los perfiles de intensidad. En este sentido, debemos decir, que el análisis es laborioso, se aumenta el tiempo que hay que dedicar al análisis de la muestra, y hay que tener cuidado con las *indels* polimórficas de las zonas intrónicas, que pueden inducir confusión. Un ejemplo es el *rs138193280*, descrito en la figura 21 de los resultados, es una deleción polimórfica en el intrón 11 del gen *BRCA2.*

La **presencia de CNVs** también fue investigada en este trabajo, mediante un método basado en la **normalización de la profundidad de lectura media por amplicón**, entre las muestras pertenecientes a la misma carrera. Existe mucho ruido de fondo tras la normalización, y aunque las deleciones presentes en la muestra control se pueden intuir, es necesario, la confirmación por otra técnica alternativa, por ejemplo, en nuestro caso el MLPA. Es de esperar, que, en un futuro cercano, con la mejora de las metodologías de secuenciación y de procesamiento de datos se pueda llegar a una mayor especificidad e identificar de una manera fiable los CNVs en un entorno diagnóstico.

Los resultados obtenidos mediante **la técnica de pirosecuenciación**, en comparación de la **técnica de secuenciación Sanger**, permite establecer los siguientes puntos:

- La pirosecuenciación es un método rápido, más barato y precisión bastante alta.

- La profundidad de lectura recomendada para alcanzar el 100% de sensibilidad es de aproximadamente 50x, si se toma una frecuencia alélica de al menos 0,2[145].

- Como desventaja, destacar que no se pueden generar secuencias tan largas como en la secuenciación Sanger (menores de 1kb), presenta problemas con los homopolímeros, observándose falsos positivos.

En el presente trabajo se ha puesto de manifiesto la necesidad de confirmación mediante otras técnicas diagnósticas, el ejemplo más claro, lo tenemos en la descripción de las variantes de la línea *HapMap* 12144 mediante genotipado de *arrays*, la confirmación mediante secuenciación Sanger del SNP *rs1801406* (no descrito inicialmente para esta línea) y la no confirmación del *rs4987046.*

En el entorno diagnóstico, es necesario estudiar los genes con una buena cobertura para toda la región codificante y región intrónica adyacente, con una sensibilidad igual o mayor que la secuenciación Sanger. Nuestro algoritmo de trabajo utilizado para los genes *BRCA1/2* cumple con los requisitos de sensibilidad y especificidad para el diagnóstico genético de CMOH, por lo que es factible y rentable en comparación con los estándares actuales.

En cuanto a la comparación de los resultados del análisis de datos de secuenciación obtenidos mediante el *pipeline* de desarrollo propio y el soporte informático *AVA,* podemos extraer las siguientes ventajas e inconvenientes:

- Ventajas. El *pipeline* propio, ofrece información más completa de las variantes detectadas ya que: 1) identifica si la variante es de nueva descripción, o si ya está descrita en las bases de datos (HGMD)[146], así como el número *rs* asociado a cada variante (Número de ID de la base de datos de SNPs: dbSNP[147], 2) localiza la posición en el ADN, en el cDNA y en el aminoácido; 3) nos indica el tipo de variante: SNV, *indel*; 4) nos indica la consecuencia en el transcrito: *frameshift*, *splicing*, *missense*, intrónica, 5´o 3´UTR; 5) se visualiza de forma sencilla el *depth* asociado a cada variante; 6) describe el efecto funcional de la variante a nivel de la proteína evaluando varios algoritmos *in silico*, lo cual es interesante a la hora de interpretar su posible efecto clínico e 7) incluye una columna de datos que refleja la frecuencia descrita de cada variante a nivel de poblaciones.
- Inconvenientes: En cuanto a los inconvenientes, es destacable el mayor ruido de fondo, es decir, detecta bastantes variantes (*indels*) que son meros artefactos de la tecnología (zonas de homopolímeros) (>78%), hay que validar muchos más cambios por secuenciación Sanger para descartar los falsos positivos y desde el punto de vista del análisis, es menos manejable ("visualmente") que el *AVA* de *Roche* ya que el manejo es sobre tablas Excel.

Como conclusión, desde un punto de vista analítico, **ambos *softwares* son complementarios para realizar el análisis de los *BRCA1/2*.**

También hemos presentado un **flujo de trabajo completo para el análisis de los genes *BRCA1/2*,** basado en el uso de la estrategia de PCR multiplex (*Multiplicom*) para generar bibliotecas de ADN del paciente, seguido por secuenciación por terminador reversible cíclico, utilizando la **plataforma *MiSeq*** y el análisis bioinformático basado en un *software* propio. El análisis de las *indels*, aunque no es necesario, se siguió utilizando, mediante un análisis de fragmentos de las PCR multiplex. Utilizando este flujo de trabajo, logramos una sensibilidad óptima para el diagnóstico.

Aunque todavía se necesita mejoras en el análisis, los resultados de la presencia de CNVs abren una puerta a la identificación de grandes deleciones/duplicaciones mediante esta tecnología, ya que, en ocasiones, cuando se observan alteraciones en la cobertura de los fragmentos, permiten identificar la presencia de una alteración.

En general el ensayo permite la generación de una biblioteca robusta para los pacientes en estudio, y establecimos un número de muestras medio por carrera de 8 a 12 muestras. Este tamaño muestral es adecuado para un laboratorio de tamaño medio.

La comparación de los resultados obtenidos mediante **la técnica de terminación cíclica reversible con respecto a los resultados detectados mediante la tecnología de pirosecuenciación**, permite destacar que:

- Con la pirosecuenciación aparecen una media de aproximadamente 26 variantes/muestras mientras que con la terminación reversible una media de 20 variantes/muestras. Esto se debe al cambio en la química de secuenciación, con la secuenciación de terminación reversible se eliminan los falsos positivos, que se observaban en las zonas de homopolímero, introducidos por la pirosecuenciación del *454 GS Junior* (salvo excepciones). El cambio de tecnología para la secuenciación, presenta una mejora considerable.

- El Var/*Depth* más bajo detectado para una variante en heterocigosis para un SNV conocido es de 0,29. Por este motivo, se usó como frecuencia mínima un 0,20 para la detección de una variante en heterocigosis.

Como conclusiones a destacar, debemos decir: 1) Se detectaron todas las variantes puntuales e *indels* en las muestras control, previamente detectadas mediante pirosecuenciación; 2) No se observan errores en la zona de homopolímero; 3) El *pipeline* propio detecta todos los cambios puntuales y pequeñas deleciones/inserciones de las muestras control. Este *pipeline* resulta más completo que el *software MiSeqReporter* (*Illumina*), por la anotación más detallada de las variantes y por permitir la evaluación de todas las muestras de la carrera en paralelo; 4) Las grandes deleciones/duplicaciones se pueden detectar mediante el *depth* por amplicón, pero dado el elevado ruido de fondo que se observa, es necesario completar la evaluación mediante una tecnología complementaria, como por ejemplo el MLPA. No se ha evaluado el comportamiento de esta tecnología en la detección de grandes duplicaciones.

La llegada de los mini-secuenciadores con tecnología NGS, hacen que la adopción de esta tecnología sea de preferencia en el diagnóstico genético de rutina en un futuro cercano, en pequeños y medianos laboratorios debido a:

1) El continuo descenso de los costes de la secuenciación por base
2) Menor coste en la adquisición de los equipos, comparando con las grandes plataformas NGS
3) Reducción del número de muestras necesarias para comenzar una carrera completa, dada su menor capacidad de producción de datos

A modo de resumen, mostramos una comparativa de las técnicas de secuenciación empleadas para el diagnóstico genético de los genes *BRCA1/2*, tal y como se han llevado a cabo en este trabajo, comparativa de las tres técnicas, tabla 29.

Tabla 29. Técnicas de secuenciación empleadas. Modificado de Liu et al (2012)[94].

	Plataforma		
	SANGER 3730xl	***454 GS Junior***	***MiSeq***

Tiempo de diagnóstico	12 semanas	5 semanas	5 semanas
Tiempo de preparación muestras	7 horas	30 horas	30 horas
Estudio genes	Secuencial	Simultáneo	Simultáneo
Análisis de fragmentos	NO	SI	SI (OPCIONAL)
CNVs	NO	NO (Se pueden intuir)	NO (Se pueden intuir)
Nº muestra/carrera	1	6-8	8-12
Coste aparato aproximadamente	84.000 €	95.500 €	110,000 €
Coste por muestra aproximado	1264 €	325 €	283 €
Ventajas	Gold Standard	Menor coste comparado con la secuenciación Sanger, lecturas largas	Menor coste de los tres, flujo de trabajo automatizado
Desventajas	Alto coste, bajo rendimiento	Error en las zonas de homopolímero y hacia el final de las lecturas. Requiere análisis de fragmentos. Mayor tiempo de análisis por muestra Más laborioso en el laboratorio	Mayor error en zonas ricas en GC y hacia el final de las lecturas

Nº: número

Es de destacar, el menor tiempo empleado para obtener los resultados mediante el uso de las técnicas NGS y el estudio simultáneo de ambos genes, diferencias importantes con respecto a la secuenciación Sanger, lo cual hace disminuir el tiempo de diagnóstico. Además de tener unos menores costes.

Con respecto a los resultados observados en las diferentes series de casos, los porcentajes de diagnóstico globales son muy similares un 15,95% en metodología Sanger frente a un 13,69% de secuenciación masiva.

Nuestro algoritmo de diagnóstico para testar los genes *BRCA1/2*, nos permite identificar todo el espectro de variantes previamente descritas. Presenta unos requerimientos de sensibilidad y especificidad adecuados para el diagnóstico genético de familias con sospecha de CMOH, haciendo factible un coste efectivo en comparación con las técnicas estandarizadas.

21. Caracterización del espectro de variantes en los genes *BRCA1* y *BRCA2*.

La prevalencia de VP en *BRCA1/2* en familias españolas con síndrome de CMOH se han identificado en diversas regiones, mostrando datos heterogéneos tanto a nivel global como a nivel de VP recurrentes, según el área geográfica analizada[148]. Esta prevalencia, que está directamente relacionada con los criterios de selección del estudio, las características demográficas y étnicas de las poblaciones analizadas y las técnicas de detección empleadas puede variar en nuestro país en un rango del 20-40%[104,149,150].

En la literatura la frecuencia de VP, ha sido estimada entre un 5,6-37%[151,152] de los casos, con un amplio rango registrado, sobre todo en familias con tres o más casos de cáncer de mama/cáncer de ovario, existiendo una menor prevalencia en familias con solo 2 casos de cáncer de mama (10-15%) o en pacientes muy jóvenes sin antecedentes familiares (<5%). Estos datos coinciden con lo que hemos identificado en nuestra serie de casos atendiendo a los fenotipos familiares. En nuestra serie de casos obtenemos un porcentaje diagnóstico del 26,6% cuando tenemos más de tres casos con cáncer de mama/ovario, un 12,28% cuando tenemos 2 casos de cáncer de mama, y un porcentaje mayor del esperado según la bibliografía en pacientes jóvenes de un único caso con un 11,88%.

De igual forma, la presencia de un caso de cáncer de ovario se relaciona con una elevada probabilidad de VP en *BRCA1/2*, sobre todo en el gen *BRCA1*[43]. Estos datos coinciden con lo que hemos identificado en nuestra serie, detectando más de un 30% de VP en los casos de cáncer de ovario, y 6 de los 7 casos positivos fueron VP en *BRCA1*. Además, estos pacientes con cáncer de ovario y VP pueden beneficiarse clínicamente del empleo de PARPi, si cumplen con los requisitos.

Cabe destacar que, más de la mitad de las familias con un cáncer de mama en varón presentan VP en el gen *BRCA2*[127,128]. En nuestra serie de casos solo hemos identificado una VP en un varón con cáncer de mama, c.2808_2811del (p.Ala938Profs*21) en el gen *BRCA2*, coincidiendo con la bibliografía. Esta baja incidencia, puede ser debida al escaso número de varones que hemos analizado en el que solo analizamos 14 casos en nuestra cohorte.

Nuestro estudio, registró una **prevalencia de VP del 15,38%** (70 casos índices), es una cifra bastante similar a la registrada en el trabajo de Gabaldó et al[153], del 18,1% siendo el número de familias estudiadas también similar 592 frente a 455 familias de nuestro estudio y siguiendo la selección de criterios recomendado por la SEOM.

Estudios como el de Miramar et al[150] encuentran una tasa del 33,33%, testando a 60 familias españolas previamente seleccionadas siguiendo los criterios de la SEOM, tasas diagnósticas mayores que en nuestro estudio, aunque el número de familias es mucho menor que el

nuestro. Por otro lado, un estudio de pacientes no seleccionados en población China, Zhong et al[154] detectan un porcentaje diagnóstico del 9,9% (50/507), para variantes somáticas y germinales, y un 7,9% (40/507) fueron variantes en línea germinal.

A la luz de estos resultados, es importante definir unos criterios de selección adecuados. Hay que poner de manifiesto la importancia de la **historia familiar**, la **edad de diagnóstico** y el **tipo de tumor**, como factores limitantes para obtener un alto porcentaje diagnóstico.

Si lo comparamos con el trabajo de Azzollini et al[155], en el que encuentran un porcentaje de diagnóstico del 29,3% empleando unos criterios de selección de familias estrictos en 1854 familias testadas, este porcentaje es muy superior al nuestro, y la diferencia se debe mayoritariamente a unos criterios estrictos de selección de las familias, siendo nuestros criterios mucho más laxos.

Hemos observado en nuestros datos, que cuando se presenta un solo caso familiar de cáncer de mama y cáncer de ovario, la edad de diagnóstico temprana está relacionada con la detección de una VP en estos individuos, se podría aplicar un sesgo para realizar el test diagnóstico de cáncer de mama<40 años y lo mismo para el cáncer de ovario, aunque hemos detectado un caso con 42 años y otro con 50 que se estarían escapando al acortar la edad de diagnóstico a menores de 40 años.

Por tanto, el porcentaje de probandos portadores de VP obtenido en nuestra serie, es menor y también similar al determinado para otras poblaciones seleccionadas, aunque superior a la población no seleccionada. En general, y teniendo en cuenta los estudios publicados en la literatura, podemos decir, que el porcentaje diagnóstico obtenido en el análisis de variantes en el síndrome de CMOH, depende de varios aspectos que creemos han influido en nuestro resultado:

1) **Selección de las familias**. Esta menor tasa podría explicarse debido a que este estudio se empezó en el año 2005, con los consiguientes cambios en los criterios clínicos de selección de pacientes, los casos índices estudiados proceden de distintos hospitales pertenecientes a diferentes comunidades autónomas, y cada uno de ellos ha seguido sus propios criterios de selección. En este estudio solo se ha incluido casos índices, no se han incluido a sus familiares en el estudio. Además, hay que añadir, otra limitación importante, no poder clasificar a un 26% de las familias de estudio (119 casos índice) por no disponer de información clínica completa de estas familias (edad de diagnóstico e historia familiar). Al considerar criterios poco restrictivos, vemos reducida nuestra probabilidad de detectar una VP. Los criterios de selección de las familias presentan una gran relevancia clínica, ya que las tasas de la detección de VP varían considerablemente dependiendo de estos criterios. En familias con más de tres familiares afectos, encontramos una mayor tasa diagnóstica.
2) **Los métodos empleados para el análisis de variantes.** Se han empleado varias técnicas para la detección de variantes germinales tipo SNV y CNVs, que

nos han permitido detectar todo el espectro conocido de variantes en los genes *BRCA1/2*, como han sido: la secuenciación Sanger, la secuenciación masiva, el empleo de análisis de fragmentos y el MLPA. Pensamos que la secuenciación no ha sido un factor limitante en el porcentaje de diagnóstico, ya que los métodos han sido debidamente validados, hemos sido capaces de identificar variantes de nueva descripción y se han secuenciado todos los pacientes incluidos en el estudio. En cambio, el estudio de CNVs sí que ha sido un factor limitante, ya que no disponemos del estudio de grandes deleciones de todos los casos negativos, solo disponemos de 308 estudios completos (67,69% del total de casos), no hemos estudiado las CNVs en 147 casos. Estos casos no han sido estudiados, debido a que el facultativo especialista no lo ha considerado oportuno. Tomando como referencia los datos de la bibliografía[156], hubiéramos encontrado aproximadamente un 2% más de casos diagnosticados.

3) **Porcentaje de casos índice con VSI**. Inicialmente identificamos un 13,63% de VSI, que estaría dentro del rango descrito para estos genes en la bibliografía (hasta un 20%)[157] y este porcentaje bajó a un 3,95% de VSI, una vez hecha la reclasificación, un porcentaje similar, si comparamos con el estudio de Azollini et al[155] ya que ellos encuentran 4.5% de variantes VSI. Para poder catalogar de forma definitiva estas variantes es necesario la realización de estudios complementarios, como son los estudios de cosegregación, estudios funcionales, estudios de *splicing*, así como la actualización constante de la clasificación de estas variantes en las bases de datos públicas. A todo esto, también hay que sumarle un problema importante, como es, la falta de estandarización en la clasificación clínica de variantes. En general, distintas variantes genéticas se asocian con distintos niveles de evidencia apoyando o descartando su potencial relevancia clínica. Esto hace, que una misma variante genética pueda asociarse a clasificaciones clínicas distintas (a veces incluso pueden ser contradictorias), en diferentes bases de datos o informes genéticos. Existen guías o documentos consenso, y su empleo está siendo a esta fecha muy extendido, resolviendo estos problemas[52]. Creemos necesario una continua revisión de estas variantes, de forma periódica, por ejemplo, anual para ver si se pueden reclasificar, sabemos por nuestro estudio, que la mayor parte de estas variantes se reclasifican como VPB, aunque también hemos tenido una variante reclasificada como patogénica, por eso, creemos que realizar una continua revisión de estas variantes es importante.

Relación fenotipo-genotipo

Se han descrito correlaciones **genotipo-fenotipo en los genes *BRCA1/2*** en función de la **localización de la VP en el gen**[158]. Los fenotipos de los casos índices portadores de VP de nuestro estudio, se muestran representados en función de la localización de la VP en el gen (figuras 42 y 43). Al analizarlos observamos que una proporción importante de casos de cáncer de mama, un 32% (9/28) con VP en el gen *BRCA1*, se encuentran en la región descrita por Tompson *et al*[159] como región de bajo riesgo de cáncer de mama. El resto de casos con VP

estaba fuera de esta región 68% (19/28). Con respecto al cáncer de mama y ovario y cáncer de ovario, un 33% (3/9) de los casos con VP se encuentran dentro de la región y el resto de casos 67% (6/9) se encontraban fuera de esta región.

Se ha comprobado en diferentes estudios, que existe una correlación genotipo-fenotipo más clara para las variantes del gen *BRCA2,* ya que en la parte central del gen (nucleótidos 3035-6629 en el exón 11), hay un mayor número de variantes asociadas a cáncer de ovario *versus* cáncer de mama[158,160], a esta región se le ha denominado región del grupo de cáncer de ovario (OCCR). Un estudio posterior definió dos regiones más concretas OCCR: OCCR1 (nucleótidos 2831-3847) y OCCR2 (nucleótidos 6275-6401)[159]. Estas correlaciones genotipo-fenotipo han sido confirmadas por otros estudios independientes[161,162]. Aunque otros estudios no muestran diferencias significativas en cuanto a riesgo asociado con estas regiones[48,163,164].

