



**Universidad de Chile**

Facultad de Ciencias Químicas y Farmacéuticas

Departamento de Bioquímica y Biología Molecular

# **Variantes en los genes *RECQL* y *MCPH1* y su asociación con el riesgo a desarrollar cáncer de mama hereditario en población chilena**

Memoria para optar al título profesional de  
Bioquímico

**Emilio Andrés Sepúlveda Sáez**

**Directora**

**Dra. Sandra Ampuero Llanos**

Programa de Virología, ICBM

Facultad de Medicina

Universidad de Chile

**Patrocinante**

**Dr. Dante Miranda Wilson**

Dpto. Bioquímica y Biología Molecular

Facultad de Ciencias Químicas y

Farmacéuticas

Universidad de Chile

**Agosto 2023**

## ii. Financiamiento

Esta Memoria de Título fue realizada al interior del Programa de Virología de la Facultad de Medicina de la Universidad de Chile y financiada por el Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT), Número 1200049.

### iii. Agradecimientos

Primero quiero agradecer a mi familia quienes con mucho esfuerzo me dieron la posibilidad de estudiar. A mi mamá por su inmenso amor y que sin ella nada de esto hubiese sido posible. A mi papá por siempre estar ahí para todo en cada momento. A mi hermana por su cariño y ridiculeces que me alegran los días. A mi perrito que aún sin saberlo con su ternura llena mis días.

A mis amigos de la vida misma Darri, Eduardito, Jona, Javito, Luciano, Miki, Manuel, Martin, Matías, Saulo y Tato formamos un grupo en que podemos ser nosotros mismos sin miedo a que nos juzguen y siempre están presente, ya sea en los buenos momentos como en los malos.

A mis amigos que conocí durante la universidad Ali, Cami, Coni y Matías. Gracias por las risas y los momentos vividos, además de esos interminables días en odonto sobreviviendo a base de cafecitos.

A la natación y las personas que he conocido en la piscina, son de gran ayuda para despejar la mente en días complicados.

A las personas que de una u otra forma conocí en el laboratorio, Don Luis y Don Cristian gracias por el apoyo y enseñanzas entregadas. Así como también a la Dani y Gustavo que ayudaron a sobrellevar este periodo.

Por último, quisiera agradecer a mi directora de tesis, la profe Sandra por aguantarme a pesar de mis fallos durante los años desde que llegue a realizar la unidad de investigación y ser un pilar fundamental para lograr terminar esta tesis. Gracias por los consejos, enseñanzas y engaños dulces a lo largo de estos años.

## iv. Índice de contenidos

<b>Financiamiento</b> .....	ii
<b>Agradecimientos</b> .....	iii
<b>Índice de contenidos</b> .....	iv
<b>Índice de tablas</b> .....	vi
<b>Índice de figuras</b> .....	vii
<b>Abreviaturas</b> .....	viii
<b>Resumen</b> .....	ix
<b>Abstract</b> .....	x
<b>1. Introducción</b> .....	1
1.1 Cáncer.....	1
1.2 Cáncer de mama.....	1
1.3 Cáncer de mama hereditario.....	3
1.4 Gen <i>RECQL</i> .....	5
1.5 Gen <i>MCPH1</i> .....	7
<b>2. Hipótesis</b> .....	10
<b>3. Objetivos general y específicos</b> .....	10
<b>4. Materiales y métodos</b> .....	14
4.1 Diseño experimental.....	14
4.2 Grupos muestrales.....	15
4.2.1 Obtención de muestras.....	15
4.2.2 Casos y controles.....	15
4.3 Extracción de material genético.....	15
4.4 Cuantificación de material genético.....	15
4.5 Búsqueda en bases de datos sobre <i>RECQL</i> y <i>MCPH1</i> .....	16
4.6 Diseño de partidores qPCR-HRM.....	16
4.7 Optimización qPCR-HRM.....	17
4.8 Diseño de partidores PCR convencional.....	18
4.9 Optimización PCR convencional.....	19
4.10 Electroforesis en geles de agarosa.....	19
4.11 Extracción y purificación material genético pos PCR.....	20
4.12 Secuenciación de producto amplificado.....	20
4.13 Análisis bioinformático.....	20
4.13.1 Análisis de secuencias.....	20
4.13.2 Variación en la proteína.....	20
4.14 Análisis estadístico.....	21
<b>5. Resultados</b> .....	22
5.1 Características demográficas de los casos con CM.....	22
5.2 Búsqueda de regiones y variantes de interés <i>in silico</i> .....	23
5.3 Búsqueda de variantes mediante perfil HRM gen <i>RECQL</i> .....	24

5.4 Búsqueda de variantes mediante perfiles HRM gen <i>MCPH1</i> .....	25
5.4.1 Casos CMH y perfiles HRM para el gen <i>MCPH1</i> .....	28
5.4.2 Controles y perfiles HRM para el gen <i>MCPH1</i> .....	28
5.4.3 Comparación entre casos CMH y controles en el gen <i>MCPH1</i> .....	28
5.5 Identificación de variantes en el gen <i>MCPH1</i> por cromatograma.....	29
5.6 Asociación de los haplotipos encontrados en <i>MCPH1</i> con los perfiles HRM.....	31
5.7 Distribución de casos con CMH según distintas categorías en presencia o ausencia de las variantes encontradas en el gen <i>MCPH1</i> .....	32
5.7.1 Distribución por edad del paciente.....	32
5.7.2 Distribución por tipo histológico de CM del paciente.....	33
5.7.3 Distribución por historial familiar del paciente.....	34
5.8 Posible impacto de los SNP a nivel de proteínas <i>in silico</i> .....	35
<b>6. Discusión</b> .....	38
<b>7. Conclusiones</b> .....	43
<b>8. Bibliografía</b> .....	44

## v. Índice de tablas

<b>Tabla 1.</b> Secuencia de partidores para la qPCR-HRM.....	16
<b>Tabla 2.</b> Concentraciones de reactivos usados qPCR-HRM con kit Kapa HRM Fast.....	17
<b>Tabla 3.</b> Concentraciones de reactivos usados qPCR-HRM con kit Type-it HRM-PCR.....	17
<b>Tabla 4.</b> Protocolo de reacción para qPCR-HRM con kit Kapa HRM Fast.....	17
<b>Tabla 5.</b> Protocolo de reacción para qPCR-HRM con kit Type-it HRM-PCR.....	18
<b>Tabla 6.</b> Secuencia de partidores para PCR convencional.....	19
<b>Tabla 7.</b> Concentraciones de reactivos usados PCR convencional.....	19
<b>Tabla 8.</b> Protocolo de amplificación para PCR convencional.....	19
<b>Tabla 9.</b> Características de los 200 casos chilenas con CM y 250 controles analizadas en este estudio.....	22
<b>Tabla 10.</b> Mutaciones de interés en el gen <i>MCPH1</i> .....	23
<b>Tabla 11.</b> Mutaciones de interés en el gen <i>RECQL</i> .....	24
<b>Tabla 12.</b> Comparación de los perfiles HRM obtenidos en casos CMH y controles.....	30
<b>Tabla 13.</b> Distribución de haplotipos para los 3 SNP encontrados en <i>MCPH1</i> y su relación con el perfil HRM.....	33
<b>Tabla 14.</b> Resumen de los haplotipos posibles encontrados y su frecuencia en muestras secuenciadas y totales.....	33
<b>Tabla 15.</b> Comparación de los perfiles de HRM obtenidos en casos menores o igual a 45 años y mayores de 45 años.....	34
<b>Tabla 16.</b> Comparación de los perfiles de HRM obtenidos entre CM de tipo ductal y CM de tipo lobulillar.....	35
<b>Tabla 17.</b> Comparación de los perfiles de HRM obtenidos en casos CMH en ausencia o presencia de familiares con CM.....	36
<b>Tabla 18.</b> Predicciones resultantes de los cambios en la proteína según plataformas bioinformáticas.....	37

## vi. Índice de figuras

<b>Figura 1.</b> Incidencia de cáncer en mujeres de todas las edades a nivel mundial.....	3
<b>Figura 2.</b> Ilustración de las contribuciones proporcionales de los genes de riesgo de susceptibilidad al CM.....	5
<b>Figura 3.</b> Mapa de los estudios de mutaciones en <i>BRCA1/2</i> en los países de América Central y del Sur.....	6
<b>Figura 4.</b> Representación de los exones y la proteína RECQL junto a sus dominios.....	8
<b>Figura 5.</b> Esquema resumen de las funciones de RECQL.....	8
<b>Figura 6.</b> Representación de los exones y la proteína RECQL junto a sus dominios.....	10
<b>Figura 7.</b> Esquema resumen de las funciones biológicas de MCPH1.....	11
<b>Figura 8.</b> Gráfico modelo de la amplificación qPCR-HRM.....	18
<b>Figura 9.</b> Gráfico modelo presentando la fluorescencia normalizada del perfil HRM para el gen <i>RECQL</i> .....	25
<b>Figura 10.</b> Gráfico modelo presentando la fluorescencia derivada del perfil HRM para el gen <i>RECQL</i> .....	26
<b>Figura 11.</b> Gráfico modelo presentando la fluorescencia normalizada del perfil HRM tipo A para el gen <i>MCPH1</i> .....	27
<b>Figura 12.</b> Gráfico modelo presentando la fluorescencia normalizada del perfil HRM tipo B para el gen <i>MCPH1</i> .....	27
<b>Figura 13.</b> Gráfico modelo presentando la fluorescencia normalizada del perfil HRM tipo A y B para el gen <i>MCPH1</i> .....	28
<b>Figura 14.</b> Gráfico modelo presentando la fluorescencia derivada del perfil HRM tipo A para el gen <i>MCPH1</i> .....	28
<b>Figura 15.</b> Gráfico modelo presentando la fluorescencia derivada del perfil HRM tipo B para el gen <i>MCPH1</i> .....	29
<b>Figura 16.</b> Alineamiento de las muestras secuenciadas para el gen <i>MCPH1</i> .....	31
<b>Figura 17.</b> Cromatograma del SNP rs35590577 en una región de <i>MCPH1</i> .....	31
<b>Figura 18.</b> Cromatograma del SNP rs2083914 en una región de <i>MCPH1</i> .....	32
<b>Figura 19.</b> Cromatograma del SNP rs930557 en una región de <i>MCPH1</i> .....	32
<b>Figura 20.</b> Distribución de los perfiles HRM de <i>MCPH1</i> según rango etario de los casos con CM.....	34
<b>Figura 21.</b> Distribución de los perfiles HRM de <i>MCPH1</i> según tipo histológico de los casos con CM.....	35
<b>Figura 22.</b> Distribución de los perfiles HRM de <i>MCPH1</i> según historial familiar de los casos con CM.....	36

## vii. Abreviaturas

aa: Aminoácidos  
ADN: Ácido desoxirribonucleico  
ARN: Ácido ribonucleico  
BRCA1: Breast cancer 1  
BRCA2: Breast cancer 2  
BRCT: Breast cancer Carboxy-Terminal  
BRIT1: BRCT-repeat inhibitor of hTERT expression  
CDC27: Cell division cycle protein 27  
CDI: Cáncer ductal infiltrante  
CDIS: Cáncer ductal *in situ*  
CHK1: Checkpoint Kinase 1  
CLI: Cáncer lobulillar infiltrante  
CLIS: Cáncer lobulillar *in situ*  
CM: cáncer de mama  
CMH: cáncer de mama hereditario  
COSMIC: Catalogue of Somatic Mutations in Cancer  
Ct: Cycle threshold  
del: Delección  
DMC1: DNA Meiotic Recombinase 1  
ER: Receptor de estrógeno  
FEN-1: Flap endonuclease 1  
GLOBOCAN: Global Cancer Observatory  
H2AX: Histone  
HRM: High Resolution Melt  
HER2: Receptor 2 del factor de crecimiento epidérmico humano  
hTERT: Telomerase Reverse Transcriptase  
ins: Inserción  
Ku70/80: ATP-dependent DNA helicase II subunit 1/2  
MCPH1: Microcephalin 1  
MCPH: Microcefalia Primaria Autosómica Recesiva  
NCBI: National Center for Biotechnology Information  
NGS: Secuenciación de nueva generación  
OMS: Organización Mundial de la Salud  
PARP-1: Poly [ADP-ribose] polymerase 1  
Pb: Pares de bases  
PCR: Reacción en Cadena de la Polimerasa  
PhD-SNP: Predictor of human Deleterious Single Nucleotide Polymorphisms  
PolyPhen-2: Polymorphism Phenotyping v2  
PR: Receptor de progesterona  
qPCR: Reacción en Cadena de la Polimerasa cuantitativa  
RPA: Replication protein A  
TNBC: Cáncer de mama triple negativo  
Tm: Temperatura de melting  
TBE: Tris-borato-EDTA



## viii. Resumen

El cáncer de mama (CM) representa el 24,5% de los casos totales de cáncer y el 15,5% de las muertes por cáncer. En Chile, es la principal causa de muerte por cáncer en mujeres.

La etiología exacta del CM es desconocida, aunque existen factores de riesgo como sexo, edad y predisposición genética. El 10% clasifica como CM hereditario (CMH). Se han estudiado genes de alta, moderada y baja penetrancia asociados a la susceptibilidad a CMH. Estudios en población chilena y latinoamericana son escasos.

Entre los genes estudiados con baja penetrancia se encuentran *RECQL*, helicasa implicada en reparación de ADN y *MCPH1* asociada a respuestas celulares al daño del ADN.

El objetivo general fue determinar si variaciones en regiones específicas de los genes *RECQL* y *MCPH1* en población chilena con CMH están asociadas a un mayor riesgo de desarrollar este cáncer.

Se obtuvo ADN genómico de 200 casos de CMH chilenos caracterizados por edad, tipo histológico e historial familiar, y 250 controles. Se realizó la búsqueda de variantes ("screening") dirigidas al exón 6 de *RECQL* y al exón 8 de *MCPH1* mediante reacción en cadena de la polimerasa-"high resolution melting" (PCR-HRM). Se secuenciaron 9 casos y 6 controles con un perfil HRM característico. Se evaluaron frecuencias de variantes y haplotipos (utilizando MEGA y Chromas), y su asociación al riesgo de CM. Además se analizó *in silico* el posible impacto de las variantes en la estructura de la proteína.

La edad promedio de casos y controles fue 45 y 56 años, respectivamente. El tipo histológico más común fue ductal infiltrante (61%). En *RECQL*, se encontró el mismo perfil HRM en todas las muestras. En *MCPH1* se identificaron dos perfiles (A y B), representando el 71% y 29% de los casos CM, y el 72% y 28% de los controles, respectivamente. No hubo diferencias significativas entre casos y controles según el perfil HRM ( $p=0,8231$ ). Tampoco se encontraron diferencias de los perfiles HRM según edad ( $p=0,2635$ ), tipo histológico ( $p=0,5023$ ) o presencia o ausencia de familiares con CM ( $p=0,1336$ ). La secuenciación vinculó el perfil HRM B de *MCPH1* con heterocigotos para 3 polimorfismos (c.863C>A, c.911G>T y c.940C>G). El análisis *in silico* de estas variantes indicó que afectarían la estructura, función y estabilidad de la proteína, aunque no hay datos descritos en CMH.

En conclusión, no se detectaron variantes en las regiones investigadas de *RECQL* y *MCPH1* que aumenten el riesgo de CM en la población chilena.

