



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

**Epidemiología molecular de alelos patogénicos en
genes asociados a síndrome hereditario de cáncer
de mama y ovario en población latina**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN BIOLOGÍA

P R E S E N T A:

Luis Enrique Romero Cruz



DIRECTOR DE TESIS:

Dr. Felipe Vaca Paniagua

Cd. Mx. 2024



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas

Tesis Digitales

Restricciones de uso

DERECHOS RESERVADOS ©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México, por haberme ofrecido una educación de calidad excepcional y una formación académica integral.

A la Facultad de Estudios Superiores Iztacala, por tener un colegiado con una gran vocación y fomentar en mí el estudio de las ciencias y la biología.

Al doctor Felipe Vaca Paniagua por recibirme en su grupo de trabajo a muy temprana edad y brindarme las herramientas y tutela para la realización de este trabajo.

A la doctora Clara Estela Díaz Velásquez, por compartir conmigo su extenso conocimiento conmigo.

Agradecimientos personales

A mis padres y hermanos, por su apoyo incondicional.

A Ernesto, Rosie y Héctor, quienes fueron mis mentores cuando apenas incursionaba en el mundo de la investigación y a quienes admiró.

Al laboratorio 13 y a todos sus integrantes.

Dedicatoria

Dedico esta tesis a mi madre quien siempre cree en mí y a mi padre y su amor incondicional.

CONTENIDO

LISTA DE FIGURAS Y TABLAS	5
LISTA DE ABREVIATURAS	6
RESUMEN.....	9
ABSTRACT	10
1. INTRODUCCIÓN.....	11
1.1. Definición de cáncer	11
1.2. Epidemiología del cáncer	11
1.3. Factores de riesgo.....	14
1.3.1. Factores demográficos	14
1.3.2. Factores reproductivos	15
1.3.3. Factores Hormonales	15
1.3.4. Factores genéticos	17
1.4. Enfermedades y síndromes asociados.....	19
1.4.1. Síndrome Hereditario de Cáncer de Mama y Ovario	19
1.5. Clasificación de variantes genéticas	23
2. ANTECEDENTES	26
3. JUSTIFICACIÓN	28
4. HIPÓTESIS	29
5. OBJETIVOS	29
5.1. General.....	29
5.2. Específicos	29
6. MATERIALES Y MÉTODOS	30
6.1. Diseño de investigación.....	30
6.1.1. Criterios de inclusión	30
6.1.2. Criterios de exclusión	32
6.2. Flujo de trabajo experimental	33
6.2.1. Obtención de muestras	33
6.2.2. Extracción de DNA	35
6.2.3. Preparación de bibliotecas	37
6.2.4. Secuenciación.....	43
6.2.5. Análisis bioinformático	46
6.2.6. Selección de variantes patogénicas y VUS.....	49

6.2.7.	Análisis de la mutación mexicana fundadora (Deleción exón 9-12)....	52
6.2.8.	Validación de resultados.....	52
7.	RESULTADOS.....	53
8.	DISCUSIÓN.....	62
9.	CONCLUSIONES.....	68
10.	PERSPECTIVAS.....	69
11.	BIBLIOGRAFÍA	70
12.	OTROS RECURSOS	79

LISTA DE FIGURAS Y TABLAS

Figura 1. Incidencia y mortalidad de cáncer a nivel mundial.	12
Figura 2. Incidencia de cáncer en Latino América y el Caribe.....	13
Figura 3. Incidencia y mortalidad de cáncer de mama en países de Latino América.	13
Figura 4. Composición geográfica de las 944 muestras analizadas.....	34
Figura 5. Extracción de DNA mediante columnas de afinidad.....	36
Figura 6. Composición de algunos genes incluidos en el panel de secuenciación y su implicación en padecimientos oncológicos	38
Figura 7. Enriquecimiento de la región de interés mediante una PCR multiplex. .	39
Figura 8. Descripción gráfica de una biblioteca de DNA.	42
Figura 9. Secuenciación por síntesis de Illumina.	45
Figura 10. Flujo de trabajo bioinformático.	48
Figura 11. Criterios de filtrado de variantes genéticas.....	51
Figura 12. Diseño de PCR para la detección de la delección 9-12.	52
Figura 13. Proporción de muestras por país y porcentaje de variantes halladas. .	54
Figura 14. Distribución de variantes patogénicas en genes de alto riesgo.....	56
Figura 15. Distribución de variantes patogénicas en genes de riesgo moderado. .	58
Figura 16. Distribución de variantes de significado clínico incierto.....	61
Figura 17. Mutaciones recurrentes.....	61
Tabla 1. Subtipos moleculares de cáncer de mama.....	16
Tabla 2. Condiciones hereditarias de alto y moderado riesgo que predisponen al cáncer de mama.....	18
Tabla 3. Mutaciones asociadas con el síndrome hereditario de cáncer de mama y ovario.	21
Tabla 4. Criterios de inclusión para realización de pruebas genéticas en personas con sospecha de cáncer de mama y ovario hereditario.	22
Tabla 5. Clasificación de variantes germinales en asociadas a cáncer.....	24
Tabla 6. Clasificación de variantes genéticas somáticas unificadas de la AMP, ACGM y ASCO propuestas por el instituto de patología del hospital universitario de Heidelberg.	25
Tabla 7. Comparación de plataformas de secuenciación Ion Torrent e Illumina. ..	44
Tabla 8. Composición de población analizada y las 3 cohortes que la componen.	53

LISTA DE ABREVIATURAS

ABREVIATURA	SIGNIFICADO (Inglés)	(Español)
3' UTR	3' Untranslated Region	Región no traducida 3'
5' UTR	5' Untranslated Region	Región no traducida 5'
ACGM	American College of Genomic Medicine	Colegio Americano de Medicina Genómica
AD	Autosomal Dominant	Autosómico Dominante
AMP	Association of Molecular Pathology	Asociación de Patología Molecular
ANNOVAR	ANNOtate VARIation	Anotar Variación
APC	Adenomatous polyposis coli	Poliposis Coli adenomatosa
AR	Autosomal recessive	Autosómico Recesivo
ASCO	American Society of Clinical Oncology	Sociedad Americana de Oncología Clínica
ATM	Ataxia Telangiectasia Mutated	Ataxia Telangiectasia Mutado
ATP	Adenosine Triphosphate	Adenosín Trifosfato
ATR	Ataxia Telangiectasia and Rad3 related	Ataxia Telangiectasia y Rad3 relacionado
BAI	Indexed Binary alignment	Alineamiento Binario Indexado
BAM	Binary Alignment Mapped	Alineamiento Binario Mapeado
BP	base pairs	Pares de bases
BRCA1	Early Onset Breast Cancer 1	Cáncer de mama de aparición temprana 1
BRCA2	Early Onset Breast Cancer 2	Cáncer de mama de aparición temprana 2
BRIP1	BRCA1 Interacting Protein C-Terminal Helicase 1	Proteína C-Terminal que interactúa con BRCA1 Helicasa 1
BWA	Burrows-Wheeler Aligner	Alineador Burrows-Wheeler
CDH1	Cadherin 1	Caderina 1
CHEK2	checkpoint kinase 2	Punto de control quinasa 2
CNV	Copy Number Variation	Variación en el número de copias
CPT	Camptothecin	Camptotecina
DICER1	DICER1 Ribonuclease III	Ribonucleasa III <i>DICER1</i>
DNA	Deoxyribonucleic acid	Ácido Desoxirribonucleico
EDTA	Ethylenediaminetetraacetic acid	Ácido Etilendiaminetetraacético
EPCAM	Epithelial cell adhesion molecule	Molécula de adhesión celular epitelial
ER	Estrogen receptors	Receptor de Estrógenos
ERCC2	ERCC excision repair 2	Reparación por escisión 2 ERCC
ERCC3	ERCC excision repair 3	Reparación por escisión 3 ERCC
ESCO2	Establishment of sister chromatid cohesion N-acetyltransferase 2	Establecimiento de cohesión de la cromátida hermana N-acetiltransferasa 2
EX9-12DEL	Exonic 9-12 deletion	Delección del exón 9-12
FANCB	Fanconi Anemia Complementation Group B	Grupo de complementación B de la anemia de Fanconi
FANCC	Fanconi Anemia Complementation Group C	Grupo de complementación C de la anemia de Fanconi
FANCF	Fanconi Anemia Complementation Group F	Grupo de complementación F de la anemia de Fanconi

FANCI	Fanconi Anemia Complementation Group I	Grupo de complementación I de la anemia de Fanconi
FANCL	Fanconi Anemia Complementation Group L	Grupo de complementación L de la anemia de Fanconi
FANCM	Fanconi Anemia Complementation Group M	Grupo de complementación M de la anemia de Fanconi
GATK	Genome Analysis Tool Kit	Kit de herramientas para análisis de Genoma
HBOC	Hereditary Breast and Ovarian Cancer Syndrome	Síndrome Hereditario de cáncer de mama y ovario
HER2	Human Epidermic Growth Factor Receptor 2	Receptor 2 del factor de crecimiento epidérmico humano
HG19	Human Genome 19	Genoma Humano 19
HOZ	Homozygous	Homocigoto
HPHBC	highly penetrant hereditary breast cancer	Cáncer de mama con alta penetrancia
HRR	Homologous recombinational repair of DNA	Reparación por recombinación homóloga del DNA
INDELS	Insertion-Deletion (DNA)	Inserción-Delección (DNA)
KBP	Kilo base pairs	Kilo pares de bases
LFS	Li-Fraumeni Syndrome	Síndrome de Li-Fraumeni
LIG4	DNA Ligasa 4	DNA ligasa 4
MLH1	MutL Homolog 1	Homólogo 1 de la proteína MutL
MLH3	MutL Homolog 3	Homólogo 3 de la proteína MutL
MSH2	MutS Homolog 2	Homólogo 2 de la proteína MutL
MSH6	MutS Homolog 6	Homólogo 6 de la proteína MutS
MSR1	Macrophage Scavenger Receptor 1	Receptor 1 del eliminador de macrófagos
MUTYH	MutY Homolog	Homólogo de la proteína MutY
NBN	Nibrine	Nibrina
NCCN	National Cancer Comprehensive Network	Red Nacional Integral del Cáncer
NF1	Neurofibblastome 1	Neurofibroblastoma 1
NGS	Next Generation Sequencing	Secuenciación de siguiente generación
NM	Nanometers	Nanómetros
PAHO	PanaAmerican Health Organization	Organización Panamericana de Salud
PALB2	Partner and Localizer of BRCA2	Acompañante y localizador de BRCA2
PBS	Phosphate buffered saline	Buffer de Fosfato salino
PCR	Polimerase Chain Reaction	Reacción en cadena de Polimerasa
PDE11A	Phosphodiesterase 11A	Fosfodiesterasa 11A
PDGFB	Platelet derived Growth Factor subunit B	Subunidad B de Plaquetas derivadas del Factor de crecimiento
PMS2	PMS1 Homolog, Mismatch Repair System	Homólogo de la proteína PSM1, Sistema de reparación de errores de combinación
POLH	Polymerase H	Polimerasa H
POU6F2	Pou class 6 homeobox 2	Homeobox Clase 6 Pou

PR	Progesterone Receptor	Receptor de progesterona
PTCH1	Patched 1Homolog 1	Homólogo 1 de la Proteína Patched 1
PTEN	Phosphate and Tesin Homolog	Homólogo de la proteína fosfato tesina
RAD51C	RAD51 Paralog C	Parálogo C de RAD51
RAD51B	RAD51 Paralog B	Parálogo B de RAD51
RAD51D	RAD51 Paralog D	Parálogo D de RAD51
RAS	RAS protein	Proteína RAS
RECQL1	RecQ Like Helicase 1	Helicasa 1 similar a la proteína RecQ
RECQL4	RecQ Like Helicase 4	Helicasa 4 similar a la proteína RecQ
RECQL5	RecQ Like Helicase 5	Helicasa 5 similar a la proteína RecQ
RNA	Ribonucleic Acid	Ácido Ribonucleico
RNASEL	Ribonuclease L	Ribonucleasa L
RPM	Revolutions per minute	Revoluciones por Minuto
RR	Relative Risk	Riesgo Relativo
SAM	Sorted Aligned Mapped	Alineamiento Mapeado Ordenado
SDHA	Succinate Deshydrogenase A	Succinato Deshidrogenasa A
SDHB	Succinate Deshydrogenase B	Succinato Deshidrogenasa B
SNP	Single Nucleotide Polymorfism	Polimorfismo de un solo nucleótido
SNV	Single Nucleotide Variant	Variante de un solo nucleótido
STK11	Serine/Threonine Kinase 11	Quinasa 11 Serina/Treonina
TP53	Tumor Protein 53	Proteína Tumoral 53
UROD	Uroporphyrinogen Decarboxylase	Descarboxilasa de Uroporfirinógeno
VCF	Variant Calling Format	Formato de llamado de Variante
VUS	Variant unknow significant	Variante con significado desconocido
WRN	Werner Syndrome Helicase	Helicasa del Síndrome de Werner
XRCC3	X ray repair cross complementing 3	Reparación de rayos x por complementación cruzada 3

RESUMEN

El cáncer de mama es la principal causa de muerte en mujeres de todo el mundo. Se estima que alrededor del 10%-15% de los casos de cáncer de mama se deben a una predisposición genética debido a una alteración en genes asociados, lo que se conoce como síndrome hereditario de cáncer de mama y ovario (HBOC). Actualmente se conocen varios genes que pueden estar implicados en esta enfermedad, más allá de los genes *BRCA1* y *BRCA2*. No obstante, existen muy pocos estudios que evalúen la diversidad de alelos patogénicos en poblaciones minoritarias, como es el caso de los países de América Latina. En este trabajo, analizamos la región codificante de 143 genes asociados a cáncer en 944 pacientes provenientes de Argentina, Colombia, Guatemala, México y Perú utilizando tecnología de secuenciación de siguiente generación (NGS) mediante las plataformas Illumina® y IonTorrent®. Nuestros resultados mostraron que el 9.95% (94/944) de las pacientes tuvo una variante patogénica en genes de alto riesgo, mientras el 6.46% (61/944) tuvo estas variantes en genes de bajo riesgo o moderado. También se detectaron alteraciones en genes que no habían sido reportados para esta enfermedad como *WRN*, *MSR1* y *LIG4*, así como un gran número de variantes con significado clínico incierto (209/944). Este trabajo aporta nuevos datos sobre la composición genética del HBOC en países poco estudiados, permitiendo hacer un contraste con otros estudios internacionales.

ABSTRACT

Breast cancer is the leading cause of death in women worldwide. It is estimated that around 10%-15% of breast cancer cases are due to a genetic predisposition due to an alteration in associated genes, which is known as hereditary breast and ovarian cancer syndrome (HBOC). Several genes are currently known that may be involved in this disease, beyond the BRCA1 and BRCA2 genes. However, there are very few studies that evaluate the diversity of pathogenic alleles in minority populations, such as in Latin American countries. In this work, we analyzed the coding region of 143 cancer-associated genes in 944 patients from Argentina, Colombia, Guatemala, Mexico, and Peru using next-generation sequencing (NGS) technology using the Illumina® and IonTorrent® platforms. Our results showed that 9.95% (94/944) of the patients had a pathogenic variant in high-risk genes, while 6.46% (61/944) had these variants in low-risk or moderate-risk genes. Alterations were also detected in genes that had not been reported for this disease, such as WRN, MSR1, and LIG4, as well as a large number of variants with uncertain clinical significance (209/944). This work provides new data on the genetic composition of HBOC in understudied countries, allowing for a contrast with other international studies.

1. INTRODUCCIÓN

1.1. Definición de cáncer

El cáncer puede ser definido como una enfermedad causada por un cambio en la expresión genética que produce desregulación de los procesos de proliferación y muerte celular, lo que conlleva al crecimiento anormal de las células somáticas hasta formar tumores, los cuales pueden invadir otros tejidos y hacer metástasis a órganos distantes, causando morbilidad y posteriormente la muerte del hospedero [1-2].

1.2. Epidemiología del cáncer

Junto con otras enfermedades no infecciosas, el cáncer es una de las principales causas de muerte en todo el mundo. Su mortalidad sólo es superada por enfermedades cardiovasculares como infarto, cardiopatía isquémica y accidentes cerebro vasculares [3].

El cáncer de pulmón es la neoplasia maligna con mayor mortalidad alrededor del mundo tanto en hombres como mujeres (Figura 1). En 2018 se reportaron 2,093,876 nuevos casos de cáncer de pulmón y 1,761,007 defunciones debido a este padecimiento. Sin embargo, el cáncer de mama ocupa el puesto número uno en cuanto a incidencia y es la principal causa de muerte para las mujeres de todo el mundo. Se estima que en 2018 hubo más de 2 millones de casos nuevos y 626,679 muertes [4]. Existe una proyección hacia el incremento en la incidencia de cáncer de mama, por el contrario, la mortalidad de esta enfermedad ha disminuido en los últimos años. Esta tendencia es más evidente en los países desarrollados [5].

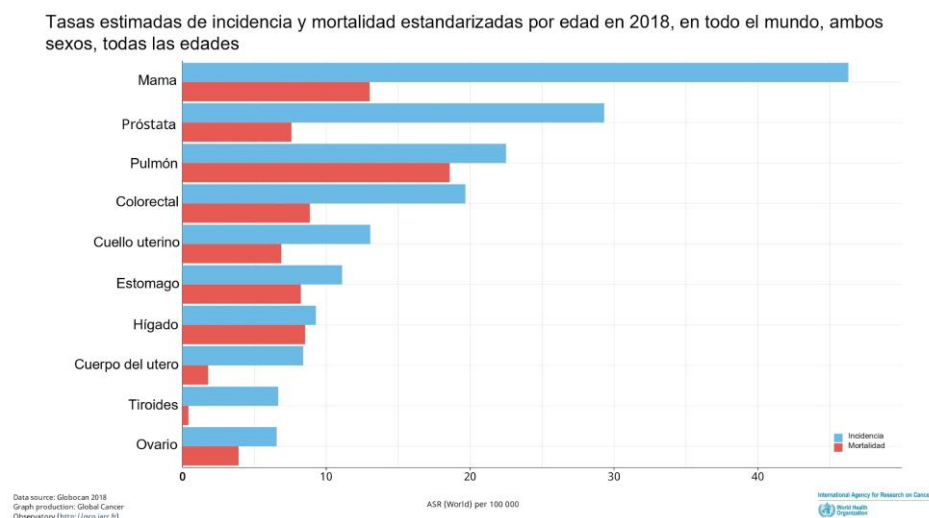
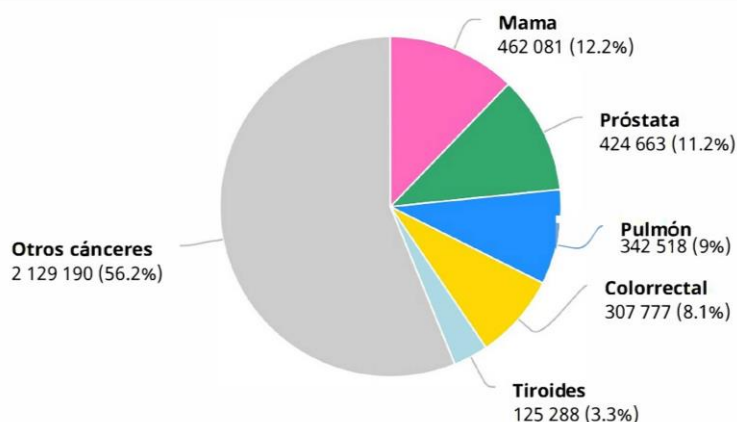


Figura 1. Incidencia y mortalidad de cáncer a nivel mundial. Se muestra el número de nuevos casos y defunciones por cada 100,000 habitantes, ambos sexos (Modificada de GLOBOCAN, 2018).