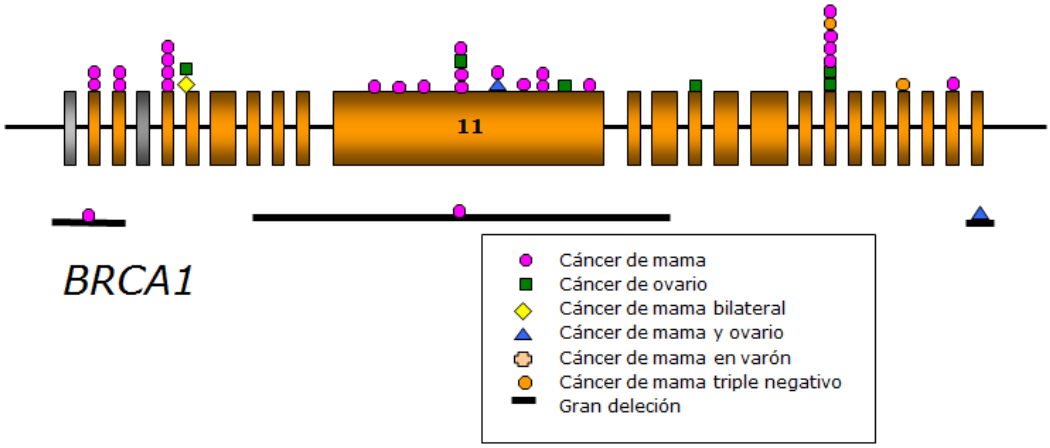


Figura 41. Representación en el gen BRCA1 de los tipos de cáncer de los casos índice portadores de VP en este gen.

El exón 11 es el exón más frecuentemente mutado en el gen *BRCA1*, en un 40% (15/37) de VP está implicado este exón, este hecho puede ser debido a que es el exón más grande del gen. El siguiente exón con más VP ha sido el exón 18, con un 19% (7/37) y el exón 5 con un 16% (6/37).

Con respecto al gen *BRCA2*, se han encontrado que un 77% (25/33) de las VP detectadas se localizaban en el exón 11, este exón es el más largo de gen, el siguiente exón con más VP ha sido el exón 10, con un 9% (3/33). Hemos observado que los dos casos de cáncer de mama y ovario y uno con cáncer de ovario se encuentran dentro de la región OCCR, pero quizá no sean representativos, por ser un número de casos tan bajo.

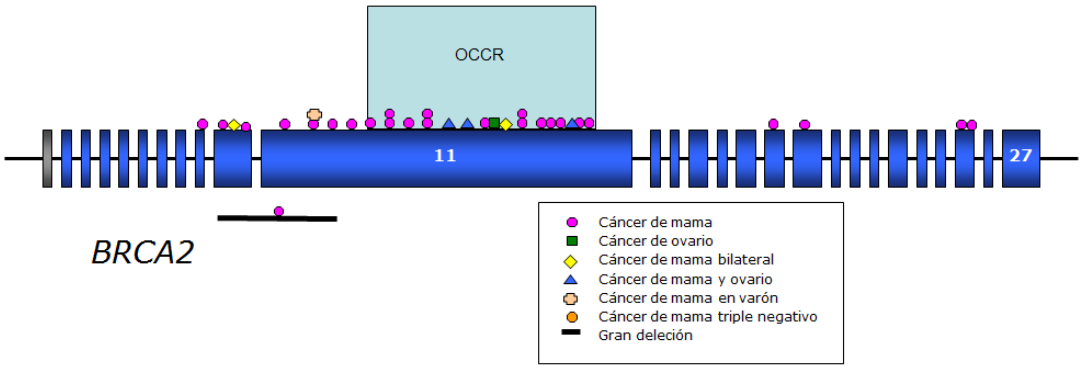


Figura 42. Representación en el gen BRCA2 de los tipos de cáncer de los casos índice portadores de VP en este gen.

Estudios predictivos en familiares de riesgo

A pesar de recomendar siempre **estudios predictivos en familiares de riesgo**, al identificar VP y VPP, han sido pocos los estudios llevados a cabo durante este estudio, vamos a indicar los estudios más frecuentes de VP con más portadores detectados. En la tabla 30, se representan las VP en las que se han realizado más estudios predictivos. Gracias a los estudios predictivos se puede llevar a cabo un seguimiento/prevención a edades más tempranas de los portadores asintomáticos.

Tabla 30. Estudios predictivos en las familias

Gen	Variante	Casos afectos /portadores	CM	Edad diagnóstica (rango)	CO	Edad diagnóstica
BRCA1	c.68_69del (p.Glu23Valfs*17)	3/5	1	35 (24-47)	-	-
	c.5123C>A (p.Ala1708Glu)	9/12	7	25 (18-47)	2	-
	Deleción exones 8 al 13	3/4	1	44 (26-57)	-	-
BRCA2	c.5350_5351del (p.Asn1784Hisfs*2)	3/3	2	-	1	-
	c.5720_5723del (p.Ser1907*)	3/7	3	40,5 (37-44)	-	-
	c.6034del (p.Ser2012Profs*28)	2/2	1	38 (34-42)	-	-

CM: Cáncer de mama; CO: Cáncer de ovario

Grandes deleciones y duplicaciones

En nuestro estudio, mediante la tecnología del MLPA, hemos detectado un total de 4 CNVs en los genes *BRCA1/2*. Todos las CNVs han sido grandes deleciones. Tres grandes deleciones han sido detectadas en el gen *BRCA1*, en tres familias distintas, mientras que se ha detectado una gran deleción en el gen *BRCA2*, en una familia. Del total de VP en el gen *BRCA1*, las CNVs suponen un 8% del total de VP detectadas en este gen. Tenemos una prevalencia para las CNVs del 0,97% para el gen *BRCA1*.

Estos resultados son similares a las frecuencias de CNVs en población australiana 2,2%[165], y la población española 1,4%[166], en cambio son menores a las detectadas en población norteamericana 10,3%[167].

Del total de VP en el gen *BRCA2*, las CNVs suponen un 3% del total de VP detectadas en este gen. Tenemos una prevalencia para las CNVs de 0,32% para el gen *BRCA2.* Estos resultados contrastan con las frecuencias de CNVs en *BRCA2* en población española 1.5%[168], en población norteamericana 5%[167], y en población australiana 2%[165].

Variantes de significado incierto (VSI)

Las VSI son difíciles de clasificar en el entorno clínico, no existe un sistema de clasificación consensuado a nivel internacional para estas variantes. Debería existir un sistema de clasificación consensuado, para que al encontrarnos con este tipo de variantes en un resultado genético, se pueda minimizar el riesgo de una mala interpretación en la práctica clínica[169]. Se han propuestos múltiples sistemas, pero no ofrecen una clasificación clínica relevante. La Organización Mundial de la Salud financió la Agencia Internacional para la Investigación del Cáncer (IARC) en 2008[170], recomendando un modelo de probabilidad multifactorial para estimar la patogenicidad de la variante[171]. Este es el único sistema de clasificación que une recomendaciones clínicas en cada clase[169].

Se han detectado 18 VSI distintas en nuestra cohorte de pacientes, representan el 3,95% de casos índice portadores de este tipo de variantes. Esta prevalencia es baja si comparamos nuestros resultados con los publicados en el trabajo de Lee et al[172], ya que ellos encuentran un 8.9% de VSI haciendo también una reclasificación de variantes. Con su estrategia reducen la proporción de VSI del 19% al 8,9% y nosotros logramos reducir del 13,63% al 3,95%, la proporción obtenida a nivel global es similar con ambos métodos.

La mayoría de las VSI detectadas en nuestro estudio han sido de tipo *missense*, 6 variantes *missense* detectadas en el gen *BRCA1* y 8 variantes *missense* en el gen *BRCA2*.

Para ayudar a interpretar las VSI, además de consultar las diferentes bases de datos y los estudios *in silico* de los predictores Bioinformáticos, se recomienda realizar estudios de cosegregación en la familia, conservación evolutiva entre especies del aminoácido, estudios de ARN para evaluar si la VP altera el *splicing*, ensayos funcionales, así como estudios del fenotipo[173,174].

Se han detectado 14 variantes de nueva descripción, representan un 3% de familias portadoras de este tipo de variantes en nuestra cohorte, de ellas 6, se han clasificado como VPP debido al efecto deletéreo que producen en la proteína, representando un 8,5% de las VP detectadas. Una variante la hemos clasificado como VPB, ya que en nuestra cohorte co-ocurre con una VP en ese mismo gen y el resto están clasificadas como VSI, representando un 50% del total de VSI.

En nuestro estudio, que empezó en el año 2005 y concluyó en el año 2014, hemos visto cómo variantes consideradas como VSI, han podido ser reclasificadas debido a: nuevos estudios internacionales, ensayos funcionales, actualización de las bases de datos, nuestra propia base de datos, estudios de *splicing,* entre otros. Han sido reclasificadas un total de 64 variantes, la mayoría de las variantes han pasado a ser VPB coincidiendo con el trabajo de Ecles et al[169], aunque una de ellas ha pasado a ser patogénica. Por lo tanto, hay que tener especial atención sobre ellas. Es necesario hacer una re-evaluación de las variantes, y emitir informes adicionales genéticos, si se ha encontrado alguna discrepancia entre la evaluación inicial de la variante y la posterior reevaluación de esa misma variante.

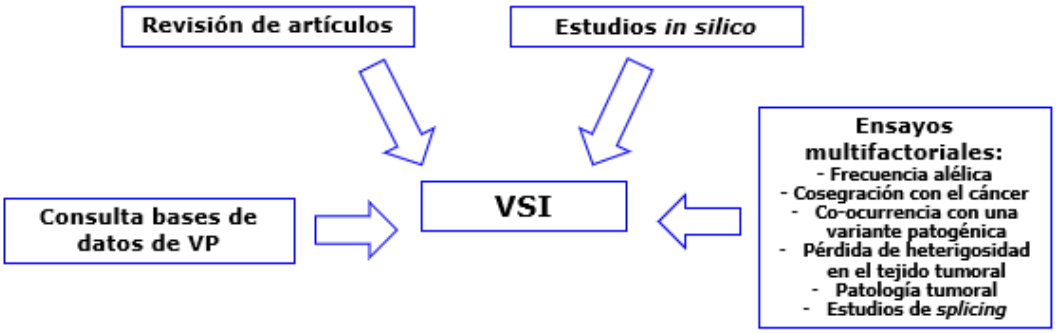


Figura 43. Reclasificación de VSI. Tarea compleja que envuelve a muchos elementos. Figura modificada del trabajo de Wallace et al[175].

También hemos observado una disparidad de clasificaciones de variantes entre diferentes bases de datos, para una misma variante. Como conclusión fundamental, es necesario centralizar toda la información en una base de datos única y pública, para así lograr un rápido avance en la minimización de la incertidumbre que ocasionan las VSI, para un mejor manejo clínico de las familias con riego de cáncer. Un intento para la unificación de la información, es el *Human Project Variome* bajo los auspicios de la UNESCO, es un portal de información unificada internacional no gubernamental que unifica la información y enumera cuatro bases de datos separadas de los genes *BRCA1/2* e intenta llegar a un consenso en la terminología. Otra base de datos que busca la unificación es ENIGMA, consorcio internacional multidisciplinar de investigadores que colaboran para la clasificación de VSI en los genes *BRCA1/2*, así como otros genes relacionados con el cáncer de mama[176].

Debe haber un consenso a la hora de la interpretación de las variantes, ya que hay interpretaciones conflictivas en determinadas variantes y esto se ve reflejado en las bases de datos consultadas, por ejemplo, ClinVar (*Conflicting interpretations of pathogenicity*). Para evitar esto, se debería realizar un registro *on line*, en una base de datos pública, donde se

puedan identificar las variantes con interpretaciones conflictivas, llevadas a cabo por diferentes laboratorios, así veríamos la frecuencia y el tipo de discrepancias y las razones de esas discrepancias, para así poder realizar una guía específica, para la curación de estas variantes. Con esto, se aumentaría la coherencia de la interpretación de variantes entre laboratorios. Es importante compartir datos, y evidenciar en qué se basan las interpretaciones. Esto permitirá promover estrategias para estandarizar un algoritmo de curación de variantes, un claro ejemplo es la base de datos *on line* PROMPT (*Prospective Registry of Multiplex Testing*)[177].

22. Conclusiones de la discusión

- El empleo de mini-secuenciadores NGS es adecuado para el diagnóstico genético de los genes *BRCA1/2* junto con técnicas de análisis de fragmentos, siendo eficiente, fiable, con un coste efectivo y rápido.
- La sensibilidad de los mini-secuenciadores NGS para el diagnóstico genético de los genes *BRCA1/2* junto con técnicas de análisis de fragmentos, es sumamente elevada, ya que alcanza el 100% de las variantes testadas, detectándose todo tipo de variantes (*frameshift*, *missense*, *nonsense* y *splicing*, e incluso variantes complejas como *indels*). Para evitar falsos positivos, todas las VP, VPP deben ser validadas por secuenciación Sanger, en una segunda muestra de la paciente. La detección de CNVs debe ser testada mediante otro tecnología, como por ejemplo MLPA.
- El 15,38% de la población de nuestro estudio, es portadora de alguna VP o VPP en los genes *BRCA1/2.* Un 8,13% de las familias resultó ser portadora de una VP o VPP en el gen *BRCA1* y un 7,25% de las familias en el gen *BRCA2*.
- Para obtener un alto porcentaje diagnóstico en la población de estudio, es importante definir unos criterios adecuados de selección, tales como la historia familiar, la edad de diagnóstico y el tipo de tumor.
- Seis variantes nuevas VPP han sido descritas en el presente estudio. Una variante en el gen *BRCA1* (c.4106_4107delinsGCT) y cinco en el gen *BRCA2* (c.1792del, c.4263dup, c.4677del, c.6034del y deleción parcial del final del exón 10 y principio del exón 11).
- Las VSI exigen una constante revisión de las bases de datos, debido al aumento de publicaciones y a los nuevos estudios funcionales realizados, estudios de *splicing*, así como estudios de cosegregación en las familias con este tipo de variantes. La actualización de la información en los pacientes con un test genético con este tipo de variantes, deberá ser un proceso activo y dinámico.

23. Perspectiva futura

El estudio simultáneo de ambos genes mediante secuenciación masiva y CNVs, conlleva una serie de ventajas, tales como:

- La detección de todo el espectro posibles de VP, además de poder detectar los posibles casos de dobles heterocigotos en los genes *BRCA1/2*, que es un fenómeno raro, con una incidencia estimada de entre 0,22-0,87% entre portadores de VP en los genes *BRCA1/2*[178,179].

- La obtención de resultados en menor tiempo (de 12 semanas en un estudio mediante secuenciación Sanger y MLPA, a 4-5 semanas)

- Un menor coste

- Una sensibilidad mayor[91]

Han salido ya guías para la determinación de variantes en los genes *BRCA1/2* a nivel germinal para cáncer de mama, cáncer de ovario, cáncer de próstata y cáncer de páncreas para terapia dirigida con PARPi. También se recomienda el estudio de estos dos genes a nivel somático para cáncer de ovario y próstata para terapia dirigida con PARPi[32].

Al análisis de los genes *BRCA1/2* en sangre, se sumará el análisis de estos genes en muestras parafinadas tumorales, aumentando el número de muestras a analizar, y cada vez se demanda una bajada de los tiempos de respuesta para poder aplicar estas terapias lo más rápido posible.

El algoritmo diagnóstico que se está imponiendo en la actualidad, para la detección de VP asociadas al síndrome de CMOH, es el empleo de paneles de secuenciación dirigida de genes de susceptibilidad al cáncer, previamente validados. Estos estudios estarían enfocados a familias de alto riesgo negativas para VP en los genes *BRCA1/2*. Para realizar estos estudios, se requieren de una serie de pasos; 1) selección biomédica de los genes asociados a la patología, 2) diseño bioinformático del panel, 3) validación bioinformática y 4) validación diagnóstica del panel que permita evaluar los parámetros de calidad del panel como son: la reproducibilidad, cobertura media, sensibilidad, especificidad, detección de *indels*, confirmación de variantes previamente secuenciadas por secuenciación Sanger, detección de CNVs mediante algoritmos Bioinformáticos y su confirmación mediante la técnica del MLPA de los casos positivos que se encuentran cerca del límite de detección de la técnica de cribado de CNVs, previamente a su aplicación en el diagnóstico genético[180].

CONCLUSIONES:

1. El estudio de los genes *BRCA1/2* mediante el kit *BRCA MASTR Assay (Multiplicom)* secuenciando con *454 SG Junior* (Roche) y el *software AVA*, da una buena fiabilidad a los estudios para variantes puntuales, pero es necesario el análisis de fragmentos para dar una mayor fiabilidad para identificar variantes en la zona de homopolímeros.
2. El estudio de los genes *BRCA1/2* mediante el kit *BRCA MASTR Assay (Multiplicom)* secuenciando con *MiSeq* (*Illumina*) y el *pipeline* propio dan la mayor fiabilidad a los estudios, sin necesidad de utilizar el análisis de fragmentos, debido a la mejora en la química de la secuenciación masiva.
3. La secuenciación masiva consigue mejorar costes y tiempos de entrega de resultados con un rendimiento diagnóstico similar al de la secuenciación Sanger.
4. La caracterización molecular de los genes *BRCA1/2* en pacientes con sospecha clínica de CMOH obtuvo rendimientos diagnósticos similares a lo descrito previamente en la bibliografía. La mayoría de las variantes clasificadas como VP y VPP fueron de tipo *nonsense* o *frameshift*, mientras que las VSI han sido mayoritariamente variantes *missense*. Este estudio ha aportado la identificación de 14 variantes previamente no descrita en el momento en el que se identificaron (6 clínicamente relevantes y 8 VSI).
5. Se considera necesaria una revisión periódica de las variantes clasificadas como VSI, debido a la nueva evidencia científica que se genera continuamente. La actualización de la información en los pacientes con un test genético con este tipo de variantes, deberá ser un proceso activo y dinámico.

BIBLIOGRAFÍA

1. Cifras del cáncer en España | SEOM - Día Mundial del Cáncer 2023. https://seom.org/dmcancer/cifras-del-cancer/.

2. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).

3. Stanisławek, A. Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies— An Updated Review. *Cancers (Basel).* **13**, 1–30 (2021).

4. Hunn, J. & Rodriguez, G. C. Ovarian cancer: Etiology, risk factors, and epidemiology. *Clin. Obstet. Gynecol.* **55**, 3–23 (2012).

5. Momenimovahed, Z. & Salehiniya, H. Breast Cancer-Targets and Therapy Dovepress epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer* **11**, 151–164. (2019).

6. Gage, M., Wattendorf, D. & Henry, L. R. Translational advances regarding hereditary breast cancer syndromes. *Journal of Surgical Oncology* vol. 105 444–451 (2012).