Se requiere estudiar otros sectores en *MCPH1* y *RECQL* para establecer la presencia de variantes asociadas con CMH.

## ix. Abstract

Breast cancer (BC) accounts for 24.5% of total cancer cases and 15.5% of cancer deaths. In Chile, it is the leading cause of cancer death in women.

The exact etiology of breast cancer is unknown, although there are risk factors such as sex, age and genetic predisposition. Ten percent are classified as hereditary breast cancer (HBC). Genes of high, moderate and low penetrance associated with susceptibility to HCM have been studied. Studies in Chilean and Latin American population are scarce.

Among the genes studied with low penetrance are *RECQL*, helicase involved in DNA repair and *MCPH1* associated with cellular responses to DNA damage.

The overall objective was to determine whether variations in specific regions of the *RECQL* and *MCPH1* genes in the Chilean population with HCM are associated with an increased risk of developing this cancer.

Genomic DNA was obtained from 200 Chilean HBC cases characterized by age, histologic type and family history, and 250 controls. The screening for variants targeting exon 6 of *RECQL* and exon 8 of *MCPH1* was performed by polymerase chain reaction-high resolution melting (PCR-HRM). Nine cases and six controls with a characteristic HRM profile were sequenced. Variant and haplotype frequencies (using MEGA and Chromas) and their association with CM risk were evaluated. In addition, the possible impact of variants on protein structure was analyzed *in silico*.

The mean age of cases and controls was 45 and 56 years, respectively. The most common histologic type was infiltrating ductal (61%). In *RECQL*, the same HRM profile was found in all samples. In *MCPH1*, two profiles (A and B) were identified, representing 71% and 29% of CM cases, and 72% and 28% of controls, respectively. There were no significant differences between cases and controls according to HRM profile ( $p=0.8231$ ). There were also no differences in HRM profiles according to age ( $p=0.2635$ ), histological type ( $p=0.5023$ ) or presence or absence of family members with CM ( $p=0.1336$ ). Sequencing linked the HRM B profile of *MCPH1* with heterozygotes for 3 polymorphisms (c.863C>A, c.911G>T and c.940C>G). *In silico* analysis of these variants indicated that they would affect the structure, function and stability of the protein, although there are no data described HBC. In conclusion, no variants were detected in the investigated regions of *RECQL* and *MCPH1* that increase the risk of CM in the Chilean population.

It is required to study other sectors in *MCPH1* and *RECQL* to establish the presence of variants associated with HBC.

# 1. INTRODUCCIÓN

## 1.1 Antecedentes generales sobre cáncer

El cáncer se ubica como una de las principales causas de muerte y una barrera importante para aumentar la esperanza de vida en todos los países del mundo. Según estimaciones de la Organización Mundial de la Salud (OMS) en 2019, el cáncer es la primera o segunda causa de muerte en 112 de 183 países mayoritariamente desarrollados o en vías de desarrollo antes de la edad de 70 años y ocupa el tercer o cuarto lugar en otros 23 países principalmente subdesarrollados. En general, la incidencia y mortalidad por cáncer está creciendo rápidamente en todo el mundo; esto refleja tanto el envejecimiento y crecimiento de la población como cambios en la prevalencia y distribución de los principales factores de riesgo de cáncer, varios de los cuales están asociados con el desarrollo socioeconómico (Sung & cols., 2021). En Chile, de acuerdo con el Departamento de Estadísticas e Información de Salud (DEIS), el cáncer corresponde a la segunda causa de muerte y se proyecta que para el año 2023 o más tardar el año 2024 se convierta en la primera causa de muerte a nivel nacional, afectando principalmente a mujeres (DEIS Minsal, 2022).

A nivel molecular el cáncer se desarrolla por la acumulación de alteraciones genéticas y epigenéticas, que permiten a las células sobrevivir, replicarse y evadir mecanismos reguladores de apoptosis, proliferación y del ciclo celular. El cáncer es impulsado por alteraciones del ADN, que incluyen reordenamientos cromosómicos, mutaciones y cambios epigenéticos, como la hipermetilación del promotor que da como resultado la activación de genes promotores del crecimiento (oncogenes) o la supresión de genes inhibidores del crecimiento (genes supresores de tumores). Las mutaciones pueden ser hereditarias o esporádicas. A nivel de nucleótidos, una mutación puede corresponder a una sustitución, adición o delección, y alterar la fisiología celular induciendo a la transformación de las mismas células. Mutaciones puntuales pueden llevar al cambio de aminoácidos en porciones críticas de la proteína (Meza-Junco & cols., 2006).

En mujeres, los cánceres más frecuentes son cáncer de mama, pulmón y bronquios, colon y recto, cérvico uterino y tiroides. En el caso de los hombres, los cánceres más prevalentes corresponden a próstata, pulmón y bronquios, colon, recto y vejiga (Hassanpour & Dehghani, 2017).

## 1.2 Cáncer de mama: antecedentes epidemiológicos

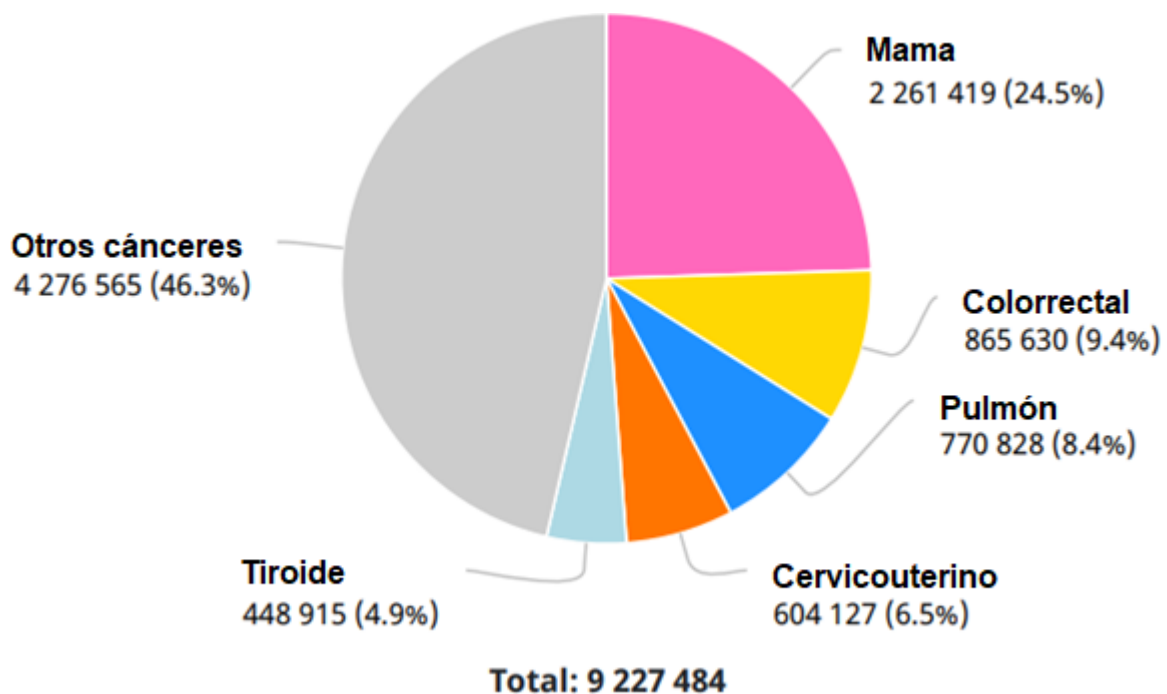
El CM es una neoplasia maligna que se origina en las células del revestimiento (epitelio) de los conductos o lóbulos del tejido glandular de los senos. En la etapa inicial del cáncer, el tumor está confinado al conducto o lóbulo (*in situ*), donde generalmente no causa síntomas y tiene un potencial mínimo de diseminación. Con el paso del tiempo, este cáncer *in situ* puede progresar e invadir el tejido mamario circundante, y a continuación propagarse a los ganglios linfáticos cercanos u a otros órganos del individuo. De acuerdo con lo anterior el CM se puede clasificar por tipos dependiendo de su origen en el seno, si el cáncer se origina en los conductos que llevan la leche desde los lóbulos hasta el pezón se denomina cáncer ductal que puede ser infiltrante (CDI) o *in situ* (CDIS); si el cáncer se origina en los lóbulos que producen la leche se denomina cáncer lobulillar infiltrante (CLI) o *in situ* (CLIS); existen otros carcinomas infiltrantes en esta región que representan menos del 5% del total, los que corresponden a medular, papilar, mucinoso, tubular, inflamatorio, adenoide quístico, metaplásico, enfermedad de Paget del seno, angiosarcoma, tumores filoides. Existe otra clasificación para el CM según el subtipo molecular que depende de la presencia o no de tres receptores en las células cancerosas: el receptor 2 del factor de crecimiento epidérmico humano (HER2); receptor para estrógeno (ER); y receptor para progesterona (PR). En esta clasificación el CM triple negativo (TNBC) se considera el más agresivo debido a que las células no presentan ninguno de los tres receptores (American Cancer Society (ACS), 2021).

Las tasas de supervivencias promedio a 5 y 10 años de las mujeres con CM infiltrante no metastásico son 90 % y 84 %, respectivamente. Es importante considerar que el 65% de los casos de CM infiltrante se diagnostica precozmente cuando este se encuentra solo en la mama, la tasa de supervivencia a 5 años de mujeres en este estadio es del 99%. Por otro lado, las mujeres adolescentes y adultas jóvenes de 15 a 39 años tienen menos probabilidades de ser diagnosticadas en un estadio temprano del CM (47% de los casos en este grupo etario) en comparación con las mujeres mayores de 65 años (68% de los casos en este grupo etario). El 6% de las mujeres tienen cáncer que se ha diseminado fuera de la mama y los ganglios linfáticos regionales (cáncer metastásico) en el momento en que se les diagnostica CM por primera vez, para este estadio la tasa de supervivencia a 5 años es solo del 30% (ACS, 2022).

El diagnóstico del CM generalmente se lleva a cabo mediante pruebas de detección de tumores como el ultrasonido mamario, mamografía, imagen por resonancia magnética y biopsia para lograr caracterizar el tumor de forma más específica, todo esto para lograr un

tratamiento lo más adecuado y rápido. Sin embargo, es necesaria la incorporación de nuevos diagnósticos, como las imágenes moleculares y perfiles de expresión genómica, que permiten tomar decisiones más certeras sobre tratamientos y ayuda a evitar efectos secundarios no deseados de un tratamiento no requeridos. Además, estos avances permiten la detección de tumores de menor tamaño lo que conlleva a una probabilidad menor de metástasis, mayor viabilidad de conservar la mama y axila, y es menos probable que requieran quimioterapia. Estos diagnósticos más actuales, combinados con tratamientos modernos que incluyen nuevas técnicas quirúrgicas, terapias químicas y de radiación, dan como resultado un enfoque multidisciplinario y colaborativo para reducir la morbilidad asociada a esta enfermedad (McDonald & cols., 2016).

El CM tiene presenta la incidencia más alta entre los cánceres que afectan a las mujeres a nivel mundial, en la **figura 1** se muestra el estudio realizado en 2020 a nivel mundial donde el CM representa un 24,5% de todos los casos de cáncer diagnosticados en mujeres que ascienden cerca de los 9.2 millones. Por otro lado, en el caso de hombres la incidencia es menor llegando a 2800 casos. Además de su alta incidencia, el CM también es la principal causa de muerte relacionada con el cáncer en mujeres alrededor del mundo llegando al 15,5% del total de muertes por cáncer, en cuanto a los hombres la mortalidad alcanza los 530 casos (Global Cancer Observatory (GLOBOCAN), 2020).



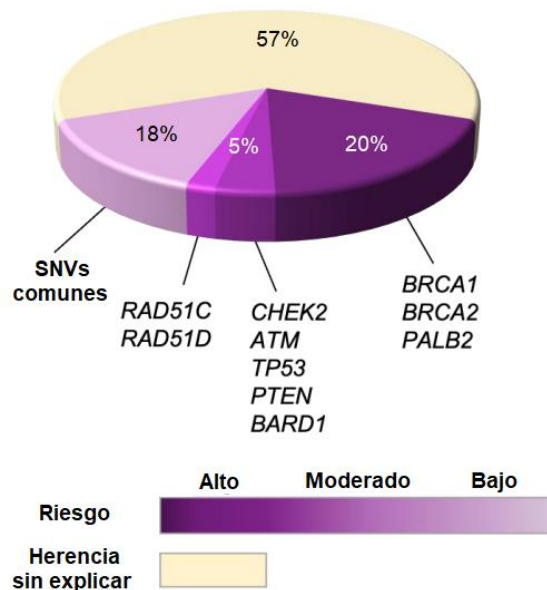
**Figura 1. Incidencia de cáncer en mujeres de todas las edades a nivel mundial.** El CM representa el 24,5% de casos totales

Chile no es una excepción a estas estadísticas globales, el CM tiene la mayor incidencia entre los cánceres que afectan a las mujeres a nivel nacional según GLOBOCAN representando el 21% del total de casos que sobrepasa las 25.000. El CM también es la mayor causa de muertes en mujeres relacionado al cáncer en Chile. La mortalidad ha ido en aumento en las últimas tres décadas, en 1990 se registraba una tasa de mortalidad de 12,4 por cada 100.000 mujeres, mientras que en el año 2020 causó 1674 muertes, con una tasa de mortalidad de 17,27 casos por cada 100.000 mujeres (GLOBOCAN, 2020). El 40% del total de los CM se diagnostican antes de los 45 años en Chile (Prieto, 2011).

La etiología exacta de este cáncer es desconocida, pero existen diversas causas y factores de riesgo que predisponen a tener CM entre los que se encuentran: el sexo de la persona, el 99% de los casos de CM se diagnostican en mujeres (Giordano & cols., 2002); la edad de la persona también es un factor de riesgo importante, ya que generalmente el CM de aparición esporádica se presenta en una edad avanzada. Se estima que el riesgo de desarrollar CM para una mujer de 30 años es de 1 entre 258, mientras que, a los 50 años, es de 1 entre 38 y a los 70 años es de 1 entre cada 25 mujeres (Singletary, 2003). También existen factores asociados a aspectos reproductivos, como lo es una menarquia temprana y una menopausia tardía, ambos se asocian con mayor riesgo de desarrollar CM (Kim & cols., 2015); el uso de anticonceptivos y terapia hormonal postmenopausia suelen aumentar el riesgo (Beaber & cols., 2014). Por otro lado, existe un tipo de CM que representa aproximadamente del 5 al 10% de los casos totales que se clasifica como CM de tipo hereditario y afecta principalmente a mujeres menores de 45 años (Frank & cols., 2014). Estos cánceres de desarrollo a edad temprana suelen tener una menor posibilidad de supervivencia (Thakur & cols., 2017). Por último, hay que destacar los factores hereditarios que representan un elemento de gran importancia a desarrollar CM, como son los genes de alta penetrancia, los genes de moderada penetrancia y los alelos de baja penetrancia (Cobain & cols., 2016) y el historial familiar de CM de las personas, se ha visto que el riesgo relativo aumenta, mientras más familiares han presentado la enfermedad, como también con el grado de parentesco (Bravi & cols., 2018).