En el caso particular de América Latina y el Caribe el cáncer de mama también es la principal causa de muerte en mujeres y mantiene su tendencia como el cáncer con mayor incidencia. El último reporte de la Organización Panamericana de la Salud (PAHO) señaló 462 081 nuevos casos de cáncer de mama para esta región, lo que representa el 12.2% de todos los casos de cáncer en América latina (Figura 2) [6]. En 2013 los países con mayor incidencia fueron Puerto Rico, Argentina y Uruguay, comparable con las tasas de incidencia de Países Europeos o Estados Unidos (Figura 3). Por otra parte, países como México, Perú, Colombia y Brasil registran diagnósticos y defunciones a edades muy tempranas, siendo la edad promedio de diagnóstico a los 54.8 años. Así mismo la correlación entre una alta mortalidad y las tasas de incidencia señalan una baja supervivencia para estos países latinos [7].

Diferentes estudios proyectan un incremento en la incidencia y mortalidad de esta enfermedad, para el año 2030 se espera que haya el doble de casos y defunciones por cáncer de mama en América Latina [8]. Este panorama supone diferentes retos y áreas de oportunidad para mejorar a fin de poder brindar mejores servicios de salud que ayuden en la prevención, detección y tratamiento del cáncer de mamá.

Número de nuevos casos en 2018, ambos sexos, todas las edades



Total: 3 791 517

Figura 2. Incidencia de cáncer en Latino América y el Caribe. Se muestra el número de nuevos casos de cáncer de mama y el porcentaje que representan en comparación con otros tipos de cáncer sin considerar la edad o el sexo (Modificada de GLOBOCAN, 2018).

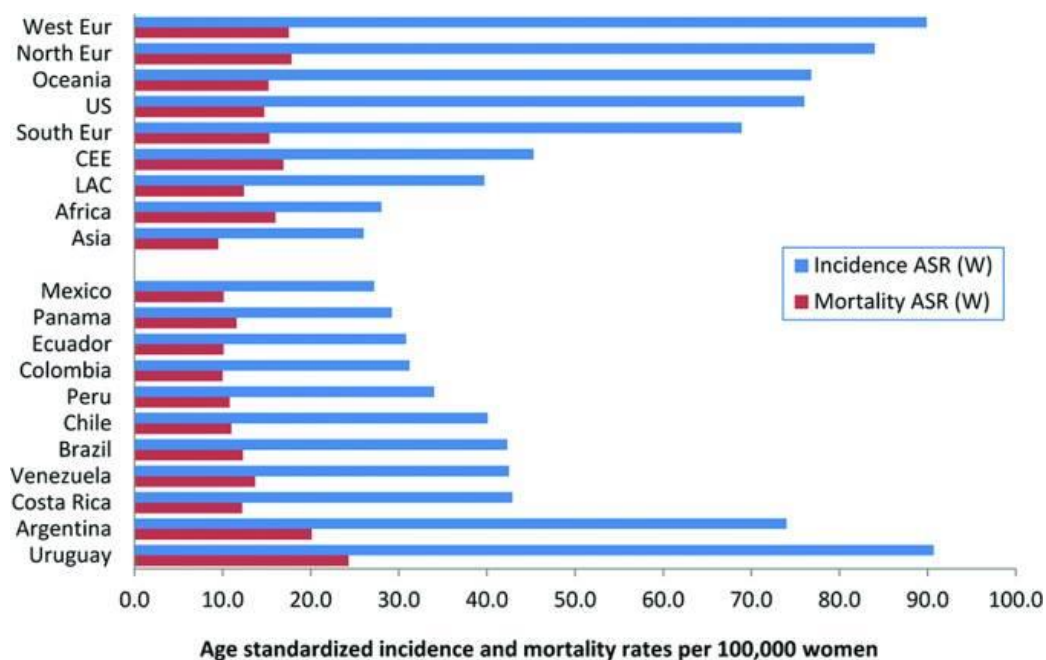


Figura 3. Incidencia y mortalidad de cáncer de mama en países de Latino América. Se observa una comparación de los países con mayor incidencia y mortalidad de cáncer de mama en Latino América con respecto a otras regiones del mundo (Tomada de Justo, Et al, 2013)

1.3. Factores de riesgo

Existen diferentes factores de riesgo que explican la presencia de cáncer de mama en la región de Latino América y el Caribe. De entre todos se destacan los factores demográficos, genéticos, reproductivos, así como de estilo de vida. Es importante señalar que todos estos factores de riesgo son muy similares en toda esta región y éstos no interactúan de forma aislada ya que el riesgo que cada factor aporta es distinto, por lo que se reconoce al cáncer de mama como una enfermedad multifactorial.

1.3.1. Factores demográficos

1.3.1.1. Edad

El cáncer de mama tiene una distribución específica en los distintos grupos de edades que componen una población. El 80% de los casos de cáncer de mama son diagnosticados en mujeres mayores de 50 años. Por otra parte, el 50% de los casos reportados, según datos epidemiológicos, se encuentra en el rango de edades de los 50 a los 69 años, siendo este segmento de la población el de mayor vulnerabilidad [9].

1.3.1.2. Etnicidad

Latino América es una región con una gran diversidad étnica y cultural. Históricamente ha existido una mezcla de indígenas nativos con otros grupos con ascendencia africana, europea y en menor medida asiática y nativos norteamericanos [10]. Actualmente no existen estudios epidemiológicos que consideren la diversidad étnica de esta región. Sin embargo, aunque la incidencia en mujeres latinas es menor a otras regiones como Europa, existe un mayor riesgo en estas mujeres de desarrollar subtipos más agresivos de cáncer de mama [11]. Estudios realizados en población de Estados Unidos demuestran que las mujeres blancas tienen mayor riesgo a desarrollar cáncer de mama, seguido de mujeres afroamericanas, asiáticas, hispanas y por último mujeres nativo-americanas [12].

De entre toda la población de mujeres que componen a Latino América, se ha observado una mayor predisposición de esta enfermedad en mujeres latinas con ascendencia caucásica-europea, seguido de mujeres afrolatinas y latino-asiáticas, este riesgo parece variar dependiendo también de la nacionalidad [11-13]. Finalmente, en otras regiones del mundo como Australia se ha observado que las mujeres pertenecientes a grupos autóctonos y aborígenes tienen el menor riesgo [14], esta tendencia podría repetirse en el caso de América Latina, aunque se necesitan estudios que comprueben esto último.

1.3.2. Factores reproductivos

Se ha documentado que diversos ámbitos reproductivos influyen en el riesgo a padecer cáncer de mama. El ejemplo más claro fue documentado por Fraumeni, quien descubrió que las monjas tienen un riesgo mayor al promedio de desarrollar cáncer de mama. A pesar de que se sabe que la paridad nula es un factor de riesgo y que cada término de embarazo reduce el riesgo de padecer esta enfermedad, existen otros factores relacionados con la maternidad; como, tiempo entre cada embarazo, lactancia, edad del primer y último embarazo, etc. cuyo riesgo no está claro y existe discusión sobre el papel que estos factores representan [16].

Por otra parte, como ya se mencionó, baja paridad y menarca a una edad temprana, así como menopausia a una edad tardía destacan como factores que aumentan el riesgo de cáncer de mama. Todos estos factores parecen estar fuertemente asociados a la exposición y sensibilidad a hormonas como el estrógeno [17].

1.3.3. Factores Hormonales

Las hormonas endógenas tienen un papel fundamental como factor de riesgo en cáncer de mama, principalmente el estrógeno, el cual tiene un efecto estimulante en los tumores con receptores hormonales y permite su clasificación en los subtipos moleculares Luminal A y Luminal B (Tabla 1). Esto recobra importancia ya que dos

terceras partes de los casos de cáncer de mama son positivos para receptores hormonales.

Existen dos hipótesis para explicar el mecanismo por el cual el estrógeno aumenta el riesgo a padecer cáncer de mama. El primero está relacionado con la vía de señalización de los receptores de estrógeno (ER), la cual es activada por esta hormona y en consecuencia modifica la expresión de genes asociados a la proliferación celular y por lo tanto la probabilidad de que ocurran mutaciones deletéreas durante los ciclos de división celular. El segundo mecanismo propone que el estrógeno al ser metabolizado puede formar aductos de estrógeno-quinona con el DNA, estos metabolitos pueden transformarse en catecol, creando especies reactivas de oxígeno y provocando daño oxidativo al DNA [18].

Otros receptores hormonales como los receptores de progesterona (PR) ayudan en la diferenciación del tejido mamario a través de la regulación de esta hormona. Se ha observado que la deficiencia de esta hormona incrementa considerablemente el riesgo de cáncer de endometrio. De forma antagónica la progesterona es un agente proliferativo en el tejido epitelial mamario y por lo tanto es considerado como un agente mitogénico [19].

Subtipo molecular	Receptores	Prevalencia (%)
Luminal A	ER y/o PR positivos Her2 Negativo	30-70
Luminal B	ER y/o PR positivos Her2 Positivo	10-20
Triple Negativo	ER y/o PR Negativo Her2 Negativo	15-20
Her2	ER y PR Negativo Her2 Positivo	5-15

Tabla 1. Subtipos moleculares de cáncer de mama (modificado de Rojas & Stuckey 2016).

1.3.4. Factores genéticos

El cáncer de mama es una enfermedad compleja causada por factores genéticos y no genéticos. Se tiene registro desde 1914 de que existe una relación entre cáncer y alteraciones genéticas [20]. Alrededor del 10% de casos de cáncer de mama se deben a una predisposición genética caracterizada por una heredabilidad autosómica dominante con penetrancia incompleta [21]. A pesar de esto, el cáncer de mama no tiene un patrón claro de heredabilidad.

Menos del 25% del total de casos de cáncer de mama hereditario son positivos a alteraciones en los genes *BRCA1* y *BRCA2*. Existen muchos otros genes que pueden contribuir al desarrollo de esta enfermedad. La susceptibilidad y el riesgo depende de qué gen se encuentre afectado e incluso del tipo de mutación o la región afectada del gen [22].

Gene Herencia	Susceptibilidad (%) de otros tumores	Frecuencia poblacional (%)	Proporción de cáncRRer de mama (%)	Proporción de HPHBC (%)	Proporción de cáncer Familiar (%)	Riesgo en la vida de una mujer (%)
<i>BRCA1</i> (AD)	Ovario	0.1	1.5	40	5 - 10	60 - 85
<i>BRCA2</i> (AD)	Ovario/Próstata Páncreas [HoZ-Fanconi (AR)]	0.1	1.5	40	5 - 10	40 - 85
<i>TP53</i> [LFS (AD)]	Sarcoma, Glioma, Tiroides	0.0025	0.02	2	0.1	80 - 90
<i>PTEN</i> [Síndrome de Cowden(AD)]	Adrenal, colorectal	0.0005	0.004	0.3	0.02	25 - 50
<i>CHEK2</i>	Colorectal, Próstata	0.5	0.5	0	2	18 - 20 (2.0)
<i>ATM</i> (AD & AR)	Linfoma, Leucemia [HoZ (AR)]	0.5	0.5	0	2	20 (2.3)
<i>STK11</i> [Peutz-Jeghers(AD)]	Colorectal	0.001	0.001	0.6	0.04	50
<i>BRIP1</i> (AD & AR)	HoZ-Fanconi (AR)	0.1	0.1	0	0.4	20 (2.0)
<i>PALB2</i> (AD & AR)	HoZ-Fanconi (AR)	0.1	0.1	0	0.4	20 (2.0)
<i>CDH1</i> (AD)	Cáncer gástrico difuso heredable	0.005	0.01	0.2 - 1	0.1	40 - 60
<i>NF1</i> (AD)	Neurofibroma, Glioma, MPNST	0.04	0.01	0	0.1	18

AD=Autosomal dominante; AR=Autosomal recesivo; Hoz=Homozigoto; HPHBC= Cáncer de mama con alta penetrancia (e.g >3 familiares afectados); LFS= Síndrome de Li-Fraumeni; RR= Riesgo Relativo

Tabla 2. Condiciones hereditarias de alto y moderado riesgo que predisponen al cáncer de mama (Modificado de Lalloo & Evans 2012).

1.4. Enfermedades y síndromes asociados

Un historial familiar de cáncer es considerado como un factor de riesgo importante y está frecuentemente vinculado a la predisposición genética. Actualmente, se conocen más de 20 síndromes hereditarios asociados a cáncer. Estos síndromes afectan a aproximadamente al 1% de todos los pacientes con cáncer. Las personas afectadas por alguno de estos síndromes son susceptibles a un tipo de tumor específico dependiendo del síndrome que padecen y son portadores de una o más copias de genes mutados (también conocidos como alelos mutantes). Estas mutaciones germinales pueden persistir en el linaje familiar y pasar desapercibidas por varias generaciones. Los síndromes asociados a cáncer suelen favorecer la aparición de tipos poco común de cáncer, subtipos agresivos y diagnósticos a edades tempranas. Sin embargo, cuando varios miembros de una familia desarrollan cáncer a una edad temprana o un solo individuo tiene múltiples cánceres primarios, incluso si son tipos de cáncer comunes, se sospecha de síndromes de cáncer hereditario [23].

El estudio de los síndromes hereditarios asociados a cáncer ha permitido conocer a los genes implicados y las vías de señalización que comúnmente se alteran durante el desarrollo de los diferentes tumores malignos [24].

1.4.1. Síndrome Hereditario de Cáncer de Mama y Ovario

El síndrome hereditario de cáncer de mama y ovario (HBOC por sus siglas en inglés) es un desorden en el que las personas que lo padecen tienen un alto riesgo de desarrollar cáncer de mama u ovario, mayor al promedio (principalmente antes de los 50 años). El HBOC tiene un patrón de herencia autosómico dominante con penetrancia incompleta y está asociado principalmente a alteraciones en los genes *BRCA1* Y *BRCA2*. Estos genes fueron los primeros en describirse como causantes de esta enfermedad. *BRCA1* Y *BRCA2* juegan un papel fundamental en la reparación del DNA. De manera específica, estos genes se encargan de reparar las rupturas de doble cadena del DNA causadas por agentes clastogénicos mediante el mecanismo de reparación por recombinación homóloga (HRR). Por esta razón

cuando uno de los dos alelos, ya sea en *BRCA1* o *BRCA2*, tienen una mutación germinal deletérea, existe pérdida de la heterocigosidad. Posteriormente una mutación somática puede inactivar al otro alelo restante, provocando un déficit en la actividad de ese gen (hipótesis del “*Double Hit*”) [25]. Dado que el mecanismo de reparación por recombinación homóloga es vital para mantener la estabilidad genómica, la nula expresión de cualquiera de estos dos genes en la célula conlleva a una tasa de mutación más alta. Por otra parte, la segregación cromosómica, la citocinesis y el correcto funcionamiento del ciclo celular durante la replicación del DNA también dependen de la HRR. Sumado a esto, un ambiente carcinogénico dependiente de hormonas, con señales constantes de proliferación y ausencia de estos supresores de tumores, contribuyen a la formación de un fenotipo maligno en las células [26].

En las últimas décadas se ha descubierto que otros genes además de *BRCA 1* y *BRCA2* están implicados en el desarrollo del síndrome hereditario de cáncer de mama y ovario. Los genes “NO BRCA” representan hasta el 25% de todos los casos de HBOC [27] y el riesgo que supone tener una alteración en alguno de estos genes es muy variable entre un gen u otro (Tabla 3). La “*National Comprehensive Cancer Network*” (NCCN) emite un informe anual en el que evalúa cada uno de los genes asociados a cáncer, también los clasifica de acuerdo con el riesgo que cada uno confiere [28].

Gene	Riesgo de cáncer de mama	Riesgo de cáncer de ovario *	Riesgo de otro cáncer
<i>ATM</i>	Aumentado	Sin Aumento	Evidencia Insuficiente
<i>BRCA1</i>	Aumentado	Aumentado	Próstata
<i>BRCA2</i>	Aumentado	Aumentado	Melanoma, Páncreas, Próstata
<i>BRIP1</i>	Sin Aumento	Aumentado	Evidencia Insuficiente
<i>CDH1</i>	Aumentado	Sin Aumento	Estómago
<i>CHEK2</i>	Aumentado	Sin Aumento	Colon
Genes del síndrome de Lynch: <i>MSH2</i> , <i>MLH1</i> , <i>MSH6</i> , <i>PMS2</i> , <i>EPCAM</i>	Evidencia Insuficiente	Aumentado	Colon, Útero, Pelvis Renal, Intestino Delgado Y Otros
<i>PALB2</i>	Aumentado	Sin Aumento	Desconocido
<i>PTEN</i>	Aumentado	Sin Aumento	Síndrome De Cowden
<i>RAD51C</i>	Sin Aumento	Aumento	Desconocido
<i>RAD51D</i>	Sin Aumento	Aumento	Desconocido
<i>STK11</i>	Aumentado	Aumento De Tumores De Los Cordones Sexuales Y El Estroma	Síndrome De Peutz-Jehger
<i>TP53</i>	Aumentado	Sin Aumento	Síndrome De Li-Fraumeni

*Se incluye cáncer de trompas de falopio y peritoneal.

Datos de National Comprehensive Cancer Network. Genetic/familial high-risk assessment: breast and ovarian. Version 2.2017. NCCN

Clinical Practice Guidelines in Oncology. Fort Washington (PA): NCCN; 2016. Available at: https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf.

Tabla 3. Mutaciones asociadas con el síndrome hereditario de cáncer de mama y ovario. (Modificado de ACOG, 2017)

La NCCN recomienda realizar una prueba genética para confirmar HBOC cuando el paciente reúne alguno de los siguientes criterios:

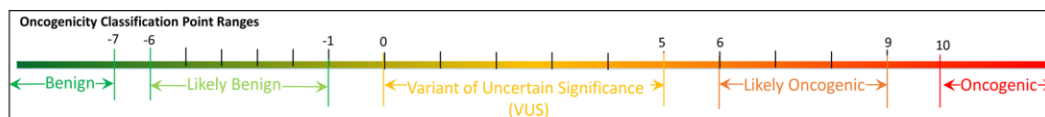
Antecedentes personales de cáncer de mama más uno o más de los siguientes criterios	Antecedentes heredofamiliares de cáncer
● Diagnóstico antes de los 45 años ● Diagnóstico entre 45-50 años y: ○ Cáncer de mama adicional a cualquier edad ○ Familiar consanguíneo con cáncer de mama a cualquier edad ○ Familiar consanguíneo con cáncer de próstata de alto grado ● Diagnóstico antes de los 60 años y: ○ Cáncer de mama triple negativo ● Diagnóstico a cualquier edad y: ○ uno o más familiares consanguíneos con: ■ cáncer de mama antes de los 50 años y/o : ● carcinoma de ovario ● cáncer de mama en hombre ● cáncer de próstata metastásico o cáncer pancreático ○ 2 o más diagnósticos de cáncer de mama en el paciente o familiar consanguíneo ● Ancestría judía Ashkenazi	● Paciente sin diagnóstico de cáncer, pero con alguno de los siguientes criterios ○ Familiar consanguíneo con ■ Mutación conocida en algún gen de susceptibilidad al cáncer ■ 2 o más tumores primarios de mama ■ 2 o más familiares con cáncer de mama del mismo lado familiar con diagnóstico antes de los 50 años ■ cáncer de ovario ■ cáncer de mama en hombre ○ Familiar consanguíneo con cáncer de mama antes de los 45 años

Tabla 4. Criterios de inclusión para realización de pruebas genéticas en personas con sospecha de cáncer de mama y ovario hereditario (Datos tomados de NCCN, 2022) [29].