7. Prat, J., Ribé, A. & Gallardo, A. Hereditary ovarian cancer. *Hum. Pathol.* **36**, 861–70 (2005).

8. Hollestelle, A., Wasielewski, M., Martens, J. W. M. & Schutte, M. Discovering moderate-risk breast cancer susceptibility genes. *Curr. Opin. Genet. Dev.* **20**, 268–76 (2010).

9. Yanes, T., Young, M. A., Meiser, B. & James, P. A. Clinical applications of polygenic breast cancer risk: A critical review and perspectives of an emerging field. *Breast Cancer Res.* **22**, 1–10 (2020).

10. Sud, A., Kinnersley, B. & Houlston, R. S. Genome-wide association studies of cancer: current insights and future perspectives. *Nat. Rev. Cancer* **17**, 692–704 (2017).

11. Dorling, L. *et al.* Breast Cancer Risk Genes — Association Analysis in More than 113,000 Women. *N. Engl. J. Med.* **384**, 428–439 (2021).

12. Foulkes, W. D. Inherited susceptibility to common cancers. *New England Journal of Medicine* vol. 359 2143 (2008).

13. Apostolou, P. & Fostira, F. Hereditary breast cancer: the era of new susceptibility genes. *Biomed Res. Int.* **2013**, 747318 (2013).

14. Bahcall, O. Common variation and heritability estimates for breast, ovarian and prostate cancers. *Nat. Genet.* (2013) doi:10.1038/ngicogs.1.

15. Imyanitov, E. N. *et al.* Hereditary cancer syndromes. **14**, 40–69 (2023).

16. Sokolova, A., Johnstone, K. J., McCart Reed, A. E., Simpson, P. T. & Lakhani, S. R. Hereditary breast cancer: syndromes, tumour pathology and molecular testing. *Histopathology* **82**, 70–82 (2023).

17. Rich, T. A., Woodson, A. H., Litton, J. & Arun, B. Hereditary breast cancer syndromes and genetic testing. *J. Surg. Oncol.* **111**, 66–80 (2015).

18. Yang, X. *et al.* Cancer risks associated with germline PALB2 pathogenic variants: An international study of 524 families. *J. Clin. Oncol.* **38**, 674–685 (2020).

19. Tischkowitz, M. *et al.* Management of individuals with germline variants in PALB2: a clinical practice resource of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.* **23**, 1416–1423 (2021).

20. Song, H. *et al.* Contribution of germline mutations in the RAD51B, RAD51C, and RAD51D genes to ovarian cancer in the population. *J. Clin. Oncol.* **33**, 2901–2907 (2015).

21. Andrews, L. & Mutch, D. G. Hereditary Ovarian Cancer and Risk Reduction. *Best Practice and Research: Clinical Obstetrics and Gynaecology* vol. 41 31–48 (2017).

22. Suszynska, M., Ratajska, M. & Kozlowski, P. BRIP1, RAD51C, and RAD51D mutations are associated with high susceptibility to ovarian cancer: Mutation prevalence and precise risk estimates based on a pooled analysis of ~30,000 cases. *J. Ovarian Res.* **13**, (2020).

23. Economopoulou, P., Dimitriadis, G. & Psyrri, A. Beyond BRCA: New hereditary breast cancer susceptibility genes. *Cancer Treatment Reviews* vol. 41 1–8 (2015).

24. Couch, F. J. *et al.* Associations between cancer predisposition testing panel genes and breast cancer. *JAMA Oncol.* **3**, 1190–1196 (2017).

25. Suszynska, M., Klonowska, K., Jasinska, A. J. & Kozlowski, P. Large-scale meta-analysis of mutations identified in panels of breast/ovarian cancer-related genes — Providing evidence of cancer predisposition genes. *Gynecologic Oncology* vol. 153 452–462 (2019).

26. Hu, C. *et al.* A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N. Engl. J. Med.* **384**, 440–451 (2021).

27. Lee, K. *et al.* Clinical validity assessment of genes frequently tested on hereditary breast and ovarian cancer susceptibility sequencing panels. *Genet. Med.* **21**, 1497–1506 (2019).

28. Spurdle, A. B. *et al.* Towards controlled terminology for reporting germline cancer susceptibility variants: An ENIGMA report. *J. Med. Genet.* **56**, (2019).

29. Triviño, J. C. *et al.* Combination of phenotype and polygenic risk score in breast cancer risk evaluation in the Spanish population: a case-control study. *BMC Cancer* **20**, (2020).

30. Nielsen, F. C., Van Overeem Hansen, T. & Sørensen, C. S. Hereditary breast and ovarian cancer: New genes in confined pathways. *Nature Reviews Cancer* vol. 16 599–612 (2016).

31. Molina-Zayas, M. *et al.* Identification of hereditary breast and ovarian cancer germline variants in Granada (Spain): NGS perspective. *Mol. Genet. Genomics* **297**, 859 (2022).

32. Pujol, P. *et al.* Clinical practice guidelines for BRCA1 and BRCA2 genetic testing. *European Journal of Cancer* vol. 146 30–47 (2021).

33. Hall, J. M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science (80-.).* **250**, 1684–1689 (1990).

34. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (80-.).* **266**, 66–71 (1994).

35. Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).

36. Venkitaraman, A. R. Linking the Cellular Functions of BRCA Genes to Cancer Pathogenesis and Treatment . *Annu. Rev. Pathol. Mech. Dis.* **4**, 461–487 (2009).

37. Takaoka, M. & Miki, Y. BRCA1 gene: function and deficiency. *Int. J. Clin. Oncol.* **23**, 36–44 (2018).

38. Huen, M. S. Y., Sy, S. M. H. & Chen, J. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat. Rev. Mol. Cell Biol.* **11**, 138–48 (2010).

39. Pfeffer, C. M., Ho, B. N. & Singh, A. T. K. The evolution, functions and applications of the

breast cancer genes BRCA1 and BRCA2. *Cancer Genomics and Proteomics* vol. 14 293–298 (2017).

40. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer* **12**, 68–78 (2012).

41. Cline, M. S. *et al.* BRCA Challenge: BRCA Exchange as a global resource for variants in BRCA1 and BRCA2. *PLoS Genet.* **14**, 1–17 (2018).

42. BIOBASE product login page. https://portal.biobase-international.com/cgi-bin/portal/login.cgi.

43. Petrucelli, N., Daly, M. B. & Pal, T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. *Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. 1998 Sep 4 [Updated 2022 May 26]. In: Adam MP, Mirzaa GM, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle;* http://www.ncbi.nlm.nih.gov/pubmed/20301425 (2022).

44. Pócza, T. *et al.* Germline Structural Variations in Cancer Predisposition Genes. *Front. Genet.* **12**, 1–14 (2021).

45. Thompson, D. & Easton, D. The genetic epidemiology of breast cancer genes. *Journal of Mammary Gland Biology and Neoplasia* vol. 9 221–236 (2004).

46. Maxwell, K. N., Domchek, S. M., Nathanson, K. L. & Robson, M. E. Population Frequency of Germline BRCA1/2 Mutations. *J. Clin. Oncol.* **34**, 4183–4185 (2016).

47. Simchoni, S. *et al.* Familial clustering of site-specific cancer risks associated with BRCA1 and BRCA2 mutations in the Ashkenazi Jewish population. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3770–3774 (2006).

48. Díez, O. *et al.* Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: A high proportion of mutations unique to Spain and evidence of founder effects. *Hum. Mutat.* **22**, 301–312 (2003).

49. Monteiro, A. N. *et al.* Variants of uncertain clinical significance in hereditary breast and ovarian cancer genes: Best practices in functional analysis for clinical annotation. *Journal of Medical Genetics* vol. 57 509–518 (2020).

50. Scambia, G. BRCA1/2 variants of unknown significance in hereditary breast and ovarian cancer (HBOC) syndrome: Looking for the hidden meaning. *Crit Rev Oncol Hematol* **172**,

(2022).

51. Colombo, M. *et al.* Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Hum. Mol. Genet.* **23**, 3666–80 (2014).

52. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).

53. Ellard, S. *et al. ACGS Best Practice Guidelines for Variant Classification 2018*. https://www.clinicalgenome.org/working-groups/sequence-variant-interpretation/ (2018).

54. Kitts, A., Phan, L., Ward, M. & Holmes, J. B. *The Database of Short Genetic Variation (dbSNP)*. https://www.ncbi.nlm.nih.gov/books/.

55. NHGRI: Breast Cancer Information Core. https://research.nhgri.nih.gov/bic/.

56. BRCA2 gene homepage - Global Variome shared LOVD. https://databases.lovd.nl/shared/genes/BRCA2.

57. BRCA1 gene homepage - Global Variome shared LOVD. https://databases.lovd.nl/shared/genes/BRCA1.

58. Plon, S. E. *et al.* Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human Mutation* vol. 29 1282–1291 (2008).

59. Lindor, N. M., Goldgar, D. E., Tavtigian, S. V., Plon, S. E. & Couch, F. J. BRCA1/2 Sequence Variants of Uncertain Significance: A Primer for Providers to Assist in Discussions and in Medical Management. *Oncologist* **18**, 518–524 (2013).

60. Kuchenbaecker, K. B. *et al.* Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA - J. Am. Med. Assoc.* **317**, 2402–2416 (2017).

61. Milne, R. L. *et al.* The average cumulative risks of breast and ovarian cancer for carriers of mutations in BRCA1 and BRCA2 attending genetic counseling units in Spain. *Clin. Cancer Res.* **14**, 2861–2869 (2008).

62. (SEOM), .Sociedad Española de Oncología Médica. *CÁNCER HEREDITARIO 3ª Edición Con*

la colaboración de. (2019).

63. Pregnancies, breast-feeding, and breast cancer risk in the International BRCA1/2 Carrier Cohort Study (IBCCS). *J Natl Cancer Inst* **98**, 535–44. (2006).

64. Lilyquist, J., Ruddy, K. J., Vachon, C. M. & Couch, F. J. Common genetic variation and breast cancer Risk—Past, present, and future. *Cancer Epidemiology Biomarkers and Prevention* vol. 27 380–394 (2018).

65. Antoniou, A. C. *et al.* RAD51 135G→C modifies breast cancer risk among BRCA2 mutation carriers: Results from a combined analysis of 19 studies. *Am. J. Hum. Genet.* **81**, 1186–1200 (2007).

66. Antoniou, A. C. *et al.* Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers. *Breast Cancer Res.* **14**, R33 (2012).

67. Karakasis, K., Burnier, J. V., Bowering, V., Oza, A. M. & Lheureux, S. Ovarian cancer and BRCA1/2 testing: Opportunities to improve clinical care and disease prevention. *Frontiers in Oncology* vol. 6 1 (2016).

68. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).

69. Turk, A. & Wisinski, K. B. PARP Inhibition in BRCA-Mutant Breast Cancer. *Cancer* **124**, 2498 (2018).

70. KRISTIN N. TAYLOR, MD, AND MARLA SCOTT, M. PARP inhibitors: Choosing what to use in epithelial ovarian cancer | Contemporary OBGYN. **65**, 30–36 (2020).

71. Tangutoori, S., Baldwin, P. & Sridhar, S. PARP inhibitors: A new era of targeted therapy. *Maturitas* vol. 81 5–9 (2015).

72. Peixoto, A. *et al.* Tumor Testing for Somatic and Germline BRCA1/BRCA2 Variants in Ovarian Cancer Patients in the Context of Strong Founder Effects. *Front. Oncol.* **10**, 1–9 (2020).

73. González-Santiago, S. *et al.* SEOM clinical guidelines in hereditary breast and ovarian cancer (2019). *Clin. Transl. Oncol.* **22**, 193–200 (2020).

74. Soto, J. L. *et al.* Documento de consenso sobre la implementación de la secuenciación masiva de nueva generación en el diagnóstico genético de la predisposición hereditaria

al cáncer. *Med. Clin. (Barc).* **151**, 80.e1-80.e10 (2018).

75. Vail, P. J. *et al.* Comparison of locus-specific databases for BRCA1 and BRCA2 variants reveals disparity in variant classification within and among databases. *J. Community Genet.* **6**, 351–359 (2015).

76. Kwong, A. *et al.* How does re-classification of variants of unknown significance (VUS) impact the management of patients at risk for hereditary breast cancer? *BMC Med. Genomics* **15**, 1–8 (2022).

77. Møller, P. *et al.* Genetic epidemiology of BRCA mutations – family history detects less than 50% of the mutation carriers. *Eur. J. Cancer* **43**, 1713–1717 (2007).

78. De Sanjosé, S. *et al.* Prevalence of BRCA1 and BRCA2 germline mutations in young breast cancer patients: A population-based study. *Int. J. Cancer* **106**, 588–593 (2003).

79. Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 560–564 (1977).

80. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).

81. Voelkerding, K. V., Dames, S. A. & Durtschi, J. D. Next-generation sequencing:from basic research to diagnostics. *Clinical Chemistry* vol. 55 641–658 (2009).

82. Liu, W., Smith, D. I., Rechtzigel, K. J., Thibodeau, S. N. & David James, C. *Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations*. *Nucleic Acids Research* vol. 26 http://lotka.stanford.edu/ (1998).

83. De Leeneer, K., Coene, I., Poppe, B., De Paepe, A. & Claes, K. Rapid and sensitive detection of BRCA1/2 mutations in a diagnostic setting: Comparison of two high-resolution melting platforms. *Clin. Chem.* **54**, 982–989 (2008).

84. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005).

85. Schadt, E. E., Turner, S. & Kasarskis, A. A window into third-generation Schadt, E. E., Turner, S., & Kasarskis, A. (2010). A window into third-generation sequencing. Human Molecular Genetics, 19(R2), 227–240. https://doi.org/10.1093/hmg/ddq416sequencing. *Hum. Mol. Genet.* **19**, 227–240 (2010).

86. Harismendy, O. *et al.* Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol.* **10**, R32 (2009).

87. Mattocks, C. J. *et al.* A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur. J. Hum. Genet.* **18**, 1276–1288 (2010).

88. Nicolussi, A. *et al.* Next-generation sequencing of BRCA1 and BRCA2 genes for rapid detection of germline mutations in hereditary breast/ovarian cancer. *PeerJ* **7**, 1–22 (2019).

89. Chirlaque, M. D. *et al.* Cancer survival in adult patients in Spain. Results from nine population-based cancer registries. *Clin. Transl. Oncol.* **20**, 201–211 (2018).

90. Md., S. S.-G. *et al.* DIAGNÓSTICO MOLECULAR DE ENFERMEDADES GENÉTICAS: DEL DIAGNÓSTICO GENÉTICO AL DIAGNÓSTICO GENÓMICO CON LA SECUENCIACIÓN MASIVA. *Rev. Médica Clínica Las Condes* **26**, 458–469 (2015).

91. D'Argenio, V. *et al.* The molecular analysis of BRCA1 and BRCA2: Next-generation sequencing supersedes conventional approaches. *Clin. Chim. Acta* **446**, 221–225 (2015).

92. Marcus, R. K. *et al.* Challenges to clinical utilization of hereditary cancer gene panel testing: perspectives from the front lines. *Fam. Cancer* **14**, 641–649 (2015).

93. Escalante, A. E., Jardón Barbolla, L., Ramírez-Barahona, S. & Eguiarte, L. E. The study of biodiversity in the era of massive sequencing El estudio de la biodiversidad en la era de la secuenciación masiva. *Rev. Mex. Biodivers.* **85**, 1249–1264 (2014).

94. Liu, L. *et al.* Comparison of Next-Generation Sequencing Systems. *J. Biomed. Biotechnol.* **2012**, 11 (2012).

95. Schouten, J. P. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* **30**, e57 (2002).

96. De Schrijver, J. M. *et al.* Analysing 454 amplicon resequencing experiments using the modular and database oriented Variant Identification Pipeline. *BMC Bioinformatics* **11**, (2010).

97. Ensembl Release 62 – Ensembl Blog. http://www.ensembl.info/2011/04/14/ensembl-release-62/.

98. Zou, D., Ma, L., Yu, J. & Zhang, Z. Biological databases for human research. *Genomics, Proteomics and Bioinformatics* vol. 13 55–63 (2015).

99. Enyedi, M. Z. *et al.* Simultaneous detection of BRCA mutations and large genomic rearrangements in germline DNA and FFPE tumor samples. *Oncotarget* **7**, 61845–61859 (2016).

100. Den Dunnen, J. T. *et al.* HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat* **37**, 564–569 (2016).

101. Craig, D. W. *et al.* Identification of genetic variants using bar-coded multiplexed sequencing. *Nat. Methods* **5**, 887–893 (2008).

102. de Leeneer, K. *et al.* Practical tools to implement massive parallel pyrosequencing of PCR products in next generation molecular diagnostics. *PLoS One* **6**, (2011).

103. Minucci, A. *et al.* Clinical impact on ovarian cancer patients of massive parallel sequencing for BRCA mutation detection: The experience at Gemelli hospital and a literature review. *Expert Rev. Mol. Diagn.* **15**, 1383–1403 (2015).

104. Díez, O. *et al.* Analysis of BRCA1 and BRCA2 Genes in Spanish Breast/Ovarian Cancer Patients: A High Proportion of Mutations Unique to Spain and Evidence of Founder Effects. *Hum Mutat* **22**, 301–12 (2003).

105. Uhrhammer, N. *et al.* BRCA1 mutations in Algerian breast cancer patients: High frequency in young, sporadic cases. *Int. J. Med. Sci.* **5**, 197–202 (2008).

106. Stoppa-Lyonnet, D. *et al.* BRCA1 sequence variations in 160 individuals referred to a breast/ovarian family cancer clinic. Institut Curie Breast Cancer Group. *Am. J. Hum. Genet.* **60**, 1021–30 (1997).

107. Blay, P. *et al.* Mutational analysis of BRCA1 and BRCA2 in hereditary breast and ovarian cancer families from Asturias (Northern Spain). *BMC Cancer* **13**, 243 (2013).

108. Houdayer, C. *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.* **33**, 1228–1238 (2012).

109. Walsh, T. *et al.* Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18032–18037 (2011).

110. Marroni, F. *et al.* Penetrances of breast and ovarian cancer in a large series of families tested for BRCA1/2 mutations. *Eur. J. Hum. Genet.* **12**, 899–906 (2004).

111. Lindor, N. M. *et al.* A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Human Mutation* vol. 33 8–21 (2012).

112. Díez, O. *et al.* BRCA1 mutation analysis in 83 Spanish breast and breast/ovarian cancer families. *Int. J. Cancer* **83**, 465–469 (1999).

113. Hernan, I. *et al.* Two novel frameshift mutations in BRCA2 gene detected by next generation sequencing in a survey of Spanish patients of breast cancer. *Clin. Transl. Oncol.* **17**, 576–580 (2015).

114. Caputo, S. *et al.* Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. *Nucleic Acids Res.* **40**, (2012).

115. Håkansson, S. *et al.* Moderate frequency of BRCA1 and BRCA2 germ-line mutations in Scandinavian familial breast cancer. *Am. J. Hum. Genet.* **60**, 1068–1078 (1997).