### 1.3 Cáncer de mama hereditario (CMH)

En este tipo de cáncer existe una predisposición genética de desarrollar la enfermedad basada en la composición de los genes de una persona. Esta predisposición resulta de variaciones genéticas específicas que se heredan de uno o ambos de los padres. Estos cambios genéticos contribuyen al desarrollo de una enfermedad, pero no la causan directamente. Las variaciones en genes específicos conocidos como genes de susceptibilidad al CM implican un incremento en el riesgo de desarrollar la enfermedad, estos se clasifican dentro de tres categorías: 1. Genes de alta penetrancia, los cuales son raros en la población pero con un riesgo relativo alto entre 10 a 20 veces; 2. Genes de moderada penetrancia, los cuales también son raros en la población aunque más comunes que los de alta penetrancia, son genes que tienen relación con mecanismos de reparación del ADN y su riesgo relativo es moderado entre 2 a 4 veces; y 3. Alelos de baja penetrancia, estos son comunes en la población, ya que tienen una frecuencia mayor al 5%, y están asociados a un riesgo relativo bajo, en el caso de los homocigotos el riesgo es de 1,65 veces, mientras que para heterocigotos es de 1,25 veces (Torabi & cols., 2021). A pesar del bajo riesgo que se les atribuyen, el hecho de ser frecuentes a nivel poblacional y que han quedado muchos por descubrir, podría llegar a explicar una proporción mayor del CM que hasta los mismos genes BRCA1/2. En la **figura 2** se muestra la distribución porcentual de los genes de contribución de los distintos genes de susceptibilidad.



**Figura 2. Ilustración de las contribuciones proporcionales de los genes de riesgo de susceptibilidad al CM.** Los genes de alta penetrancia llegan al 20%, mientras que los genes de riesgo moderado aportan cerca del 5%, y por último los alelos de baja penetrancia contribuyen con aproximadamente el 18% (Torabi & cols., 2017).

Se conoce que las mutaciones en los genes de alta penetrancia como *BRCA1/2* y *TP53* son responsables de aproximadamente del 16 al 20% del total de CMH (Stratton & Rahman, 2008). Los genes *BRCA1* y *BRCA2* se identificaron en 1994 y 1996 respectivamente, son los genes supresores de tumores más estudiados relacionados con el CMH. La deficiencia de *BRCA1* conduce a la desregulación de puntos de control del ciclo celular, duplicación anormal del centrosoma, inestabilidad genética y finalmente, apoptosis (Deng, 2006). *BRCA2* regula la reparación recombinante en las roturas de la cadena doble del ADN al interactuar con *RAD51* y *DMC1* (Sánchez & cols., 2017). En la **figura 3** se muestra la distribución de los estudios asociados a variaciones genéticas en CMH en 12 países de Centro y Sudamérica (Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Paraguay, Perú, Uruguay y Venezuela). Entre 2002 y 2017 se han realizado 28 informes evaluando a 5956 personas e identificando 190 mutaciones patogénicas en los genes *BRCA1/2* y otros genes de susceptibilidad, observando que en nueve países si se realizaron estudios para *BRCA1/2* y otros genes, en un país solo estudios en genes distintos de *BRCA1/2* y por último que en los dos países restantes no hubo informes (Jara & cols., 2017).



**Figura 3. Estudios de mutaciones en *BRCA1/2* en los países de América Central y del Sur.** Doce países fueron evaluados. Nueve países tienen estudios de *BRCA1/2* y tres países no cuentan con estudios sobre *BRCA1/2* (Jara & cols., 2017).



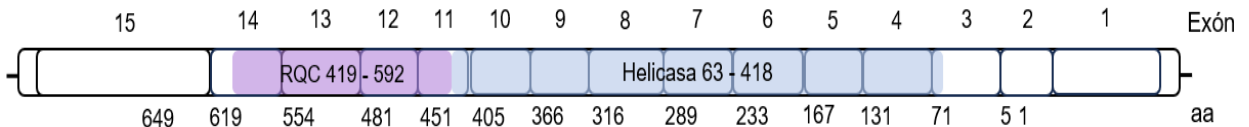
Debido a que el CMH abarca aproximadamente un 10% del total de casos de CM a nivel nacional e internacional. Sumado esto a la baja cantidad de estudios realizados a nivel nacional, así como en Centro y Sudamérica. A nivel mundial, únicamente se han identificado mutaciones en los genes BRCA1/2 y otros genes de alta o moderada penetrancia en alrededor del 25% de todos los casos de CMH , dejando un 75% restante sin explicación genética.

Por lo tanto, surge la necesidad de explorar nuevos genes o alelos de susceptibilidad al CM que puedan proporcionar información de esta parte aún desconocida del CMH. La identificación de genes o alelos contribuiría a mejorar la evaluación del riesgo, proporcionar conocimiento sobre los mecanismos de la enfermedad y para el desarrollo de terapias más eficaces y enfocadas.

Durante los últimos años en estudios de los investigadores Cybulski, Sun, Tervasmäki y Mantere se ha utilizado secuenciación de nueva generación (NGS) en genes involucrados en diversas vías de señalización de reparación del ADN en individuos con indicación de susceptibilidad a enfermedades hereditarias como personas menores de 40 años con CM y casos con historial familiar, basados en evidencia previa de asociación con el cáncer de mama, análisis de mutaciones de truncamiento de proteínas y la asociación con el cáncer mediante comparaciones de casos y controles. De este modo identificar nuevos genes relacionados con el CMH, entre los que se encuentran dentro de poblaciones europeas, asiáticas y oceánicas los genes seleccionados en este estudio *RECQL* y *MCPH1* que no han sido estudiados en poblaciones chilenas o latinoamericanas.

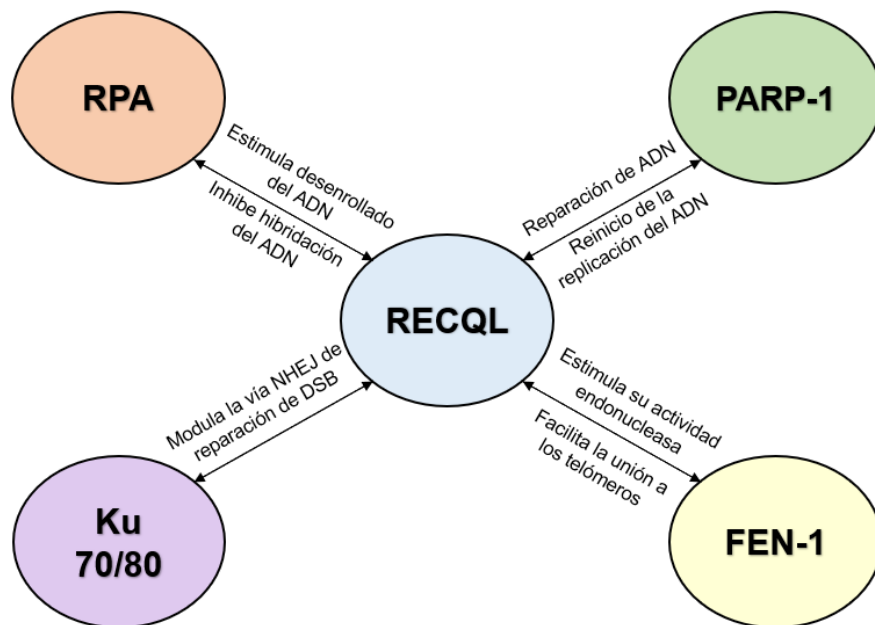
#### **1.4 Gen *RECQL***

El gen RecQLikeHelicase humano (*RECQL*), también conocido como *RECQ1*, ubicado en el cromosoma 12p12.1, contiene 15 exones codificantes que dan un transcrito de 3745 pares de bases (pb) y codifica para una enzima helicasa de 649 aminoácidos (aa) (Puranam & Blackshear, 1994) llamada *RECQL* (ATP-dependent DNA helicase Q1) (**figura 4**) que se expresa de forma ubicua en tejido humano adulto, no obstante, se observa una reducción en su nivel de expresión en tumores. Esta proteína representa el homólogo de RecQ más abundante presente en los seres humanos. Contiene dos dominios importantes, el dominio helicasa central altamente conservado entre el aa 63 al 418 y el dominio C-terminal específico de RecQ (RQC) desde el aa 419 al 592 (Pike & cols., 2009).



**Figura 4. Representación de los exones y la proteína RECQL junto a sus dominios.** Se muestran los 15 exones y su posición con relación a la proteína resultante y los dos dominios conservados, en azul se presenta el dominio helicasa del aminoácido 63 al 418 y en morado se presenta el dominio RecQ desde el aminoácido 419 al 592.

En la **figura 5** se muestran las múltiples funciones catalíticas e interacciones de RECQL con proteínas asociadas específicas para garantizar la estabilidad del genoma como lo son replication protein A (RPA), poly [ADP-ribose] polymerase 1 (PARP-1), Ku70/80, flap endonucleasa 1 (FEN-1) (Sami & Sharma, 2013). La deficiencia de *RECQL* en ratones (Sharma & cols., 2007) y células humanas (Sharma & Brosh, 2007) causa inestabilidad cromosómica constitutiva, intercambios de cromátidas hermanas, aumento de roturas de doble hebra del ADN y acumulación de intermediarios de recombinación no resueltos. Las células deficientes en *RECQL* acumulan daño en el ADN y muestran una mayor sensibilidad a los agentes que dañan el ADN que inducen horquillas de replicación estancadas y colapsadas, daño oxidativo y roturas del ADN de doble hebra (Popuri & cols., 2012).



**Figura 5. Esquema resumen de las funciones de RECQL.** 1) RECQL interactúa con RPA permitiendo la replicación, reparación y recombinación del ADN. 2) RECQL interactúa con PARP-1 siendo crítico para la reparación y replicación de ADN. 3) RECQL interactúa con Ku70/80 regulando su unión y modulando la vía de reparación de doble hebras del ADN. 4) RECQL interactúa con FEN-1 estimulando su actividad y la unión a los telómeros.

En relación con el cáncer en general, existen 509 mutaciones en este gen descritas en la base de datos COSMIC (Forbes & cols., 2011). RECQL se sobreexpresa en células transformadas en una variedad de cánceres como en glioblastoma, mieloma múltiple, ovario, hematológicos, cabeza y cuello, carcinoma hepatocelular, y carcinoma de células escamosas de la lengua (Debnath & Sharma, 2020).

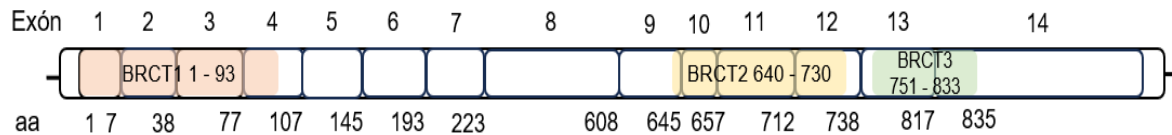
Con relación al CMH, *RECQL* fue descubierto como posible gen de predisposición en el año 2015 en familias polacas y francocanadienses, se detectó que las mutaciones (c.643C>T [p.R215] y c.1667\_1667+3delAGTA [p.K555 del ins MYKUHYSFR]) presentes en este gen provocan alteraciones a nivel génico (Cybulski & cols., 2015). También, en un estudio del norte de China, se encontraron 9 pacientes no relacionados entre ellos de los 541 pacientes con CM familiar que portaban tres mutaciones sin sentido (L128X, W172X y Q266X), una mutación en un sitio de empalme (395-2 A>G) y cinco mutaciones con cambio de sentido (A195S, R215Q, R455C, M458K and T562I) (Sun & cols., 2015). Así mismo en el estudio de Kwong en 2016, se identificaron cuatro mutaciones (c.974\_977delAAGA, c.796C>T, c.394+1G>A, c.867+1G>A) diferentes de *RECQL* en 6 de los 1110 (0,54%) pacientes analizados con CM en el sur de China (Kwong & cols., 2016).

Además de desempeñar un papel fundamental en la síntesis y reparación del ADN bajo condiciones de estrés, así como también su actividad helicasa interactuando con diversas proteínas necesarias para la replicación del ADN, la creciente evidencia apunta hacia una función de *RECQL* en la regulación de la expresión génica y la respuesta del paciente a la terapia contra el cáncer, esto debido a que se ha visto en estudios una correlación entre el mal pronóstico de CM con niveles menores del ARNm y también de expresión de *RECQL* en diversos tejidos. Por lo tanto, es importante dilucidar las funciones de *RECQL* en la reparación del ADN y la supresión de tumores, para así comprender la propuesta de *RECQL* como posible gen de susceptibilidad al CM (Benedict & cols., 2020).

### **1.5 Gen *MCPH1***

El gen microcephalin humano (*MCPH1*), también conocido como BRCT-repeat inhibitor of hTERT expresión (*BRIT1*), se ubica en el cromosoma 8p23.1, contiene 14 exones que dan un transcrito de 8008 pb. y codifica para una proteína de 835 aa llamada Microcefalina 1 (**figura 6**) que se expresa de manera ubicua en tejidos humanos adultos, incluidos el cerebro, los testículos, el páncreas y el hígado, además existe una disminución en la expresión en tumores. Esta proteína tiene tres dominios Breast cancer gene Carboxy-Terminal (BRCT), los dominios BRCT están presentes en proteínas que participan en la

señalización del daño en el ADN y median interacciones de tipo proteína-proteína y proteína-ADN (Gavvovidis & cols., 2012). MCPH1 contiene un dominio BRCT N-terminal que interactúa con el complejo SWI-SNF, mientras que los dos dominios BRCT C-terminal median las interacciones con péptidos fosforilados (Pulvers & cols., 2015).

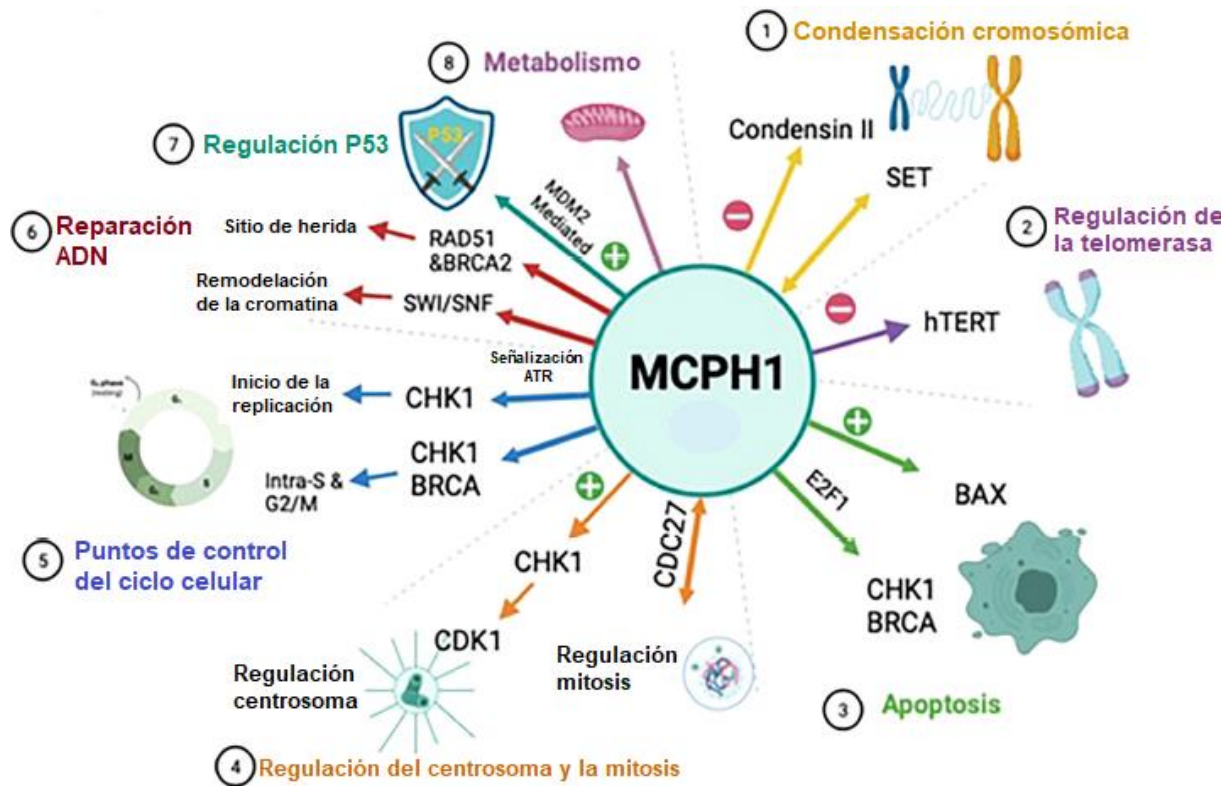


**Figura 6. Representación de los exones y la proteína RECQL junto a sus dominios.** Se muestran los 14 exones y su posición en relación con la proteína resultante y los tres dominios BCRT, en rosado se presenta el dominio BCRT1 del aminoácido 1 al 93, en naranja se presenta el dominio BCRT2 desde el aminoácido 640 al 730 y en azul se presenta el dominio BCRT3 desde el aminoácido 751 al 833.