1.5. Clasificación de variantes genéticas

La finalidad de interpretar las variantes obtenidas mediante las pruebas genéticas es aportar información útil que pueda dirigir las estrategias de prevención, tratamiento y pronóstico. La clasificación clínica de variantes genéticas en cáncer ha sido una tarea ardua y en la actualidad aún no existe un sistema que sea aceptado de forma unánime. Las dificultades que enfrentan los genetistas para lograr dicha clasificación es que hay un enorme catálogo de variantes genéticas debido a las emergentes tecnologías de secuenciación de última generación.

En 2015 el Colegio Americano de Genética y Genómica Médica (ACGM) implementó una serie de recomendaciones para la interpretación de variantes genómicas. Estas recomendaciones son aplicables a muchos tipos de alteraciones genéticas tales como variantes de un solo nucleótido (SNV por sus siglas en inglés), deleciones o inserciones (indels) y variaciones en el número de copias (CNV). Este sistema aportó la terminología para designar a las variantes en genes que causan desórdenes mendelianos. La ACGM propuso cinco categorías para mutaciones germinales que son: Patogénicas, Probablemente patogénicas, Variantes con significado incierto (VUS), Probablemente benignas y Benignas (Tabla 5). Los criterios de clasificación son explicados como un sistema de puntajes que recapitulan toda la información disponible acerca de la variante a analizar [30].



	Benign			Oncogenic			
Evidence Strength	Very Strong	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
POINTS	-8	-4	-1	+1	+2	+4	+8
Population Data	MAF is >5%	MAF is >1%		Absent in population databases			
Functional Data		Well-established functional studies show no oncogenic effects				Well-established functional studies supportive of an oncogenic effect	
Predictive Data			Silent mutation (no predicted impact on splicing)		Missense change at an amino acid residue where a different missense change determined to be oncogenic has been documented	Same amino acid change as a previously established oncogenic mutation	Null variant in tumor suppressor
Cancer Hotspots				Cancer hotspots with low frequency of recurrence	Cancer hotspot with moderate frequency of recurrence	Cancer hotspot with high frequency of recurrence	
Computational Evidence			All utilized lines of computational evidence suggest no impact of a variant	All utilized lines of computational evidence support oncogenicity			

Tabla 5. Clasificación de variantes germinales en asociadas a cáncer (Tomada de Horak, et al, 2022).

Actualmente se han propuesto otros dos sistemas para la clasificación de variantes genéticas de tipo somáticas en cáncer, propuestos por La Asociación de Patología Molecular (AMP) y La Sociedad Americana de Oncología Clínica (ASCO) (Tabla 6). Todos estos sistemas se basan en las frecuencias alélicas poblacionales proporcionadas por las bases de datos internacionales, los estudios clínicos de correlación, la evidencia experimental (ensayos funcionales), los programas computacionales predictivos y la interpretación por un panel de varios expertos.

De forma conjunta la clasificación puede resumirse de la siguiente manera:

Categoría		Explicación
Nivel 1	Biomarcadores que predicen la respuesta o resistencia a una terapia aprobada por la FDA para un tumor específico. Biomarcadores incluidos en manuales profesionales que predicen la respuesta o resistencia a terapias para un tipo de tumor en específico. Biomarcadores que predicen la respuesta o resistencia a terapias para un tipo específico de tumor basados en estudios bien fundamentados con consenso de expertos en el campo.	Variantes con fuerte significado clínico
Nivel 2	Biomarcadores que predicen la respuesta o resistencia a terapias aprobadas por la FDA o sociedades especializadas para un tipo diferente de tumor. Biomarcadores que sirven como criterio de inclusión para ensayos clínicos. Biomarcadores que muestran un significado terapéutico plausible basado en estudios pre-clínicos.	Variantes con potencial significado clínico
Nivel 3	Variantes no observadas con una frecuencia alélica significativa en la base de datos de la subpoblación general o específica, o en bases de datos de variantes de pan-cáncer o tumores específicos. Evidencia publicada de asociación de cáncer no convincente.	Variantes con significado clínico desconocido
Nivel 4	Variantes observadas con una frecuencia alélica significativa en las bases de datos de la subpoblación. Sin evidencia publicada de asociación de cáncer.	Variantes benignas o posiblemente benignas

Tabla 6. Clasificación de variantes genéticas somáticas unificadas de la AMP, ACGM y ASCO propuestas por el instituto de patología del hospital universitario de Heidelberg (Modificada de Leichsering, et al, 2019).

2. ANTECEDENTES

En 2017 Quezada y colaboradores realizaron un estudio en 300 pacientes mexicanas que cumplían los criterios para HBOC, el objetivo de este estudio fue encontrar nuevos alelos patogénicos que no hubieran sido descritos anteriormente en esta población. Para ello se analizaron la región codificante de 143 genes asociados a enfermedades oncológicas. En este estudio se detectaron 46 variantes patogénicas en 23 genes distintos, los genes con el mayor número de mutaciones fueron *BRCA1* y *BRCA2*; 33 variantes con significado incierto también fueron halladas entre la población. La mutación más frecuente fue la mutación mexicana fundadora ex9-12del [31] y la mutación p.G228fs en *BRCA1*. Los resultados también mostraron una mayor diversidad de genes afectados (54%) en comparación con *BRCA1/2*, los cuales solo concentraron 46% de las variantes detectadas en esta cohorte [32].

Por otra parte, en 2018 el equipo de Caswell evaluó las variantes patogénicas presentes en 1483 personas en la universidad de Stanford, agrupándolas en grupos étnicos y raciales. Los datos de dicho estudio se obtuvieron a través de la recolección de resultados de pruebas genéticas de predisposición a cáncer que los pacientes donaron de forma anónima a la universidad de Stanford. Para este estudio se incluyeron 14 genes relacionados a diferentes síndromes asociados a cáncer (*BRCA1*, *BRCA2*, *MLH1*, *MSH2*, *MSH6*, *PMS2*, *EPCAM*, *APC*, *MYH*, *TP53*, *CDH1*, *CDKN2A*, *PTEN* y *STK11*). Los pacientes con ascendencia asiática e hispana con variantes patogénicas tuvieron un rango de edades menor en comparación con los caucásicos. El número de variantes patogénicas (15%) no tuvo diferencias entre los grupos raciales. Por otra parte, un mayor porcentaje de persona caucásicas presentó mutaciones en el gen *CHEK2* en comparación con las no caucásicas (3.8% vs. 1.0%). Finalmente, el número de VUS fue más abundante en los grupos no caucásicos [33].

En 2019 el consorcio LACAM recién consolidado emitió un estudio en el que se analizaron a 222 pacientes que reunían los criterios para HBOC provenientes de

México, Argentina, Colombia, Guatemala y Perú. 143 genes fueron seleccionados para secuenciar su región codificante, mediante las plataformas MiSeq (Illumina, Inc.) y PGM (Ion Torrent-Thermo Fisher Scientific). La prevalencia de variantes patogénicas fue del 17% (38/222) distribuidas en 14 genes. La incidencia por países fue la siguiente: Argentina (25%, 14/57), México (18%, 12/68), Guatemala (16%, 3/19) y Colombia (13%, 10/78). En este trabajo se encontraron variantes patogénicas en genes de riesgo moderado como *MSH6*, *MSH2*, *MUTYH* y *PALB2*; También se detectaron alteraciones en genes no asociados a HBOC (*DCLRE1C*, *WRN*, *PDE11A* y *PDGFB*) [34]. Este estudio fue el primero en caracterizar el Síndrome Hereditario de Cáncer de Mama y Ovario en una región poco estudiada como lo es Latinoamérica.

3. JUSTIFICACIÓN

El Síndrome Hereditario de Cáncer de Mama y Ovario (HBOC) es responsable de hasta el 10% de todos los casos de cáncer de mama. Los pacientes con este síndrome son particularmente propensos a desarrollar cáncer a una edad temprana, tener altas tasas de reincidencia y en general un peor pronóstico. El HBOC se atribuye a alteraciones en genes asociados a cáncer, conocidas como variantes patogénicas. A pesar de ser un problema de salud en toda la región de Latinoamérica, existe poca información sobre la composición y estructura genómica de esta enfermedad en dicha región. La mayoría de los estudios solo se enfocan en un reducido número de genes, los cuales no satisfacen todos los casos de HBOC y subestiman la diversidad de alelos que pudieran estar asociados a esta enfermedad. Por otra parte, todos los estudios que se han hecho en esta región son de carácter nacional, lo cual no permite valorar las diferencias genéticas a nivel internacional entre los países de esta misma región, asociadas a este síndrome.

En este proyecto se propuso analizar la región codificante de 143 genes asociados al síndrome hereditario de cáncer de mama y ovario y otras enfermedades oncológicas en 944 pacientes provenientes de Argentina, Colombia, Guatemala, México y Perú. Este estudio tiene como objetivo detectar alelos que no se hayan reportado anteriormente y aumentar el catálogo de variantes genéticas que puedan ser de interés por su relevancia en esta enfermedad o su importancia a nivel clínico. En resumen, el análisis genómico permitirá enriquecer la información que se tiene disponible respecto al HBOC en estos países.

4. HIPÓTESIS

Existen variantes patogénicas y de significado clínico incierto asociadas al HBOC que no han sido descritas en pacientes de América Latina las cuales pueden ser detectadas mediante la secuenciación de la región exónica de 143 genes.

5. OBJETIVOS

5.1. General

Conocer la diversidad de alelos patogénicos presentes en pacientes con síndrome hereditario de cáncer de mama y ovario de Argentina, México, Guatemala, Colombia y Perú.

5.2. Específicos

- Identificar las alteraciones patogénicas recurrentes en nuestra cohorte de 944 pacientes.
- Reportar las variantes genéticas con significado clínico incierto (VUS) en nuestra cohorte.
- Describir los genes frecuentemente mutados en pacientes de América Latina.
- Determinar la presencia de posibles mutaciones fundadoras en esta población.

6. MATERIALES Y MÉTODOS

6.1. Diseño de investigación

Para este proyecto se planteó hacer un estudio exploratorio, transversal y de tipo descriptivo.

6.1.1. Criterios de inclusión

Los criterios de inclusión para todas las cohortes fueron establecidos en su momento para cada estudio, con base en la información disponible de la NCCN. Para la cohorte nueva de 422 pacientes se utilizó la versión más reciente en ese momento [29]. No se detectó ninguna discrepancia a lo largo de las diferentes versiones de la guía propuesta por la NCCN [35,36].

A continuación, se describen las variables de mayor relevancia consideradas en este estudio.

- I. Un individuo a cualquier edad con una variante patogénica o probablemente patogénica en algún gen de susceptibilidad al cáncer dentro de la familia del paciente.
- II. Un individuo diagnosticado a cualquier edad con cualquiera de los siguientes criterios:
 - A. Cáncer de ovario.
 - B. Cáncer pancreático.
 - C. Cáncer de próstata metastásico.
 - D. Cáncer de mama o cáncer de próstata con alto grado (Puntaje de Gleason ≥ 7) y ancestría judía askenazi.
- III. Un individuo con diagnóstico de cáncer de mama que reúna cualquiera de los siguientes criterios:
 - A. Cáncer de mama diagnosticado antes de los 50 años

- B. Cáncer de mama triple negativo (ER-, PR, HER-) diagnosticado antes de los 60 años.
- C. Dos cánceres de mama primarios.
- D. Cáncer de mama a cualquier edad y:
 - 1. Un familiar consanguíneo cercano con:
 - a) Cáncer de mama
 - b) Cáncer de ovario invasivo
 - c) Cáncer de mama masculino
 - d) Cáncer pancreático
 - e) Cáncer de próstata metastásico o con alto grado (Puntaje de Gleason ≥ 7)
 - 2. 2 familiares consanguíneos cercanos con cáncer de mama a cualquier edad.
- IV. Un individuo que no reúna los criterios mencionados previamente pero que tenga un familiar de primer o segundo grado con cualquiera de las siguientes características:
 - A. Cáncer de mama antes de los 45 años
 - B. Cáncer de ovario
 - C. Cáncer de mama masculino
 - D. Cáncer pancreático
 - E. Cáncer de próstata metastásico
 - F. Dos o más cánceres de mama primarios en una sola persona
 - G. Dos o más personas con cáncer de mama primario en el mismo linaje familiar (materno/paterno) con al menos uno de ellos diagnosticado antes de los 50 años.
- V. Un individuo con un historial personal o familiar, en el mismo lado de la familia con tres o más de los siguientes criterios (especialmente si fueron diagnosticados antes de los 50 años, o si un individuo tuvo más de un cáncer primario):
 - A. Cáncer de mama, sarcoma, carcinoma adrenocortical, tumor cerebral, leucemia.

- B. Cáncer de colon, cáncer endometrial, cáncer de tiroides, cáncer de riñones, manifestaciones dermatológicas asociadas al síndrome de Cowden, macrocefalia, pólipos hamartomatosos del tracto intestinal.
- C. Cáncer de mama lobular, cáncer gástrico difuso.
- D. Cáncer de mama, cáncer gastrointestinal, pólipos hamartomatosos, tumores en las trompas de Falopio, cáncer pancreático, tumores testiculares en las células de Sertoli o pigmentación infantil de piel.

6.1.2. Criterios de exclusión

Para este estudio se excluyeron las muestras que se encontraran en cualquiera de las siguientes condiciones.

1. Material biológico insuficiente, degradado o no rastreable.
2. Pacientes sin datos epidemiológicos asociados a factores de riesgo.
3. Pacientes sin el consentimiento informado escrito.
4. Cualquier muestra que no cumpliera con los estándares de calidad en algún punto de flujo de trabajo experimental. (Extracción de DNA, Preparación de bibliotecas, secuenciación.)

6.2. Flujo de trabajo experimental

6.2.1. Obtención de muestras

La recolección de muestras se hizo a través de diversos centros de salud de México, Guatemala, Colombia, Perú y Argentina.

En el estudio de Quezada, *et al*, 2017 se reunieron un total de 300 muestras con el esfuerzo de 4 fundaciones y centros de investigación. En esta cohorte las muestras provinieron de los siguientes institutos:

- Instituto de salud del estado de México
- Instituto estatal de cancerología de Guerrero
- Centro de investigación biomédica, Torreón, Coahuila.
- Hospital de oncología del centro médico nacional siglo XXI

Para el estudio realizado en 2019 por Olivier y colaboradores se contó con la participación de 11 centros, no sólo de México, si no de Argentina, Colombia y Guatemala.

- Centro oncológico Estatal del Estado de Toluca
- Centro médico nacional Siglo XXI
- Clínica universitaria Colombia
- Hospital universitario San Ignacio
- UPB Clínica Universitaria Bolivariana
- IMAT- Oncomédica S.A.
- Centro médico Imbanaco
- Fundación cardiovascular de Colombia
- Instituto Nacional de enfermedades Neoplásicas (Perú)
- Hospital italiano de Buenos Aires
- Instituto para la Investigación Científica y la Educación Acerca de las Enfermedades Genéticas y Metabólicas Humanas (Guatemala)

Posteriormente, en este trabajo de tesis se analizaron 422 nuevas muestras provenientes de México, Perú y Guatemala.

Distribución geográfica y número de muestras

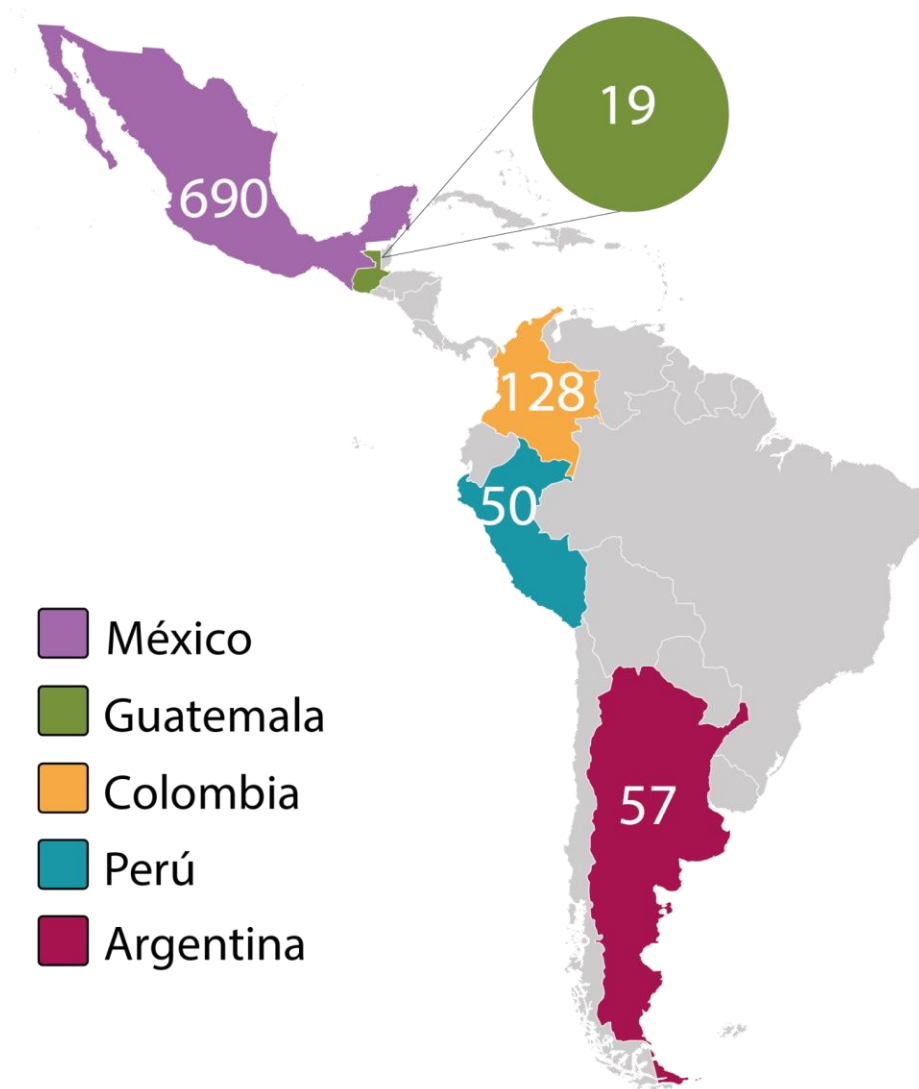


Figura 4. Composición geográfica de las 944 muestras analizadas. Se muestra el país de procedencia de las 944 muestras. El número de muestras por país está indicado en números blancos.

A cada paciente se le extrajo un total de 10 mL de sangre total. La extracción fue realizada por personal capacitado de cada centro participante en cualquiera de las 3 etapas de este trabajo recopilatorio. Las muestras sanguíneas fueron depositadas en tubos vacutainer con EDTA para evitar la coagulación de las muestras y posteriormente se almacenaron en refrigeración a -20°C hasta el momento de la extracción de DNA.