116. Van Der Hout, A. H. *et al.* A DGGE system for comprehensive mutation screening of BRCA1 and BRCA2: Application in a Dutch cancer clinic setting. *Hum. Mutat.* **27**, 654–666 (2006).

117. Risch, H. A. *et al.* Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *Am. J. Hum. Genet.* **68**, 700–710 (2001).

118. Seong, M. *et al.* Comprehensive mutational analysis of BRCA1/BRCA2 for Korean breast cancer patients: Evidence of a founder mutation. *Clin. Genet.* **76**, 152–160 (2009).

119. Serova, O. M. *et al.* Mutations in BRCA1 and BRCA2 in breast cancer families: are there more breast cancer-susceptibility genes? *Am. J. Hum. Genet.* **60**, 486–95 (1997).

120. De Juan Jiménez, I. *et al.* Novel and recurrent BRCA1/BRCA2 mutations in early onset and familial breast and ovarian cancer detected in the Program of Genetic Counseling in Cancer of Valencian Community (eastern Spain). Relationship of family phenotypes with mutation prevalence. *Fam. Cancer* **12**, 767–777 (2013).

121. Phelan, C. M. *et al.* A low frequency of non-founder BRCA1 mutations in Ashkenazi Jewish breast-ovarian cancer families. *Hum. Mutat.* **20**, 352–357 (2002).

122. Rodríguez, A. O. *et al.* BRCA1 and BRCA2 mutations among ovarian cancer patients from

Colombia. *Gynecol. Oncol.* **124**, 236–243 (2012).

123. Pajares, B. *et al.* Hereditary breast and ovarian cancer in Andalusian families: A genetic population study. *BMC Cancer* **18**, 1–12 (2018).

124. Meindl, A. Comprehensive analysis of 989 patients with breast or ovarian cancer providesBRCA1 andBRCA2 mutation profiles and frequencies for the German population. *Int. J. Cancer* **97**, 472–480 (2002).

125. Lecarpentier, J. *et al.* Variation in breast cancer risk associated with factors related to pregnancies according to truncating mutation location, in the French National BRCA1 and BRCA2 mutations carrier cohort (GENEPSO). *Breast Cancer Res.* **14**, R99 (2012).

126. Zhang, S. *et al.* Frequencies of BRCA1 and BRCA2 mutations among 1,342 unselected patients with invasive ovarian cancer. *Gynecol. Oncol.* **121**, 353–357 (2011).

127. Esteban Cardeñosa, E. *et al.* Twenty-three novel BRCA1 and BRCA2 sequence alterations in breast and/or ovarian cancer families of Eastern Spain. *Breast Cancer Res. Treat.* **112**, 69–73 (2008).

128. Infante, M. *et al.* Two founder BRCA2 mutations predispose to breast cancer in young women. *Breast Cancer Res. Treat.* **122**, 567–571 (2010).

129. Peixoto, A. *et al.* The c.156-157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. *Breast Cancer Res. Treat.* **114**, 31–38 (2009).

130. Sluiter, M. D. & Van Rensburg, E. J. Large genomic rearrangements of the BRCA1 and BRCA2 genes: Review of the literature and report of a novel BRCA1 mutation. *Breast Cancer Res. Treat.* **125**, 325–349 (2011).

131. Arnold, A. G. *et al.* Assessment of individuals with BRCA1 and BRCA2 large rearrangements in high-risk breast and ovarian cancer families. *Breast Cancer Res. Treat.* **145**, 625–634 (2014).

132. Meyer, P., Voigtlaender, T., Bartram, C. R. & Klaes, R. Twenty-three novelBRCA1 andBRCA2 sequence alterations in breast and/or ovarian cancer families in Southern Germany. *Hum. Mutat.* **22**, 259–259 (2003).

133. Tsai, G. J. *et al.* Outcomes of 92 patient-driven family studies for reclassification of variants of uncertain significance. *Genet. Med.* **21**, 1435–1442 (2019).

134. Zuntini, R. *et al.* Dealing With BRCA1/2 unclassified variants in a cancer genetics clinic: Does cosegregation analysis help? *Front. Genet.* **9**, 378 (2018).

135. Exome Variant Server. https://evs.gs.washington.edu/EVS/.

136. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

137. Stratton, M. R. & Rahman, N. The emerging landscape of breast cancer susceptibility. *Nature Genetics* vol. 40 17–22 (2008).

138. Chan, M. *et al.* Development of a next-generation sequencing method for brca mutation screening: A comparison between a high-throughput and a benchtop platform. *J. Mol. Diagnostics* **14**, 602–612 (2012).

139. Ruiz, A. *et al.* Genetic Testing in Hereditary Breast and Ovarian Cancer Using Massive Parallel Sequencing. *Biomed Res Int.* **2014**, (2014).

140. Feliubadaló, L. *et al.* Next-generation sequencing meets genetic diagnostics: Development of a comprehensive workflow for the analysis of BRCA1 and BRCA2 genes. *Eur. J. Hum. Genet.* **21**, 864–870 (2013).

141. Recomendaciones del Estudio Genético de BRCA1 y BRCA2 en Cáncer de Ovario. 1–6 (2016).

142. Tung, N. M. & Garber, J. E. BRCA1/2 testing: therapeutic implications for breast cancer management. *Br. J. Cancer* **119**, 141–152 (2018).

143. Rehm, H. L. *et al.* ACMG clinical laboratory standards for next-generation sequencing. *Genet. Med.* **15**, 733–747 (2013).

144. Nielsen, R., Paul, J. S., Albrechtsen, A. & Song, Y. S. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics* vol. 12 443–451 (2011).

145. De Leeneer, K. *et al.* Massive parallel amplicon sequencing of the breast cancer genes BRCA1 and BRCA2: Opportunities, challenges, and limitations. *Hum. Mutat.* **32**, 335–344 (2011).

146. Cooper, D. N., Ball, E. V. & Krawczak, M. The human gene mutation database. *Nucleic Acids Res.* **26**, 285–287 (1998).

147. Sherry, S. T. *et al.* DbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**,

308–311 (2001).

148. Diez, O., Gutiérrez-Enríquez, S. & Balmaña, J. Heterogeneous prevalence of recurrent BRCA1 and BRCA2 mutations in Spain according to the geographical area: Implications for genetic testing. *Fam. Cancer* **9**, 187–191 (2010).

149. de la Hoya, M. *et al.* Association between *BRCA1* and *BRCA2* mutations and cancer phenotype in Spanish breast/ovarian cancer families: Implications for genetic testing. *Int. J. Cancer* **97**, 466–471 (2002).

150. Miramar, M. D. *et al.* Genetic analysis of BRCA1 and BRCA2 in breast/ovarian cancer families from Aragon (Spain): Two novel truncating mutations and a large genomic deletion in BRCA1. *Breast Cancer Res. Treat.* **112**, 353–358 (2008).

151. Martínez-Ferrandis, J. I. *et al.* Mutational analysis of BRCA1 and BRCA2 in Mediterranean Spanish women with early-onset breast cancer: identification of three novel pathogenic mutations. *Hum. Mutat.* **22**, 417–418 (2003).

152. Llort, G. *et al.* Low frequency of recurrent BRCA1 and BRCA2 mutations in Spain. *Hum. Mutat.* **19**, 307–307 (2002).

153. Gabaldó Barrios, X. *et al.* Molecular characterization and clinical interpretation of BRCA1/BRCA2 variants in families from Murcia (south-eastern Spain) with hereditary breast and ovarian cancer: clinical–pathological features in BRCA carriers and non-carriers. *Fam. Cancer* **16**, 477–489 (2017).

154. Zhong, X. *et al.* Prevalence and Prognostic Role of BRCA1/2 Variants in Unselected Chinese Breast Cancer Patients. *PLoS One* **11**, e0156789 (2016).

155. Azzollini, J. *et al.* Mutation detection rates associated with specific selection criteria for BRCA1/2 testing in 1854 high-risk families: A monocentric Italian study. *Eur. J. Intern. Med.* **32**, 65–71 (2016).

156. Judkins, T. *et al.* Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer* **118**, 5210–5216 (2012).

157. Federici, G. & Soddu, S. Variants of uncertain significance in the era of high-throughput genome sequencing: A lesson from breast and ovary cancers. *J. Exp. Clin. Cancer Res.* **39**, 1–12 (2020).

158. Lubinski, J. *et al.* Cancer variation associated with the position of the mutation in the

BRCA2 gene. *Fam. Cancer* **3**, 1–10 (2004).

159. Thompson, D. & Easton, D. Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *Am. J. Hum. Genet.* **68**, 410–419 (2001).

160. Gayther, S. A. *et al.* Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. *Nat. Genet.* **15**, 103–105 (1997).

161. Risch, H. A. *et al.* Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: A kin-cohort study in Ontario, Canada. *J. Natl. Cancer Inst.* **98**, 1694–1706 (2006).

162. Neuhausen, S. L. *et al.* Haplotype and phenotype analysis of nine recurrent BRCA2 mutations in 111 families: Results of an international study. *Am. J. Hum. Genet.* **62**, 1381–1388 (1998).

163. Meric-Bernstam, F. *et al.* Genotype in BRCA-associated breast cancers. *Breast J.* **19**, 87–91 (2013).

164. Pal, T. *et al.* BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. *Cancer* **104**, 2807–2816 (2005).

165. Woodward, A. M., Davis, T. A., Silva, A. G., Kirk, J. A. & Leary, J. A. Large genomic rearrangements of both BRCA2 and BRCA1 are a feature of the inherited breast/ovarian cancer phenotype in selected families. *J. Med. Genet.* **42**, (2005).

166. De La Hoya, M. *et al.* Genomic rearrangements at the BRCA1 locus in Spanish families with breast/ovarian cancer. *Clin. Chem.* **52**, 1480–1485 (2006).

167. Walsh, T. *et al.* Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *J. Am. Med. Assoc.* **295**, 1379–1388 (2006).

168. Gutiérrez-Enríquez, S. *et al.* Screening for large rearrangements of the BRCA2 gene in Spanish families with breast/ovarian cancer. *Breast Cancer Res. Treat.* **103**, 103–107 (2007).

169. Eccles, E. B. *et al.* BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann. Oncol.* **26**, 2057–2065 (2015).

170. Tavtigian, S. V., Greenblatt, M. S., Goldgar, D. E. & Boffetta, P. Assessing pathogenicity: Overview of results from the IARC unclassified genetic variants working group. *Human Mutation* vol. 29 1261–1264 (2008).

171. Plon, S. E. *et al.* Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human Mutation* vol. 29 1282–1291 (2008).

172. Lee, J. S. *et al.* Reclassification of BRCA1 and BRCA2 variants of uncertain significance: A multifactorial analysis of multicentre prospective cohort. *J. Med. Genet.* **55**, (2018).

173. Pérez-Cabornero, L. *et al.* Evaluating the effect of unclassified variants identified in MMR genes using phenotypic features, bioinformatics prediction, and RNA assays. *J. Mol. Diagnostics* **15**, 380–390 (2013).

174. Radice, P., de Summa, S., Caleca, L. & Tommasi, S. Unclassified variants in BRCA genes: Guidelines for interpretation. *Ann. Oncol.* **22**, (2011).

175. Wallace, A. J. New challenges for BRCA testing: A view from the diagnostic laboratory. *European Journal of Human Genetics* vol. 24 S10–S18 (2016).

176. Spurdle, A. B. *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.* **33**, 2–7 (2012).

177. Balmaña, J. *et al.* Conflicting interpretation of genetic variants and cancer risk by commercial laboratories as assessed by the prospective registry of multiplex testing. *J. Clin. Oncol.* **34**, 4071–4078 (2016).

178. Lavie, O. *et al.* Double heterozygosity in the BRCA1 and BRCA2 genes in the Jewish population. *Ann. Oncol.* **22**, 964–966 (2011).

179. Leegte, B. *et al.* Phenotypic expression of double heterozygosity for BRCA1 and BRCA2 germline mutations. *Journal of medical genetics* vol. 42 (2005).

180. Kang, H. P. *et al.* Design and validation of a next generation sequencing assay for hereditary BRCA1 and BRCA2 mutation testing. *PeerJ* **2016**, e2162 (2016).

181. Hunter, S. *et al.* InterPro: The integrative protein signature database. *Nucleic Acids Res.* **37**, (2009).

182. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence Project: update and current status. *Nucleic Acids Res.* **31**, 34–37 (2003).

183. Pruitt, K. D. *et al.* The consensus coding sequence (CCDS) project: Identifying a common

protein-coding gene set for the human and mouse genomes. *Genome Res.* **19**, 1316–1323 (2009).

184. Mazoyer, S. *et al.* A polymorphic stop codon in BRCA2. *Nat. Genet.* **14**, 253–254 (1996).

185. de la Hoya, M. *et al.* Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum. Mol. Genet.* **25**, 2256–2268 (2016).

ANEXOS

Anexo I: Secuencias de los oligos usados para la secuenciación Sanger

Secuencias de los oligos para la amplificación de cada exón de *BRCA1*.			
OLIGO	**SECUENCIA**	**OLIGO**	**SECUENCIA**
B1E2F	GACGTTGTCATTAGTTCTTTGG	B1E2R	TAAGGTCAATTCTGTTCATTTGC
M13_B1E3F	TGT AAA ACG ACG GCC AGTTGACTCAGTCATAACAGCTC	M13_B1E3R'	CAGGAAACAGCTATGACCGAGAAAGAATGAAA TGGAGTTGG
B1E5F	CTCTTAAGGGCAGTTGTGAG	B1E5R	GAATGGTTTTATAGGAACGCTATG
B1E6F	GTTTTCTACTGTTGCTGCATC	B1E6R	CTTATCACCACGTCATAGAAAG
nB1_E7F	TCACAGGTAACCTTAATGCATTGTCTTAAC	B1E7R'	AGGACTGCTTCTAGCCTGGGCC
B1E8F	TGTTAGCTGACTGATGATGG	B1E8R	CCACAAATACAAATTATGACCAAG
B1E9F	CTGCCACAGTAGATGCTCAG	B1E9R	TTTAAAAAGAGAGAAACATCAATCC
B1E10F	TGGTCAGCTTTCTGTAATCG	B1E10R	GTTGTAAAGGTCCCAAATGG
B1E11aR	AGTCCGCCTATCATTACATG	nB1E11aF	AGTTGGTTGATTTCCACCTCC
B1E11bF	AAACAGCCTGGCTTAGCAAG	B1E11bR	CTCTACTGATTTGGAGTGAAC
nB1E11cF	AGAGAAAATAGACTTACTGG	B1E11cR	TTCTCATTCTGAATAGAATCACC
B1E11dF	AGGGAACTAACCAAACGGAG	B1E11dR	CATGAGTTGTAGGTTTCTGCTG
nB1E11eF	TGCAAATTGATAGTTGTTCTAGC	nB1E11eR	GATACTTTCCTGAGTGCCATAATC
B1E11fF	AAGTGGAGAAAGGGTTTTGC	B1E11fR	TCTTTAAGGACCCAGAGTGG
nB1E11gF	TTGCTCCGTTTTCAAATCCAGG	nnB1E11gR	AGGTGACAATGAATGTTCCTC
B1E11hF	ACCACTTTTTCCCATCAAGTC	B1E11hR	ACCCCTAATCTAAGCATAGC
B1E11iF	GAAAACATTCAAGCAGAACTAGG	B1E11iR	GAGGACTCTAATTTCTTGGC
B1E11jF	CCTAGCCCTTTCACCCATAC	B1E11jR	CTTCCAATTCACTGCACTGTG
B1E11kF	GGAACATCACCTTAGTGAGG	B1E11kR	TGCTCCCCAAAAGCATAAAC
B1E12sigF	GTTGCAGCGTTTATAGTCTG	B1E12sigR	TGTCAGCAAACCTAAGAATGT
B1E13F	GGTGATTTCAATTCCTGTGC	B1E13R	TGTTGGAGCTAGGTCCTTAC
B1E14F	CTAACCTGAATTATCACTATCAG	B1E14R	AAAGTGTATAAATGCCTGTATGC
B1E15F	TGGCTGCCCAGCAAGTATG	B1E15R	AACCAGAATATCTTTATGTAGGA
B1E16F'	AATTCTTAACAGAGACCAGAACTTTG	B1E16R'	AACTCTTTCCAGAATGTTGTTAAGTC
B1E17F	TGTAGAACGTGCAGGATTGC	B1E17R	CATGCGCCACCGTGCCTCGC
B1E18F	AGGCTCTTTAGCTTCTTAGG	nB1E18R	GCATCACCAGCTTATCTGAAC
B1E19F	CTGTCATTCTTCCTGTGCTC	B1E19R	CATTGTTAAGGAAAGTGGTGC
B1E20F	ATATGACGTGTCTGCTCCAC	B1E20R	CTGCAAAGGGGAGTGGAATAC
B1E21F	AAGCTCTTCCTTTTTGAAAGTC	B1E21R	GGTAGAGAAATAGAATAGCCTC
B1E22F	TCTCCCATTGAGAGGTCTTG	B1E22R	GAAGAAGACTTCTGAGGCTAC
B1E23F	CAGAGCAAGACCCTGTCTC	B1E23R	ACTGTGCTACTCAAGCACCA