Las mutaciones en *MCPH1* son la tercera causa más frecuente de microcefalia primaria autosómica recesiva (MCPH) y representan aproximadamente el 8% de los casos (Zaqout & cols., 2017). También se ha visto que los niveles de expresión del gen *MCPH1* disminuyen en el ADN, ARN y proteína en una serie de tipos de cáncer, incluido el cáncer de mama, pulmón, cuello uterino, próstata y de ovario en comparación con el tejido normal (Alsolami & cols., 2023).

Estudios han demostrado que está involucrado en varias funciones en el ciclo celular. Es necesario para la localización del punto de control quinasa1 (CHK1); en ausencia de MCPH1, se observa una entrada prematura a la mitosis que conduce a defectos en la organización del huso y desalineación cromosómica (Lin & cols., 2005) (Xu & cols., 2004). Actúa como represor transcripcional de hTERT y forma el complejo MCPH1-TRF2 aumentando la localización de factores de daño del ADN y promueve la reparación de telómeros defectuosos (Cicconi & cols., 2020). MCPH1 interactúa con la proteína componente del complejo promotor de la anafase CDC27 que controla la progresión del ciclo celular marcando proteínas específicas para su degradación permitiendo la entrada a la anafase desde la metafase (Singh & cols., 2012a). La pérdida de MCPH1 provoca una disminución en la relajación de la cromatina, a través de la reducción de la interacción MCPH1-SWI/SNF que podría conducir al desarrollo de cáncer (Peng & cols., 2009). Además, interactúa con la histona H2AX, que está involucrada en los mecanismos de reparación del daño de ADN actuando como supresor de tumores (Singh & cols., 2012b). En ratones se ha visto que la eliminación de *MCPH1* conduce a defectos en la condensación cromosómica, deficiencias en la reparación del ADN, desorientación del huso y, finalmente,

a una reducción del tamaño del cerebro que imita la enfermedad humana (Zhou & cols., 2013). En la **figura 7** se muestran las funciones biológicas de MCPH1 que apoyan su función como supresor de tumores.



**Figura 7. Esquema resumen de las funciones biológicas de MCPH1.** 1) MCPH1 afecta la condensación cromosómica. 2) MCPH1 ayuda a regular la función de las telomerasas. 3) MCPH1 promueve la apoptosis e inhibe la proliferación. 4) MCPH1 está involucrado en la regulación del centrosoma y la mitosis. 5) MCPH1 es necesario para los puntos de control del ciclo celular. 6) MCPH1 juega un papel en varios mecanismos relacionados con la reparación del ADN. 7) MCPH1 estabiliza la función de p53. 8) MCPH1 es crucial para controlar la actividad mitocondrial. (Alsolami & cols., 2023).

Con relación al cáncer en general, existen 1640 mutaciones en este gen descritas en la base de datos COSMIC (Forbes & cols., 2011). Se ha demostrado que *MCPH1* disminuye sus niveles de expresión en los cánceres de mama, próstata y ovario, y además está mutado en 1 de cada 10 CM y 5 de cada 41 tumores de endometrio, lo que sugiere que tiene función como un gen supresor de tumores (Venkatesh & cols., 2013).

En relación con el CMH, la expresión de este gen aparece significativamente reducida en muestras de CM (Rai & cols., 2008), también existen dos polimorfismos de un solo nucleótido (SNP) descritos para *MCPH1* asociados con un incremento en el grado de patogénesis del tumor de mama, esto al tener presente el alelo T en ambos rs2912010 y

rs1057090 (Jo & cols., 2013). En 2016, *MCPH1* fue descubierto como potencial gen de susceptibilidad en familias con CM en el norte de Finlandia, se observó que en pacientes con CM había una prevalencia del 3,4% de portadores de mutaciones a nivel del genoma (c.904\_916del) que conllevan a cambios a nivel de la proteína (p.Arg304ValfsTer3) demostrando un aumento significativo en la inestabilidad genómica (Mantere & cols., 2016).

La interacción de *MCPH1* con proteínas que se encargan de la reparación del ADN, así como también su participación en diversos puntos de control del ciclo celular, podría explicar su contribución al desarrollo de un tumor, como se ha visto en otros genes relacionados con el funcionamiento del ciclo celular. Esto ayudaría a entender mejor la propuesta de *MCPH1* como posible gen de susceptibilidad al CM.

La mayoría de los estudios descritos se han realizado en poblaciones caucásicas (Polonia, Canadá-Francia, Australia, Italia, Países Bajos, Finlandia y España) (Gracia-Aznarez & cols., 2013) (Cybulski & cols., 2015) (Mantere & cols., 2016) y no hay estudios concluyentes en poblaciones latinoamericanas, incluida la población chilena. Considerando que la variación intrapoblación es específica y en individuos chilenos con CMH, más de la mitad de la susceptibilidad genética permanece sin explicar se plantea la siguiente hipótesis.

## 2. HIPÓTESIS

Existen variantes en los genes *RECQL* y *MCPH1* que se asocian a mayor riesgo a desarrollar cáncer de mama hereditario en población chilena

## 3. OBJETIVOS

### 3.1 Objetivo general

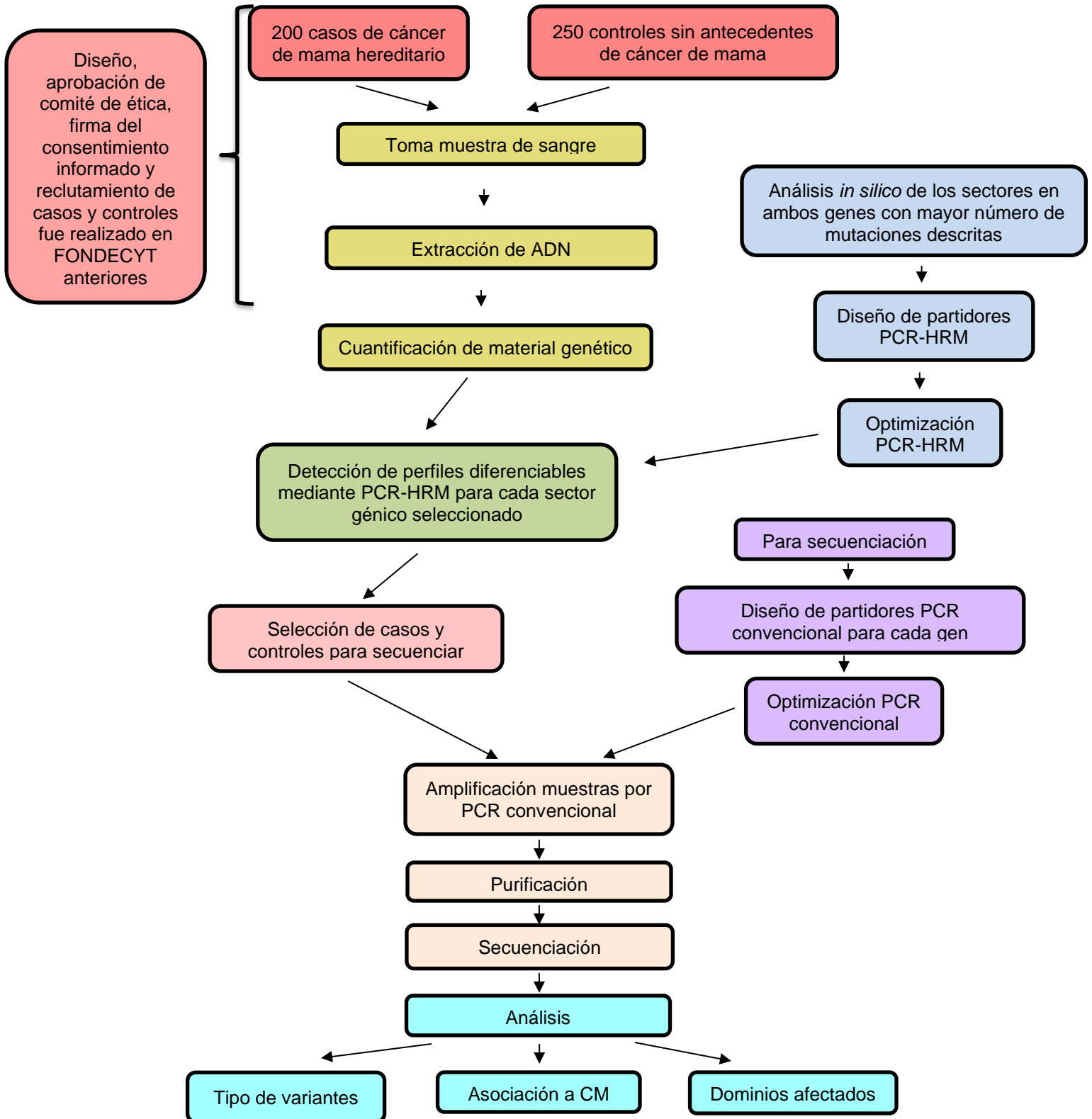
Determinar si variaciones en regiones específicas de los genes *RECQL* y *MCPH1* en población chilena con cáncer de mama hereditario están asociadas a un mayor riesgo de desarrollar este cáncer

### 3.2 Objetivos específicos

1. Investigar en bases bibliográficas acerca de variantes descritas en determinadas regiones de los genes *RECQL* y *MCPH1* con posible impacto en el desarrollo de cáncer de mama
2. Determinar la presencia de variantes dentro de las regiones de los genes *RECQL* y *MCPH1* en una población chilena con cáncer de mama hereditario y población sin antecedentes de cáncer de mama hereditario
3. Analizar *in silico* el impacto de las variantes detectadas en *RECQL* y *MCPH1* que podrían aumentar la susceptibilidad al desarrollo de cáncer de mama hereditario

## 4. MATERIALES Y MÉTODOS

### 4.1 Diseño experimental





## **4.2 Grupos muestrales**

### **4.2.1 Obtención de muestras**

Se reclutaron 200 casos con cáncer de mama seleccionados de la Corporación Nacional del Cáncer (CONAC), Fundación Arturo López Pérez y otros lugares de la Región Metropolitana que han sido parte de los proyectos FONDECYT (1200049, 1150117, 1110081, 1061078), relacionados con la temática. También se recolectaron 250 muestras de controles sin antecedentes de cáncer. A todos los individuos seleccionados se les dio a conocer el contenido del estudio y se les pidió firmar un consentimiento informado, aceptando ser parte de la investigación voluntariamente. El estudio fue aprobado por el Comité de Ética (Facultad de Medicina y de cada servicio hospitalario) y la información de cada participante se mantiene confidencial.

### **4.2.2 Casos y controles**

Los casos incluyen 200 muestras de individuos enfermos con cáncer de mama hereditario pertenecientes a proyectos FONDECYT anteriores relacionados con el tema. De estos individuos se tiene registrado los siguientes datos: edad al momento de la toma de la muestra, el tipo histológico del cáncer detectado, el historial familiar de cáncer de mama y otros tipos acerca de parientes en primer y segundo grado. Por otro lado, los controles incluyen 250 casos de controles sin cáncer de mama a los cuales se les realizó una entrevista en proyectos FONDECYT anteriores para corroborar la ausencia de antecedentes de cáncer de mama en su historial familiar. Además, se conoce la edad de los participantes a la fecha de la toma de muestra.

## **4.3 Extracción de material genético**

Se tomaron muestras de sangre periférica de los 450 individuos mediante punción venosa. Desde estas muestras se extrajo ADN genómico en proyectos FONDECYT anteriores utilizando el método de Chomczynski (Chomczynski & Sacchi, 1987).

## **4.4 Cuantificación de material genético**

Para establecer la concentración y pureza del ADN obtenido se midió la absorbancia a 230nm, 260nm y 280nm en el espectrofotómetro EPOCH (Biotek™, Winooski, Vermont, Estados Unidos). Las muestras de ADN cuantificadas fueron diluidas a 10ng/μL para ser

usadas en la técnica de PCR en tiempo real, mientras que se utilizó una dilución de 50ng/μL para la técnica de PCR convencional.

#### 4.5 Búsqueda en bases de datos sobre *RECQL* y *MCPH1*

Se analizaron las características de los genes *RECQL* y *MCPH1* descritas en las bases de datos de GeneCards (<https://www.genecards.org/>), así como también de sus respectivas proteínas por medio de la base de datos de UniProt (<https://www.uniprot.org/>). Se realizó una búsqueda bibliográfica en la base de PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) sobre alteraciones descritas en estos genes, como mutaciones genómicas o funcionales a nivel de proteínas publicadas en los últimos 20 años. Todo esto con el fin de encontrar las regiones de los genes donde sea más probable la aparición de mutaciones que los afecten y puedan contribuir a catalogarlos como genes de susceptibilidad al CM.

#### 4.6 Diseño de partidores qPCR-HRM

Los partidores fueron diseñados con el programa Beacon Designer 8.21 (Premier Biosoft®, San Francisco, California, Estados Unidos), a partir de la secuencia NC\_000012.12 para el gen *RECQL* y NC\_000008.11 en el caso del gen *MCPH1*, estas son recolectadas del GenBank perteneciente al sitio del National Center for Biotechnology Information (NCBI). Los partidores fueron sintetizados por Integrated DNA Technologies (IDT) (Coralville, Iowa, Estados Unidos). Las secuencias de partidores para el gen *RECQL* se ubican en la zona intrónica cercana al exón 6, incluyéndolo en su totalidad. En el gen *MCPH1* la secuencia del partidador directo se ubica en la zona intrónica cercana al exón 8, mientras que la secuencia del partidador reverso se ubica dentro del exón 8. En la **tabla 1** se detallan las secuencias para cada uno de ellos.