6.2.2. Extracción de DNA

La extracción del DNA se hizo mediante el kit DNeasy® Blood and Tissue de Qiagen [37] en todas las muestras, salvo las muestras provenientes de argentina (Figura 5).

Para la extracción de DNA genómico se utilizó 100 µL de sangre total, posteriormente se agregaron 20 µL de proteinasa K y finalmente se agregaron 100 µL de buffer PBS para tener un volumen total de 220 µL. También se agregaron 8 µL de RNasa A. Después se incubaron las muestras a 56° C por 10 minutos. Una vez terminada la lisis de las células sanguíneas, se agregó 200 µL de etanol absoluto (96-100%). La mezcla se agitó utilizando un vortex y después se transfirió a una columna DNeasy unida a un tubo de recolección de 2 mL. Las muestras se centrifugaron a 8000 rpm durante 1 minuto. El sobrenadante que terminó en el tubo de recolección fue desechado.

La columna de afinidad fue puesta en un nuevo tubo de recolección, se agregaron 500 µL de buffer AW1 y se volvió a centrifugar a 8000 rpm durante 1 minuto. Nuevamente se desechó el sobrenadante hallado en el tubo de recolección.

Un nuevo ciclo de lavado se llevó a cabo, para esto se transfirió la columna de afinidad a un nuevo tubo de recolección y se agregaron 500 µL de buffer AW2 para centrifugar a 14000 rpm durante 3 minutos. El sobrenadante volvió a ser descartado. Finalmente, se eluyó el DNA colocando la columna de afinidad en un tubo de microcentrífuga de 2 mL y agregando 100 µL de agua bidestilada libre de nucleasa para después centrifugar a 8000 rpm durante 1 minuto. Con el fin de aumentar el rendimiento de la extracción se repitió el último paso, haciendo pasar el DNA purificado por la columna de afinidad de nuevo.

La concentración de DNA de cada muestra fue cuantificada mediante Fluorometría. Para ello se utilizó el equipo *Qubit*™ siguiendo el protocolo proporcionado por el fabricante [46]; La pureza se midió utilizando un espectrofotómetro con el coeficiente 260/280 nm; La integridad del DNA Genómico de alto peso molecular se comprobó mediante un gel de agarosa al 0.8%, utilizando un marcador de 1 kbp. Todas las

muestras por encima de 10,000 bp se consideraron como DNA con alta integridad [47].

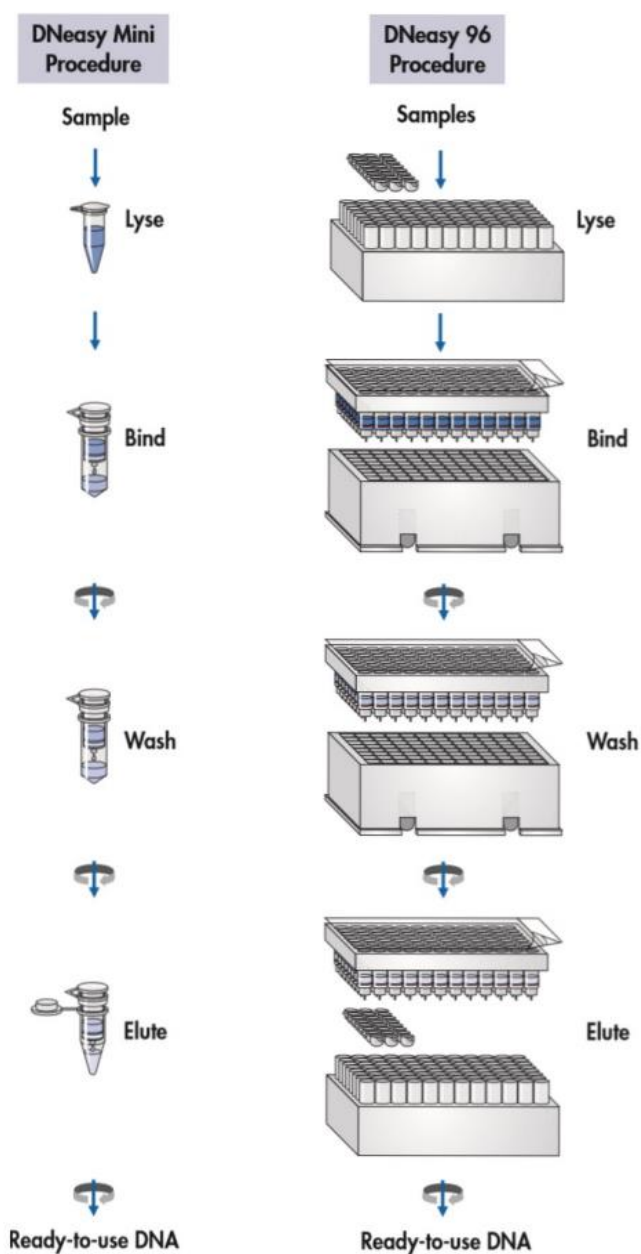


Figura 5. Extracción de DNA mediante columnas de afinidad. El sistema DNeasy® Blood & Tissue permite una extracción eficiente con un gran número de muestras de forma simultánea (Tomada de Qiagen, 2011).

El DNA de las muestras provenientes de Argentina fue extraído utilizando la plataforma Magna Pure System (Roche) siguiendo las instrucciones del fabricante.

6.2.3. Preparación de bibliotecas

6.2.3.1. Selección del panel

Para la preparación de la biblioteca de secuenciación se utilizó el Panel Human *Cancer predisposition panel* de Qiagen. Este kit contiene el catálogo más amplio de genes relacionados con enfermedades oncológicas (Figura 6). El panel se basa en una PCR multiplex de enriquecimiento de la región codificante de 143 genes de riesgo para 88 enfermedades oncológicas. Otra de las ventajas del kit es su practicidad, ya que las regiones blanco a amplificar se engloban en un total de 6582 pares de primers, distribuidos en 4 reacciones, lo cual permite optimizar los tiempos de preparación de las muestras [48].

Human cancer predisposition panel

Se basa en el enriquecimiento por PCR multiplex de la región codificante de 143 genes en 88 enfermedades oncológicas

Algunas de las enfermedades relacionadas con genes incluidos en el panel



Figura 6. Composición de algunos genes incluidos en el panel de secuenciación y su implicación en padecimientos oncológicos.

Este panel (Human cancer predisposition panel) se basa en el enriquecimiento por PCR multiplex de la región codificante de 143 genes asociados a diversos tipos de cáncer.

Para comenzar la preparación de bibliotecas de secuenciación, todas las muestras fueron diluidas a una concentración de 2.5 ng/μL. Posteriormente se hizo el enriquecimiento de la región de interés la cual comprende la región codificante de los 143 genes de susceptibilidad a cáncer incluidos en este panel. Para esto se llevó a cabo 4 reacciones de PCR multiplex para cada muestra. En total, los 4 lotes de reacciones comprenden 6582 pares de primers. El diseño de los oligos en este panel está diseñado de tal forma que cada amplicón queda solapado de forma contigua con los demás y de esta manera logran cubrir en su totalidad la región de interés (Figura 7).

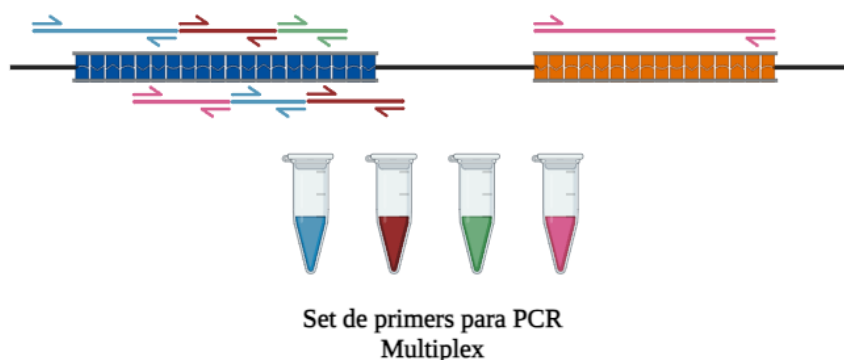


Figura 7. Enriquecimiento de la región de interés mediante una PCR multiplex. El panel GeneRead está diseñado con un sofisticado arreglo de primers para maximizar la cobertura de la región de interés y minimizar la formación de productos inespecíficos durante la amplificación.

Una vez hecho el enriquecimiento, se mezclaron los 4 conjuntos de reacciones de PCR y se purificaron los productos de las 4 reacciones para retirar los primers y productos no específicos que pudieran haberse formado durante la PCR. La purificación se hizo mediante perlas paramagnéticas de la marca AMPure XP beads™ con una proporción de 0.9 μL de perlas por cada microlitro de reacción de PCR.

Los siguientes pasos en la preparación de las bibliotecas de secuenciación fueron: La reparación de los extremos protuyentes, adenilación de los amplicones y la adición de adaptadores universales.

Para la reparación de los extremos romos, se adicionó 20.5 μL de DNA enriquecido previamente, 2.5 μL de buffer End-Repair (10X) y 2.0 μL del mix de T4 DNA polimerasa, después se incubó a 25° C por 30 minutos, seguido de otra incubación a 75° por 20 minutos. La adenilación se hizo agregando 25 μL del DNA reparado, 3 μL de A-addition buffer y 3 μL del fragmento klenow luego se incubó a 37° C por 30 minutos, seguido de otra incubación a 75° C por 10 minutos. Finalmente, la unión de adaptadores se hizo añadiendo 31 μL de DNA previamente adenilado, 45 μL de buffer de ligación (2X), 1 μL de adaptador, 4 μL de T4 DNA ligasa y 9 μL de agua libre de DNasa. Esta mezcla se incubó a 25° C por 10 minutos. El adaptador está compuesto por el primer de secuenciación de Illumina™ y una combinación única de 24 pares de bases que sirven como una etiqueta molecular, la cual permitió distinguir cada una de las muestras posterior a la secuenciación masiva.

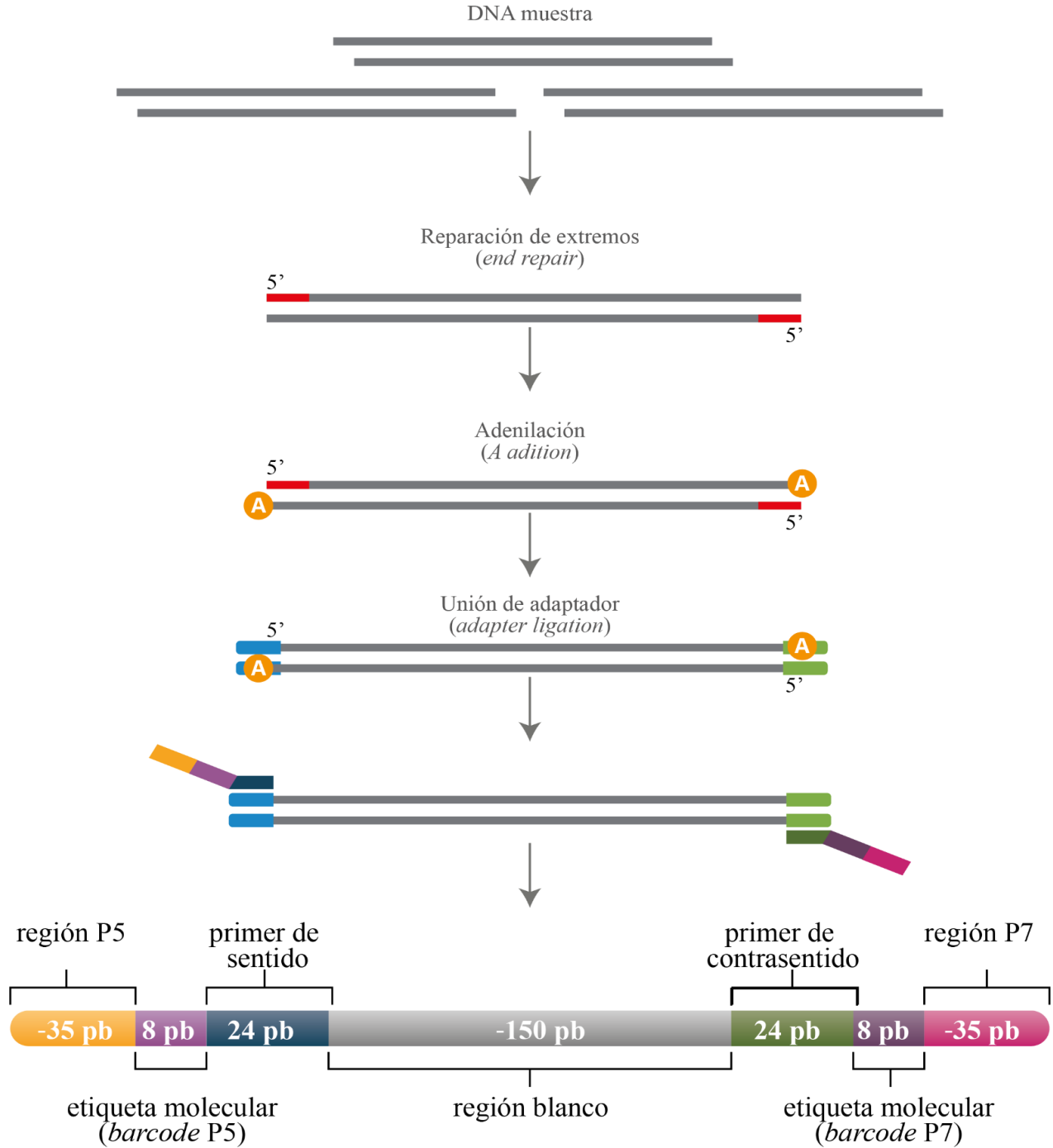
Una vez terminada la unión de adaptadores, es necesario hacer una purificación para seleccionar todos los fragmentos finos que tengan un tamaño aproximado de 200 bp (150 bp de amplicón + 48 bp de adaptadores). Para esta purificación se utilizaron las perlas paramagnéticas a una concentración de 1.2 μL por cada microlitro de reacción ya ligado a sus respectivos adaptadores. El resto de la purificación se hizo siguiendo las instrucciones del fabricante.

Por último, se amplificaron las bibliotecas purificadas agregando 25 μL de HiFi PCR master Mix (2X), 1.5 μL de cada uno de los primers, 17 μL de la biblioteca de DNA purificada y 6.5 μL de agua libre de RNasa. La PCR de esta amplificación constó de una desnaturalización inicial a 98° C por 2 minutos; 4 ciclos de elongación compuestos de la siguiente manera, 98°C por 20 segundos, 60°C por 30 segundos, 72°C por 30 segundos; La extensión final se hizo a 72°C por 1 minuto (Figura 8).

La biblioteca de DNA amplificada pasó por una purificación doble, utilizando las perlas paramagnéticas con una relación de 0.8 μL por cada microlitro de biblioteca de DNA y posteriormente una concentración de 0.4 μL para únicamente conservar los fragmentos de 200 bp.

El tamaño de los fragmentos de las bibliotecas de DNA fue evaluado mediante electroforesis de alta resolución para corroborar que las secuencias se encontraban en el rango de 200 bp.

Biblioteca de DNA Illumina



Fragmento listo para secuenciación

Figura 8. Descripción gráfica de una biblioteca de DNA. A partir del DNA amplificado se reparan los extremos protuyentes, seguido de la adenilación de dichos extremos y para finalmente unir dos adaptadores universales. El resultado es una secuencia que incluye la región P5, la etiqueta molecular P5, el primer de sentido, la región de interés o blanco, el primer de contrasentido, la etiqueta molecular P7 y la región P7.

6.2.4. Secuenciación

Para el proceso de secuenciación, todas las bibliotecas previamente preparadas, fueron ecualizadas para asegurar que todas las muestras tuvieran la misma molaridad. Este paso es fundamental en las plataformas de secuenciación masiva en paralelo para evitar que alguna muestra esté sobre representada durante el proceso de secuenciación.

Las 944 muestras fueron secuenciadas usando las plataformas de Illumina Miseq y Hiseq. Salvo las 50 muestras provenientes de Argentina, las cuales fueron secuenciadas con la tecnología de Ion Torrent.

Los secuenciadores de la marca Illumina son equipos de secuenciación masiva en paralelo. El principio mediante el cual funcionan es la secuenciación por síntesis con terminadores reversibles.

Todas las muestras fueron depositadas en el cartucho de secuenciación siguiendo las instrucciones del fabricante. El cartucho de secuenciación contiene todos los reactivos para llevar a cabo la secuenciación. Posteriormente el equipo depositó las muestras en una lámina de flujo (*Flowcell*) que se encuentra dentro del secuenciador. Esta lámina de flujo contiene secuencias complementarias a los adaptadores P5 y P7, unidas mediante enlaces covalentes a la lámina. De esta forma las moléculas de DNA desnaturalizadas de las bibliotecas quedaron ancladas a la lámina de flujo mediante la hibridación por secuencias complementarias. Posteriormente el secuenciador comenzó a hacer una amplificación por puente de cada una de las moléculas ancladas para formar un grupo de clones (clúster) en la lámina de flujo. Durante la amplificación por puente se sintetiza la cadena complementaria en sentido 3' a 5', las hebras sintetizadas se doblan y se unen mediante el adaptador complementario P5 o P7 según sea el caso. Este proceso se repite hasta que no hay más fragmentos P5 ni P7 disponibles en la lámina de secuenciación. Dado que la secuenciación por síntesis de los equipos Illumina utilizan fluorescencia, la formación de "clusters" es importante para amplificar la

intensidad de las señales de fluorescencia que serán captadas por la cámara del equipo y de esta forma leer las moléculas de DNA (Tabla 7).

Durante cada uno de los ciclos de los equipos Miseq o Hiseq, una DNA polimerasa incorpora un nucleótido con un fluoróforo y un terminador reversible asociado, el cluster emite una señal de fluorescencia que el equipo registra y la síntesis de DNA se detiene debido al terminador reversible. Posteriormente se hace un ciclo de limpieza para retirar el terminador reversible y el proceso se repite de nuevo hasta terminar la secuencia (Figura 9).

Tipo de secuenciador	Tamaño de lectura (bp)	Técnica de secuenciación	Detección de señal	Notas acerca de la tecnología	Lecturas por corrida (bp)	Tiempo por corrida
Secuenciación por síntesis (Illumina)	30 - 300	Síntesis	Florescencia	Los clusters se generan en una lámina de flujo mediante amplificación por puente con bases con terminadores reversibles	Hasta 1000 millones	1 a 11 días
Semiconductor de iones (Ion torrent)	Hasta 600	Síntesis	Flujo de iones a través de un semiconductor	Las secuencias blanco se colocan en un micro pocillo. Los nucleótidos se inundan e incorporan mediante la polimerasa. La adición de cada nucleótido genera un ion hidrógeno que es detectado por el semiconductor.	Hasta 80 millones	2 a 5 horas

Tabla 7. Comparación de plataformas de secuenciación Ion Torrent e Illumina. (Modificada de Church, 2020)

Todas las muestras se secuenciaron usando un formato de corrida de 2 por 150 bp dado que el tamaño promedio de los amplicones generados durante la preparación de bibliotecas es de 200 bp.