B1E24F	TGGAGTCGATTGATTAGAGC	nB1E24R	ATGGCAGATTTCCAAGGGAGAC

Secuencias de los oligos para la amplificación de cada exón de *BRCA2*.			
OLIGO	**SECUENCIA**	**OLIGO**	**SECUENCIA**
B2E2F	CAGGAGATGGGACTGAATTAG	B2E2R	GTGACGTACTGGGTTTTTAGC
B2E3F	TGGGTCACAAATTTGTCTGTC	B2E3F	TGGGTCACAAATTTGTCTGTC
B2E4F	AGAATGCAAATTTATAATCCAGAG	B2E4R	AATCAGATTCATCTTTATAGAAC
B2E5_6F	CCAACAATTTATATGAATGAGAATC	B2E5_6R	TCAGGGCAAAGGTATAACGC
B2E7F	CCTTAATGATCAGGGCATTTC	B2E7R	CCTTTTAACAGAAGTATTAGAGATG
B2E8F	GTAGATGTGCTTTTTGATGTCTG	B2E8R	CATATAGGACCAGGTTTAGAG
B2E9F	GATAACTGAAATCACCAAAAGTG	B2E9R	CGGGTGACAGAGCAAGACTC
B2E10aF	CTATGAGAAAGGTTGTGAG	B2E10aR	GTACAACTTCCTTGGAGATTTTG
B2E10bF	CTGAAGTGGAACCAAATGATAC	B2E10bR	GTGGCAAAGAATTCTCTGAAG
B2E10cF	AAGACCTATTAGACACAGAG	B2E10cR	TTAGGTGATTCTCTTATTCTG
B2E10dF	CAGGCAATATCTGGAACTTCTC	B2E10dR	AAGTGGATATTAAACCTGCATTC
B2E10eF	GAAGGAGGACTCCTTATGTC	B2E10eR	AACACAGAAGGAATCGTCATC
B2E11aF	AACACTACCTTTTTAACTTAGTG	B2E11aR	AGCCAAGACCTCTTCTTTTATATCTG
B2E11bF	TACCCCAGAAGCTGATTCTC	B2E11bR	ATTTGTGTTTTGGTTGAATTGTACC
B2E11cF	CCCATGGAAAAGAATCAAGATG	B2E11cR	ATGGGTTCGTTTACACAAGTC
B2E11dF	GAAGAACTTTTCTCAGACAATGAG	B2E11dR	GAAGCTGTTCTGAAGCTACC
nnB2E11eF	TGGTTTATGTTCTTGCAGAGG	nnB2E11eR	AGATCAGCATCTCTGCATTCC
B2E11fF	CTGCACATTTACAGAGTAGTG	B2E11fR	TTGCTGCTGTCTACCTGACC
nB2_E11GF	GAATGCAGAGATGCTGATCTTCA	nnB2_E11GR	CAGAATCATGACATTTACTTGAAGA
B2E11hF	ACTTCTGCAGAGGTACATCC	B2E11hR	TCCCTCCTTCATAAACTGGC
nnB2E11iF	CGGAAGTTTGCTGGCCTGTTGAAA	nnB2E11iR	TTTCATCACGTTCGGGTTGTCCCT
B2E11jF	TTTTTTCAGACTGCAAGTGGG	B2E11jF	TTTTTTCAGACTGCAAGTGGG
B2E11kF	AAAGAAATCTTTGGACAAAGTG	B2E11kR	CTTGTGTAAAAAGCTAAGGCTG
B2E11lF	GTCCTGCAACTTGTTACAC	B2E11lR	TACTGGCTCAATACCAGAATC
B2E11mF	TAGCATGTCTAACAGCTATTCC	B2E11mR	CTGTCTGTAAATATGTCTTTCAC
M13_B2E11nF	TGT AAA ACG ACG GCC AGTGCATTTAGGATAGCCAGTGG	M13_B2E11nR	CAG GAA ACA GCT ATG ACCGTATTTGCAGATGAGACTGAC
B2E11oF	CACCTTGTGATGTTAGTTTGG	B2E11oR	GAAACTTGCTTTCCACTTGC
B2E11pF	CAAAAAGGCTTTTCATATAATGTGG	B2E11pR	GTGAGACTTTGGTTCCTAATAC
B2E11qF	ATGTTGAAGGTGGTTCTTCAG	B2E11qR	CCTCTTCTTTTTCCAATTCTTG
B2E11rF	GAAACAGAAGCAGTAGAAATTGC	B2E11rR	ACCATACTCCCCCAAACTGAC
B2E12F	CATTAGGTCACTATTTGTTGTAAG	nB2E12R	CCAGCACTTTGGAGAGGCAG
B2E13F	GAGCATCCGTTACATTCACTG	B2E13R	GGGAAGTGTTAACTTCTTAACG
B2E14aF	TGTAGCAAATGAGGGTCTGC	B2E14aR	TGCTTTTTGTCTGTTTTCCTCC
B2E14bF	TTTAGCAGTTTCAGGACATCC	B2E14bF	TTTAGCAGTTTCAGGACATCC

nB2e15F	GATTACAGGCGTGAGCCACT	nB2e15R	TCATTCATCCATTCCTGCAC
B2E16F	TTTGGTAAATTCAGTTTTGGTTTG	B2E16R	GCCAACTTTTTAGTTCGAGAG
B2E17F	CAGAGAATAGTTGTAGTTGTTG	B2E17R	TCACAGACTACACAGAAACC
B2E18aF	GATCCACTATTTGGGGATTGC	B2E18aR	ATCTAACTGGGCCTTAACAGC
B2E18bF	GCAGATACCCAAAAAGTGGC	B2E18bR	TCTGGACCTCCAAAAAACTG
M13_B2E19F	TGT AAA ACG ACG GCC AGTGAAAACTCTTATGATATCTGTAATA G	M13_B2E19R	CAG GAA ACA GCT ATG ACCAGAGACCGAAACTCCATCTC
B2E20F	ACTGTGCTTGGCCTGATAC	B2E20R	ATGTTAAATTCAAAGTCTCTAAGAC
B2E21F	GGGTGTTTTATGCTTGGTTC	B2E21R	GCATTTCAACATATTCCTTCCTG
B2E22F	AACCACACCCTTAAGATGAGC	B2E22R	GGGCATTAGTAGTGGATTTTGC
B2E23F	GAGAATAATATAAAAGAGGATCTG	B2E23R	TCAACATATGGAGATTCCATAAAC
B2E24F	GCAGCGACAAAAAAAACTCAG	B2E24R	ATTTGCCAACTGGTAGCTCC
B2E25F	GAGCTTTCGCCAAATTCAGC	B2E25R	TTCCTTGATACTGGACTGTC
B2E26F	GTCCCAAACTTTTCATTTCTGC	B2E26R	AGGAGCCACATAACAACCACA
B2E27aF	TTAGGGGAGGGAGACTGTG	B2E27aR	GGAGTCATCTGAGGAGAATTC
nnB2E27bF	AAGAGAAGAGCCTTGGATTTCTTGAG	nB2E27bR	GATCAGAGATGTAGTACAACGTCG
B2E27cF	CTGTCAGTGAATCCACTAGGA	B2E27cR	ATTTCTTTGCTCATTGTGCAAC

B1: gen *BRCA1*; B2: gen *BRCA2*; E: exón*; F: Forward;* R*: Reverse*

Anexo II: Concentraciones de reactivos y condiciones de PCR para generar amplicones para los genes *BRCA1* y *BRCA2*.

Concentraciones finales de cada uno de los reactivos en la PCR		
	Concentración del reactivo en el stock	**Concentración en la PCR**
Desoxirribonucleótidos	100 mM de cada: dATP, dGTP, dCTP y dTTP	200 μM de cada: dATP, dGTP, dCTP y dTTP
Tampón de PCR	10X/5X	1X
Solución $MgCl_2$	25 mM	2.5 mM
Oligos (F+R)	100 μM	20 μM
Taq gold	5U/μl	1U
H_2O	Ultrapura estéril desionizada	q.s.

Condiciones de PCR y tamaño del fragmento que generan para el gen *BRCA1*:				
OLIGOS		**TAMAÑO**	**PROGRAMA**	**MIX**
B1E2F	B1E2R	318	60TD55	2,5mM $MgCl_2$
M13_B1E3F	M13_B1E3R	307	60TD55	2,5mM $MgCl_2$
B1E5F	B1E5R	280	60TD55	2,5mM $MgCl_2$

B1E6F	B1E6R	298	60TD55	2,5mM $MgCl_2$
nB1E7F	B1E7R	376	60TD55	2,5mM MgCl2
B1E8F	B1E8R	318	60TD55	2,5mM $MgCl_2$
B1E9F	B1E9R	254	60TD55	2,5mM $MgCl_2$
B1E10F	B1E10R	288	60TD55	2,5mM $MgCl_2$
nB1E11aF	B1E11aR	424	60TD55	2,5mM $MgCl_2$
B1E11bF	B1E11bR	435	60TD55	2,5mM $MgCl_2$
nB1E11cF	B1E11cR	420	60TD55	2,5mM $MgCl_2$
B1E11dF	B1E11dR	407	60TD55	2,5mM $MgCl_2$
nB1E11eF	nB1E11eR	434	60TD55	2,5mM $MgCl_2$
B1E11fF	B1E11fR	425	60TD55	2,5mM $MgCl_2$
nB1E11gF	nnB1E11gR	400	60TD55	2,5mM $MgCl_2$
B1E11hF	B1E11hR	321	60TD55	2,5mM $MgCl_2$
B1E11iF	B1E11iR	440	60TD55	2,5mM $MgCl_2$
B1E11jF	B1E11jR	346	60TD55	2,5mM $MgCl_2$
B1E11kF	B1E11kR	330	60TD55	2,5mM $MgCl_2$
B1E12Fsig	B1E12R	233	60TD55	2,5mM $MgCl_2$
B1E13F	B1E13R	370	60TD55	2,5mM $MgCl_2$
B1E14F	B1E14R	315	60TD55	2,5mM $MgCl_2$
B1E15F	B1E15R	338	60TD55	2,5mM $MgCl_2$
B1E16F	B1E16R	448	60TD55	2,5mM $MgCl_2$
B1E17F	B1E17R	278	60TD55	2,5mM $MgCl_2$
B1E18F	nB1E18R	337	60TD55	2,5mM $MgCl_2$
B1E19F	B1E19R	249	60TD55	2,5mM $MgCl_2$
B1E20F	B1E20R	232	60TD55	2,5mM $MgCl_2$
B1E21F	B1E21R	299	60TD55	2,5mM $MgCl_2$
B1E22F	B1E22R	300	60TD55	2,5mM $MgCl_2$
B1E23F	B1E23R	400	60TD55	2,5mM $MgCl_2$
B1E24F	nB1E24R	409	60TD55	2,5mM $MgCl_2$

Condiciones de PCR y tamaño del fragmento que generan para el gen *BRCA2*:

OLIGOS		TAMAÑO	PROGRAMA	MIX
B2E2F	B2E2R	308	60TD55	2,5mM $MgCl_2$
B2E3F	B2E3R	409	60TD55	2,5mM $MgCl_2$
B2E4F	B2E4R	249	60TD55	2,5mM $MgCl_2$
B2E5_6F	B2E5_6R	359	60TD55	2,5mM $MgCl_2$

B2E7F	B2E7R	247	60TD55	2,5mM $MgCl_2$
B2E8F	B2E8R	280	60TD55	2,5mM $MgCl_2$
B2E9F	B2E9R	400	60TD55	2,5mM $MgCl_2$
B2E10aF	B2E10aR	451	60TD55	2,5mM $MgCl_2$
B2E10bF	B2E10bR	276	60TD55	2,5mM $MgCl_2$
B2E10cF	B2E10cR	262	60TD55	2,5mM $MgCl_2$
B2E10dF	B2E10dR	292	60TD55	2,5mM $MgCl_2$
B2E10eF	B2E10eR	333	60TD55	2,5mM $MgCl_2$
B2E11aF	B2E11aR	383	60TD55	2,5mM $MgCl_2$
B2E11bF	B2E11bR	463	60TD55	2,5mM $MgCl_2$
B2E11cF	B2E11cR	323	60TD55	2,5mM $MgCl_2$
B2E11dF	B2E11dR	401	60TD55	2,5mM $MgCl_2$
nnBRCA2E11e	nnB2E11eR	652	60TD55	2,5mM $MgCl_2$
B2E11fF	B2E11fR	334	60TD55	2,5mM $MgCl_2$
nB2_E11GF	nnB2_E11GR	334	60TD55	2,5mM MgCl2
B2E11hF	B2E11hR	378	60TD55	2,5mM $MgCl_2$
nnB2E11iF	nnB2E11iR	976	60TD55	2,5mM $MgCl_2$
B2E11jF	nB2E11jR	430	60TD55	2,5mM $MgCl_2$
B2E11kF	B2E11kR	417	60TD55	2,5mM $MgCl_2$
B2E11lF	B2E11lR	377	60TD55	2,5mM $MgCl_2$
B2E11mF	B2E11mR	378	BM50-40	2,5mM $MgCl_2$
M13_B2_E11nF	M13_B2_E11nF	395	60TD55	2,5mM $MgCl_2$
B2E11oF	B2E11oR	364	60TD55	2,5mM $MgCl_2$
B2E11pF	B2E11pR	394	60TD55	2,5mM $MgCl_2$
B2E11qF	B2E11qR	412	60TD55	2,5mM $MgCl_2$
B2E11rF	B2E11rR	400	60TD55	2,5mM $MgCl_2$
B2E12F	nB2E12R	400	60TD55	2,5mM $MgCl_2$
B2E13F	B2E13R	400	60TD55	2,5mM $MgCl_2$
B2E14aF	B2E14aR	389	60TD55	2,5mM $MgCl_2$
B2E14bF	B2E14bR	424	60TD55	2,5mM $MgCl_2$
B2E15F	nB2E15R	323	60TD55	2,5mM $MgCl_2$
B2E16F	B2E16R	379	60TD55	2,5mM $MgCl_2$

B2E17F	B2E17R	318	60TD55	2,5mM $MgCl_2$
B2E18aF	B2E18aR	456	60TD55	2,5mM $MgCl_2$
B2E18bF	B2E18bR	400	60TD55	2,5mM $MgCl_2$
M13_B2E19F	M13_B2E19R	321	60TD55	2,5mM $MgCl_2$
B2E20F	B2E20R	400	60TD55	2,5mM $MgCl_2$
B2E21F	B2E21R	400	60TD55	2,5mM $MgCl_2$
B2E22F	B2E22R	455	60TD55	2,5mM $MgCl_2$
B2E23F	B2E23R	332	50-40	2,5mM $MgCl_2$
B2E24F	B2E24R	365	60TD55	2,5mM $MgCl_2$
B2E25F	B2E25R	461	60TD55	2,5mM $MgCl_2$
B2E26F	B2E26R	380	60TD55	2,5mM $MgCl_2$
B2E27aF	B2E27aR	446	60TD55	2,5mM $MgCl_2$
nnB2E27bF	nB2E27bR	375	60TD55	2,5mM $MgCl_2$
B2E27cF	B2E27cR	277	60TD55	2,5mM $MgCl_2$

Nota: El programa 60TD55, se basa en comenzar la reacción de PCR a temperaturas alta de 60ºC, durante 23 clicos, y se baja a 55ºC durante 14 ciclos. El programa BM50-40, la temperatura de extensión es de 50ºC durante 39 ciclos.

Anexo III: Reactivos y condiciones de PCR para la secuenciación:

Reacción de secuenciación	
REACTIVOS	**Volúmen**
DNA molde	Máximo 14.4 µl (cantidad correspondiente según cuantificación)
Cebador (1 µM)	3.2 µl
Kit secuenciación	0,5 µl
Tampón TOM (10X)	1.9 µl
Agua ultrapura estéril	Hasta 20 µl

Programa de amplificación de los genes *BRCA1* y *BRCA2*.		
Temperatura	**Tiempos**	**Ciclos**
96º C	60 segundos	

96º C	10 segundos	30 ciclos
50º C	5 segundos	
60º C	4 minutos	
4º C	pausa	

Anexo IV: Campos contenidos en los ficheros bioinformáticos, así como la clasificación de las variantes en diferentes categorías dependiendo de la frecuencia

Anotaciones del *pipeline* propio
HGNC_symbol: símbolo HGNC del gen.
Chr: cromosoma.
Pos: posición genómica.
Ref_Allele: alelo en la referencia.
Var_Allele: alelo en la muestra.
Sample_Genotype: genotipo identificado en la muestra en función de la frecuencia del alelo alternativo
Sample_Depth: profundidad de lectura alcanzada en esta posición.
Sample_Var/Depth: frecuencia del alelo alternativo.
Gene_ID: código de identificación del gen en *Ensembl*.
Gene_description: descripción del gen.
HGVSc_name: nomenclatura de la variante siguiendo los estándares de la HGVS a nivel de cDNA.
Intrón: número de intrón.
Exón: número de exón.
HGVSp_name: nomenclatura de la variante siguiendo los estándares de la HGVS a nivel proteico.
Variant_effect: efecto de la variante a nivel transcripcional.
ALL_MAF: frecuencia alélica global menor obtenida a partir de los datos de la fase 1 de "1000 Genomes".
AFR_MAF: frecuencia alélica menor en la población africana.
AMR_MAF: frecuencia alélica menor en la población americana.
ASN_MAF: frecuencia alélica menor en la población del este asiático.
EUR_MAF: frecuencia alélica menor en la población europea.
InterPro_IDs: código del dominio proteico afectado en InterPro[181].

InterPro_descriptions: descripción del dominio proteico en InterPro.
HGMD_info: código de identificación de la HGMD Professional, nomenclatura HGNC del cambio con el transcrito según la base de datos RefSeq[182] y nombre de la enfermedad asociada.
Related_publication: artículos relacionados con la variante.
Existing_variation: código de identificación de la variante en caso de existir.
Transcript_ID: código de identificación del transcrito en *Ensembl*.
RefSeq_ID: código de identificación del transcrito en RefSeq.
CCDS_ID: código de identificación del transcrito en CCDS[183].
Canonical_isoform: indica si la isoforma a la que afecta la variante es la isoforma canónica o no.
Conservation_score: valor de conservación GERP de las bases afectadas por la variante.
Grantham_distance: distancia de Grantham.
Condel_prediction: predicción del efecto de la variante a nivel proteico según el predictor Condel.
SIFT_prediction: predicción del efecto de la variante a nivel proteico según SIFT.
PolyPhen_prediction: predicción del efecto de la variante a nivel proteico según PolyPhen.
Flanking_sequence: secuencia a ambos lados de la variante, 50pb upstreams y 50pb downstreams.

Clasificación de las variantes atendiendo a su frecuencia alélica		
Categoría	**Frecuencia en SNVs**	**Frecuencia en *indels***
Homo_ref	0	0
P_Homo_ref	>0-0,12	NA
UNC_Hetero	>0,12-0,35	>0-0,3
P_Hetero	>0,35-0,65	>0,3-0,6
UNC_Homo_var	>0,65-0,85	NA
P_Homo_var	>0,85-1	>0,6-1

Anexo V: Descripción de los criterios de clasificación de variantes según el grado de evidencia de patogenicidad y de benignidad, así como el algoritmo que emplea la combinación de ambos para clasificar las variantes.

Criterios de clasificación de VP	
Evidencia de patogenicidad	**Categoría**

Muy fuerte	PVS1: variante nula (sin sentido, con alteración del marco de lectura, se alteran ±1 sitios de *splicing*, codón de inicio, deleción de uno o varios exones) en un gen en el que la pérdida de función es un mecanismo conocido de enfermedad Advertencias: - Atención a los genes en los que la pérdida de función no es un mecanismo conocido de enfermedad. - Precaución a la hora de interpretar las variantes con pérdida de función en el extremo 3´ o al final del gen. - Atención en las variantes de *splicing* que predicen la pérdida de un exón pero que dejan intacta el resto de la proteína. - Precaución cuando haya presencia de múltiples transcritos. Por ejemplo, la variante *BRCA2*: c.9976A>T (p.Lys3326*), no confiere un alto riesgo de cáncer de mama familiar[184]. Y tampoco confiere riesgo la variante *BRCA1* c.594-2A>C[185].
Fuerte	PS1: El mismo cambio de aminoácido ha sido establecido como VP independientemente del cambio de nucleótido. Advertencia: precaución con los cambios que afectan el *splicing* en lugar de un cambio a nivel proteico. PS2: Variante *de novo* en una paciente con la enfermedad, pero sin historia familiar (se debe confirmar tanto la maternidad como la paternidad). PS3: Estudios funcionales *in vitro* o *in vivo* bien establecidos que apoyen el efecto dañino en el gen o su producto. PS4: La prevalencia de la variante en individuos afectados está incrementada significativamente al compararla con la prevalencia en los controles
Moderado	PM1: Localizado en un punto caliente mutacional y/o un dominio funcional bien establecido. PM2: ausente en controles en el proyecto *1000 Genomas*, *Exome Sequencing Project, Exome Aggregation Consortium*. Advertencia: los datos poblacionales para inserciones/deleciones pueden ser poco adecuados si se obtienen por secuenciación de nueva generación. PM3: para enfermedades recesivas, detectada en *trans* con una VP. PM4: la longitud de la proteína cambia como resultado de inserción/deleción en zonas no repetidas o por variantes con pérdida del codón de parada. PM5: Nueva variante de tipo *missense* en el mismo residuo don de un cambio *missense* distinto se ha definido como patogénico. PM6: La variante se asume como *de novo*, pero sin confirmación de la paternidad y maternidad.
Secundario	PP1: Co-segregación de la variante con la enfermedad en varios familiares afectados en un gen relacionado con la enfermedad. PP2: Variante *missense* en un gen con una baja tasa de variantes missense benignas y en el que las variantes tipo missense son un mecanismo común de enfermedad. PP3: Varias líneas de evidencia computacional apoyan un efecto deletéreo en un gen o su producto. Advertencia: dado que muchos algoritmos *in silico* usan la misma entrada o parecidas cada algoritmo no debe contabilizarse como un criterio independiente. PP4: El fenotipo del paciente o la historia familiar es altamente específica para una enfermedad con una única etiología genética. PP5: Fuentes reconocidas informaron de manera reciente dicha variante como patogénica, pero no hay evidencia disponible para que el laboratorio realice una evaluación independiente.