**Tabla 1. Secuencia de partidores para la qPCR-HRM**

Partidor qPCR-HRM	Secuencia	Tamaño fragmento	Exón
<i>RECQL</i> Directo	GATAACGGATATTCAATGTGTAAC	390pb	6
<i>RECQL</i> Reverso	ACCTAGTATTTTCAGTCTGCTA		
<i>MCPH1</i> Directo	AAGTTAGCTCTCCGTTAA	445pb	8
<i>MCPH1</i> Reverso	AGCCATGTGATACTCTTT		

## 4.7 Optimización PCR tiempo real con “High Resolution Melting” (qPCR-HRM)

Se optimizaron las reacciones para ambos genes con el fin de realizar la búsqueda de mutaciones, se utilizaron 20ng de material genético por reacción en un volumen final de 10µL. Se emplearon dos tipos de reactivos para PCR-HRM debido a la disponibilidad de los proveedores, este cambio fue verificado con muestras que se ensayaron con ambos reactivos encontrando resultados similares en cuanto al perfil de HRM. Las reacciones se optimizaron siguiendo las indicaciones de cada proveedor. Las concentraciones utilizadas se detallan en la **tabla 2** para el reactivo Kapa HRM Fast (KAPABIOSYSTEMS, San Luis, Misuri, Estados Unidos) y en la **tabla 3** para el reactivo Type-it HRM-PCR (QIAGEN, Hilde, Alemania). Las reacciones de amplificación se realizaron en el termociclador ECO real time PCR (Illumina®, San Diego, California, Estados Unidos). Los protocolos utilizados se indican en la **tabla 4** para el reactivo Kapa HRM Fast y en la **tabla 5** para el reactivo Type-it HRM-PCR. En la **figura 8** se muestra un ejemplo de la amplificación obtenida a partir de este último protocolo.

**Tabla 2. Concentraciones de reactivos usados qPCR-HRM con kit Kapa HRM Fast**

Reactivo qPCR-HRM	Concentración inicial	Concentración final	Volumen final
Kapa HRM Fast	2x	1X	5µL
MgCl <sub>2</sub>	25mM	2.5mM	1µL
Partidor directo	10µM	0.2µM	0.2µL
Partidor reverso	10µM	0.2µM	0.2µL
H <sub>2</sub> O	-	-	1.6 µL

**Tabla 3. Concentraciones de reactivos usados qPCR-HRM con kit Type-it HRM-PCR**

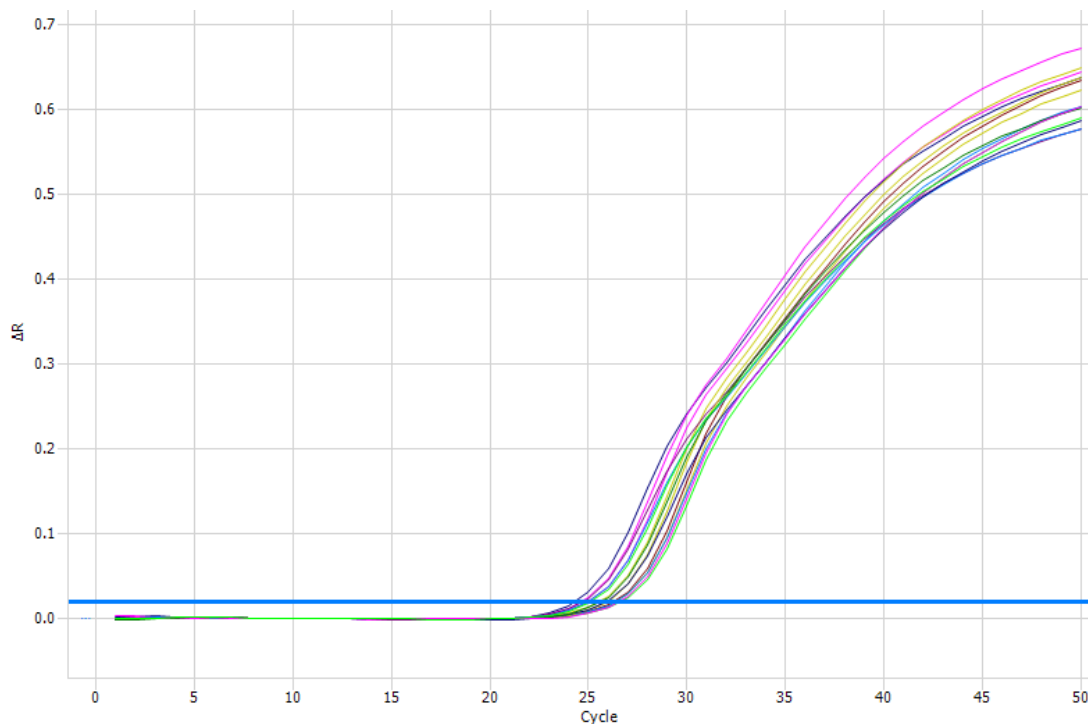
Reactivo qPCR-HRM	Concentración inicial	Concentración final	Volumen final
Type-it HRM-PCR kit	2X	1X	5µL
Partidor directo	10µM	0.7µM	0.7µL
Partidor reverso	10µM	0.7µM	0.7µL
H <sub>2</sub> O	-	-	1.6µL

**Tabla 4. Protocolo de reacción para qPCR-HRM con kit Kapa HRM Fast**

Protocolo qPCR-HRM KAPA	Tiempo (s)	Temperatura (°C)	Ciclos
Activación de la enzima	180	95	1
Denaturación	15	95	45
Alineamiento	60	54	
HRM	45	55-95	1

**Tabla 5. Protocolo de reacción para qPCR-HRM con kit Type-it HRM-PCR**

Protocolo qPCR-HRM QIA	Tiempo (s)	Temperatura (°C)	Ciclos
Activación de la enzima	300	95	1
Denaturación	15	95	50
Alineamiento	40	54	
Extensión	20	72	
HRM	45	55-95	1



**Figura 8. Gráfico modelo de la amplificación qPCR-HRM para el gen *MCPH1*.** El eje “x” muestra la cantidad de ciclos de amplificación y el eje “y” señala la magnitud de la señal fluorescente. La línea celeste indica la línea base de fluorescencia. El cycle threshold (Ct) se ubica cercano a los 25 ciclos de amplificación en las muestras analizadas en este ensayo. Cada curva de color representa una muestra individual.

#### 4.8 Diseño de partidores PCR convencional para secuenciar

Para diseñar estos partidores, el fragmento a amplificar para la secuenciación debía incluir en el sector central el fragmento que se analizó en el qPCR-HRM. Los partidores fueron diseñados con el programa Beacon Designer 8.21 (Premier Biosoft®), a partir de la secuencia NC\_000012.12 para el gen *RECQL* y NC\_000008.11 en el caso del gen *MCPH1*, estas son recolectadas del GenBank perteneciente al sitio del NCBI. Los partidores fueron sintetizados por IDT. Las secuencias de partidores para el gen *RECQL* se ubican en la zona intrónica cercana al exón 6, incluyendo en su totalidad la secuencia estudiada en PCR-HRM. En el gen *MCPH1* la secuencia del partidor directo se ubica en la zona intrónica

cercana al exón 8, mientras que la secuencia del partidor reverso se ubica dentro del exón 8, incluyendo en su totalidad la secuencia analizada en PCR-HRM. En la **tabla 6** se detallan las secuencias para cada uno de ellos.

**Tabla 6. Secuencia de partidores para PCR convencional**

Partidor PCR convencional	Secuencia	Tamaño fragmento	Exón
<i>RECQL</i> _Directo	TCAGTCCTCCTCATAATG	664pb	6
<i>RECQL</i> _Reverso	TTAGTCATCAACAATGTAGA		
<i>MCPH1</i> _Directo	TTACACCTCTTCTCTATCC	904pb	8
<i>MCPH1</i> _Reverso	TCTCTGAATACCTTTCCTTA		

#### 4.9 Optimización PCR convencional

Se optimizaron las reacciones para ambos genes con el fin de amplificar completamente el fragmento obtenido por el qPCR-HRM, se usaron 100ng de material genético por reacción en un volumen final de 25µL en cuadruplicado, las concentraciones de los reactivos se detallan en la **tabla 7**. El protocolo de la amplificación se muestra en la **tabla 8** para el PCR convencional realizado en el termociclador Veriti (Applied Biosystems™, Waltham, Massachusetts, Estados Unidos)

**Tabla 7. Concentraciones de reactivos usados PCR convencional**

Reactivo PCR convencional	Concentración inicial	Concentración final	Volumen
Buffer	10x	1x	2.5µL
dNTPs	2mM	0.2mM	2.5µL
MgCl <sub>2</sub>	50mM	3mM	1.5µL
Partidor directo	10µM	0.3µM	0.75µL
Partidor reverso	10µM	0.3µM	0.75µL
Polimerasa TaqBiotools	5U/µL	2U	0.4µL
H <sub>2</sub> O	-	-	14.6µL

**Tabla 8. Protocolo de amplificación para PCR convencional**

Protocolo PCR convencional	Tiempo (s)	Temperatura (°C)	Ciclos
Activación de la enzima	300	95	1
Denaturación	15	95	35
Alineamiento	15	55	
Extensión	60	72	
Extensión final	300	72	1

#### **4.10 Electroforesis en geles de agarosa**

Para evaluar los resultados de la PCR convencional se realizaron electroforesis en geles de agarosa al 1.5% utilizando el buffer Tris-borato-EDTA (TBE) con 4  $\mu$ L del intercalante SafeView™ (Applied Biological Materials, Richmond, Canadá). La electroforesis se realizó a un voltaje de 100 volts durante 45 minutos. En cada uno de los carriles se agregan 5  $\mu$ L de producto amplificado de la PCR convencional con 2  $\mu$ L de buffer de carga. Se usó el estándar de ADN (Accuruler, Maestrogen, Hsinchu, Taiwán) de 100 pb como referencia.

#### **4.11 Extracción y purificación material genético pos PCR**

Se realizó la extracción y purificación de ácidos nucleicos mediante columnas de centrifugación con el sistema comercial de FavorPrep™ (Favorgen®, PingTung, Taiwán) siguiendo las indicaciones del fabricante. El producto purificado se evaluó mediante electroforesis en gel de agarosa al 1,5% y se cuantificó usando el equipo EPOCH corroborando la obtención de una concentración mínima necesaria de 15 ng/ $\mu$ L de material genético.

#### **4.12 Secuenciación del producto amplificado**

Se secuenciaron 20 muestras en total que representan los distintos perfiles en ambos genes, usando los servicios de secuenciación de la Unidad de Secuenciación y Espectrometría de Masas (Pontificia Universidad Católica), quienes ocupan el secuenciador ABI PRISM 3500 xl (AppliedBiosystems™) junto con el kit BigDye™ Terminator v3.1 CycleSequencing (AppliedBiosystems™). Se secuenció solo una hebra en sentido 5 'a 3' con el mismo partidor utilizado en la PCR convencional para cada gen.

#### **4.13 Análisis bioinformático**

##### **4.13.1 Análisis de secuencias**

Para el análisis de las secuencias se usó el programa MEGA 11 (Tamura, K. & cols., 2021) para realizar alineamiento de las secuencias y ver posibles cambios a nivel nucleotídico, y también el programa Chromas 2.6 (Technelysium Pty. Ltd.) para examinar la secuencia en cromatograma y ver las variaciones detectadas en las muestras según perfil HRM.

#### 4.13.2 Variación en la proteína

Para evaluar *in silico* el posible impacto de variantes nucleotídicas que producen cambios aminoacídicos en la función y estabilidad de las proteínas derivadas de los genes *MCPH1* y *RECQL* se usaron una varios programas entre los que se encuentran Polymorphism Phenotyping v2 (PolyPhen-2), Sorting Intolerant From Tolerant (SIFT), Predictor of human Deleterious Single Nucleotide Polymorphisms (PhD-SNP), DisEMBL e I-Mutant 2.0.

#### 4.14 Análisis estadístico

Se realizó un análisis descriptivo de las características de la población chilena estudiada (edad, tipo histológico, historial familiar) y frecuencias genotípicas a través de promedio como medida de tendencia central y de dispersión (desviación estándar, rangos) para las variables cuantitativas y distribución de frecuencias para las variables categóricas.

En todos los SNPs se analizó el equilibrio Hardy-Weinberg (HWE) usando la prueba de chi-cuadrado ( $\chi^2$ ), obteniéndose un  $p > 0.05$  (población en equilibrio HWE).

Para las variables cuantitativas no paramétricas se utilizó la prueba de Mann-Whitney y para las variables cualitativas entre los casos con CM y controles mediante se usó pruebas de  $\chi^2$ , con el programa GraphPad Prism 8.0 (GraphPadSoftware, Inc). Se consideró una diferencia estadísticamente significativa si  $p < 0,05$ . La asociación con aumento del riesgo para CM se determinó calculando OR, con un IC al 95%. La interpretación estadística para los valores obtenidos de OR fue la siguiente:

- OR > 1.00 Asociación positiva (Aumenta el riesgo para CM)
- OR = 1.00 Sin asociación
- OR < 1.00 Asociación negativa (Disminuye el riesgo para CM)

## 5. RESULTADOS

### 5.1 Características demográficas de los casos con CM

En la **tabla 9** se muestra la distribución de los casos con CM y controles presentes en este estudio según los criterios de edad, y solo para los casos con CM se muestra el tipo histológico de cáncer e historial familiar.

**Tabla 9. Características de los 200 casos chilenos con CM y 250 controles analizadas en este estudio.**

<b>Mediana de edad</b>	<b>N (años)</b>	<b>IQR (años)</b>
Casos CM	45	36 - 52
Controles	57	44 - 69
p	<0,0001	
<b>Intervalos de edades casos CM (años)</b>	<b>N</b>	<b>Proporción (%)</b>
<30	18	9
30-39	54	27
40-49	57	28.5
50-60	50	25
>60	21	10.5
<b>Intervalo de edades controles (años)</b>	<b>N</b>	<b>Proporción (%)</b>
<30	3	1.2
30-39	36	14.4
40-49	54	21.6
50-60	48	19.2
>60	109	43.6
<b>Tipo histológico del CM</b>	<b>N (casos CM)</b>	<b>Proporción (%)</b>
Ductal infiltrante	121	60.5
Ductal <i>in situ</i>	29	14.5
Lobulillar infiltrante	18	9
Lobulillar <i>in situ</i>	4	2
Ductal + Lobulillar	7	3.5
Otros	10	5
S/I	11	5.5
<b>Historial familiar CM</b>	<b>N (casos CM)</b>	<b>Proporción (%)</b>
Tres o más familiares	31	15.5
Dos familiares de 1º grado	19	9.5
Dos familiares de 2º grado	16	8
Un familiar en 1º grado y otro en 2º grado	13	6.5
Un familiar de 1º grado	44	22
Un familiar de 2º grado	23	11.5
Sin familiares afectados	42	21
S/I	12	6

IQR, rango intercuartil; CM, cáncer de mama. El valor p para edades entre casos y controles se obtuvo con la prueba estadística de Mann-Whitney.