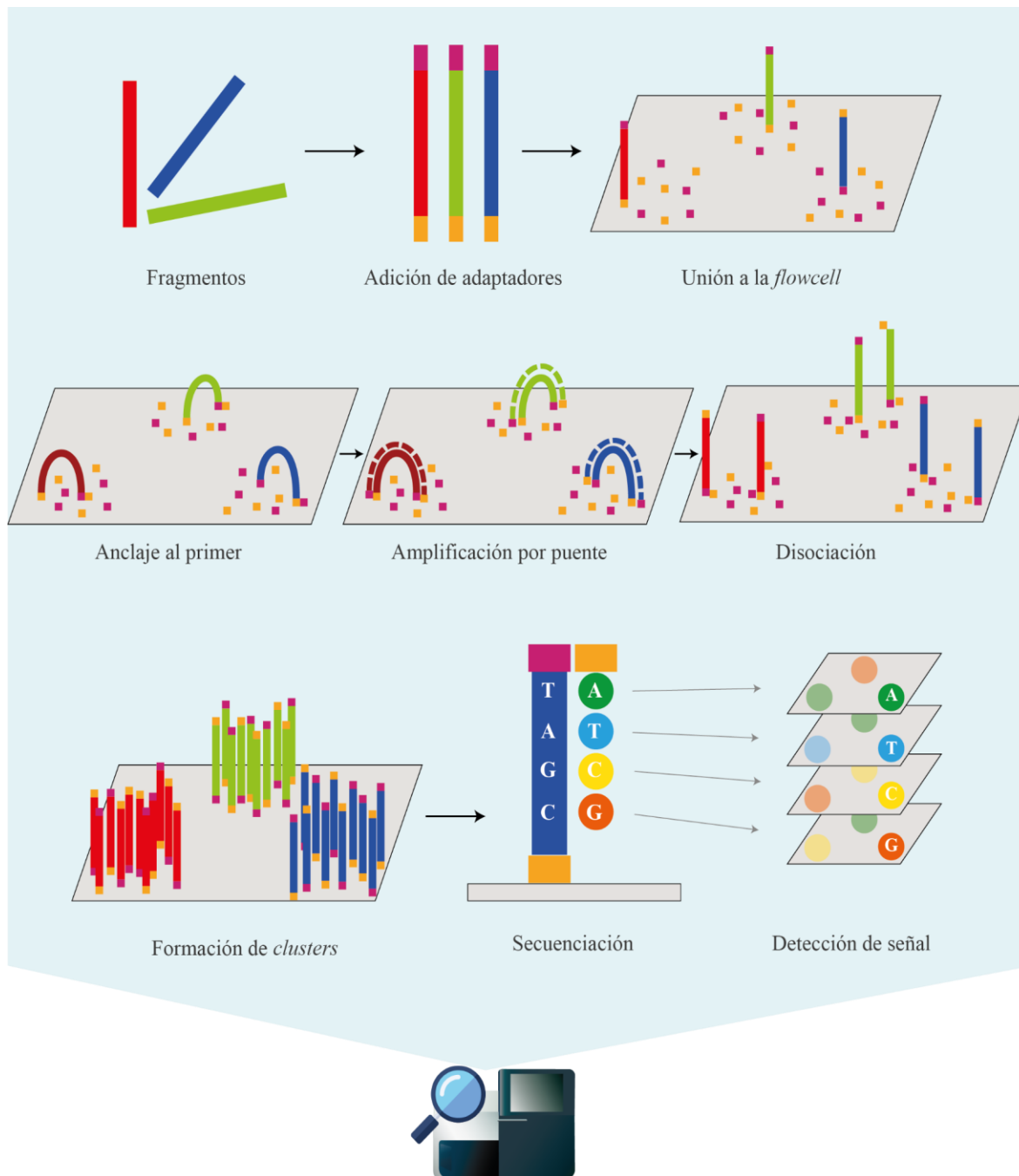


Figura 9. Secuenciación por síntesis de Illumina. Durante la secuenciación masiva en paralelo cada molécula con un adaptador es unido a una lámina de flujo mediante un enlace covalente. Posteriormente cada una de las secuencias se amplifica en un proceso de amplificación por puente, para formar grupos locales que contienen miles de clones de la misma molécula. Finalmente, el equipo registra la señal de fluorescencia cuando cada nucleótido es incorporado.

6.2.5. Análisis bioinformático

Una vez secuenciadas todas las muestras, el equipo proporciona varios archivos en formato FASTQ, los cuales contienen las secuencias y valores de calidad de dicha secuencia [38]. En nuestro caso particular, los secuenciadores emiten dos archivos para cada una de las muestras, estos archivos corresponden a la cadena con sentido y contrasentido de toda la región de interés, es decir, todos los amplicones que contemplan los paneles de secuenciación empleados en este proyecto.

El primer paso en el análisis bioinformático de datos genómicos es asegurar la calidad de las lecturas. Para esto se revisó que todas las secuencias tuvieran un puntaje de calidad Q mayor a 30. Este valor corresponde a la escala PHRED y es asignado por el propio equipo y refleja la probabilidad de que el equipo haya cometido un error ($Q30 = 99.9$ de precisión).

El segundo paso es hacer un alineamiento de cada una de las secuencias obtenidas usando el software BWA [39-40]. Para este alineamiento se empleó una referencia genómica genérica de *Homo Sapiens*, en concreto hg19, la cual es un estándar para este tipo de análisis [41]. Las lecturas sentido y contrasentido de cada paciente fueron unidas y ubicadas genéticamente con este proceso. El resultado de este paso es un archivo denominado SAM. Luego el archivo SAM es convertido en su homólogo binario con finalidad de que sea más sencillo el manejo y almacenamiento de toda la información genómica que contiene, con este paso obtenemos un nuevo archivo denominado BAM.

Dado que varias de las lecturas obtenidas son clones, producto de las múltiples PCR's de la preparación de bibliotecas. Fue necesario marcar los duplicados con el fin de sobre evidenciar cualquier hallazgo en estos datos. Finalmente, en el preprocesamiento de los datos se hace una recalibración de los valores de calidad que emitió el secuenciador, esto para asegurar la calidad de los futuros análisis sobre estos datos [42].

En este tipo de estudios, un paso clave es el llamado de variantes. En este proceso un potente algoritmo (HaplotypeCaller GATK) se encargó de detectar todas las variaciones de un solo nucleótido (SNV) así como deleciones e inserciones cortas (Indels). Los criterios y algoritmos que utiliza este programa para identificar una verdadera variante genómica, de falsos positivos (error de secuenciación o alineamiento) van más allá de esta tesis, por tal motivo no profundizaré en ello, sin embargo, pese a que este software ha sido probado y existe literatura que confirma su exactitud [44] Los datos que el programa Haplotype Caller arrojó, también fueron validados de manera experimental.

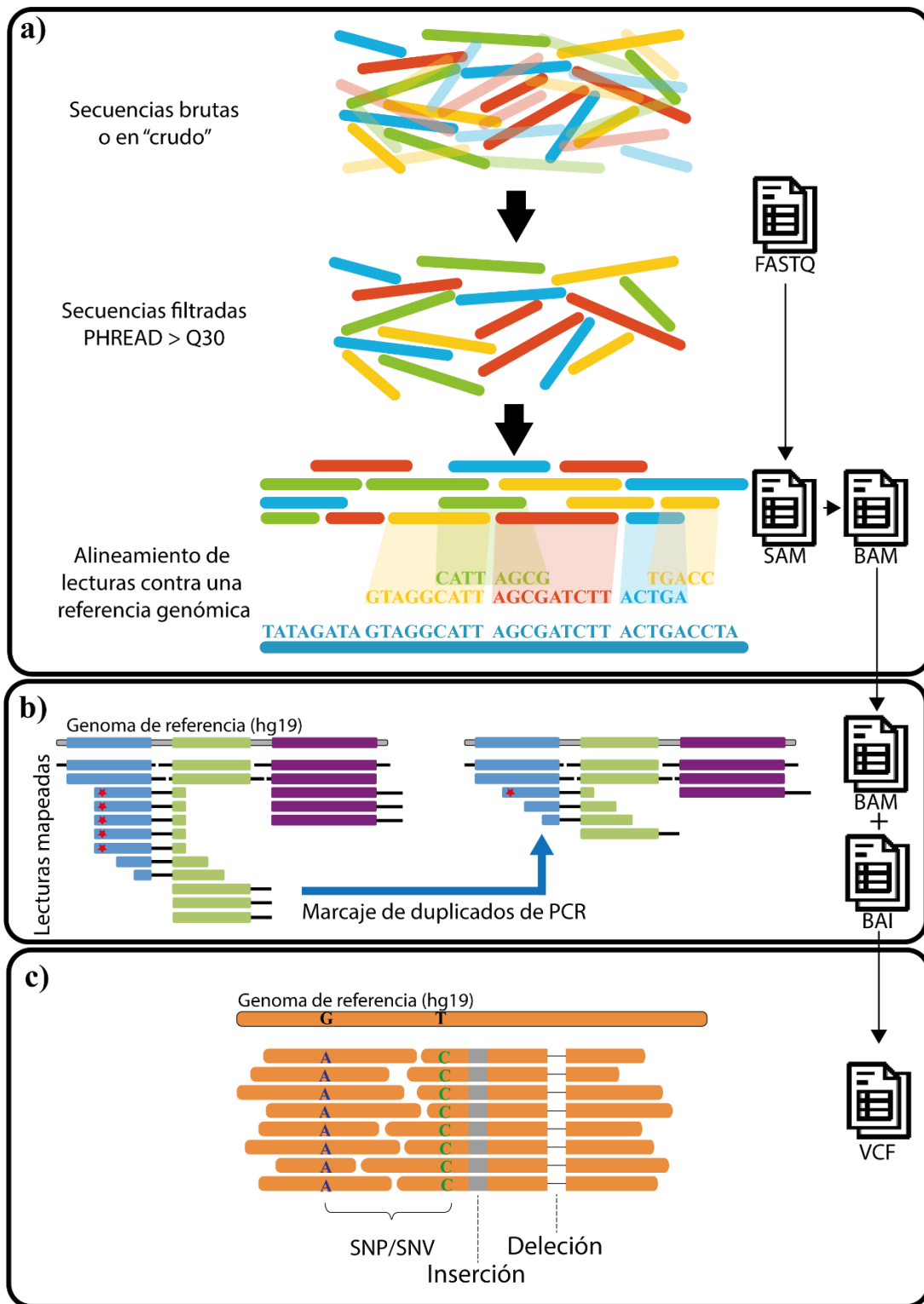


Figura 10. Flujo de trabajo bioinformático. a) Los datos de secuenciación son filtrados usando los valores de calidad que proporciona el secuenciador, posteriormente cada lectura fue alineada contra un genoma de referencia y la secuencia de cada paciente es reensamblada. b) Una vez la secuencia de cada paciente fue reconstruida, se eliminan los duplicados de PCR para evitar posibles falsos positivos debido a la propagación de errores durante las múltiples amplificaciones durante el proceso de preparación de bibliotecas. c) El llamado de variante permite evidenciar todos los cambios en los datos genómicos de los pacientes respecto al genoma de referencia en un proceso llamado "variant calling".

6.2.6. Selección de variantes patogénicas y VUS

El primer paso en el análisis de variantes genómicas fue la depuración de datos. A pesar de que existe una variación aproximada del 0.1% entre el genoma de dos individuos no mellizos homocigóticos [50], esto da lugar a un estimado de más de un millón de nucleótidos distintos distribuidos a lo largo de todo el genoma. Los cálculos más conservadores reportan alrededor de 1.4 millones de SNPs [51].

Al realizar el llamado de variantes (*variant calling*) se obtuvo un extenso catálogo de polimorfismos de un solo nucleótido (SNP), variantes de un solo nucleótido (SNV) y deleciones e inserciones cortas (Indels) de cada paciente. Para reducir el número de variantes y seleccionar las mutaciones asociadas al síndrome hereditario de cáncer de mama y ovario el primer paso fue realizar la anotación genómica de cada una de estas variantes a través del software Annovar.

Annovar es una herramienta que permite conocer el contexto funcional a nivel biológico de SNV e INDELS, así como el impacto clínico de dichas variantes y la incidencia y frecuencia alélica poblacional con la que se reportan dichas variantes en divergentes bases de datos que comprenden a diversas poblaciones de todo el mundo [49]. Este programa integra bases de datos como *1000 Genomes* [52], *dbSNP* [53], *GenomeAD* [54], entre otras, las cuales sirven de repositorios de variantes genéticas a nivel mundial. Estas bases de datos llevan un registro de la variación genómica en diferentes poblaciones y subpoblaciones, indicando la frecuencia con la que una variante puede aparecer en una población específica. Por otra parte, Annovar también integra bases de datos clínicas como: *ClinVar* [55], *COSMIC* [56], *InterVar* [57] y *VarSome* [58]; las cuales clasifican las variantes genéticas con base en la relación que existe entre alguna variante y un padecimiento o desorden médico. Algunas bases de datos de este estilo, como COSMIC también sirven de catálogo de mutaciones somáticas halladas en pacientes con cáncer. Finalmente, el anotado también integra algunos programas predictivos que estiman el impacto del cambio de nucleótidos de un gen o región

específica en la secuencia de aminoácidos de dicho gen como es el caso de *SIFT* [59] y Polyphen [60].

Una vez anotados los datos de cada paciente, se procedió a filtrar los datos de cada uno de ellos. El primer criterio de filtrado fueron los valores numéricos generados durante el variant calling. Los parámetros seleccionados fueron un valor de calidad mayor a Q30. A su vez, la alteración de nucleótidos debía de ser observada en al menos 30 lecturas (Profundidad $\geq 30X$) y estar presente en ambas cadenas de nucleótidos, sentido y contrasentido (no strand bias).

Los siguientes criterios de selección de variantes, fueron con base a la ubicación de la variante en el contexto génico. De esta forma se excluyeron todas las variantes halladas en regiones Inter génicas, regiones intrónicas, sitios 5'UTR y 3'UTR, así como en zonas de homopolímeros.

Posteriormente se filtraron las variantes restantes haciendo uso de las bases de datos poblacionales. De esta manera se descartaron las variantes “comunes” o frecuentes. Para tal propósito solo se seleccionaron las variantes con una frecuencia alélica poblacional menor 0.01 en las compendios de 1000Genomes, dbSNP, Exac y GnomeAD. Esto debido a que las variantes patológicas o con efectos deletéreos a nivel biológico son raras y poco observables en poblaciones sanas.

Finalmente se eligieron las variantes que cumplieron todos los criterios anteriores y que además habían sido descritas en genes de alto riesgo asociados a cáncer, reportadas en las bases de datos de ClinVar, COSMIC o VarSome y a su vez los algoritmos predictivos (*SIFT* y Polyphen) que reportaban un impacto negativo a nivel funcional en la secuencia de aminoácidos o proteína (Figura 11).

Las variantes con significado clínico incierto fueron (VUS) analizadas aplicando los mismos criterios, salvo que de las bases de datos clínicas (ClinVar, VarSome o COSMIC) se seleccionaron únicamente las variantes en las cuales la evidencia no

era clara, había conflictos en la interpretación de la alteración genética o dicha alteración había sido catalogada como probablemente patogénica. Para refinar los datos de las variantes en esta categoría se seleccionaron las mutaciones en genes de “alto riesgo” asociados a cáncer (*BRCA1*, *BRCA2*, *TP53*, *PTEN*, *STK11*, *CDH1*, *MLH1*, *MSH2*, *MSH5*, *EPCAM*, *PMS2*, *CDKN2A*, *PALB2*) [61.]

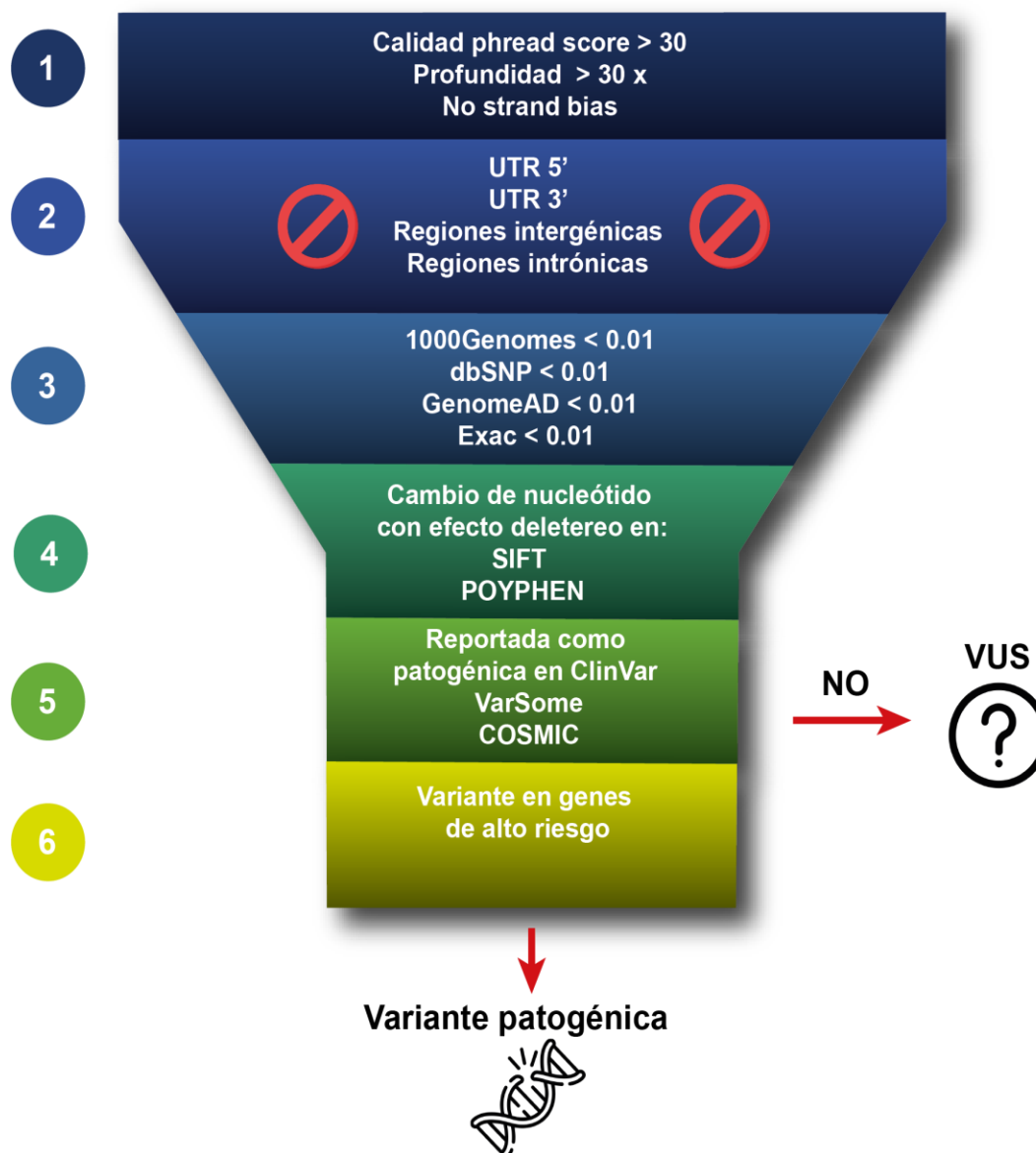


Figura 11. Criterios de filtrado de variantes genéticas. Los datos obtenidos mediante el “variant calling” fueron filtrados siguiendo el orden jerárquico mostrado en el esquema.