Criterios de clasificación de VB	
Evidencia de Patogenicidad	**Categoría**
Independiente	BA1: La frecuencia alélica es >5% en el proyecto 1000 genomas, *Exome Sequencing Project* o en el *Exome Aggregation Consortium*
Fuerte	BS1: La frecuencia alélica es mayor que la esperada para dicha enfermedad. BS2: Se ha observado en adultos sanos para enfermedades recesivas (en homocigosis), dominantes (en heterocigosis) o ligadas al cromosoma X (en hemicigosis) para las que se espera una penetrancia completa temprana. BS3: Estudios funcionales in vitro o in vivo muestran que no se produce un efecto dañino en la función de la proteína o *splicing*. Advertencia: la presencia de fenocopias en fenotipos comunes puede simular una ausencia de segregación entre individuos afectos. Además, puede haber familias con más de una VP que contribuyan a una enfermedad de herencia autosómica dominante, lo que también puede confundir sobre una aparente ausencia de segregación. BS4: Ausencia de segregación en miembros afectados de una familia

Secundario	BP1: Variante *missense* en un gen en el que las variantes truncadas son las que principalmente causan enfermedad. BP2: Se observa en *trans* con una VP en una enfermedad dominante de penetrancia completa o se observa en *cis* con una VP en una enfermedad con cualquier tipo de herencia. BP3: Inserciones/deleciones en regiones repetitivas sin función conocida. BP4: Múltiples líneas de evidencia computacional sugieren que no se altera el gen ni su producto (conservación, evolución, impacto en el *splicing*, etc.) BP5: La variante se ha hallado en un individuo con una base molecular alternativa para la enfermedad. BP6: Fuentes de confianza clasifican la variante como benigna, pero no hay evidencia disponible para realizar una evaluación independiente. BP7: Es una variante sinónima y los algoritmos de predicción del *splicing* predicen que no se alteran los sitios consenso ni se crea un nuevo sitio de *splicing* ni se trata de un nucleótido altamente conservado.

Reglas de combinación de criterios para la clasificación de variantes	
Evidencia de Patogenicidad	**Categoría**
VP	1) Muy fuerte (PVS1) y a) ≥1 Fuerte (PS1-PS4) o b) 2 Moderados (PM1-PM6) o c) 1 Moderado (PM1-PM6) y un Secundario (PP1-PP5) o d) ≥2 Secundarios (PP1-PP5) 2) ≥2 Fuertes (PS1-PS4) o 3) 1 Fuerte (PS1-PS4) y a) ≥3 Moderados (PM1-PM6) o b) 2 Moderados (PM1-PM6) y ≥2 Secundarios o c) 1 Moderado (PM1-PM6) y ≥4 Secundarios
VPP	1) Muy fuerte (PVS1) y 1 Moderado (PM1-PM6) o 2) 1 Fuerte (PS1-PS4) y 1-2 Moderado (PM1-PM6) o 3) 1 Fuerte (PS1-PS4) y ≥2 Secundarios (PP1-PP5) o 4) ≥3 Moderados (PM1-PM6) o 5) 2 Moderados (PM1-PM6) y ≥2 Secundarios (PP1-PP5) 6) 1 Moderado (PM1-PM6) y ≥4 Secundarios (PP1-PP5)
VB	1) 1 Independiente (BA1) o 2) ≥2 Fuertes (BS1-BS4)
VPB	1) Fuerte (BS1-BS4) y 1 Secundario (BP1-BP7) o 2) ≥Secundarios (BP1-BP7)
VSI	1) No se cumplen los criterios mencionados anteriormente 2) Los criterios de benignidad y patogenicidad son contradictorios

Anexo VI: Descripción de las variantes detectadas en las 12 muestras utilizadas para la validación en los genes *BRCA1* y *BRCA2*.

Detección de variantes en las muestras de validación en el gen *BRCA2*

					Muestra 1						**Muestra 2**						**Muestra 3**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**
32890572	G	A	*rs1799943*	5´UTR	D	het	0,5	P_het	1972	0,5	-		-		-		D	het	0,5	P_het	1984	0,53
32893197	AT	A	*rs276174878*	variante intrónica	n/d		0,5	UNC_het	1489	0,25	n/d		0,5	UNC_het	1727	0,27	n/d		0,5	UNC_het	1233	0,23
32900933	T	A	*rs3752451*	variante intrónica	D	het	0,5	P_het	988	0,51	-		-		-		D	het	0,5	P_het	798	0,50
32903685	C	T	*rs2126042*	variante intrónica	-		-		-		-		-		-		D	het	0,5	P_het	1458	0,48
32905141	C	T	-	missense	-		-		-		D	het	0,5	P_het	990	0,53	-		-		-	
32906729	A	C	*rs144848*	missense	-		-		-		D	het	0,5	P_het	929	0,51	-		-		-	
32907535	CT	C	*rs276174816*	variante intrónica	n/d		0,5	P_het	477	0,38	n/d		0,5	P_het	707	0,50	n/d		0,5	P_het	687	0,42
32910430	C	T	*rs28897711*	sinónima	-		-		-		-		-		-		D	het	0,5	P_het	505	0,52
32911888	A	G	*rs1801406*	sinónima	D	het	0,5	P_het	1083	0,49	-		-		-		D	het	0,5	P_het	1212	0,50
32912299	T	C	*rs543304*	sinónima	D	het	0,5	P_het	1974	0,48	D	het	0,5	P_het	1980	0,49	-		-		-	
32912560	G	A	*rs28897724*	sinónima	D	het	0,5	P_het	655	0,47	-		-		-		D	het	0,5	P_het	667	0,49
32913055	A	G	*rs206075*	sinónima	D	hom	1	P_hom_var	1982	1	D	hom	1	P_hom_var	1970	1	D	hom	1	P_hom_var	1977	1
32913804	G	A	*rs80358755*	missense	D	het	0,5	P_het	1235	0,55	-		-		-		-		-		-	
32915005	G	C	*rs206076*	sinónima	D	hom	1	P_hom_var	823	1	D	hom	1	P_hom_var	770	1	D	hom	1	P_hom_var	714	1
32915410	CAATT	C	*rs138193280*	variante intrónica	D	het	0,5	P_het	802	0,5	-		-		-		D	het	0,5	P_het	799	0,79
32915429	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		-		-		-	
32929387	T	C	*rs169547*	missense	n/d		1	P_hom_var	1181	1	n/d		1	P_hom_var	1091	1	n/d		1	P_hom_var	1182	1
32936646	T	C	*rs9534262*	variante intrónica	-		-		-		-		-		-		D	het	0,5	P_het	822	0,48
32953388	T	C	*rs4942486*	variante intrónica	-		-		-		-		-		-		D	het	0,5	P_het	815	0,49
32972984	C	A	-	3´UTR	-		-		-		n/d		0,5	P_het	n/d		-		-		-	
32973012	A	C	*rs15869*	3´UTR	D	het	0,5	P_het	532	0,5	D	-	-	-	585	0,54	-		-		-	

Detección de variantes en las muestras de validación en el gen *BRCA2*

					Muestra 4						**Muestra 5**						**Muestra 6**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**
32890572	G	A	*rs1799943*	5´UTR	D	het	0,5	P_het	1858	0,52	D	het	0,5	P_het	1971	0,5	-		-		-	
32893197	AT	A	*rs276174878*	variante intrónica	n/d		0,5	UNC_het	1367	0,23	n/d		0,5	UNC_het	1768	0,25	n/d		0,5	UNC_het	1496	0,23
32900473	T	C	TMP_ESP_13_	variante intrónica	-		-		-		-		-		-		n/d		0,5	P_het	15	0,60
32900933	T	A	*rs3752451*	variante intrónica	D	het	0,5	P_het	928	0,50	D	het	0,5	P_het	972	0,49	-		-		-	

32903685	C	T	*rs2126042*	variante intrónica	-		-		-		D	het	0,5	P_het	1355	0,51	D	het	0,5	P_het	1510	0,49								
32906729	A	C	*rs144848*	missense	-		-		-		-		-		-		D	het	0,5	P_het	1172	0,48								
32907535	CT	C	*rs276174816*	variante intrónica	n/d		0,5	P_het	725	0,36	n/d		0,5	P_het	732	0,43	n/d		0,5	P_het	668	0,50								
32911888	A	G	*rs1801406*	sinónima	D	het	0,5	P_het	1102	0,48	D	het	0,5	P_het	1084	0,48	-		-		-									
32912299	T	C	*rs543304*	sinónima	D	het	0,5	P_het	1972	0,50	-		-		-		-		-		-									
32913055	A	G	*rs206075*	sinónima	D	hom	1	P_hom_var	1976	1	D	hom	1	P_hom_var	1967	1	D	hom	1	P_hom_var	1992	1								
32915005	G	C	*rs206076*	sinónima	D	hom	1	P_hom_var	1976	1	D	hom	1	P_hom_var	1140	1	D	hom	1	P_hom_var	907	1								
32915410	CAATT	C	*rs138193280*	variante intrónica	D	het	0,5	P_het	820	0,49	D	het	0,5	P_het	919	0,59	-		-		-									
32915429	C	A	-	variante intrónica	n/d		0,5		n/d		n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d									
32929232	A	G	*rs1799955*	sinónima	D	het	0,5	UNC_het	1095	0,29	D	het	0,5	P_het	1520	0,34	-		-		-									
32929387	T	C	*rs169547*	missense	n/d		1	P_hom_var	1347	1	n/d		1	P_hom_var	1019	1	n/d		1	P_hom_var	1235	1								
32936646	T	C	*rs9534262*	variante intrónica	D	het	0,5	P_het	682	0,51	D	hom	1	P_hom_var	794	1	D	het	0,5	P_het	798	0,52								
32953388	T	C	*rs4942486*	variante intrónica	D	het	0,5	P_het	414	0,5	D	hom	1	P_hom_var	803	1	D	het	0,5	P_het	628	0,51								
32973012	A	C	*rs15869*	3´UTR	D	het	0,5	P_het	555	0,5	-		-		-		-		-		-									

Detección de variantes en las muestras de validación en el gen *BRCA2*

					Muestra 7						**Muestra 8**						**Muestra 9**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var /depth**
32890572	G	A	*rs1799943*	5´UTR	D	het	0,5	P_het	1974	0,52	-		-		-		-		-		-	
32893197	AT	A	*rs276174878*	variante intrónica	n/d		0,5	UNC_het	1727	0,24	n/d		0,5	UNC_het	1305	0,23	n/d		0,5	UNC_het	1379	0,23
32900933	T	A	*rs3752451*	variante intrónica	D	het	0,5	P_het	936	0,49	-		-		-		-		-		-	
32903685	C	T	*rs2126042*	variante intrónica	-		-		-		D	het	0,5	P_het	1501	0,48	-		-		-	
32906729	A	C	*rs144848*	missense	-		-		-		D	het	0,5	P_het	1273	0,47	D	het	0,5	P_het	1266	0,45
32907535	CT	C	*rs276174816*	variante intrónica	n/d		0,5	P_het	505	0,46	n/d		0,5	P_het	744	0,43	n/d		0,5	P_het	813	0,37
32911888	A	G	*rs1801406*	sinónima	D	het	0,5	P_het	1222	0,52	-		-		-							
32912299	T	C	*rs543304*	sinónima	-		-		-		-		-		-		D	het	0,5	P_het	1977	0,51
32913055	A	G	*rs206075*	sinónima	D	hom	1	P_hom_var	1978	1	D	hom	1	P_hom_var	1982	1	D	hom	1	P_hom_var	1985	1
32914209	**ACTCT**	**A**	**-**	**frameshift**	-		-		-		**D**	**het**	**0,5**	**P_het**	**1413**	**0,68**	-		-		-	
32915005	G	C	*rs206076*	sinónima	D	hom	1	P_hom_var	852	1	D	hom	1	P_hom_var	597	1	D	hom	1	P_hom_var	718	1
32913076	C	T	*rs80359788*	sinónima	D	het	0,5	P_het	1930	0,52	-		-		-							
32915410	CAATT	C	*rs138193280*	variante intrónica	n/d		0,5	P_het	943	0,56	-		-		-							
32915429	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d	
32918675	T	A	-	variante intrónica	-		-		-		D	het	0,5	P_het	982	0,50						
32929387	T	C	*rs169547*	missense	n/d		1	P_hom_var	1181	1	n/d		1	P_hom_var	1288	1	n/d		1	P_hom_var	1275	1
32936646	T	C	*rs9534262*	variante intrónica	D	hom	1	P_hom_var	970	1	D	hom	1	P_hom_var	772	1	-		-		-	
32953388	T	C	*rs4942486*	variante intrónica	D	hom	1	P_hom_var	686	1	D	hom	1	P_hom_var	471	1	-		-		-	

3298877	**TAA**	**T**	**-**	**frameshift**	**D**	**het**	**0,5**	**P_het**	**1989**	**0,49**	-	-	-	-	-	-						
32972984	C	A	-	3´UTR	n/d		0,5	P_het	n/d		-	-	-	-	-	-						
32973012	A	C	*rs15869*	3´UTR	-		-		-		-	-	-	-			D	het	0,5	P_het	490	0,50

Detección de variantes en las muestras de validación en el gen *BRCA2*

					Muestra 10						Muestra 11						Muestra 12					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/depth**
32890572	G	A	*rs1799943*	5´UTR	-		-		-		D	het	0,5	P_het	1974	0,49	-		-		-	
32893197	AT	A	*rs276174878*	variante intrónica	n/d		0,5	UNC_het	1510	0,23	n/d		0,5	UNC_het	1641	0,24	n/d		0,5	UNC_het	1424	0,22
32900775	C	T	-	variante intrónica	-		-		-		D	het	0,5	P_het	1964	0,52	-		-		-	
32900933	T	A	*rs3752451*	variante intrónica	-		-		-		D	het	0,5	P_het	860	0,52	-		-		-	
32906565	**C**	**CA**	**-**	**frameshift**	-		-		-		**D (perfiles)**	**het**	**0,5**	**P_het**	**1972**	**0,50**	-		-		-	
32906729	A	C	*rs144848*	missense	D	het	0,5	P_het	1244	0,49	D	het	0,5	P_het	1022	0,51	D	hom	1	P_hom_var	1258	1
32907535	CT	C	*rs276174816*	variante intrónica	n/d		0,5	P_het	388	0,47	n/d		0,5	P_het	570	0,58	n/d		0,5	P_het	811	0,41
32911888	A	G	*rs1801406*	sinónima	-		-		-		D	het	0,5	P_het	1050	0,49	-		-		-	
32912299	T	C	*rs543304*	sinónima	D	het	0,5	P_het	1981	0,51	-		-		-		-		-		-	
32912751	**A**	**AT**	**-**	**frameshift**	**D**	**het**	**0,5**	**P_het**	**1318**	**0,51**	-		-		-		-		-		-	
32913055	A	G	*rs206075*	sinónima	D	hom	1	P_hom_var	1976	1	D	hom	1	UNC_het	1977	0,11	D	hom	1	P_hom_var	1982	1
32915005	G	C	*rs206076*	sinónima	D	hom	1	P_hom_var	934	1	D	hom	1	P_hom_var	662	1	D	hom	1	P_hom_var	710	1
32915410	CAATT	C	*rs138193280*	variante intrónica	-		-		-		D	het	0,5	P_het	799	0,79	-		-		-	
32915429	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		-		-		-	
32929387	T	C	*rs169547*	missense	n/d		1	P_hom_var	1170	1	n/d		1	P_hom_var	938	0,52	n/d		1	P_hom_var	865	0,49
32936646	T	C	*rs9534262*	variante intrónica	D	het	0,5	P_het	927	0,50	D	het	0,5	P_het	899	0,50	D	het	0,5	P_het	367	0,49

32953388	T	C	*rs4942486*	variante intrónica	D	het	0,5	P_het	572	0,50	D	het	0,5	P_het	574	0,51	D	het	0,5	P_het	994	0,35
32972984	C	A	-	3´UTR	-		-		-		n/d		0,5	P_het	n/d		-		-		-	
32973012	A	C	*rs15869*	3´UTR	D	het	0,5	P_het	717	0,50	-		-		-		D	het	0,5	P_het	529	0,51

Detección de variantes en las muestras de validación en el gen *BRCA1*

					Muestra 1						**Muestra 2**						**Muestra 3**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var /depth**
41215825	C	T	*rs3092994*	variante intrónica	D	het	0,5	P_het	587	0,49	D	hom	1	P_hom_var	639	0,99	-		-		-	
41219780	T	C	*rs8176234*	variante intrónica	D	het	0,5	P_het	1112	0,53	D	hom	1	P_hom_var	1007	1	-		-		-	
41219804	T	C	*rs8176233*	variante intrónica	D	het	0,5	P_het	1107	0,52	D	hom	1	P_hom_var	994	1	-		-		-	
41223094	T	C	*rs1799966*	missense	D	het	0,5	P_het	1926	0,51	D	hom	1	P_hom_var	1925	1	-		-		-	
41226601	G	C	*rs8176212*	variante intrónica	D	het	0,5	P_het	725	0,49	D	hom	1	P_hom_var	649	1	-		-		-	
41234470	A	G	*rs1060915*	sinónima	D	het	0,5	P_het	968	0,49	D	hom	1	P_hom_var	1036	1	-		-		-	
41243190	T	G	*rs799916*	variante intrónica	D	het	0,5	P_het	865	0,49	D	hom	1	P_hom_var	917	0,99	-		-	-		
41244000	T	C	*rs16942*	missense	D	het	0,5	P_het	393	0,6	D	hom	1	P_hom_var	376	1	-		-		-	
41244192	G	C	-	missense	-		-		-		D	het	0,5	P_het	1194	0,49	-		-		-	
41244429	C	T	*rs4986852*	missense	D	het	0,5	P_het	1566	0,52	-		-		-		-		-		-	
41244435	T	C	*rs16941*	missense	D	het	0,5	P_het	1560	0,48	D	hom	1	P_hom_var	1546	1	-		-		-	
41244936	G	A	*rs799917*	missense	D	het	0,5	UNC_hom	875	0,65	D	hom	1	P_hom_var	1079	1	-		-		-	
41245237	A	G	*rs16940*	sinónima	D	het	0,5	P_het	1117	0,48	D	hom	1	P_hom_var	1084	1	-		-		-	
41245466	G	A	*rs1799949*	sinónima	D	het	0,5	P_het	543	0,53	D	hom	1	P_hom_var	554	1	-		-		-	
41245471	C	T	*rs4986850*	missense	-		-		-		D	het	0,5	P_het	554	0,54	-		-		-	
41249363	TA	T	*rs8176144*	variante intrónica	D (perfiles)	het	0,5	P_het	1369	0,47	D (perfiles)	hom	1	P_hom_var	1556	0,99	-		-		-	
41251931	G	A	*rs799923*	variante intrónica	D	het	0,5	P_het	893	0,52	-		-		-		-		-		-	
41256402	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d	
41276247	A	G	*rs3765640*	variante intrónica	D	het	0,5	P_het	745	0,49	D	hom	1	P_hom_var	385	1	-		-		-	