## 5.2 Búsqueda de regiones y mutaciones de interés

Posterior a la selección de los dos genes, se inició una búsqueda de mutaciones descritas en estos genes en el contexto de CM, principalmente en las zonas exónicas, para analizar posteriormente implicancias a nivel de las proteínas y los dominios de estas. Se consideraron artículos científicos de los últimos 10 años incluidos en PubMed y bibliotecas digitales. Esta búsqueda se realizó para determinar la región del gen en donde sea más probable encontrar variantes, basado en la frecuencia indicada en los artículo y así focalizar el estudio. En el caso de *MCPH1* se encontraron 6 variantes de interés descritas de distintos tipos (delección y SNP) en estudios sobre CM a lo largo del gen, teniendo en cuenta esto, se escogió una región perteneciente al exón 8, basándose en la frecuencia, en esta región se encuentra la variante “c.904\_916del” presentada en la **tabla 10**. En el caso de *RECQL* se encontraron más de 25 variantes de interés descritas de distintos tipos (delección, inserción, duplicación, SNP) en estudios sobre CM a lo largo del gen. Considerando la frecuencia y cantidad de variantes, se seleccionó una región del exón 6 para realizar la búsqueda experimental. Esta región acumula cinco variantes que se describen en la **tabla 11**.

Tabla 10. Variantes de interés en el gen *MCPH1* estudiadas en cáncer.

Estudio	Variante	Cambio en la proteína	Frecuencia	Exón
Jo & cols. (2013)	c.2226C>T	S742S	0.3973	13
	c.2282C>T	A761V	0.4059	13
Mantere & cols. (2016)	<b>c.904_916del</b>	R304fs	0,0345	8
Woods & cols. (2016)	c.2078G>A	R693His	-	11
	c.2443T>A	W815R	-	13
Bonache & cols. (2018)	c.322-1G>C	R108Mfs*15	0,0052	5
Tervasmaki & cols. (2019)	<b>c.904_916del</b>	R304fs	-	8

del, delección; fs, frameshift; \*, no codifica para aminoácido.

En rojo se destaca la variante encontrada en la zona seleccionada para realizar PCR-HRM.

Tabla 11. Variantes encontradas en el gen *RECQL* estudiadas en cáncer.

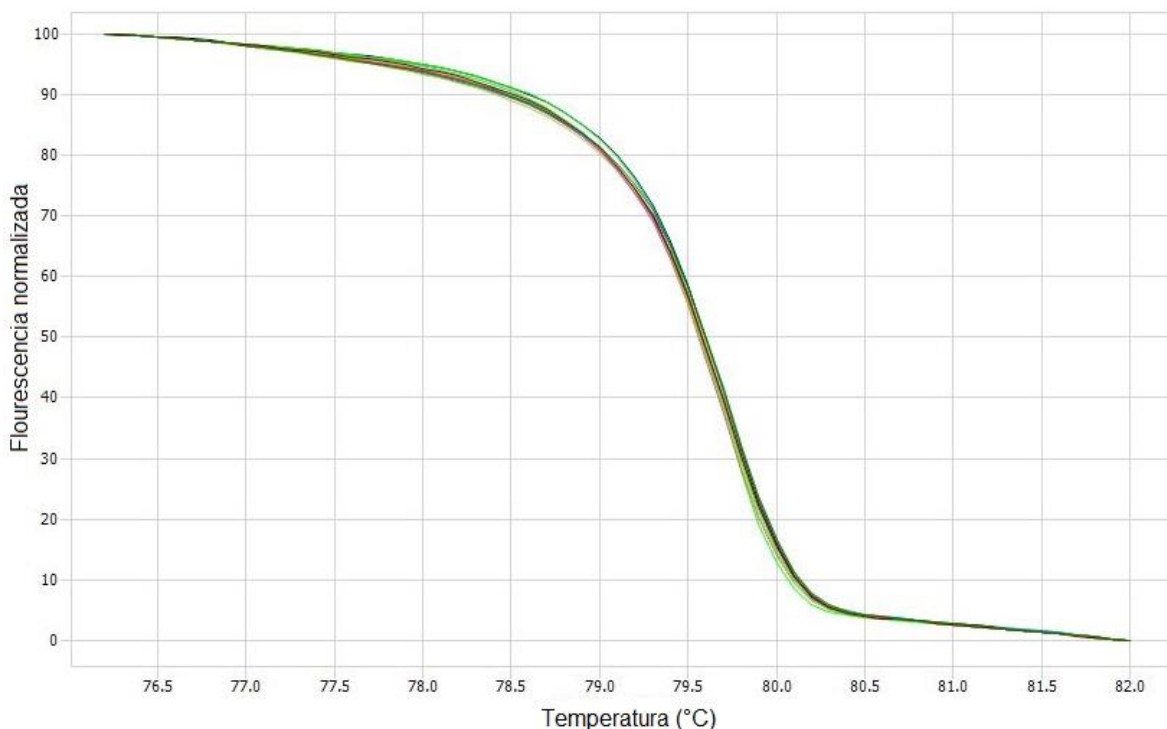
Estudio	Variante	Cambio en la proteína	Frecuencia	Exón
Cybulski & cols. (2015)	<b>c.643C&gt;T</b>	R215*	0,0069	6
	c.1667_1667+3delAGTA	K555delinsMYKLIHYSR	0,0023	13
Sun & cols. (2015)	c.383T>G	L128*	0,0022	4
	c.395-2A>G	G132fs*	0,0022	5
	<b>c.516G&gt;A</b>	W172*	0,0022	6
	<b>c.583G&gt;T</b>	A195S	0,0022	6
	<b>c.644G&gt;A</b>	R215Q	0,0022	6
	c.796C>T	Q266*	0,0022	7
	c.1363C>T	R445C	0,0022	12
	c.1373T>A	M458K	0,0022	12
c.1685C>T	T562I	0,0022	14	
Kong & cols. (2016)	c.796C>T	Q266*	0,0018	7
	c.974_977delAAGA	K325Tfs*	0,0009	9
Bogdanova & cols. (2017)	c.1667_1667+3delAGTA	K555delinsMYKLIHYSR	0,0034	13
Sun & cols. (2017)	c.120dupA	V41fs	0,0001	3
	c.180_190insATGATTCT	S64fs	0,0002	3
	c.383T>G	L128*	0,0001	4
	c.395-2A>G	G132fs*	0,0002	5
	<b>c.516G&gt;A</b>	W172*	0,0002	6
	<b>c.577A&gt;T</b>	K193*	0,0001	6
	<b>c.583G&gt;T</b>	A195S	0,0001	6
	<b>c.644G&gt;A</b>	R215Q	0,0012	6
	c.796C>T	Q266*	0,0003	7
	c.1363C>T	R445C	0,0001	12
	c.1373T>A	M458K	0,0001	12
	c.1398G>A	W466*	0,0002	12
c.1685C>T	T562I	0,0001	14	
c.1856dupA	N619fs	0,0001	14	
Tervasmaki & cols. (2018)	c.468T>G	I156M	0,0031	5
Li & cols. (2018)	c.962_965delGTTT	C321Ffs*	0,0002	9
	c.1476C>G	Y492*	0,0002	13
	c.1667_1667+3delAGTA	K555delinsMYKLIHYSR	0,0004	13
	c.1859C>G	S620*	0,0013	15

ND, no descrita; dup, duplicación; del, delección; fs, frameshift; \*, no codifica para aminoácidos.

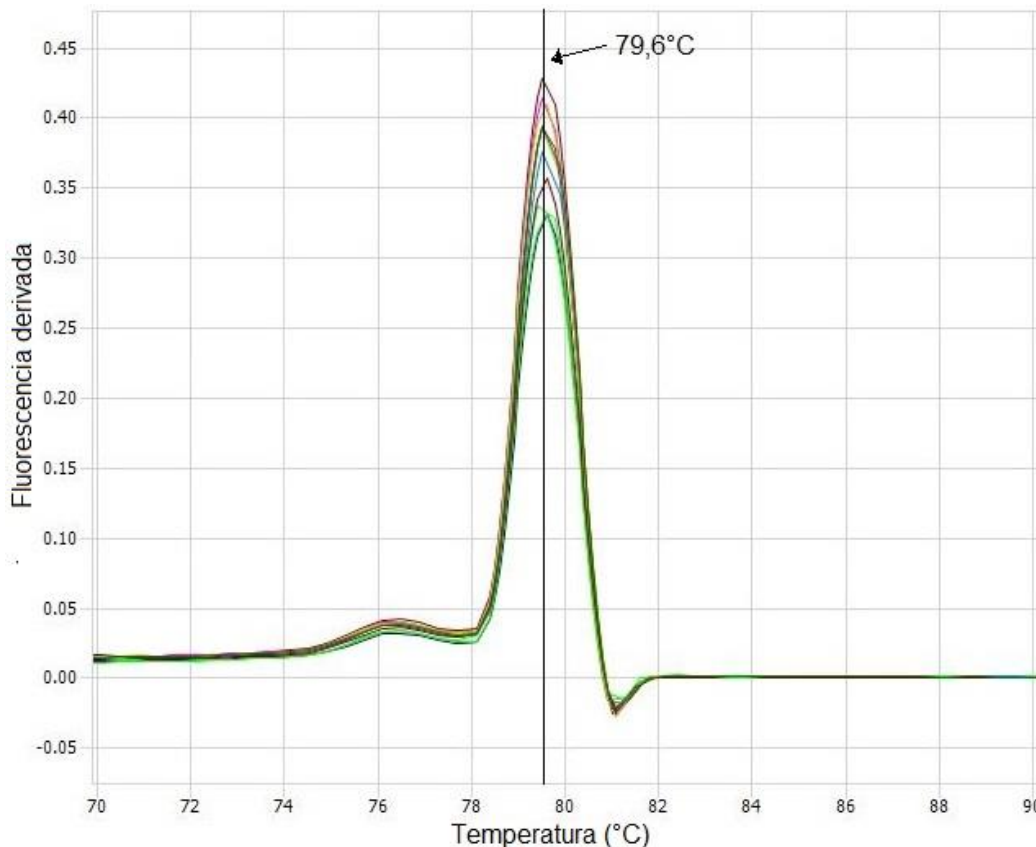
En verde se destacan las variantes encontradas en la zona seleccionada para realizar PCR-HRM.

### 5.3 Búsqueda de variantes mediante perfiles HRM gen *RECQL*

La búsqueda de variantes a nivel experimental se realizó mediante una pesquisa de variaciones en los perfiles obtenidos en las curvas disociación al obtener las temperaturas de *melting* ( $T_m$ ) en una zona seleccionada de cada gen como se describe en la metodología. Estos cambios se aprecian en las curvas de High Resolution Melt (HRM) y se denominaron como “perfiles de HRM”. Estos perfiles se obtuvieron en 200 casos con CMH y 250 casos controles dando como resultado perfiles para la región elegida del gen *RECQL* con un comportamiento similares en las 450 muestras analizadas. Estos resultados se ejemplifican en la **figura 9** mediante perfiles que muestran un solo punto de inflexión, y que se aprecia como un solo “peak” en la gráfica de la derivada de fluorescencia vs temperatura que correspondería a la temperatura de *melting* ( $T_m$ ). En la **figura 10** se presenta uno de los resultados con una  $T_m$  aproximada de 79,6°C. Este perfil y  $T_m$  es similar para todas las muestras analizadas para este gen con esta metodología.



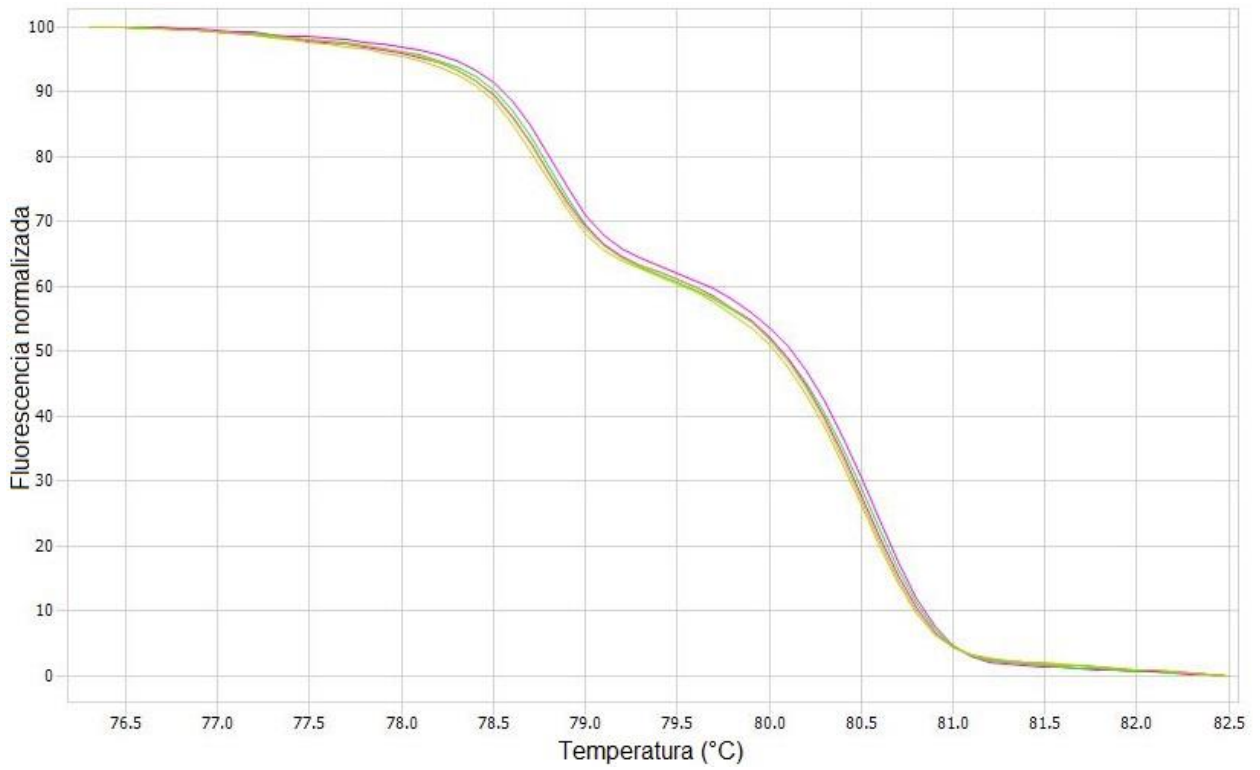
**Figura 9. Gráfico modelo del perfil de HRM para el gen *RECQL*.** El eje “x” muestra la temperatura en grados Celsius (°C) y el eje “y” señala la fluorescencia normalizada.