6.2.7. Análisis de la mutación mexicana fundadora (Delección exón 9-12)

Todas las muestras provenientes de pacientes de México fueron analizadas para la mutación fundadora BRCA1 delección exón 9-12 [62] mediante una reacción en cadena de la polimerasa. El diseño de esta PCR se basó en 3 oligonucleótidos descritos en la tesis de maestría de Quezada Urban (Figura 12) [63]. Esta mutación fundadora ha sido reportada en estudios previos realizados en pacientes mexicanas, encontrándose esta mutación hasta en el 33% de las pacientes portadoras de una mutación en BRCA1 [64-65].

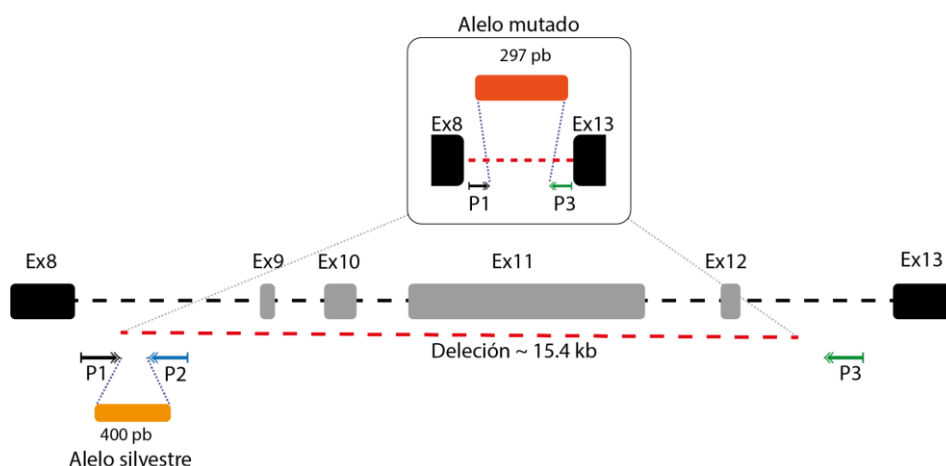


Figura 12. Diseño de PCR para la detección de la delección 9-12. El par de oligos para el alelo silvestre P1 y P2 amplifican para la región ubicada entre el exón 8 y el exón 9. La amplificación para el alelo mutante (P1, P3) sólo es posible cuando la región entre el exón 9 y 12 se encuentra ausente (~15.4 kb). (Modificada de Quezada Urban, 2017)

6.2.8. Validación de resultados

Finalmente, todos los resultados de las pacientes portadoras de una mutación patogénica fueron verificados mediante secuenciación Sanger y en análisis de la secuencia en su correspondiente electroferograma.

7. RESULTADOS

En este trabajo se analizó un total de 422 pacientes provenientes de diferentes países de Latinoamérica. También se contó con el consentimiento explícito para utilizar los datos de las 300 pacientes analizados por Quezada *et al*, 2017 (DOI: 10.3390/cancers10100361) y las 222 pacientes obtenidas en el estudio de Oliver J. *et al* 2019 (DOI: 10.3389/fonc.2019.01429). (Tabla 7.)

Cohorte	Número de pacientes y nacionalidad	
Quezada, et al 2017	México	(300)
Olivier et al 2019	México	(68)
	Guatemala	(19)
	Colombia	(78)
	Argentina	(57)
Propia	México	(323)
	Colombia	(50)
	Perú	(49)
TOTAL		944

Tabla 8. Composición de población analizada y las 3 cohortes que la componen.

Respecto a la distribución geográfica, las pacientes provenientes de México representaban el 73.2% (691/944); Colombia 13.55%(128/944); Argentina 6.03%(57/944), Perú 5.2%(49/944) y Guatemala 2.01% (19/944). Con base en estos porcentajes sólo el 17.36%(120/691) de las pacientes mexicanas dio positivo para alguna alteración patogénica. Para Colombia, solo 12.5% (16/128) de las pacientes presentaron alguna variante patogénica. En el caso de Argentina, el 17.5% (8/57) fue el porcentaje de pacientes con más alteraciones patogénicas. Finalmente, Perú tuvo un porcentaje de 12.2% (6/49) de pacientes positivas y Guatemala 15.8% (3/19) (Figura 13).

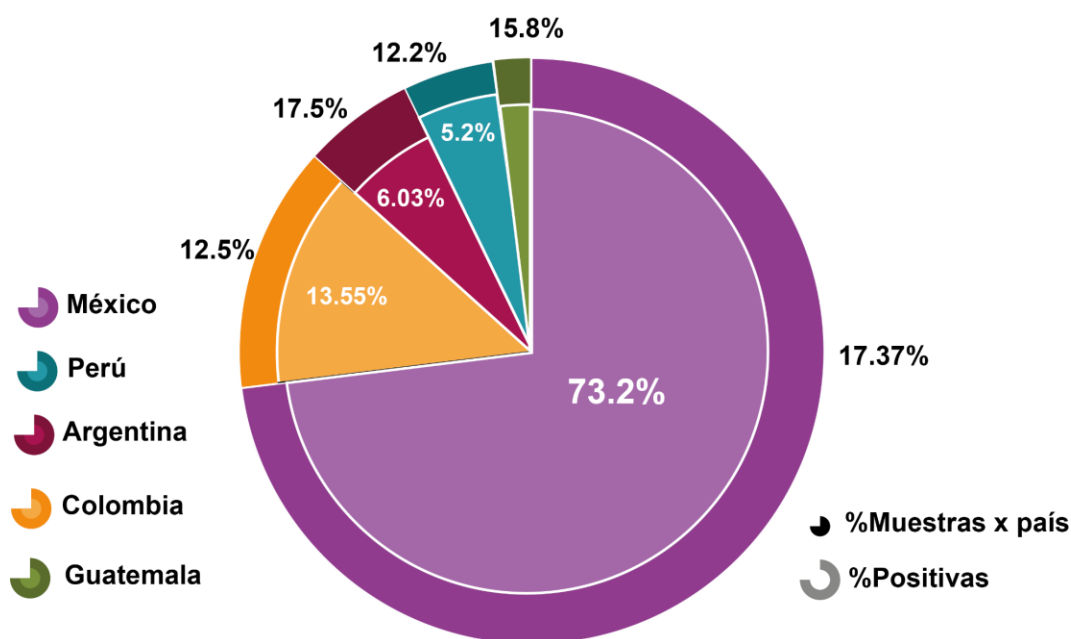


Figura 13. Proporción de muestras por país y porcentaje de variantes halladas. Se muestra la proporción de muestras por país respecto al total de 944 muestras incluidas en este estudio. El aro exterior representa el porcentaje de pacientes positivas respecto al total de muestras pertenecientes a ese país.

El análisis de la región codificante de 143 genes permitió detectar variantes patogénicas en genes de alto riesgo en el 9.95% (94/944) de las pacientes (Figura 14). Los genes de alto riesgo con variantes patogénicas mayormente afectados fueron *BRCA1* 47% (45/94) de las cuales 16 variantes corresponden a la mutación fundadora Deleción 9-12; *BRCA2* 36% (32/94), *ATM* 6% (7/94) y *PALB2* 6% (7/94). Otras variantes en genes de alto riesgo fueron *APC*, *CDKN2A*, *MLH11*, *MSH2*, *MSH6* y *TP53*, cada uno con una única alteración patogénica (1.06% 1/94).

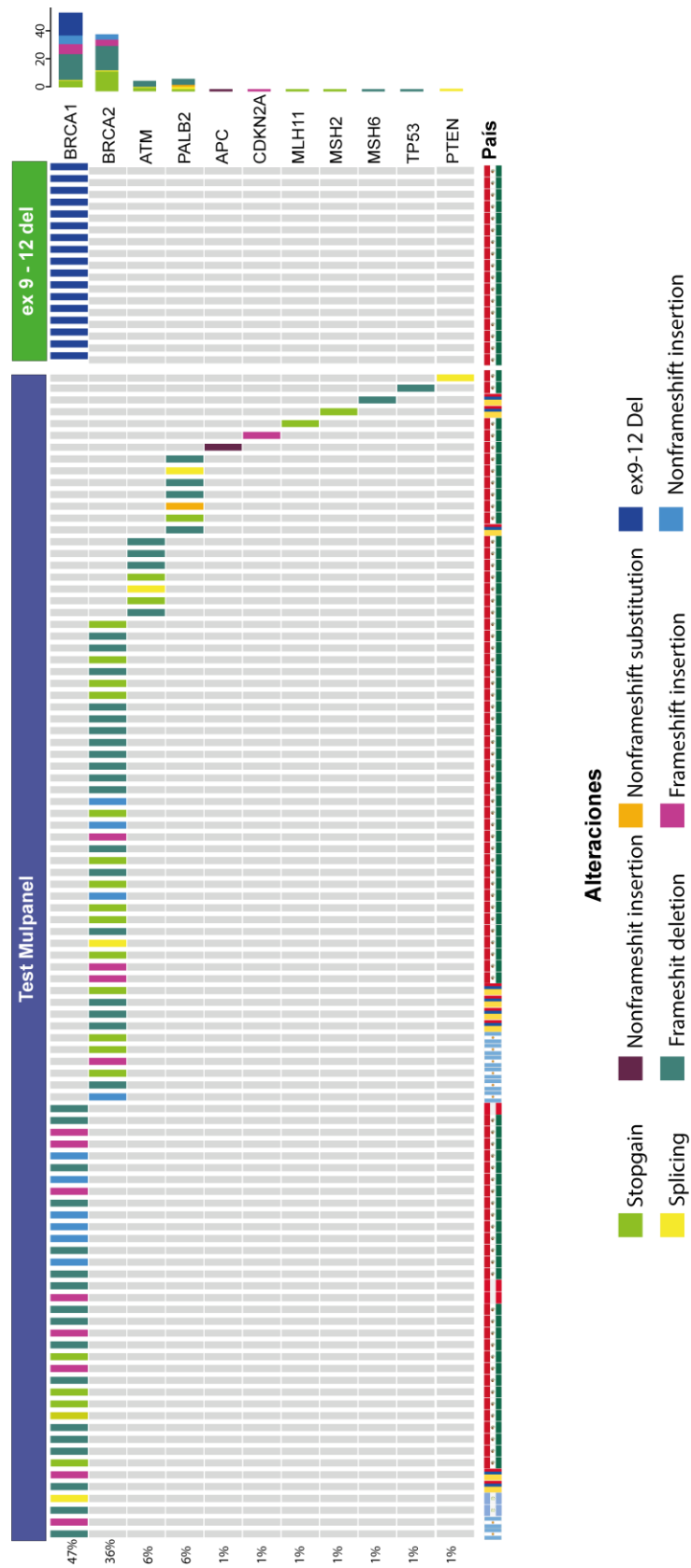


Figura 14. Distribución de variantes patogénicas en genes de alto riesgo. Se muestra las mutaciones halladas en los genes de alto riesgo (*BRCA1*, *BRCA2*, *ATM*, *PALB2*, *APC*, *CDKN2A*, *MLH1*, *MSH2*, *MSH6*, *TP53*). El código de color indica el tipo de alteración.

Asimismo, también se detectaron 61 alteraciones patogénicas en 37 genes de riesgo bajo o moderado (Figura 15). Dentro de esta categoría, los genes con mayor frecuencia de mutaciones fueron *MUTYH* y *NF1* con un 8% cada uno (5/61), seguido de *PDE11A* con un 7% (4/61) y *MSR1* con un 5% (3/61). Dentro de este grupo de genes de riesgo moderado o bajo, 10 genes dieron positivo para 20 alteraciones patogénicas, distribuyéndose 2 pacientes por gen (*ATR*, *ERCC3*, *FANCC*, *FANCI*, *FANCM*, *LIG4*, *RECQL4*, *RNASEL*, *UROD* y *WRN*). El resto de los 23 genes (*AIP*, *BIVM-ERCC5*, *BRIP1*, *CHEK2*, *DCLRE1C*, *DICER1*, *ERCC2*, *ESCO2*, *FANCB*, *FANCF*, *FANCL*, *MLH3*, *NBN*, *PDGFB*, *POLH*, *POU6F2*, *PTCH1*, *PTEN*, *RAD51B*, *RAD51C*, *SDHA*, *SDHB*, *XRCC3*) sólo presentaron una variante patogénica por paciente (1/61).

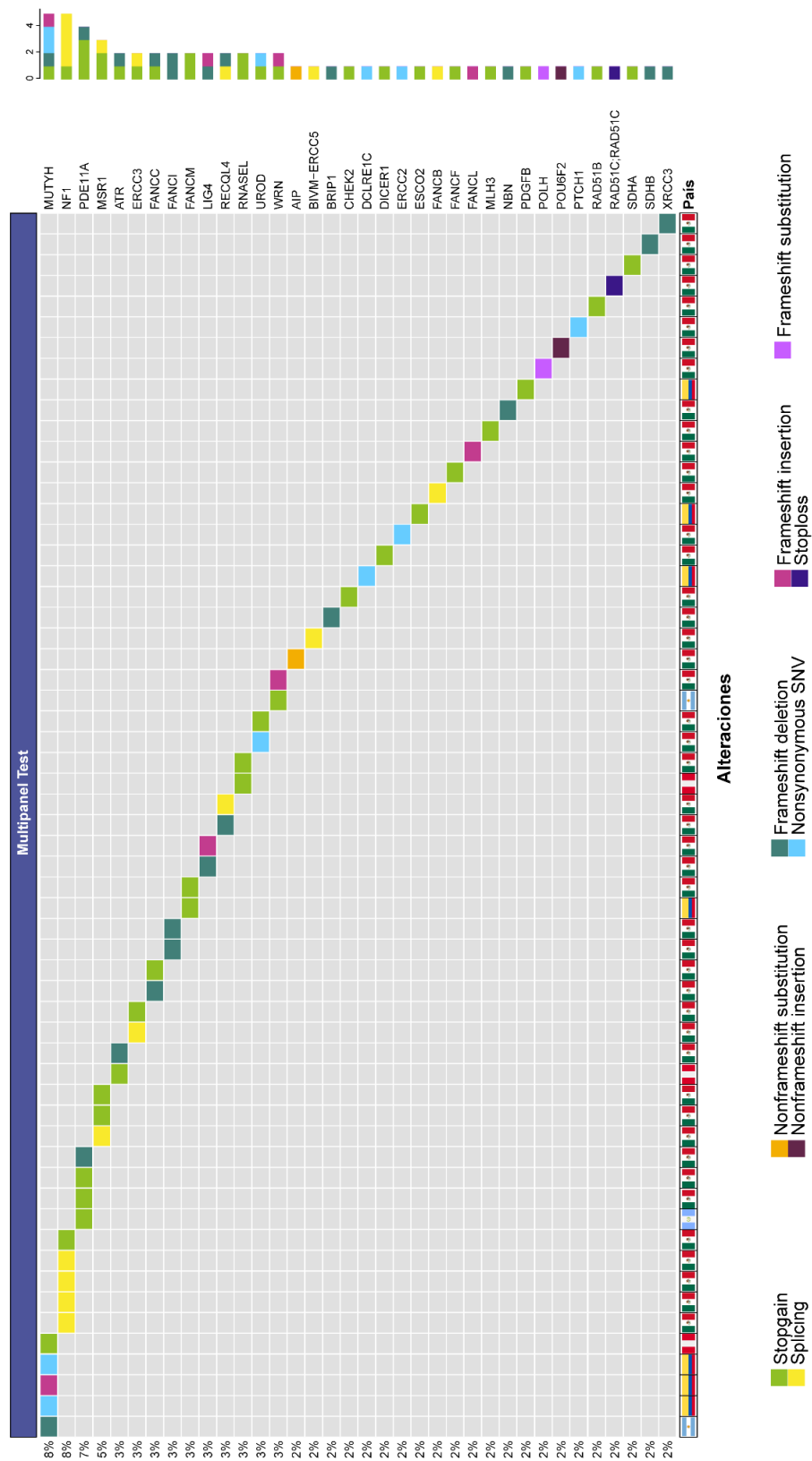


Figura 15. Distribución de variantes patogénicas en genes de riesgo moderado. Se muestra la distribución de las variantes en los pacientes (N = 60) con al menos una alteración. El tipo de alteración esta demarcada por el color de la celda

El llamado de variantes, en conjunto con el filtrado de datos, también arrojó un total de 39 variantes nuevas con significado clínico incierto (VUS) (Figura 16). Estas 39 VUS corresponden únicamente a las nuevas muestras incorporadas (323/944). El gen con mayor porcentaje de VUS fue *BRCA2* con un 44% (16/39) y *ATM* con un 23% (8/39). El resto VUS se distribuyó en *CHEK2* (3/39), *MSH6* (3/39), *PALB2* (3/39), *NF1* (2/39), *PMS2* (2/39) y *MSH2* (1/39).

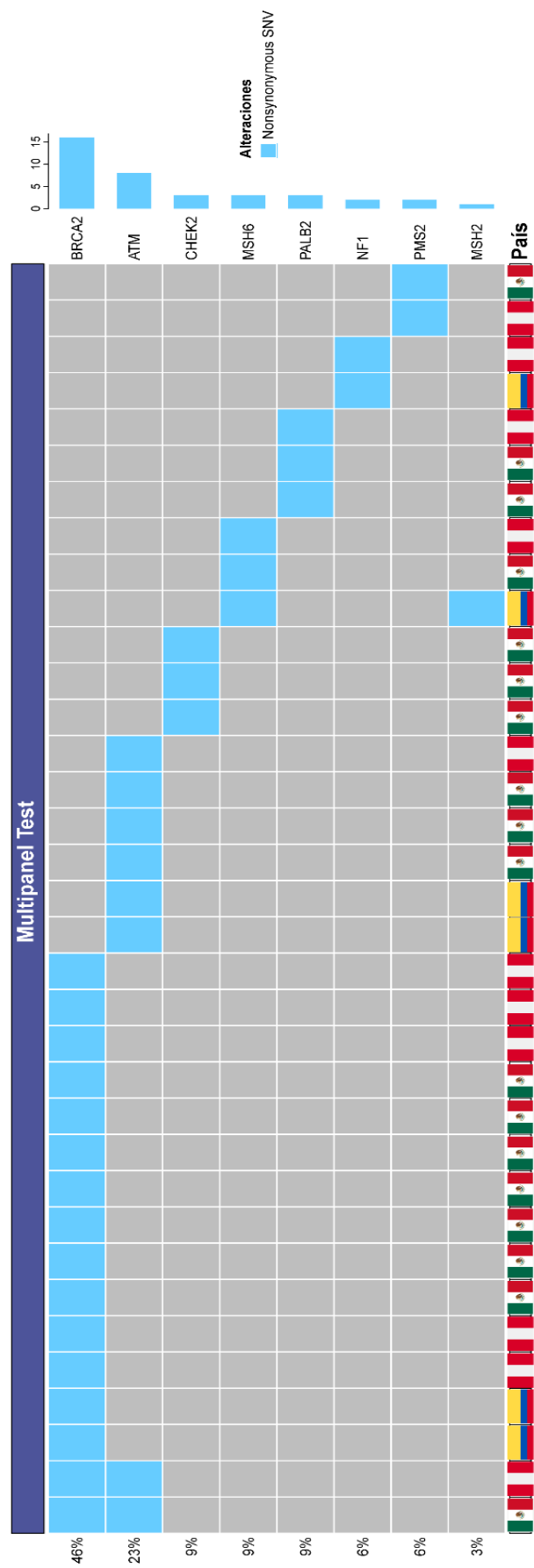


Figura 16. Distribución de Variantes de significado Clínico Incierto (VUS). Se muestra la distribución de 39 nuevas VUS detectadas en nuestra cohorte de 323 Pacientes.