Detección de variantes en las muestras de validación en el gen *BRCA1*

					Muestra 4						**Muestra 5**						**Muestra 6**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var /depth**

41215825	C	T	*rs3092994*	variante intrónica	D	het	0,5	P_het	480	0,52	D	het	0,5	P_het	563	0,54	-		-		-	
41215920	**G**	**T**	***rs28897696***	**missense**	**D**	**het**	**0,5**	**P_het**	**733**	**0,62**	-		-		-		-		-		-	
41219780	T	C	*rs8176234*	variante intrónica	D	het	0,5	P_het	1081	0,52	D	het	0,5	P_het	962	0,53	-		-		-	
41219804	T	C	*rs8176233*	variante intrónica	D	het	0,5	P_het	1073	0,52	D	het	0,5	P_het	953	0,52	-		-		-	
41223094	T	C	*rs1799966*	missense	D	het	0,5	P_het	1926	0,48	D	het	0,5	P_het	1918	0,50	-		-		-	
41226601	G	C	*rs8176212*	variante intrónica	D	het	0,5	P_het	672	0,50	D	het	0,5	P_het	565	0,54	-		-		-	
41234470	A	G	*rs1060915*	sinónima	D	het	0,5	P_het	1123	0,48	D	het	0,5	P_het	1103	0,52	-		-		-	
41243039	**T**	**A**	**-**	**sinónima**	-		-		-		-		-		-		**D**	**het**	**0,5**	**P_het**	**1621**	0,49
41243040	**G**	**GC**	**-**	**frameshift**	-		-		-		-		-		-				**0,5**	**P_het**	**1620**	0,54
41243190	T	G	*rs799916*	variante intrónica	D	het	0,5	P_het	726	0,52	D	het	0,5	P_het	798	0,47	-		-		-	
41244000	T	C	*rs16942*	missense	D	het	0,5	P_het	398	0,50	D	het	0,5	P_het	354	0,48	-		-		-	
41244435	T	C	*rs16941*	missense	D	het	0,5	P_het	1281	0,50	D	het	0,5	P_het	1292	0,50	-		-		-	
41244936	G	A	*rs799917*	missense	D	het	0,5	UNC_hom	794	0,62	D	het	0,5	UNC_hom	804	0,66	-		-		-	
41245237	A	G	*rs16940*	sinónima	D	het	0,5	P_het	937	0,50	D	het	0,5	P_het	885	0,46	-		-		-	
41245466	G	A	*rs1799949*	sinónima	D	het	0,5	P_het	558	0,49	D	het	0,5	P_het	352	0,49	-		-		-	
41249263	G	A	*rs1799965*	Región splicing	D	het	0,5	P_het	1170	0,49	-		-		-		-		-		-	
41249363	TA	T	*rs8176144*	variante intrónica	D (perfiles)	het	0,5	P_het	1186	0,50	D (perfiles)	het	0,5	P_het	1459	0,48	-		-		-	
41251931	G	A	*rs799923*	variante intrónica	D	het	0,5	P_het	893	0,52	-		-		-		-		-		-	
41256402	C	A	-	variante intrónica	D	het	0,5	P_het	n/d		n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d	
41276040	**G**	**GGACA**	**-**	**frameshift**	-		-		-		**D**	**het**	**0,5**	**P_het**	**721**	**0,45**	-		-		-	
41276247	A	G	*rs3765640*	variante intrónica	D	het	0,5	P_het	663	0,49	D	het	0,5	P_het	711	0,49	-		-		-	

Detección de variantes en las muestras de validación en el gen *BRCA1*

					Muestra 7						Muestra 8						Muestra 9					
					454 GS Junior		*MiSeqReporter*		*Pipeline* propio		*454 GS Junior*		*MiSeqReporter*		*Pipeline* propio		*454 GS Junior*		*MiSeqReporter*		*Pipeline* propio	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var /depth**
41215825	C	T	*rs3092994*	variante intrónica	D	het	0,5	P_het	630	0,51	D	het	0,5	P_het	559	0,47	D	het	0,5	P_het	597	0,49
41219780	T	C	*rs8176234*	variante intrónica	D	het	0,5	P_het	1207	0,46	D	het	0,5	P_het	866	0,52	D	het	0,5	P_het	958	0,49
41219804	T	C	*rs8176233*	variante intrónica	D	het	0,5	P_het	1202	0,45	D	het	0,5	P_het	866	0,51	D	het	0,5	P_het	935	0,48
41223094	T	C	*rs1799966*	missense	D	het	0,5	P_het	1950	0,51	D	het	0,5	P_het	1887	0,52	D	het	0,5	P_het	1897	0,51
41226601	G	C	*rs8176212*	variante intrónica	D	het	0,5	P_het	779	0,53	D	het	0,5	P_het	568	0,52	D	het	0,5	P_het	604	0,51
41234470	A	G	*rs1060915*	sinónima	D	het	0,5	P_het	1270	0,50	D	het	0,5	P_het	1081	0,52	D	het	0,5	P_het	846	0,46

41243190	T	G	*rs799916*	variante intrónica	D	het	0,5	P_het	924	0,51	D	hom	1	P_hom_var	763	1	D	het	0,5	P_het	801	0,51
41244000	T	C	*rs16942*	missense	D	het	0,5	P_het	459	0,53	D	het	0,5	P_het	361	0,62	D	het	0,5	P_het	375	0,53
41244130	T	C	*rs2227945*	missense	-		-		-		D	het	0,5	P_het	1413	0,44						
41244435	T	C	*rs16941*	missense	D	het	0,5	P_het	1364	0,47	D	het	0,5	P_het	1303	0,51	D	het	0,5	P_het	1459	0,49
41244936	G	A	*rs799917*	missense	D	het	0,5	P_het	1244	0,49	D	hom	1	P_hom_var	1011	1	D	het	0,5	P_het	953	0,51
41245237	A	G	*rs16940*	sinónima	D	het	0,5	P_het	1112	0,45	D	het	0,5	P_het	940	0,48	D	het	0,5	P_het	1094	0,53
41245466	G	A	*rs1799949*	sinónima	D	het	0,5	P_het	618	0,50	D	het	0,5	P_het	485	0,57	D	het	0,5	P_het	455	0,53
41245471	C	T	*rs4986850*	missense	-		-		-		-		-		-		D	het	0,5	P_het	488	0,49
41246481	T	C	*rs1799950*	missense	-		-		-		-		-		-		D	het	0,5	P_het	1259	0,46
41249363	TA	T	*rs8176144*	variante intrónica	D (perfiles)	het	0,5	P_het	1515	0,50	D (perfiles)	het	0,5	P_het	899	0,53	D (perfiles)	het	0,5	P_het	1137	0,51
41251931	G	A	*rs799923*	variante intrónica	D	het	0,5	P_het	893	0,52	-		-		-		D	het	0,5	P_het	729	0,48
41256402	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d	
41276247	A	G	*rs3765640*	variante intrónica	D	het	0,5	P_het	848	0,53	D	het	0,5	P_het	660	0,50	D	het	0,5	P_het	762	0,47

Detección de variantes en las muestras de validación en el gen *BRCA1*

					Muestra 10						**Muestra 11**						**Muestra 12**					
					454 GS Junior		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**		***454 GS Junior***		***MiSeqReporter***		***Pipeline* propio**	
Posición	**Ref**	**Alt**	**ID**	**Consecuencia**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var/ depth**	**D-n/d**	**Genotipo**	**Genotipo (AF)**	**Genotipo**	**Depth**	**Var /depth**
41215825	C	T	*rs3092994*	variante intrónica	D	hom	1	P_hom_var	489	1	-		-		-		-		-		-	
41219780	T	C	*rs8176234*	variante intrónica	D	hom	1	P_hom_var	1299	1	-		-		-		-		-		-	
41219804	T	C	*rs8176233*	variante intrónica	D	hom	1	P_hom_var	1291	1	-		-		-		-		-		-	
41223094	T	C	*rs1799966*	missense	D	hom	1	P_hom_var	1925	1	-		-		-		-		-		-	
41226601	G	C	*rs8176212*	variante intrónica	D	hom	1	P_hom_var	515	1	-		-		-		-		-		-	
41234470	A	G	*rs1060915*	sinónima	D	hom	1	P_hom_var	1292	1	-		-		-		-		-		-	
41243190	T	G	*rs799916*	variante intrónica	D	hom	1	P_hom_var	791	1	-		-		-		-		-	-		
41243509	G	A	-	missense	-		-		-		-		-		-		D	het	0,5	P_het	766	0,51
41244000	T	C	*rs16942*	missense	D	hom	1	P_hom_var	488	1	-		-		-		-		-		-	
41244435	T	C	*rs16941*	missense	D	hom	1	P_hom_var	1419	1	-		-		-		-		-		-	
41244936	G	A	*rs799917*	missense	D	hom	1	P_hom_var	1142	1	-		-		-		-		-		-	
41245237	A	G	*rs16940*	sinónima	D	hom	1	P_hom_var	845	1	-		-		-		-		-		-	
41245466	G	A	*rs1799949*	sinónima	D	hom	1	P_hom_var	695	1	-		-		-		-		-		-	
41245471	C	T	*rs4986850*	missense	D	het	0,5	P_het	688	0,52	-		-		-		-		-		-	
41246481	T	C	*rs1799950*	missense	-		-		-		D	het	0,5	P_het	-		-		-		-	
41249363	TA	T	*rs8176144*	variante intrónica	D (perfiles)	hom	1	P_hom_var	1243	1	-		-		-		-		-		-	
41251931	G	A	*rs799923*	variante intrónica	-		-		-		D	het	0,5	P_het	1263	0,51	D	het	0,5	P_het	1389	0,49
41256402	C	A	-	variante intrónica	n/d		0,5	P_het	n/d		n/d		0,5	P_het	n/d		-		-		-	
41276048	**T**	**AA**	**-**	**frameshift**	**-**		**-**		**-**		**-**		**-**		**-**		**D**	**het**	**0,5**	**P_het**	**574**	**0,51**
41276247	A	G	*rs3765640*	variante intrónica	D	hom	1	P_hom_var	731	1	-		-		-		-		-		-	

Het: Heterocigosis; n/d: no determinado; D: deteminado; hom: homocigoto

Anexo VII: Variantes reclasificadas a probablemente benignas o benignas detectadas en nuestra serie de pacientes en los genes *BRCA1* y *BRCA2*.

Variantes reclasificadas detectadas en nuestra serie de pacientes en los genes *BRCA1* y *BRCA2*.											
Cambio nucleotídico	**Variante informada (año)**	**dbSNP** **MAF**	**ClinVar**	**LOVD**	**HGMD**	**BRCA** **Exchange**	**Observaciones** **(Caso índice)**	**Reclasificación 2015**	**Criterios ACMG**	**Reclasificación 2020**	**Reclasificación 2023**
BRCA1: c.81-12dup	VSI (2013)	*rs273902789* dupG=0,0008 % (1/125568, TOPMED) dupG=0,0293 % (30/102400, ExAC)	125523 VB/VPB	NR	NR	NR	114 (CM) con VP c.3680_3681del en *BRCA2*	VSI	PM2, BP7, BP6	VSI	VPB
BRCA1: c.276T>G (p.Ala92Ala)	VSI (2013)	*rs2054677953* Sin frecuencias	NR	NR	NR	NR	Co-ocurre con la VP c.4456del en el gen *BRCA1* (caso índice 185)	VPB	PM2, BP7, BP5	VPB	
BRCA1: c.536A>G (p.Tyr179Cys)[Ψ]	VSI (2013)	*rs56187033* C=0,0258% (65/251464, GnomAD_exome) C=0,0358% (45/125568, TOPMED)	37661 VB (Panel de expertos)	BRCA1_000090 VB, VSI	CM030786 DM?	NR	264 (CM, co-ocurre con p.Asn550His y p.Phe486Leu)	VSI	PM2, PP3, BS3, BP6	VPB	

		C=0,0272% (33/121410, ExAC)									
BRCA1: c.591C>T (p.Cys197Cys)	VSI (2013)	*rs1799965* A=0,1468% (369/251304, GnomAD_exome) A=0,0765% (96/125568, TOPMED) A=0,1469% (178/121172, ExAC)	55642 VB	BRCA1_00096 VB	CM051394 DM?	VB	154 (CM). 216 (CMO) y 230 (CO)	VB	PM2, BP4, BP6	VB	
BRCA1: c.981A>G (p.Thr327Thr)	VSI (2012)	*rs1800063* C=0,1014% (255/251408, GnomAD_exome) C=0,0693% (87/125568, TOPMED) C=0,0848% (103/121402, ExAC)	55771 VB	BRCA1_00116 VPB	NR	VB	125 (CC Y CP)	VB	BS1, BS2, BP7	VB	
BRCA1: c.1456T>C (p.Phe486Leu)$^{\Psi}$	VSI (2013)	*rs55906931* G=0,0406% (51/125568, TOPMED) G=0,0297% (36/121388, ExAC)	54258 VB	BRCA1_00015 VPB/VSI	NR	VB	264 (CM, co-ocurre con p.Tyr179Cys y p.Asn550His)	VSI	PM2, BP6	VB	

BRCA1: c.1486C>T (p.Arg496Cys)	VSI (2012)	*rs28897676* A=0,0235% (59/251190, GnomAD_exome) A=0,0263% (33/125568, TOPMED) A=0,0173% (21/121362, ExAC)	37416 VB	BRCA1_000151 VB/VSI	CM065005 DM?	VB	87 (CM)	VSI	PM2, BP6	VB	
BRCA1: *c.1648A>C (p.Asn550His)*[ψ]	VSI (2013)	*rs56012641* G=0,0259% (65/250836, GnomAD_exome) G=0,0358% (45/125568, TOPMED) G=0,0273% (33/121032, ExAC)	37423 VB	BRCA1_005209 VSI	CM025218 DM?	VB	264 (CM, co-ocurre con p.Tyr179Cys y p.Phe486Leu)	VSI	PM2, BP6	VB	
BRCA1: c.2733A>G (p.Gly911Gly)	VSI (2012)	*rs1800740* C=0,1017% (255/250784, GnomAD_exome) C=0,0693% (87/125568, TOPMED) C=0,0849% (103/121366, ExAC)	54664 VPB	BRCA1_001116 VPB/VSI	NR	VPB	125 (CC Y CP)	VPB	BS2, BP7, BS1, BP6	VB	

BRCA1: c.3024G>A (p.Met1008Ile)	VSI (2012)	*rs1800704* T=0,1020% (256/251024, GnomAD_exom e) T=0,0693% (87/125568, TOPMED) T=0,0858% (104/121194, ExAC)	41814 VB	BRCA1_0022 4 VB/VSI	CM08671 9 DM?	VB	125 (CC Y CP)	VB	BS2, BS1, BP6	VB	
BRCA1: c.3119G>A (p.Ser1040Asn)	VSI (2009)	*rs4986852* T=1,3154% (3302/251032, GnomAD_exom e) T=2,0532% (3138/152838, ALFA Project) T=1,5713% (1973/125568, TOPMED)	17670 VB	BRCA1_0002 33 VB	CM94017 5 DM?	VB	Se presenta en 17 familias	VB	BS3, BS1, BS2, BP6	VB	
BRCA1: c.4039A>G (p.Arg1347Gly)	VSI (2011)	*rs28897689* C=0,3823% (960/251086, GnomAD_exom e) C=0,4810% (604/125568, TOPMED) C=0,3980% (483/121368, ExAC)	41821 VB	BRCA1_0002 85 VB	CM96018 1 DM?	VB	Se presenta en 4 casos índice	VB	BS1, BS2, BS3, BP6	VB	

BRCA1: c.4535G>T (p.Ser1512Ile)	VSI (2013) MAF: 0,001/1	*rs1800744* A=0,2371% (596/251366, GnomAD_exom e) A=0,2548% (320/125568, TOPMED) A=0,2151% (261/121344, ExAC)	41826 VB	BRCA1_0003 24 VB	CM96018 3 DM?	VB	Se presenta en 4 casos índice	VB	BS1, BS2, BS3, BP6	VB	
BRCA1: c.4812A>G (p.Gln1604Gln)	VSI (2012)	*rs28897693* C=0,1098% (276/251404, GnomAD_exom e) C=0,0701% (88/125568, TOPMED) C=0,1285% (156/121402, ExAC)	55293 VB	BRCA1_0003 37 VB	NR	VB	117 (NR)	VSI	BS1, BS2, BP6, BP7	VB	
BRCA1: c.4910C>T (p.Pro1637Leu)	VSI (2012)	*rs80357048* A=0,0032% (8/251356, GnomAD_exom e) A=0,0016% (2/125568, TOPMED) A=0,0025% (3/121412, ExAC)	55314 VB	BRCA1_0003 45 VB/VSI	NR	VB	69 (NR)	VSI	PM2, BS3, BP6	VB	

BRCA1: c.4956G>A (p.Met1652Ile)	VSI (2012)	*rs1799967* T=1,7993% (4515/250924, GnomAD_exome) T=0,9031% (1134/125568, TOPMED) T=1,7619% (2139/121406, ExAC)	41830 VB	BRCA1_00348 VB	CM014325 DM?	VB	Se presenta en 10 casos índices	VB	BS2, BP6	VB	
BRCA1: c.5075-53C>T	VSI (2012)	*rs8176258* A=1,7210% (2161/125568, TOPMED) A=1,825% (573/31396, GnomAD)	125760 VB	BRCA1_001032 VB/VSI	NR	VB	En nuestra cohorte co-ocurre con la VP en *BRCA1* (familia 148)	VB	BS2, BP6	VB	
BRCA1: c.5152+20T>A	VSI (2012)	*rs376836050* T=0,0159% (40/250844, GnomAD_exome) T=0,0080% (10/125568, TOPMED) T=0,0085% (10/117128, ExAC)	132704 VB/ VPB	BRCA1_001211 VB/ VSI	NR	No revisada	115 (CM)	VSI	PM2, BP6, BP7	VB	
BRCA1: *c.5158A>G (p.Thr1720Ala)*	VSI (2013)	*rs56195342* C=0,0175% (44/250984, GnomAD_exome)	55438 VB	BRCA1_000409 VB/ VSI	NR	VB	Se presenta en dos casos índices	VSI	PM2, BS3, BP6	VB	