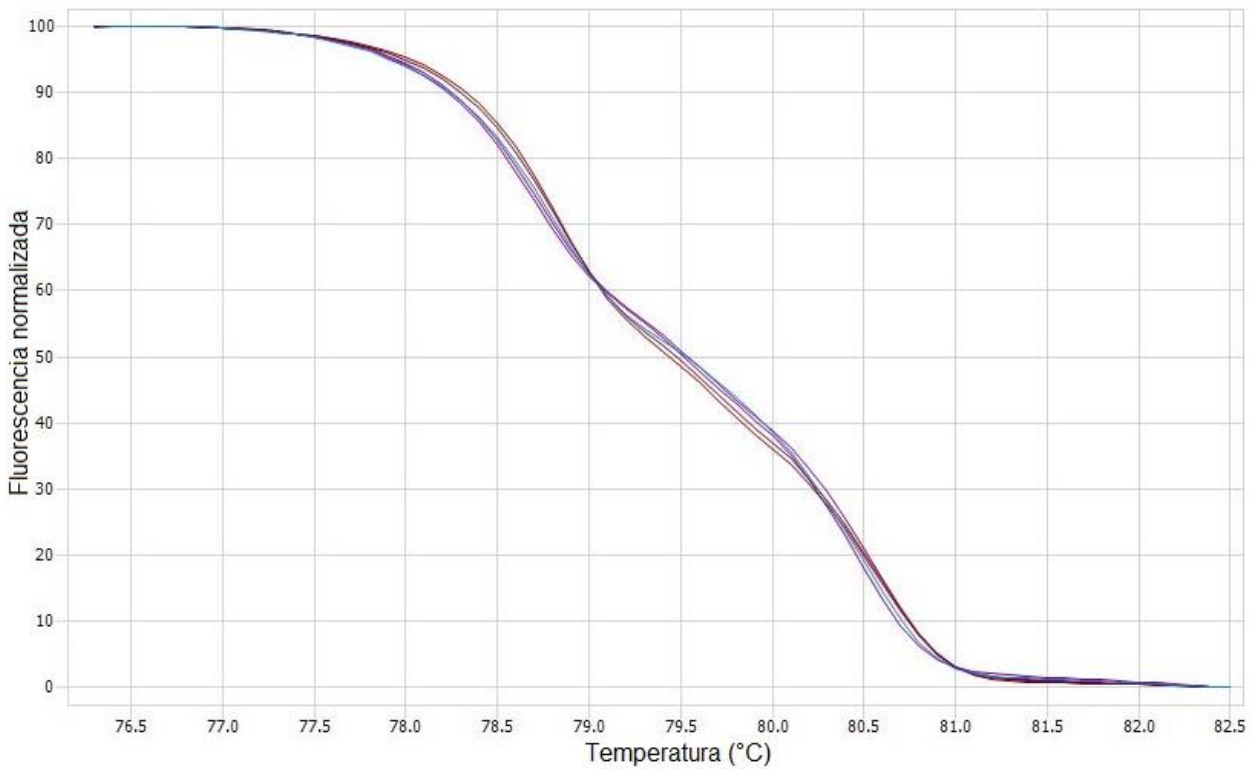
**Figura 10. Gráfico modelo del perfil de HRM tipo A para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia derivada.

#### 5.4 Búsqueda de variantes mediante perfiles HRM gen *MCPH1*

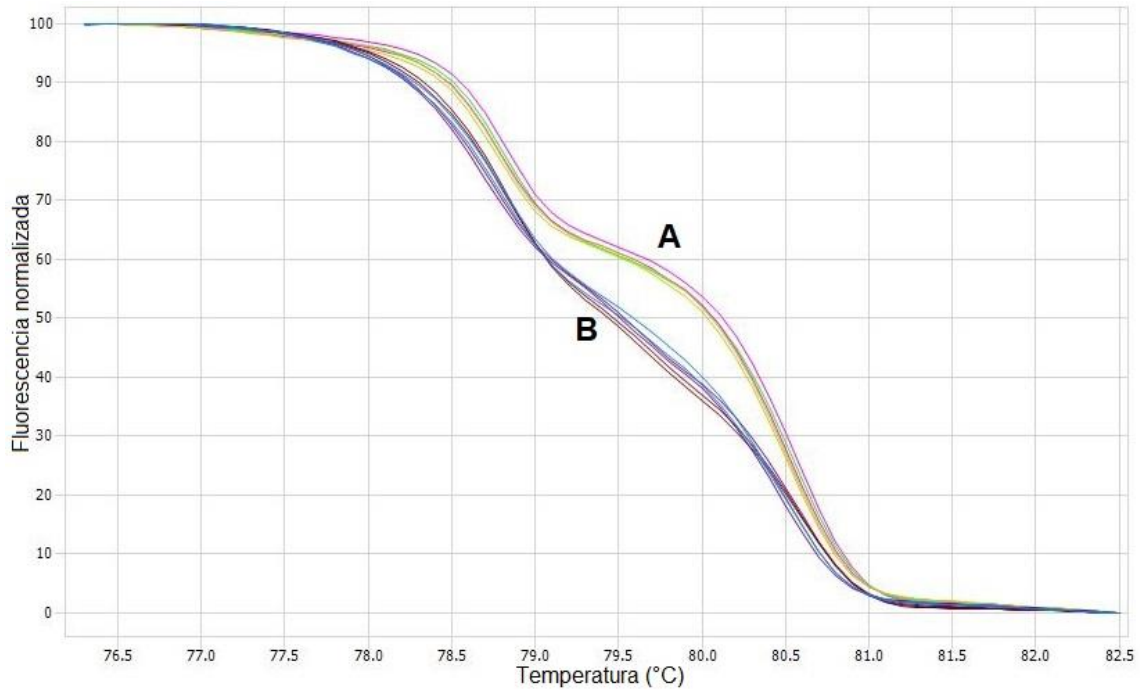
Los perfiles obtenidos en los 200 casos CMH y 250 casos controles para la región elegida del gen *MCPH1* dieron dos tipos diferenciables o perfiles HRM, los cuales se clasifican arbitrariamente como tipo A y B. En la **figura 11** se observan las perfiles de tipo A, la fluorescencia normalizada vs la temperatura indica un comportamiento del perfil en donde se tienen dos puntos de inflexión claramente diferenciables. En la **figura 12** se analizan los perfiles tipo B que se comportan de manera más plana con puntos de inflexión menos notorios; sin embargo, las  $T_m$  a la que ocurren estos puntos de inflexión son similares al caso anterior. Esta diferencia es más visible en la **figura 13** donde se comparan los perfiles de tipo A y B. Al analizar la derivada de la fluorescencia se logra ver el patrón de las  $T_m$ , donde se aprecian dos "peaks" en la **figura 14** para el tipo A y en la **figura 15** para el tipo B, en este caso, los valores de las  $T_m$  para ambos tipos de perfiles son similares 78.7°C y 80.4°C; sin embargo, la intensidad de la fluorescencia es menor en el segundo punto de inflexión para el perfil de tipo B.



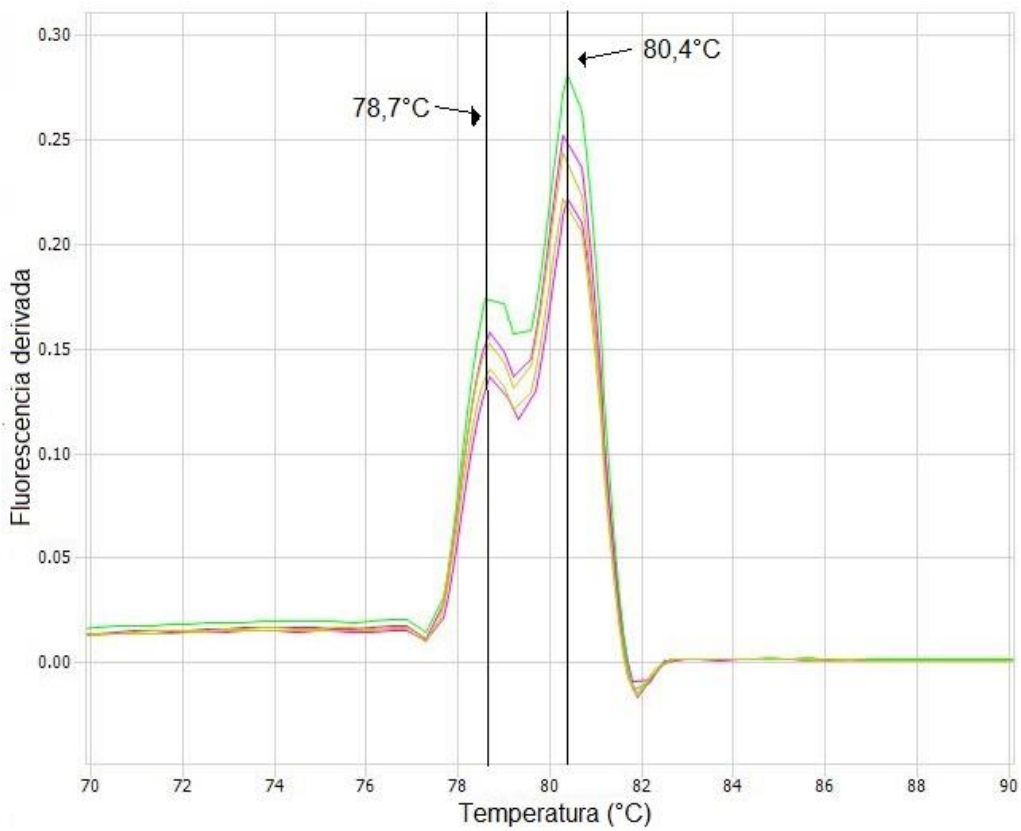
**Figura 11. Gráfico modelo del perfil HRM tipo A para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia normalizada.



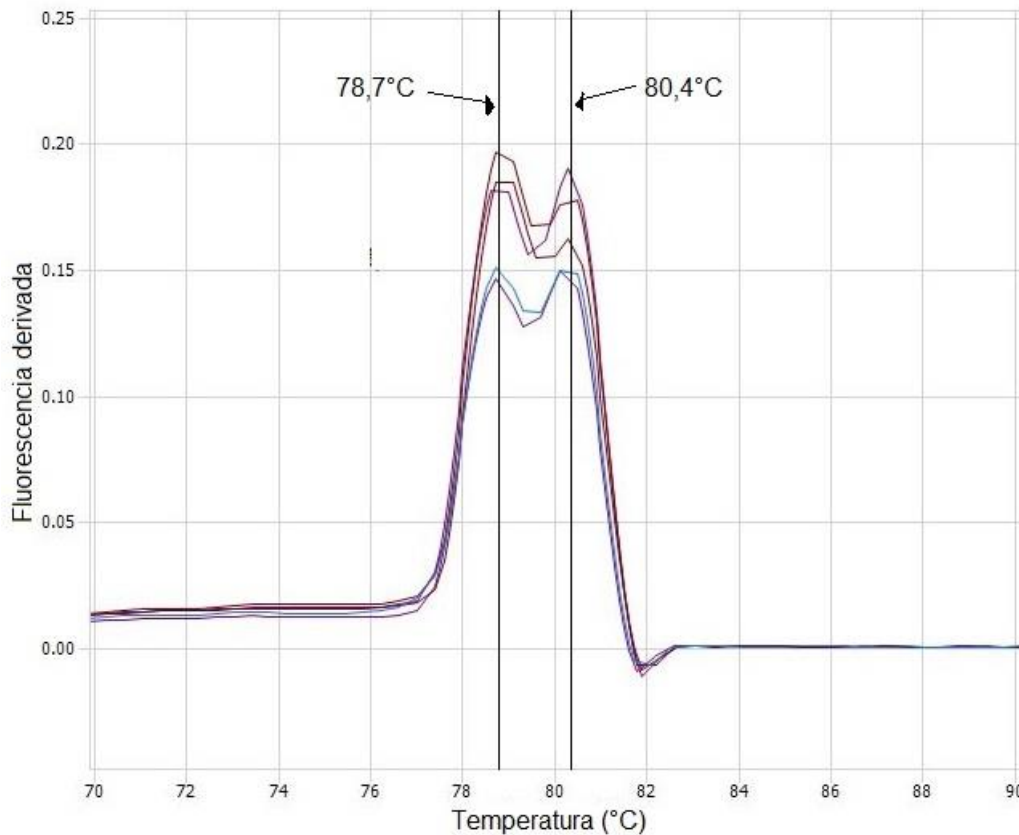
**Figura 12. Gráfico modelo del perfil HRM tipo B para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia normalizada.



**Figura 13. Gráfico modelo del perfil HRM tipo A y B para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia normalizada.



**Figura 14. Gráfico modelo del perfil HRM tipo A para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia derivada.



**Figura 15. Gráfico modelo del perfil HRM tipo B para el gen *MCPH1*.** El eje "x" muestra la temperatura en grados Celsius (°C) y el eje "y" señala la fluorescencia derivada.

#### 5.4.1 Casos CMH y perfiles HRM para el gen *MCPH1*

Los perfiles de HRM de los 200 casos con CMH en la región elegida para el gen *MCPH1* dieron dos tipos, comprobando esto al analizar las  $T_m$  obtenidas y las diferentes amplitudes que se obtuvieron en cada perfil. Un total de 142 casos positivos corresponden al perfil de tipo A representando un 71%, mientras que 58 casos positivos fueron perfiles de tipo B indicando un 29% del total.

#### 5.4.2 Controles y perfiles HRM para el gen *MCPH1*

Los perfiles de HRM de los 250 controles en la región seleccionada para el gen *MCPH1* dieron dos tipos de comportamiento, corroborando esto al observar las  $T_m$  obtenidas y además analizando la intensidad de fluorescencia. Un total de 180 controles estudiados tienen un perfil de tipo A lo cual corresponde a un 72%, en tanto 70 controles presentaron un perfil del tipo B lo que corresponde a un 28%.

### 5.4.3 Comparación entre casos CMH y controles en el gen *MCPH1*

En la **tabla 12** se comparan los 200 casos CMH con los 250 controles según el tipo de perfil HRM obtenido para la región seleccionada del gen *MCPH1*, se observa que no existe una diferencia significativa entre ambos grupos ( $p=0.8231$ ). El OR da como resultado 0,95, este valor al ser tan cercano a 1 indica que no existe un riesgo mayor ( $OR>1$ ), así como tampoco un nivel de protección ( $OR<1$ ) a desarrollar cáncer de mama, de todas formas no existe significancia dado que el 1 está incluido dentro del intervalo de confianza

**Tabla 12. Comparación de los perfiles HRM obtenidos en casos CMH y controles**

Perfil	Casos (%)	Controles (%)	OR (IC=95%)	p
B	58 (29)	70 (28)		
A	142 (71)	180 (72)	0,9521 (0,6308 – 1,4371)	0,8231

OR, odd ratio; IC, intervalo de confianza.  $p<0,05$  se considera significativo

### 5.5. Identificación de variantes en el gen *MCPH1* por cromatograma

Luego de tener los resultados de las 450 muestras con sus perfiles HRM y la posterior secuenciación de 20 muestras representativas para ambos genes (12 casos CMH y 8 controles) se realizó el análisis de los cambios nucleotídicos en los fragmentos secuenciados. En la secuenciación del fragmento del gen *MCPH1* (9 casos y 6 controles) se detectaron tres variantes que correspondieron a SNP previamente descritos. La **figura 16** muestra el alineamiento de las secuencias en la zona específica de los SNP realizado con la plataforma MEGA 11, sin embargo, no se logra determinar si las variantes son homocigotas o heterocigotas de esta forma, por lo que se ocupa el programa Chromas 2.6 para analizar las secuencias por medio de un cromatograma. La **figura 17** indica el cromatograma del polimorfismo c.863C>A (rs35590577) encontrado en uno de los casos con CMH que fueron secuenciados. La **figura 18** presenta el cromatograma del polimorfismo c.911G>T (rs2083914) detectado en 4 casos CMH y 4 controles dentro de las muestras secuenciadas. Asumiendo que el tipo de perfil HRM que tenían estas muestras presenta la forma heterocigota de este SNP, entonces 57 casos CMH y 70 controles serían heterocigotos para este polimorfismo. La **figura 19** muestra el cromatograma del polimorfismo c.940C>G (rs930557) presente en 5 casos CMH y 4 controles dentro de las muestras secuenciadas. Asumiendo que el tipo de perfil HRM que tenían estas muestras presenta la forma heterocigota de este SNP, entonces de 58 casos CMH y 70 controles en la muestra total. En la **tabla 13** se detallan las frecuencias para las muestras secuenciadas como para las muestras totales de los tres SNP.



	c.863C>A	c.911G>T	c.940C>G
	*	*	*
Caso 1	G T C C T C A	A A A G A A A	C C T C A C C
Caso 2	G T C C T C A	A A A G A A A	C C T C A C C
Caso 3	G T C C T C A	A A A G A A A	C C T C A C C
Caso 4	G T C C T C A	A A A G A A A	C C T C A C C
Caso 5	G T C C T C A	A A A T A A A	C C T C A C C
Caso 6	G T C C T C A	A A A T A A A	C C T C A C C
Caso 7	G T C C T C A	A A A T A A A	C C T C A C C
Caso 8	G T C C T C A	A A A T A A A	C C T C A C C
Caso 9	G T C A T C A	A A A G A A A	C C T C A C C
Control 1	G T C C T C A	A A A G A A A	C C T C A C C
Control 2	G T C C T C A	A A A G A A A	C C T C A C C
Control 3	G T C C T C A	A A A T A A A	C C T C A C C
Control 4	G T C C T C A	A A A T A A A	C C T G A C C
Control 5	G T C C T C A	A A A T A A A	C C T G A C C
Control 6	G T C C T C A	A A A T A A A	C C T C A C C

Figura 16. Alineamiento de las muestras secuenciadas para el gen *MCPH1*. \*, posición SNP; A, adenina; C, citosina; G, guanina; T, timina. (MEGA 11 [Tamura, K. & cols., 2021]).