En cuanto a las variantes frecuentes, se detectaron mutaciones recurrentes distribuidas en los genes *BRCA1*, *BRCA2* y *PALB2* (Figura 17). Estas mutaciones fueron detectadas en familiares no consanguíneos. En este estudio se hallaron las mutaciones *BRCA2* p.R2494X presente en dos pacientes, *BRCA1* p.A1708E presente en 4 pacientes y *PALB2* p. M723fs la cual se halló en 3 pacientes y la mutación *BRCA1* p.T276Afs presente en 3 individuos de una misma familia. La mutación *BRCA2* p.G578fs también se detectó en 5 familiares consanguíneos. Por otra parte, también se detectó de manera aislada una mutación fundadora previamente reportada en la población colombiana *BRCA1* p.Q1111fs, sin embargo esta mutación se encontró únicamente en un paciente de dicho país.

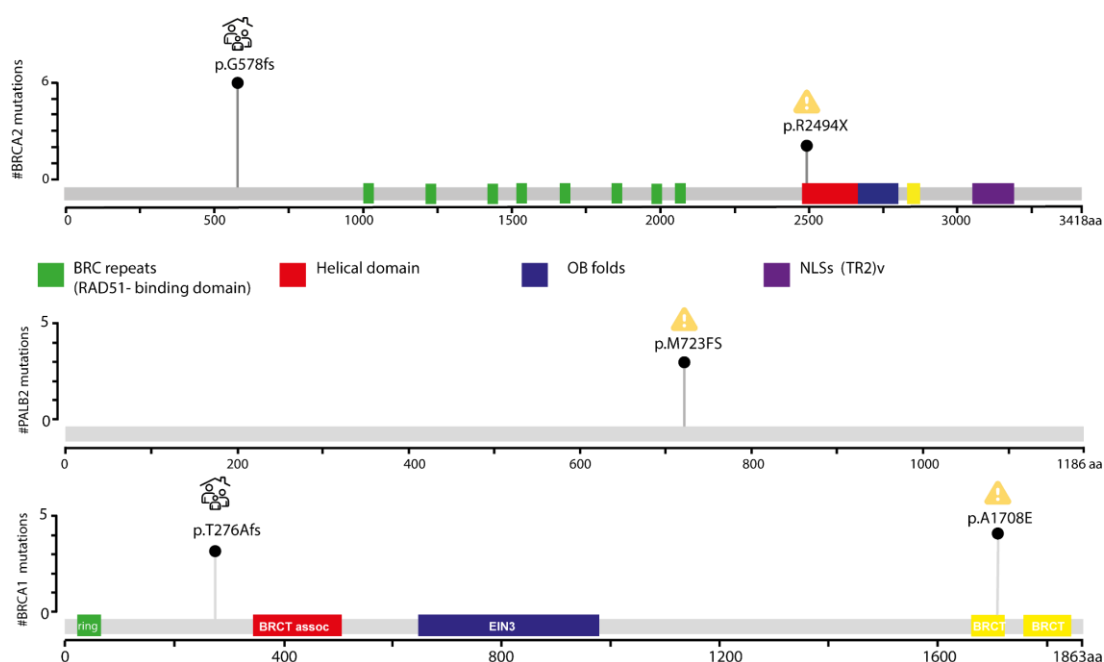


Figura 17. Mutaciones recurrentes. Se observan 4 mutaciones recurrentes en 3 genes distintos en pacientes sin ningún parentesco. También se reporta una mutación (p.G578fs) en familiares consanguíneos.

8. DISCUSIÓN

En este estudio se analizó la región codificante de 143 genes asociados a Cáncer en 944 pacientes con sospecha de síndrome hereditario de cáncer de mama y ovario. Se encontró que alrededor de 9.95% de las pacientes son portadoras de una variante patogénica en genes de alto riesgo y otro 6.46% fueron positivas para variantes en genes de riesgo bajo o moderado. En total el 16.41% de las pacientes presentó al menos una alteración en genes que pueden ser asociados al síndrome hereditario de cáncer de mama y ovario.

El alcance de este estudio abarcó a 5 países de América Latina cuyas poblaciones están subrepresentadas en estudios de talla internacional. El porcentaje de variantes positivas por país fue de 15.07% en promedio ($\sigma = 2.03$), este valor se encuentra dentro del rango esperado del 10-15% de casos de cáncer de mama con un componente hereditario [66]. Únicamente Argentina mostró tener un porcentaje mayor con un 17.5% de pacientes positivas a alguna variante patogénica, sin embargo, algunos autores reportan que hasta un 20% de los casos cáncer de mama pueden deberse a un factor genético [67], lo anterior sugiere que no existen diferencias importantes en la prevalencia de variantes patogénicas, independientemente del tamaño y la población analizada. Por otra parte, los genes mayormente mutados fueron *BRCA1/2* los cuales fueron detectados en el 8.15% de la cohorte total ($n = 994$) y representan un 49.67% de las variantes patogénicas detectadas (155/944). Cabe destacar que de las 45 variantes patogénicas detectadas en *BRCA1*, 16 corresponden a la mutación mexicana fundadora la cual fue detectada empleando una prueba dirigida a esa mutación en particular, usando un juego de oligonucleótidos diseñados exclusivamente para detectar dicha alteración y haciendo uso de una PCR de punto final. Esta delección del exón 9 al 12 representan más de una tercera parte de las mutaciones en *BRCA1* y un 10.32% de las mutaciones totales detectadas, las cuales no hubieran podido ser reportadas sin la estrategia desarrollada para detectar esta mutación fundadora. En contraste con las pruebas específicas dirigidas a una mutación o gen en particular, tenemos

las otras 78 variantes patogénicas distribuidas en 54 genes no BRCA que fueron halladas en este estudio, lo que resalta la importancia de realizar análisis en genes más allá de *BRCA1/2*. Sin embargo, también hay que decir que las pruebas genéticas basadas en tecnologías de secuenciación de siguiente generación (NGS) aún tienen limitaciones para detectar cierto tipo de mutaciones, especialmente aquellas que tienen que ver con grandes rearrreglos cromosómicos, como deleciones mayores a 1 kb, translocaciones o variaciones en el número de copias. Asimismo, el uso de NGS para la detección de variantes genéticas es más tardado, costoso y requiere de una mayor capacitación de personal para llevar a cabo el proceso de secuenciación y la interpretación de los datos, en contraste con otro tipo de pruebas más convencionales [68].

Entre los genes no BRCA de alto riesgo que se encontraron alterados en este estudio tenemos a *ATM*, *PALB2*, *APC*, *CDKN2A*, *MLH11*, *MSH2*, *MHS6* y *TP53*. Todos estos genes ya han sido reportados como genes asociados a cáncer y también han sido descritos en pacientes con síndrome hereditario de cáncer de mama y ovario a nivel mundial, con excepción de *CDKN2A* el cual no se asocia a cáncer de mama [69-70]. Por otro lado, algunos de estos genes como *APC*, *MLH11*, *MSH2* son reportados con poca frecuencia en la población latinoamericana [71-72]. Respecto a *CDKN2A*, este es un gen cuyas alteraciones principalmente se asocian a distintos tipos de melanoma y no parece haber una asociación significativa entre las mutaciones entre él y cáncer de mama. *CDKN2A* participa como un regulador negativo del ciclo celular y es un gen supresor de tumores, y pese a que no se ha reportado en cáncer de mamá, existe evidencia de que las alteraciones en este gen, sobre todo en el exón 2, podrían implicar un efecto altamente deletéreo para la célula y una asociación significativa con el cáncer de mama; sin embargo, estos datos se basan en modelos *in silico*, por lo que se requiere más evidencia experimental [73].

Para el caso de las variantes patogénicas en genes de riesgo bajo o moderado, *MUTYH* y *NF1* fueron los genes con mayor incidencia. El primero de estos, *MUTYH*,

está asociado a varios tipos de poliposis, las cuales suelen ser un factor de riesgo para otros tipos de cáncer, principalmente cáncer colorrectal. Sin embargo, está bien documentado que las alteraciones en este gen también predisponen al cáncer de mama, así como cáncer de endometrio [74]. Respecto a *NF1*, es el gen asociado a neurofibromatosis tipo 1. El producto de este gen es la neurofibromina, la cual actúa como un supresor de tumores, inhibiendo la actividad del oncogén *RAS* y aunque *NF1* se asocia principalmente a la aparición de tumores del sistema nervioso central, varios grupos han demostrado que existe una conexión directa entre las mutaciones que pueden ocurrir en este gen y el riesgo a padecer cáncer de mama, además de que los pacientes con mutaciones en *NF1* tienen un mal pronóstico y tienden a desarrollar subtipos de tumor más agresivos [75]. El gen *PDE11A* codifica para una fosfodiesterasa de nucleótidos cíclicos y cataliza la hidrólisis del 3',5'-adenosín monofosfato cíclico (cAMP) y 3',5'-guanosín monofosfato cíclico (cGMP), los cuales actúan como segundos mensajeros, regulando varios procesos celulares en diferentes tejidos [76] y hasta la fecha no parece tener una correlación reportada con cáncer de mama. En cambio, este gen se asocia a tumores adrenocorticales, tumores de células testiculares germinales y se asocia al síndrome de Cushing [77-78]. Esta pudiera ser una diferencia genética a nivel de población para México y Argentina, ya que este gen sólo está afectado en pacientes de estos países; aunque se requieren más estudios para probar esta asociación. Los 43 genes restantes del grupo de genes de riesgo bajo o moderado forman una lista muy variada, con una distribución de variantes casi uniforme de una a dos variantes por gen. Cabe resaltar que varias de estas variantes ya han sido reportadas para cáncer de mama u ovario. Sin embargo, también se detectaron algunos genes que comúnmente no se asocian con esta enfermedad como es el caso de *WRN*, *MSR1* y *LIG4*.

WRN codifica la proteína del síndrome de Werner, la cual es una RecQ helicasa que participa en la reparación del DNA, ayuda a la estabilidad genómica y senescencia celular. Junto con las otras RecQ helicasas (*RECQL1*, Bloom, *RECQL4* y *RECQL5*) juegan un rol crítico en la reparación del DNA y supervivencia

celular después de tratamientos con medicamentos anticáncer, como la Camptotecina (CPT). Recientemente se ha asociado el efecto de CPT en WRN para evitar la proliferación de células malignas en tumores de cáncer de mama [79-80]. Por otra parte, el gen *MSR1* se convierte en la proteína receptor-depuradora de macrófagos y está implicada en varios procesos fisiológicos y patológicos asociados a macrófagos, incluidos aterosclerosis, enfermedad de Alzheimer y también regula la endocitosis de lipoproteínas de baja densidad [81]. En la actualidad se sabe que variaciones en el número de copias en minisatélites dentro de este gen regulan la expresión de otros genes y se asocia a riesgo de padecer cáncer de mama y próstata [82]. Por último, *LIG4* codifica para la proteína DNA ligasa IV en humanos dependiente de ATP que une las rupturas de doble cadena (DSB) mediante la vía de reparación por unión de extremos no homólogos. Estudios recientes parecen demostrar que no existe relación entre polimorfismos en este gen y el cáncer de mama [83].

Es importante señalar que en nuestra cohorte de 944 pacientes se detectaron un total de 209 variantes con significado clínico incierto (VUS), resultando en un 22.13% de las pacientes. Este valor es superior al porcentaje de pacientes con alguna variante patogénica, lo cual corresponde con los valores presentados en otros estudios, en los cuales se reporta hasta un 23% de prevalencia de VUS en pacientes cuando se emplean paneles para evaluar múltiples genes [84]. Estos datos representan un nicho poco explorado, ya que comúnmente estas VUS no suelen ser estudiadas, ni se realizan experimentos para esclarecer el papel que estas variantes tienen en el HBOC. Asimismo, en los estudios en los que se evalúa la composición genética del HBOC considerando las diferencias étnicas y raciales, no suelen haber diferencias significativas en cuanto a los genes con variantes patogénicas y estos parecen ser los mismos sin importar la población estudiada. No obstante, cuando se comparan las VUS entre diferentes grupos raciales, suele haber diferencias respecto a la frecuencia con la que estas alteraciones aparecen y también varían los genes que comúnmente se ven afectados entre grupos [85]. Lo

anterior es muestra de la alta heterogeneidad y complejidad de esta enfermedad entre pacientes con un resultado negativo para variante patogénica.

Este estudio permitió no solo identificar los genes mayormente afectados, sino también las mutaciones puntuales que son recurrentes (*BRCA1* p.A1708E, *BRCA2* p.R2494X, *PALB2* p.M723fs), ya que en este tipo de estudios es raro que una misma mutación se presente en diferentes pacientes no emparentados y, si bien es posible que mutaciones así hayan surgido en la población más de una vez de forma independiente, es más probable que dichas mutaciones deriven de eventos mutacionales únicos y se propaguen en la población con un ancestro común, consecuencia del efecto fundador. La alteración *BRCA1* p.A1708E es considerada como una mutación fundadora presente en personas con ascendencia judía sefardí [86]. No obstante, esta mutación también ha sido reportada en poblaciones de España y Polonia [87-88]. También se ha descrito esta mutación en el 2.9% de pacientes en un estudio realizado en Colombia [89]. Sin embargo, estos estudios no detallan la ascendencia de los pacientes, por lo que no es posible determinar si se trata de una propagación de esta mutación fundadora o un evento de convergencia mutacional. Por otra parte, la mutación *BRCA2* p.R2494X detectada en este estudio se reporta como mutación fundadora con una alta incidencia en la población coreana [90], aunque también hay reportes en pacientes de Japón [91], por lo que parece ser una mutación altamente recurrente en países orientales, pero no hay reportes de ella en países de América. Respecto a la mutación p.M723fs en *PALB2*, esta alteración ha sido reportada en un estudio en pacientes con ascendencia hispana en Estados Unidos. Esta variante fue reportada con una frecuencia de 0.86% (9/1045) y también se demostró que los portadores de esta mutación compartían el mismo genotipo lo que sugiere fuertemente que podría tratarse de una mutación fundadora de la población latina [92]. En estos casos, el análisis de bloques de haplotipos permitiría determinar si se trata de una mutación fundadora. Asimismo, todas estas mutaciones recurrentes muestran la complejidad de la composición genética de los países de América Latina y similitudes con otras poblaciones. Finalmente, las mutaciones recurrentes entre familiares (*BRCA2*

p.G578fs y *BRCA1* p.T276Afs) muestran la alta frecuencia con la que este tipo de mutaciones pueden propagarse dentro de una familia afectada con HBOC.

Los resultados en este estudio no sólo presentaron datos que concuerdan con resultados previos para el HBOC, sino también muestran diferencias y nuevos hallazgos que parecen ser característicos de la población estudiada, principalmente respecto a los genes de riesgo bajo o moderado con variantes patogénicas y las variantes de significado clínico incierto. Lo que muestra la necesidad de incorporar a más pacientes con la finalidad de que todos los países se encuentren igualmente representados, así como la relevancia de los estudios genéticos en cada vez más países para tener un panorama más amplio de la diversidad genética de esta enfermedad.

9. CONCLUSIONES

En este trabajo se encontró que el 16.41% de las pacientes con sospecha de síndrome hereditario de cáncer de mama y ovario dio positivo para al menos una alteración patogénica. Los genes mayormente afectados fueron *BRCA1*, *BRCA2*, *ATM*, *PALB2*, *MUTYH* y *NF1*. También se detectaron un total de 209 variantes con significado clínico incierto. En cuanto a las mutaciones recurrentes, se hallaron genes con mutaciones recurrentes distribuidas en pacientes no relacionados, de las cuales la mutación *PALB2* p. M723fs es una fuerte candidata a ser una mutación fundadora de esta región. No fue posible encontrar diferencias significativas entre los 5 países participantes.

10. PERSPECTIVAS

Los datos presentados en esta tesis demuestran la importancia de seguir analizando la diversidad de alelos patogénicos presentes en pacientes con HBOC, por lo que se sugiere incorporar a más pacientes de Argentina, Colombia, Guatemala y Perú, con la finalidad de asegurar que cada país se encuentre igualmente representado.

Las variantes de significado clínico incierto pueden ser candidatas para realizar ensayos funcionales que aporten información para esclarecer su impacto en el desarrollo del síndrome hereditario de cáncer de mama y ovario.

Este trabajo también plantea la posibilidad de implementar pruebas específicas para la detección de alteraciones frecuentes entre las pacientes de estos países, pues si bien las técnicas de secuenciación masiva nos permiten obtener un panorama bastante amplio de las mutaciones presentes entre estos pacientes, algunas alteraciones siguen siendo indetectables, como se demostró con las pacientes portadoras de la mutación fundadora Deleción 9-12.