		C=0,0080% (10/125568, TOPMED) C=0,0092% (11/120042, ExAC)									
BRCA1: *c.5176A>G (p.Arg1726Gly)*	VSI (2012)	*rs80357501* C=0,0052% (13/251160, GnomAD_exome) C=0,0008% (1/125568, TOPMED) C=0,0058% (7/120114, ExAC)	55443 VB/ VPB	BRCA1_000411 VB/VSI	NR	No revisada	92 (CM)	VSI	PM2, BS3, BP6	VPB	
BRCA1: c.5277+48_5277+59dup	VSI (2013)	*rs273901756* NR	125812 VSI	BRCA1_000416 VB	NR	VB	269 (NR)	VSI	BP6, BP7	VB	
BRCA1: c.5411T>A (p.Val1804Asp)	VSI (2012)	*rs80356920* T=0,0099% (25/251482, GnomAD_exome) T=0,0215% (27/125568, TOPMED) T=0,0058% (7/121360, ExAC)	55568 VB	BRCA1_000470 VB/VSI	CM044859 DM?	VB	81 (NR), 274 (Estatus de portador)	VB	PM1, PM2, BP6. BP4, BP1	VB	
BRCA2: c.68-7T>A	VSI (2009)	*rs81002830* A=0,2872% (712/247874,	52187 VB	BRCA2_000013	CS033491 DFP	VB	En nuestra cohorte co-ocurre con la VP c.4456del	VB	BP4, BP6,	VB	

		GnomAD_exome) A=0,1975% (248/125568, TOPMED) A=0,2382% (282/118368, ExAC)		VB/VSI			en *BRCA1*. Se presenta en 4 casos índices		BS1, BS2		
BRCA2: c.68-7dup	VSI (2012)	*rs276174878* dupT=0,043% (13/30444, GnomAD)	38065 VB/VPB	BRCA2_003412 VPB/VSI	NR	No revisada	130 Co-ocurre con VP c.213-2A>C en *BRCA1* (CO)	VPB	BS1, BP6	VPB	
BRCA2: c.125A>G (p.Tyr42Cys)	VSI (2013)	*rs4987046* G=0,1628% (409/251278, GnomAD_exome) G=0,1585% (199/125568, TOPMED) G=0,1696% (205/120892, ExAC)	37734 VB	BRCA2_000016 VB/VSI	CM970178 DM?	VB	249, 163 (NR)	VB	BP4, BP6, BS1	VB	
BRCA2: c.198A>G (p.Gln66Gln)	VSI (2013)	*rs28897700* G=0,1082% (272/251392, GnomAD_exome) G=0,0175% (22/125568, TOPMED) G=0,1072% (130/121222, ExAC)	51234 VB	BRCA2_001203 VPB/ VSI	CM173583 DM?	VB	165 (CM) En nuestra serie de casos co-ocurre con la VP en *BRCA2* c.5621_5624del	VB	BA1, BS2, BP7, BP6	VB	

BRCA2: c.223G>C (p.Ala75Pro)	VSI (2009)	*rs28897701* C=0,0259% (65/251354, GnomAD_exom e) C=0,0271% (34/125568, TOPMED) C=0,0165% (20/121192, ExAC)	51258 VB	BRCA2_0000 20 VPB/VSI	CM04553 5 DM?	VB	33, 95, 105 (NR)	VB	PM2, BP6, BS3, BS2	VB	
BRCA2: c.425+33A>G	VSI (2013)	*rs200065709* G=0,0365% (72/197194, GnomAD_exom e) G=0,0295% (37/125568, TOPMED) G=0,052% (31/59244, ExAC)	225712 VB/VPB	BRCA2_0020 17 VB/VSI	NR	No revisada	186 (CM)	VSI	BP7, BP6	VPB	
BRCA2: *c.502C>A* *(p.Pro168Thr)*	VSI (2012)	*rs80358726* A=0,0040% (10/251180, GnomAD_exom e) A=0,0072% (9/125568, TOPMED) A=0,0008% (1/119560, ExAC)	51757 VB	BRCA2_0000 33 VB/VSI	NR	VB	108 (NR)	VB	PM2, BP4, BP6	VB	

BRCA2: c.516+14C>T	VSI (2011)	*rs182828913* T=0,0359% (90/250544, GnomAD_exome) T=0,0287% (36/125568, TOPMED) T=0,0308% (36/116930, ExAC)	132764 VB/VPB	BRCA2_000030 VSI/ VPB	NR	No revisada	42 (CM)	VPB	BS1. BP6	VPB	
BRCA2: c.517-11T>C	VSI (2013)	*rs81002828* C=0,0016% (4/251194, GnomAD_exome) C=0,0008% (1/125568, TOPMED) C=0,0033% (4/119782, ExAC)	51797 Conflicto en la interpretación de la variante: VPB (3), VSI (2)	NR	NR	No revisada	Co-ocurre con VP c.2957dup en *BRCA2* (Familia 166)	VB	PM2. BP6, BP7	VB	
BRCA2 : c.632-27C>G	VSI (2013)	*rs746474565* G=0,0021% (5/235788, GnomAD_exome) G=0,004% (3/85436, ExAC)	NR	NR	NR	NR	119 (asintomática)	VSI	PM2, BP7	VSI	VPB
BRCA2: c.707A>T (p.His236Leu)	VSI (2013)	*rs80358938* G=0,0004% (1/250320,	52263 VSI	NR	NR	NR	Co-ocurre con VP c.3680_3681del en *BRCA2* (Familia 114)	VPB	PM2, BP4, BP2	VPB	

		GnomAD_exome) G=0,0009% (1/110882, ExAC)									
BRCA2 : *c.927A>G (p.Ser309Ser)*	VSI (2013)	*rs80359806* G=0,0047% (11/233954, GnomAD_exome) G=0,0042% (5/120356, ExAC)	52804 VPB	BRCA2_001486 VPB/VSI	NR	VPB	252 (CM)	VSI	PM2, BP7, BP6	VPB	
BRCA2: c.2803G>A (p.Asp935Asn)	VSI (2012)	*rs28897716* A=0,0653% (82/125568, TOPMED) A=0,054% (17/31402, GnomAD)	51349 VB/ VPB	BRCA2_000082 VPB/ VSI	CM994285 DM?	No revisada	90 (Estatus de portador)	VPB	BP4, BS1,BP6	VPB	
BRCA2: c.2971A>G (p.Asn991Asp)	VSI (2011)	*rs1799944* G=5,4087% (13545/250428, GnomAD_exome) G=3,7694% (5605/148698, ALFA Project) G=4,6103% (5789/125568, TOPMED)	41545 VB	BRCA2_000085 VB/VSI	NR	VB	Se reportó en 13 casos índices	VB	BP4, BS2, BP6, BA1	VB	

BRCA2: c.4068G>A (p.Leu1356Leu)	VSI (2012)	*rs28897724* A=0,2986% (736/246498, GnomAD_exome) A=0,3122% (392/125568, TOPMED) A=0,3049% (368/120706, ExAC)	51589 VB	BRCA2_001186 VB/VSI	NR	VB	Se reportó en 5 casos índices	VB	BP7, BS1, BS2, BP6	VB	
BRCA2: c.4258G>T (p.Asp1420Tyr)	VSI (2012)	*rs28897727* T=0,6497% (1589/244580, GnomAD_exome) T=0,4253% (534/125568, TOPMED) T=0,6796% (813/119634, ExAC)	41549 VB	BRCA2_000115 VB	CM003133 DM?	VB	Se reportó en 4 casos índices	VB	BP4, BS2, PB6, BA1	VB	
BRCA2: c.4584C>T (p.Ser1528Ser)	VSI (2013)	*rs80359788* T=0,0319% (40/125568, TOPMED)	51677 VPB	BRCA2_000125 VPB/VSI	NR	VPB	Co-ocurre con la VP c.9310_9311del en *BRCA2* (caso índice 122)	VPB	BP7, BP6	VPB	
BRCA2: c.4681C>A (p.His1561Asn)	VSI (2012)	*rs2219594* A=0,0637% (160/251048, GnomAD_exome)	51695 Conflicto en la interpretación de la variante: VB (9), VPB (2), VSI (1)	BRCA2_001897 VSI/ VPB	CM162083 9 DM?	No revisada	Se reportó en un caso índice	VSI	BS2, BP4, BP6, BA1	VPB	

		A=0,2740% (344/125568, TOPMED) A=0,0729% (88/120662, ExAC)									
BRCA2: c.5199C>T (p.Ser1733Ser)	VSI (2013)	*rs28897734* T=0,4506% (1111/246572, GnomAD_exome) T=0,3631% (456/125568, TOPMED) T=0,4504% (538/119462, ExAC)	51813 VB	BRCA2_002206 VSI/ VPB	CM173520 DM?	VB	En nuestra serie de casos co-ocurre con la VP en *BRCA1* c.5419del	VB	BP7, BS2, BP6, BP2	VB	
BRCA2: c.5455C>T (p.Pro1819Ser)	VSI (2013)	*rs80358768* T=0,0064% (8/125568, TOPMED) T=0,019% (6/31378, GnomAD) T=0,008% (1/13006, GO-ESP)	51865 VB	BRCA2_000137 VSI/ VPB	CM161827 DM?	VB	Se reportó en un caso índice	VB	PM2, BP4, BP6	VB	
BRCA2: c.5744C>T (p.Thr1915Met)	VSI (2008)	*rs4987117* T=1,7616% (4420/250908, GnomAD_exome)	41556 VB	BRCA2_000145 VB/VSI	CM010170 DP	VB	Se reportó en 6 casos índices	VB	BP4, BS2, BP6, BA1	VB	

		T=2,6916% (4971/184684, ALFA Project) T=1.7441% (2190/125568, TOPMED)									
BRCA2: c.5882G>A (p.Ser1961Asn)	VSI (2009)	*rs80358820* A=0,0026% (7/264690, TOPMED) T=0,0008% (1/120970, ExAC)	38004 Conflicto en la interpretación de la variante: VB (1), VPB (5), VSI (4)	BRCA2_002146 VPB	NR	NR	32 (CM Bilateral)	VSI	PM2, BP4, BP6	VSO	VPB
BRCA2: c.6100C>T (p.Arg2034Cys)	VSI (2012)	rs1799954 T=0,3128% (785/250958, GnomAD_exome) T=0,4865% (662/136084, ALFA Project) T=0,3791% (476/125568, TOPMED)	41558 VB	BRCA2_000152 VB	CM994286 DP	VB	Se reportó en 7 casos índices	VB	BP4, BS1, BS2, BP6	VB	
BRCA2: c.6323G>A (p.Arg2108His)	VSI (2012)	*rs35029074* A=0,1163% (286/245876, GnomAD_exome) A=0,2357% (296/125568, TOPMED)	38037 VB	BRCA2_000159 VB/VSI	CM035693 DM?	VB	Se reportó en un caso índice	VB	BP4, BS2, BP6, BA1	VB	

		A=0,1277% (153/119850, ExAC)									
BRCA2: c.6412G>T (p.Val2138Phe)	VSI (2012)	*rs11571659* T=0,0728% (171/234900, GnomAD_exome) T=0,3385% (425/125568, TOPMED) T=0,0815% (96/117784, ExAC)	52087 Conflicto en la interpretación de la variante: VB (9), VPB (4), VSI (1)	BRCA2_000161 VB/VSI	CM162084 DM?	No revisada	77 (asintomática, dos familiares de primer grado con CM, co-ocurre con c.4681C>A (p.His1561Asn), c.7805+6C>G y c.9648+54G>A)	VSI	BP4, BS2, BP6, BA1	VPB	
BRCA2: c.6842-20T>A	VSI (2012)	*rs81002811* A=0,0243% (58/238540, GnomAD_exome) A=0,0247% (31/125568, TOPMED) A=0,0098% (10/102522, ExAC)	38069 VB	BRCA2_000169 VB/VSI	NR	VB	Se reportó en dos casos índices	VSI	BP7, BS1, BS2, BP6	VB	
BRCA2: c.7232A>C (p.Lys2411Thr)	VSI (2012)	*rs80358950* C=0,0060% (15/251296, GnomAD_exome) C=0,0024% (3/125568, TOPMED)	52294 VB	BRCA2_000182 VB/VSI	CM171864 DM?	VB	182 (CM, co-ocurre con c.4553A>G (p.Gln1518Arg) en *BRCA1*	VSI	PM2, BP6, BS3	VB	

		C=0,0049% (6/121356, ExAC)									
BRCA2: c.7397T>C (p.Val2466Ala)	VSI (2011)	*rs169547* T=0,5124% (1287/251168, GnomAD_exome) T=2.1192% (2661/125568, TOPMED) T=0,6282% (762/121304, ExAC)	133738 VB	BRCA2_000186 VB/VSI	CM960194 DM?	No revisada	Se reportó en dos casos índices	VB	BP4, BP6, BA1	VB	
BRCA2: c.7469T>C (p.Ile2490Thr)	VSI (2013)	*rs11571707* C=1,9346% (4864/251424, GnomAD_exome) C=0,9127% (1146/125568, TOPMED) C=1,4357% (1737/120986, ExAC)	96852 VB	BRCA2_000195 VB/VSI	CM065019 DM?	VB	Se reportó en dos casos índices	VB	BS2, BP6, BA1	VB	
BRCA2: c.7805+6C>G	VSI (2012)	*rs81002819* G=0,0625% (157/251168, GnomAD_exome) G=0,2732% (343/125568, TOPMED)	52412 VB/VPB	BRCA2_001208 VB/ VSI	NR	No revisada	77 (asintomática, dos familiares de primer grado con CM co-ocurre con c.9648+54G>A, c.6412G>T (p.Val2138Phe), c.7805+6C>G y c.4681C>A	VSI	BS2, BP6, BP4, BA1	VPB	

		G=0,0719% (87/121060, ExAC)					(p.His1561Asn)				
BRCA2: c.8503T>C (p.Ser2835Pro)	VSI (2013)	*rs11571746* C=0,0549% (138/251178, GnomAD_exome) C=0,1903% (239/125568, TOPMED) C=0,0725% (88/121298, ExAC)	52606 VB	BRCA2_000338 VB/VSI	CM004191 DM?	VB	Co-ocurre con la VP c.5419del en *BRCA1*, caso índice 173	VB	PM2, BP6. BP4	VB	
BRCA2: c.8850G>T (p.Lys2950Asn)	VSI (2013)	*rs28897754* T=0,0669% (84/125568, TOPMED)	52692 Conflicto en la interpretación de variante: VB (11), VPB (12), VSI (2)	BRCA2_000379 VB/ VSI	CM014327 DM?	No revisada	En nuestra serie de casos se reportó en 3 casos índices, en uno de ellos co-ocurre con la VP en *BRCA1* c.3770_3771del	VPB	BP6, BS1	VPB	
BRCA2: c.8851G>A (p.Ala2951Thr)	VSI (2012)	*rs11571769* A=0,9550% (2395/250796, GnomAD_exome) A=0,7207% (905/125568, TOPMED) A=0,7853% (948/120714, ExAC)	41570 VB	BRCA2_000380 VPB/VSI	CM970186 DM?	VB	Se reportó en tres casos índices	VB	PM5, BS2, BP6,BA1	VB	

BRCA2: c.9256+58A>T	VSI (2013)	*rs777427197* T=0,0104% (13/125568, TOPMED) T=0,016% (5/31362, GnomAD)	NR	BRCA2_001368 VB/VSI	NR	NR	En nuestra serie de casos se reportó en dos casos índices, en uno de ellos co-ocurre con la VP en *BRCA2* c.2808_2811del	VPB	PM2, BP2, BP4,	VPB	
BRCA2: c.9257-10dup	VSI (2012)	*rs276174919* delT=0,003% (1/31352, GnomAD)	91520 VPB/VB	NR	NRV	NR	51 (CM 69 años)	VSI	PM2, BP6, BP7	VSI	VPB
BRCA2: c.9257-16T>C	VSI (2012)	*rs11571818* C=0,6585% (1610/244488, GnomAD_exome) C=0,5479% (688/125568, TOPMED) C=0,7647% (812/106186, ExAC)	52790 VB	BRCA2_000409 VB/VSI	NR	VB	Se reportó en dos casos índices	VB	BP7, BS2,BP6, BA1	VB	
BRCA2: c.9292T>C (p.Tyr3098His)	VSI (2009)	*rs41293521* C=0,0291% (73/250874, GnomAD_exome) C=0,0223% (28/125568, TOPMED)	38227 VB	BRCA2_00442 VB/VSI	CM970188 DM?	VB	Se reportó en un caso índice	VB	BS1, PB6	VB	

		C=0,0264% (31/117284, ExAC)									
BRCA2: c.9375C>G (p.Leu3125Leu)	VSI (2013)	*rs276174924* Sin frecuencias	52824 VPB	NR	NR	VPB	Co-ocurre con la VPP c.793+1G>A en *BRCA2* (familia 133)	VPB	PM2, BP6, BP7, BP2	VPB	
BRCA2 : c.9648+54G>A	VSI (2012)	*rs11571823* A=0,3608% (453/125568, TOPMED) A=0,350% (110/31390, GnomAD)	52881 VPB	BRCA2_000473 VB/VSI	NR	No revisado	77 (asintomática, dos familiares de primer grado con CM), co-ocurre con otras VSI c.6412G>T (p.Val2138Phe), c.4681C>A (p.His1561Asn) y c.7805+6C>G	VSI	BS2, BP7, BP6, BA1	VPB	
BRCA2: c.10110G>A (p.Arg3370Arg)	VSI (2013)	rs28897762 A=0,1401% (352/251248, GnomAD_exome) A=0,1314% (165/125568, TOPMED) A=0,1469% (178/121148, ExAC)	51043 VB	BRCA2_000483 VB/VSI	CM173607 DM?	VB	Se reportó en dos casos índices	VB	BP7, BS1, BP6	VB	

BRCA2: c.10234A>G (p.Ile3412Val)	VSI (2011)	*rs1801426* G=2,3306% (5843/250706, GnomAD_exome) G=4,0544% (5091/125568, TOPMED)	41540 VB	BRCA2_00485 VB/VSI	CM970189 DM?	VB	Se reportó en 7 casos índices	VB	BP4, BS2, BP6, BA1	VB	
BRCA2: c.*14C>T	VSI (2011)	*rs373436334* T=0,0024% (6/248112, GnomAD_exome) T=0,0072% (19/264690, TOPMED) T=0,0026% (3/116256, ExAC)	377587 VB/ VPB	NR	NR	NR	112 (CM)	VSI	PM2, BP6, BP7	VSI	VPB

CC: cáncer de colon; CM: cáncer de mama; CO: cáncer de ovario; CP: cáncer de páncreas