11. BIBLIOGRAFÍA

1. Ruddon, R. W. (2007). Cancer biology. Oxford University Press.
2. Samuels, Y., Bardelli, A., Wolf, Y., & Lopez-Otin, C.(2009). The Cancer Genome. En: Govindan, R., & DeVita, V. T. (Eds.). DeVita, Hellman, and Rosenberg's Cancer: Principles & Practice of Oncology Review. Lippincott Williams & Wilkins.
3. World Health Organization. WHO Mortality Database; 2018.
<https://www.who.int/data/data-collection-tools/who-mortality-database>.
Consultado el 14 de septiembre 2020
4. Sancho-Garnier, H., & Colonna, M. (2019). Épidémiologie des cancers du sein. La Presse Médicale, 48(10), 1076-1084.
5. Coughlin S.S. (2019) Epidemiology of Breast Cancer in Women. In: Ahmad A. (eds) Breast Cancer Metastasis and Drug Resistance. Advances in Experimental Medicine and Biology, vol 1152. Springer, Cham. https://doi.org/10.1007/978-3-030-20301-6_2
6. PAHO
7. Justo, N., Wilking, N., Jönsson, B., Luciani, S., & Cazap, E. (2013). A review of breast cancer care and outcomes in Latin America. The oncologist, 18(3), 248.
8. Azamjah, N., Soltan-Zadeh, Y., & Zayeri, F. (2019). Global trend of breast cancer mortality rate: a 25-year study. Asian Pacific journal of cancer prevention: APJCP, 20(7), 2015.
9. Kamińska, M., Ciszewski, T., Łopacka-Szatan, K., Miotła, P., & Starosławska, E. (2015). Breast cancer risk factors. Przegląd menopauzalny= Menopause review, 14(3), 196.
10. Wade, P. (2010). Race and ethnicity in Latin America. Pluto press.
11. Fejerman, L., Serrano-Gómez, S. J., & Tamayo, L. I. (2020). Breast Cancer Risk and Mortality in Women of Latin American Origin. In Advancing the Science of Cancer in Latinos (pp. 45-55). Springer, Cham.
12. Rojas, K., & Stuckey, A. (2016). Breast cancer epidemiology and risk factors. *Clinical obstetrics and gynecology*, 59(4), 651-672.
13. Li, C. I., Malone, K. E., & Daling, J. R. (2003). Differences in breast cancer stage, treatment, and survival by race and ethnicity. Archives of internal medicine, 163(1), 49-56.
14. Supramaniam, R., Gibberd, A., Dillon, A., Goldsbury, D. E., & O'Connell, D. L. (2014). Increasing rates of surgical treatment and preventing comorbidities

- may increase breast cancer survival for Aboriginal women. *BMC cancer*, 14(1), 163.
15. Fraumeni JF. Cancer mortality among nuns: role of marital status in etiology of neoplastic disease in women. *J Natl Cancer Inst.* 1969;42: 455–468.
 16. Rojas, K., & Stuckey, A. (2016). Breast cancer epidemiology and risk factors. *Clinical obstetrics and gynecology*, 59(4), 651-672.
 17. Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., ... & Zhu, H. P. (2017). Risk factors and preventions of breast cancer. *International journal of biological sciences*, 13(11), 1387.
 18. Bernstein, L., & Ross, R. K. (1993). Endogenous hormones and breast cancer risk. *Epidemiologic reviews*, 15(1), 48-65.
 19. Kim, J. J., Kurita, T., & Bulun, S. E. (2013). Progesterone action in endometrial cancer, endometriosis, uterine fibroids, and breast cancer. *Endocrine reviews*, 34(1), 130-162.
 20. Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, 458(7239), 719-724.
 21. McPherson, K., Steel, C., & Dixon, J. M. (2000). Breast cancer—epidemiology, risk factors, and genetics. *Bmj*, 321(7261), 624-628.
 22. Lalloo, F., & Evans, D. G. (2012). Familial breast cancer. *Clinical genetics*, 82(2), 105-114.
 23. Fearon, E. R. (1997). Human cancer syndromes: clues to the origin and nature of cancer. *Science*, 278(5340), 1043-1050.
 24. Garber, J. E., & Offit, K. (2005). Hereditary cancer predisposition syndromes. *Journal of clinical oncology*, 23(2), 276-292.
 25. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 1971;68:820–3. (Level III)
 26. Yoshida, R. (2021). Hereditary breast and ovarian cancer (HBOC): review of its molecular characteristics, screening, treatment, and prognosis. *Breast Cancer*, 28(6), 1167-1180.
 27. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM, et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. 2011;108:18032–7. (Level III)
 28. Hereditary breast and ovarian cancer syndrome. Practice Bulletin No.182. American College of Obstetricians and Gynecologists. *ObstetGynecol* 2017;130:e110–26.
 29. National Comprehensive Cancer Network. (2022). Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic. Version 2.2022. Sitio Web: https://www.nccn.org/professionals/physician_gls/pdf/genetics_bop.pdf

30. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... & Rehm, H. L. (2015). 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. *Genetics in medicine*, 17(5), 405-423.
31. Weitzel, J. N., Lagos, V. I., Herzog, J. S., Judkins, T., Hendrickson, B., Ho, J. S., ... & Scholl, T. (2007). Evidence for common ancestral origin of a recurring BRCA1 genomic rearrangement identified in high-risk Hispanic families. *Cancer Epidemiology Biomarkers & Prevention*, 16(8), 1615-1620.
32. Quezada Urban, R., Díaz Velásquez, C. E., Gitler, R., Rojo Castillo, M. P., Sirota Toporek, M., Figueroa Morales, A., Moreno García, O., García Esquivel, L., Torres Mejía, G., Dean, M., Delgado Enciso, I., Ochoa Díaz López, H., Rodríguez León, F., Jan, V., Garzón Barrientos, V. H., Ruiz Flores, P., Espino Silva, P. K., Haro Santa Cruz, J., Martínez Gregorio, H., Rojas Jiménez, E. A., ... Vaca Paniagua, F. (2018). Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility. *Cancers*, 10(10), 361.
33. Caswell-Jin, J. L., Gupta, T., Hall, E., Petrovchich, I. M., Mills, M. A., Kingham, K. E., ... & Kurian, A. W. (2018). Racial/ethnic differences in multiple-gene sequencing results for hereditary cancer risk. *Genetics in Medicine*, 20(2), 234-239.
34. Oliver, J., Quezada Urban, R., Franco Cortés, C. A., Díaz Velásquez, C. E., Montealegre Paez, A. L., Pacheco-Orozco, R. A., ... & Vaca Paniagua, F. (2019). Latin american study of hereditary breast and ovarian cancer LACAM: a genomic epidemiology approach. *Frontiers in oncology*, 1429.
35. Gradishar, W. J., Anderson, B. O., Balassanian, R., Blair, S. L., Burstein, H. J., Cyr, A., ... & Kumar, R. (2017). NCCN guidelines insights: breast cancer, version 1.2017. *Journal of the National Comprehensive Cancer Network*, 15(4), 433-451.
36. Telli, M. L., Gradishar, W. J., & Ward, J. H. (2019). NCCN guidelines updates: breast cancer. *Journal of the National Comprehensive Cancer Network*, 17(5.5), 552-555.
37. Blood, D., & Kits, T. (2011). DNeasy Blood and Tissue Handbook. DNeasy Blood and Tissue Handbook.
38. Church, A. J. (2020). Next-Generation Sequencing. *Genomic Medicine: A Practical Guide*, 25-40.
39. Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic acids research*, 38(6), 1767-1771.

40. Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25:1754-60. [PMID: 19451168]
41. Ye H, Meehan J, Tong W, Hong H. Alignment of short reads: a crucial step for application of next-generation sequencing data in precision medicine. *Pharmaceutics*. 2015;7(4):523–541. doi: 10.3390/pharmaceutics7040523.
42. Lander, E. S., Linton, L. M., & Birren, B. (2001). Initial sequencing and analysis of the human genome [published correction appears in *Nature*. 2001; 411 (6838): 720]. *Nature*, 409(6822), 860-921.
43. DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., ... & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*, 43(5), 491-498.
44. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), 1297-1303.
45. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), 1297-1303.
46. Quantus(TM) fluorometer operating manual #TM396 - Promega Corporation. (n.d.). Retrieved from <https://www.promega.com/-/media/files/resources/protocols/technical-manuals/101/quantus-fluorometer-operating-manual.pdf?la=en>
47. Adkins, N. L., Hall, J. A., & Georgel, P. T. (2007). The use of quantitative agarose gel electrophoresis for rapid analysis of the integrity of protein–DNA complexes. *Journal of biochemical and biophysical methods*, 70(5), 721-726.
48. LaDuca, H., Stuenkel, A. J., Dolinsky, J. S., Keiles, S., Tandy, S., Pesaran, T., ... & Chao, E. (2014). Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genetics in medicine*, 16(11), 830-837.
49. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data *Nucleic Acids Research*, 38:e164, 2010
50. Kruglyak, L., & Nickerson, D. A. (2001). Variation is the spice of life. *Nature genetics*, 27(3), 234-236.
51. Gibbs, R. A., Belmont, J. W., Hardenbol, P., Willis, T. D., Yu, F. L., Yang, H. M., ... & Duster, T. (2003). The international HapMap project.

52. Siva, N. (2008). 1000 Genomes project. *Nature biotechnology*, 26(3), 256-257.
53. Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, 29(1), 308-311.
54. Koch, L. (2020). Exploring human genomic diversity with gnomAD. *Nature Reviews Genetics*, 21(8), 448-448.
55. Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., ... & Maglott, D. R. (2016). ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic acids research*, 44(D1), D862-D868.
56. Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., ... & Wooster, R. (2004). The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *British journal of cancer*, 91(2), 355-358.
57. Li, Q., & Wang, K. (2017). InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *The American Journal of Human Genetics*, 100(2), 267-280.
58. Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Aguilera, M. A., Meyer, R., & Massouras, A. (2019). VarSome: the human genomic variant search engine. *Bioinformatics*, 35(11), 1978.
59. Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research*, 31(13), 3812-3814.
60. Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, 76(1), 7-20.
61. Wang, Y. A., Jian, J. W., Hung, C. F., Peng, H. P., Yang, C. F., Cheng, H. C. S., & Yang, A. S. (2018). Germline breast cancer susceptibility gene mutations and breast cancer outcomes. *BMC cancer*, 18, 1-13.
62. Villarreal-Garza, C., Alvarez-Gómez, R. M., Pérez-Plasencia, C., Herrera, L. A., Herzog, J., Castillo, D., ... & Weitzel, J. N. (2015). Significant clinical impact of recurrent BRCA1 and BRCA2 mutations in Mexico. *Cancer*, 121(3), 372-378.
63. Quezada Urban, R. (2017). Identificación de mutaciones germinales en 143 genes de predisposición al cáncer en pacientes mexicanos con síndrome de cáncer de mama hereditario por secuenciación masiva [Tesis de maestría, Universidad Nacional Autónoma de México] TESIUNAM
64. Vidal-Millan, S., Taja-Chayeb, L., Gutierrez-Hernandez, O., Ugalde, M. R., Robles-Vidal, C., Bargallo-Rocha, E., ... & Dueñas-González, A. (2009).

- Mutational analysis of BRCA1 and BRCA2 genes in Mexican breast cancer patients. *Eur. J. Gynaec. Oncol.*-ISSN, 30(5), 2009.
65. Vaca-Paniagua, F., Alvarez-Gomez, R. M., Fragoso-Ontiveros, V., Vidal-Millan, S., Herrera, L. A., Cantu, D., ... & Perez-Plasencia, C. (2012). Full-exon pyrosequencing screening of BRCA germline mutations in Mexican women with inherited breast and ovarian cancer. *PloS one*, 7(5), e37432.
 66. Huber-Keener K. J. (2022). Cancer genetics and breast cancer. Best practice & research. *Clinical obstetrics & gynaecology*, 82, 3–11. <https://doi.org/10.1016/j.bpobgyn.2022.01.007>
 67. Sokolova, A., Johnstone, K. J., McCart Reed, A. E., Simpson, P. T., & Lakhani, S. R. (2023). Hereditary breast cancer: syndromes, tumour pathology and molecular testing. *Histopathology*, 82(1), 70–82. <https://doi.org/10.1111/his.14808>
 68. Castéra, L., Krieger, S., Rousselin, A., Legros, A., Baumann, J. J., Bruet, O., Brault, B., Fouillet, R., Goardon, N., Letac, O., Baert-Desurmont, S., Tinat, J., Bera, O., Dugast, C., Berthet, P., Polycarpe, F., Layet, V., Hardouin, A., Frébourg, T., & Vaur, D. (2014). Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *European journal of human genetics : EJHG*, 22(11), 1305–1313. <https://doi.org/10.1038/ejhg.2014.16>
 69. Weitzel, J.N., Slavin, T.P. (2020). Genomic Cancer Risk Assessment. In: Salgia, R. (eds) *Oncology in the Precision Medicine Era*. Springer, Cham. https://doi.org/10.1007/978-3-030-31471-2_12
 70. Iklkjñl
 71. Zavala, V. A., Serrano-Gomez, S. J., Dutil, J., & Fejerman, L. (2019). Genetic Epidemiology of Breast Cancer in Latin America. *Genes*, 10(2), 153. <https://doi.org/10.3390/genes10020153>
 72. Guindalini, R. S. C., Viana, D. V., Kitajima, J. P. F. W., Rocha, V. M., López, R. V. M., Zheng, Y., Freitas, É., Monteiro, F. P. M., Valim, A., Schlesinger, D., Kok, F., Olopade, O. I., & Figueira, M. A. A. K. (2022). Detection of germline variants in Brazilian breast cancer patients using multigene panel testing. *Scientific reports*, 12(1), 4190. <https://doi.org/10.1038/s41598-022-07383-1>

73. Aftab, A., Shahzad, S., Hussain, H.M.J. et al. *CDKN2A/P16/INK4A* variants association with breast cancer and their in-silico analysis. *Breast Cancer* **26**, 11–28 (2019). <https://doi.org/10.1007/s12282-018-0894-0>
74. Nielsen, M., Infante, E., & Brand, R. (2012). *MUTYH* Polyposis. In M. P. Adam (Eds.) et. al., *GeneReviews®*. University of Washington, Seattle.
75. Howell, S. J., Hockenhull, K., Salih, Z., & Evans, D. G. (2017). Increased risk of breast cancer in neurofibromatosis type 1: current insights. *Breast cancer (Dove Medical Press)*, 9, 531–536. <https://doi.org/10.2147/BCTT.S111397>
76. Yuasa, K., Kanoh, Y., Okumura, K., & Omori, K. (2001). Genomic organization of the human phosphodiesterase PDE11A gene: evolutionary relatedness with other PDEs containing GAF domains. *European Journal of Biochemistry*, 268(1), 168-178.
77. Libé, R., Fratticci, A., Coste, J., Tissier, F., Horvath, A., Ragazzon, B., ... & Bertherat, J. (2008). Phosphodiesterase 11A (PDE11A) and genetic predisposition to adrenocortical tumors. *Clinical Cancer Research*, 14(12), 4016-4024.
78. Kong, G., Lee, H., Vo, T. T. T., Juang, U., Kwon, S. H., Park, J., ... & Kim, S. H. (2022). Functional characteristics and research trends of PDE11A in human diseases. *Molecular Medicine Reports*, 26(4), 1-10.
79. Shamanna, R. A., Lu, H., Croteau, D. L., Arora, A., Agarwal, D., Ball, G., Aleskandarany, M. A., Ellis, I. O., Pommier, Y., Madhusudan, S., & Bohr, V. A. (2016). Camptothecin targets WRN protein: mechanism and relevance in clinical breast cancer. *Oncotarget*, 7(12), 13269–13284. <https://doi.org/10.18632/oncotarget.7906>
80. Baxter, J. S., Zatreanu, D., Pettitt, S. J., & Lord, C. J. (2022). Resistance to DNA repair inhibitors in cancer. *Molecular oncology*, 16(21), 3811–3827. <https://doi.org/10.1002/1878-0261.13224>
81. Wang, X., Wen, D., You, C., Tao, C., & Ma, L. (2023). Comprehensive analysis of immune cell infiltration and role of MSR1 expression in aneurysmal subarachnoid haemorrhage. *Cell Proliferation*, 56(6), e13379.
82. Rose, A. M., Krishan, A., Chakarova, C. F., Moya, L., Chambers, S. K., Hollands, M., Illingworth, J. C., Williams, S. M. G., McCabe, H. E., Shah, A. Z., Palmer, C. N. A., Chakravarti, A., Berg, J. N., Batra, J., & Bhattacharya, S. S. (2018). MSR1 repeats modulate gene expression and affect risk of breast and prostate cancer. *Annals of oncology : official journal of the*

- European Society for Medical Oncology*, 29(5), 1292–1303.
<https://doi.org/10.1093/annonc/mdy082>
83. Zhou, L. P., Luan, H., Dong, X. H., Jin, G. J., Man, D. L., & Shang, H. (2012). Lack of association between LIG4 gene polymorphisms and the risk of breast cancer: a HuGE review and meta-analysis. *Asian Pacific journal of cancer prevention : APJCP*, 13(7), 3417–3422.
<https://doi.org/10.7314/apjcp.2012.13.7.3417>
 84. van Marcke, C., Collard, A., Vikkula, M., & Duhoux, F. P. (2018). Prevalence of pathogenic variants and variants of unknown significance in patients at high risk of breast cancer: A systematic review and meta-analysis of gene-panel data. *Critical reviews in oncology/hematology*, 132, 138-144.
 85. Tatineni, S., Tarockoff, M., Abdallah, N., Purrington, K. S., Assad, H., Reagle, R., ... & Simon, M. S. (2022). Racial and ethnic variation in multigene panel testing in a cohort of BRCA1/2-negative individuals who had genetic testing in a large urban comprehensive cancer center. *Cancer Medicine*, 11(6), 1465-1473.
 86. Sagi, M., Eilat, A., Ben Avi, L., Goldberg, Y., Bercovich, D., Hamburger, T., ... & Lerer, I. (2011). Two BRCA1/2 founder mutations in Jews of Sephardic origin. *Familial cancer*, 10, 59-63.
 87. Infante, M., Duran, M., Esteban-Cardenosa, E., Miner, C., & Velasco, E. (2006). High proportion of novel mutations of BRCA1 and BRCA2 in breast/ovarian cancer patients from Castilla-Leon (central Spain). *Journal of human genetics*, 51(7), 611-617.
 88. Szczerba, E., Kamińska, K., Mierzwa, T., Misiek, M., Kowalewski, J., & Lewandowska, M. A. (2021). BRCA1/2 mutation detection in the tumor tissue from selected polish patients with breast cancer using next generation sequencing. *Genes*, 12(4), 519.
 89. Briceño-Balcázar, I., Gómez-Gutiérrez, A., Díaz-Dussán, N. A., Noguera-Santamaría, M. C., Díaz-Rincón, D., & Casas-Gómez, M. C. (2017). Mutational spectrum in breast cancer associated BRCA1 and BRCA2 genes in Colombia. *Colombia Médica*, 48(2), 58-63.
 90. Kim, H., Cho, D. Y., Choi, D. H., Choi, S. Y., Shin, I., Park, W., ... & Haffty, B. G. (2012). Characteristics and spectrum of BRCA1 and BRCA2 mutations in 3,922 Korean patients with breast and ovarian cancer. *Breast cancer research and treatment*, 134, 1315-1326.
 91. Hirotsu, Y., Nakagomi, H., Sakamoto, I., Amemiya, K., Mochizuki, H., & Omata, M. (2015). Detection of BRCA1 and BRCA2 germline mutations in

- Japanese population using next-generation sequencing. *Molecular Genetics & Genomic Medicine*, 3(2), 121-129.
92. Weitzel, J. N., Neuhausen, S. L., Adamson, A., Tao, S., Ricker, C., Maoz, A., ... & Ziv, E. (2019). Pathogenic and likely pathogenic variants in PALB2, CHEK2, and other known breast cancer susceptibility genes among 1054 BRCA-negative Hispanics with breast cancer. *Cancer*, 125(16), 2829-2836.
93. Horak, P., Griffith, M., Danos, A. M., Pitel, B. A., Madhavan, S., Liu, X., ... & Sonkin, D. (2022). Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): joint recommendations of Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC), and Variant Interpretation for Cancer Consortium (VICC). *Genetics in Medicine*, 24(5), 986-998.

12. OTROS RECURSOS

Algunas de las figuras y esquemas presentados en este trabajo fueron realizadas haciendo uso de iconos e imágenes cuyos créditos de autor aparecen aquí abajo.

- Imagen de `<a href="https://www.freepik.es/vector-gratis/mexico-mapa-vectorial_333986.htm">Freepik</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/tabla-de-datos" title="tabla de datos iconos">Tabla de datos iconos creados por Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/senal-de-stop" title="señal de stop iconos">Señal de stop iconos creados por Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.com/free-icons/mutation" title="Mutation icons">Mutation icons created by Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.com/free-icons/interrogation-mark" title="interrogation mark icons">Interrogation mark icons created by zky.icon - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/argentina" title="argentina iconos">Argentina iconos creados por verluk - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/guatemala" title="guatemala iconos">Guatemala iconos creados por Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/bandera-mexico" title="bandera mexico iconos">Bandera mexico iconos creados por verluk - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/colombia" title="colombia iconos">Colombia iconos creados por Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/peru" title="perú iconos">Perú iconos creados por Freepik - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/advertencia" title="advertencia iconos">Advertencia iconos creados por Design View - Flaticon</a>`
- `<a href="https://www.flaticon.es/iconos-gratis/familia" title="familia iconos">Familia iconos creados por yut1655 - Flaticon</a>`