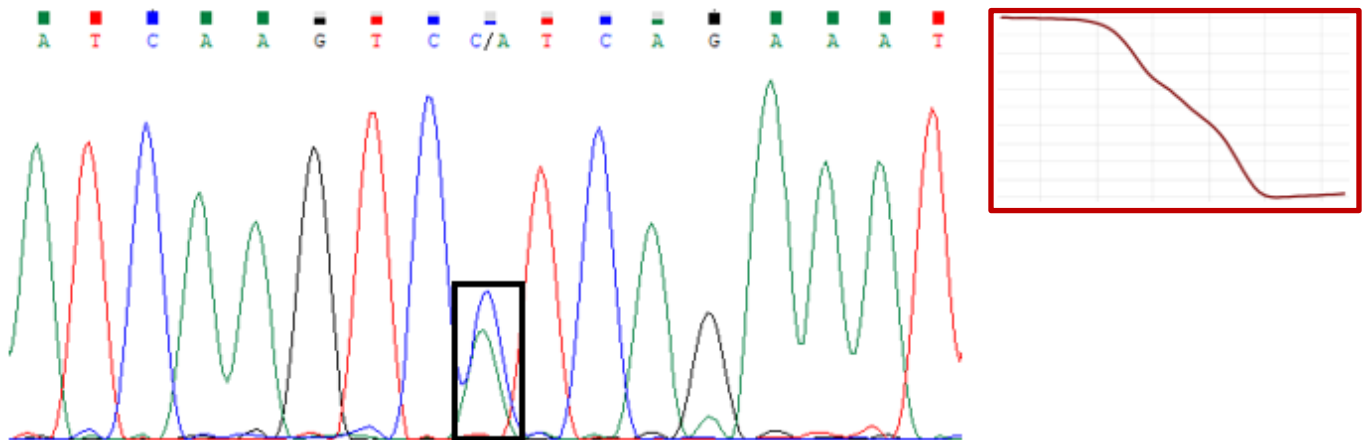
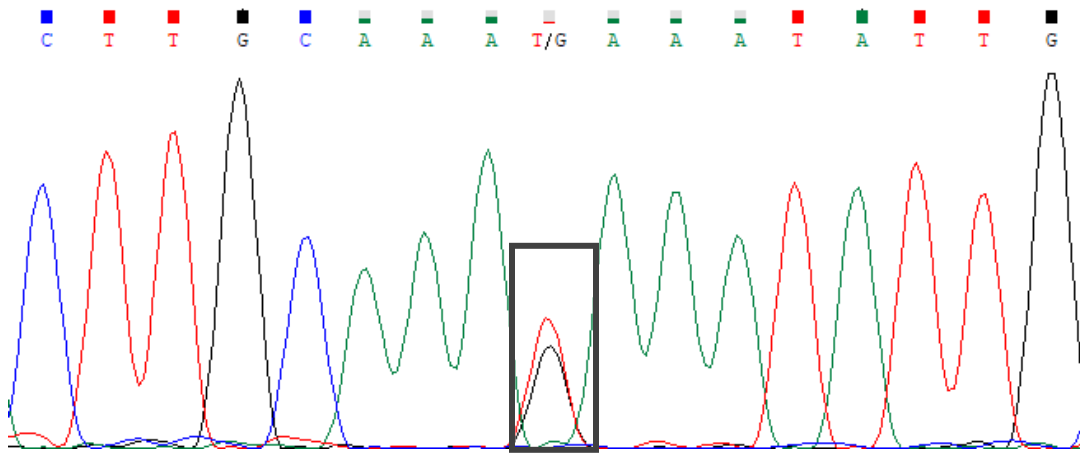
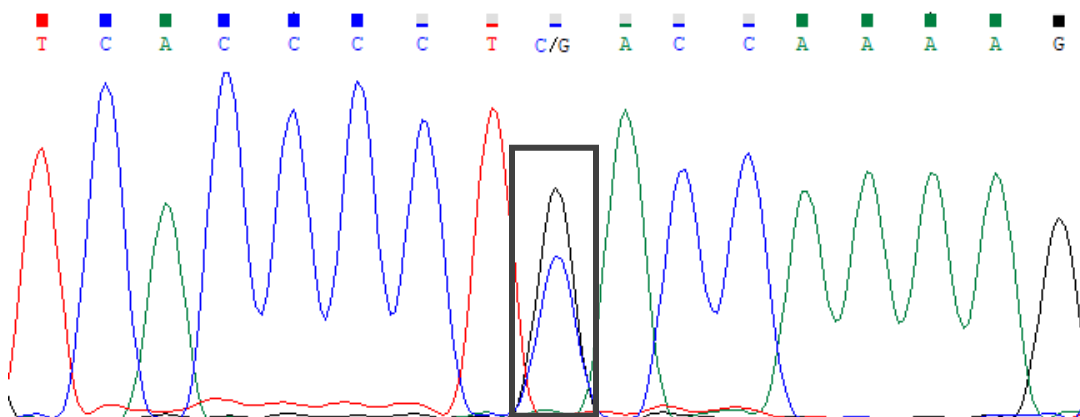


Figura 17. Cromatograma de una región del sector secuenciado del gen *MCPH1*. Se destaca el alelo heterocigoto SNP c.863C>A (rs35590577) con un rectángulo en negro. Se muestra a la derecha destacado en un rectángulo en rojo el perfil de HRM único correspondiente a este SNP. (Chromas 2.6 [Technelysium Pty. Ltd.])



**Figura 18. Cromatograma de una región del sector secuenciado del gen *MCPH1*.** Se observa el alelo heterocigoto SNP c.911G>T (rs2083914) destacado en un rectángulo. (Chromas 2.6 [Technelysium Pty. Ltd.])



**Figura 19. Cromatograma de una región del sector secuenciado del gen *MCPH1*.** Se observa el alelo heterocigoto del SNP c.940C>G (rs930557) destacado en un rectángulo. (Chromas 2.6 [Technelysium Pty. Ltd.])

## 5.6 Asociación de los haplotipos encontrados en *MCPH1* con los perfiles HRM

La **tabla 13** detalla la distribución y frecuencias de los tres haplotipos posibles encontrados en la secuenciación de *MCPH1* y como estos se relacionan con el perfil HRM en las muestras secuenciadas. Se observa que los homocigotos para el alelo C en c.863C>A, alelo G en c.911G>T y alelo C en c.940C>G muestran el perfil HRM de tipo A. Mientras que de forma independiente entre si los heterocigotos para c.863C>A y c.911G>T se ven acompañados regularmente del heterocigoto para c.940C>G representando ambas posibilidades el perfil HRM de tipo B, sin embargo, debido a la baja frecuencia encontrada en la literatura para el alelo A en c.863C>A se asume que el perfil HRM de tipo B se debe a los heterocigotos del SNP c.911G>T. También se indica el tipo histológico y el historial

familiar para determinar si hay alguna concordancia con los haplotipos detectados, se puede observar que el CM ductal infiltrante es el más común seguido del lobulillar infiltrante, lo que concuerda con los resultados generales, en cambio se observa una mayor cantidad de casos con tres o más familiares afectados con CM siendo diferente a la tendencia de las muestras totales. En la **tabla 14** muestra un resumen de los haplotipos encontrados y compara su frecuencia en los casos totales

**Tabla 13. Distribución de haplotipos para los 3 SNP encontrados en *MCPH1* y su relación con el perfil HRM, edad, tipo histológico e historial familiar de las muestras secuenciadas.**

Tipo muestra	Perfil HRM	N muestra	Edad	Tipo histológico	Historial familiar CM	c.863C>A	c.911G>T	c.940C>G
Casos	A	1	47	Lobulillar infiltrante	3 fam. o más con CM	C/C	G/G	C/C
		2	47	Ductal infiltrante	1 fam. 2º grado con CM	C/C	G/G	C/C
		3	44	Lobulillar infiltrante	3 fam. o más con CM	C/C	G/G	C/C
		4	37	Ductal <i>in situ</i>	3 fam. o más con CM	C/C	G/G	C/C
	B	5	35	Ductal infiltrante	Sin fam. con CM	C/C	G/T	C/G
		6	57	Ductal infiltrante	3 fam. o más con CM	C/C	G/T	C/G
		7	36	Ductal infiltrante	2 fam. 2º grado con CM	C/C	G/T	C/G
		8	42	Ductal infiltrante	3 fam. o más con CM	C/C	G/T	C/G
		9	53	Ductal infiltrante	1 fam. 1º grado con CM	C/A	G/G	C/G
Frecuencias casos ---->						C/C= 88,9%	G/G= 55,6%	C/C= 44,4%
						C/A= 11,1%	G/T= 44,4%	C/G= 55,6%
Controles	A	1	57	S/I	S/I	C/C	G/G	C/C
		2	61	S/I	S/I	C/C	G/G	C/C
	B	3	85	S/I	S/I	C/C	G/T	C/G
		4	85	S/I	S/I	C/C	G/T	C/G
		5	31	S/I	S/I	C/C	G/T	C/G
		6	42	S/I	S/I	C/C	G/T	C/G
Frecuencias controles ---->						C/C= 100%	G/G= 33,3%	C/C= 33,3%
							G/T= 66,7%	C/G= 66,7%
Frecuencias totales (casos y controles) ---->						C/C= 93,3%	G/G= 46,7%	C/C= 40%
						C/A= 6,7%	G/T= 53,3%	C/G= 60%
						A/A= 0%	T/T= 0%	G/G= 0%

fam, familiares; S/I, sin información.

**Tabla 14. Resumen de los haplotipos posibles encontrados y su frecuencia en muestras secuenciadas y totales**

Perfil HRM	c.863C>A	c.911G>T	c.940C>G	Controles (%)	Casos (%)	Casos totales (%)	Controles totales (%)
<b>A</b>	C	G	C	2/6 (33,3)	4/9 (44,4)	142/200 (71)	180/250 (72)
<b>B</b>	C	G/T	G/C	4/6 (66,6)	4/9 (44,4)	57/200 (28,5)	70/250 (28)
<b>B única</b>	C/A	G	G/C	0/6 (0)	1/9 (11,1)	1/200 (0,5)	0/250 (0)

## 5.7 Distribución de los casos con CMH según perfil HRM encontrado en el gen *MCPH1*.

### 5.7.1 Distribución por edad de los casos con CM

En la **figura 20** se muestra la disposición de los casos CMH según el perfil HRM, en donde se pueden observar pequeños cambios porcentuales en los diferentes rangos etarios nombrados; sin embargo, como lo indica la **tabla 15** al analizar a los casos con CM mayores de 45 años y menores o iguales a 45 años, se puede ver que no existe asociación con cáncer de mama ( $p=0,2635$ ) al comparar casos que presentan polimorfismos (heterocigotos) en la región estudiada en relación a los casos que no los presenta.

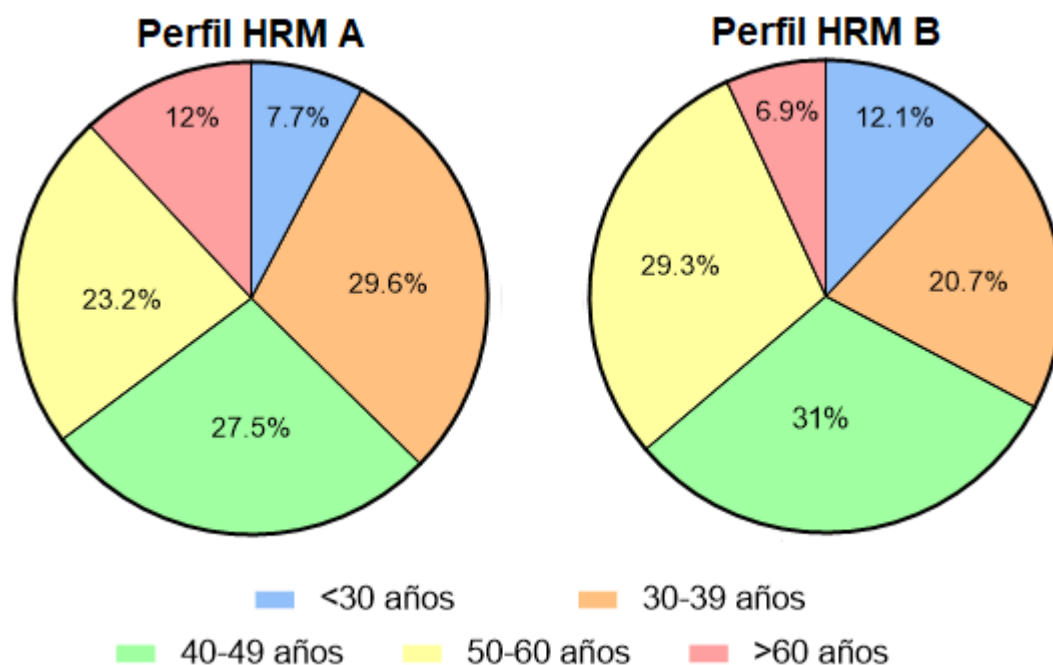


Figura 20. Distribución de los perfiles HRM de *MCPH1* según rango etario de los casos con CM.

Tabla 15. Comparación de los perfiles de HRM obtenidos en casos menores o igual a 45 años y mayores de 45 años

Perfil	≤ 45 años (%)	> 45 años (%)	OR (IC=95%)	p
B	26 (25)	32 (33)		
A	76 (75)	66 (67)	1,4172 (0,7673 – 2,6179)	0,2635

OR, odd ratio; IC, intervalo de confianza.  $p<0,05$  se considera significativa

### 5.7.2 Distribución por tipo histológico de CM del paciente

En la **figura 21** se muestra la distribución de los casos de CMH según el tipo de perfil HRM, se puede notar que la gran mayoría de los casos corresponden al cáncer de tipo ductal *in situ*, seguido en segundo lugar por el tipo lobulillar infiltrante; sin embargo, como lo indica la **tabla 16** al analizar a los casos con CM según si el cáncer se origina en los conductos o en los lóbulos, se puede ver que no existe asociación al desarrollo del cáncer mama ( $p=0.5023$ ) en los casos que presentan polimorfismos (heterocigotos) en la región estudiada en relación a los casos que no los presentan.

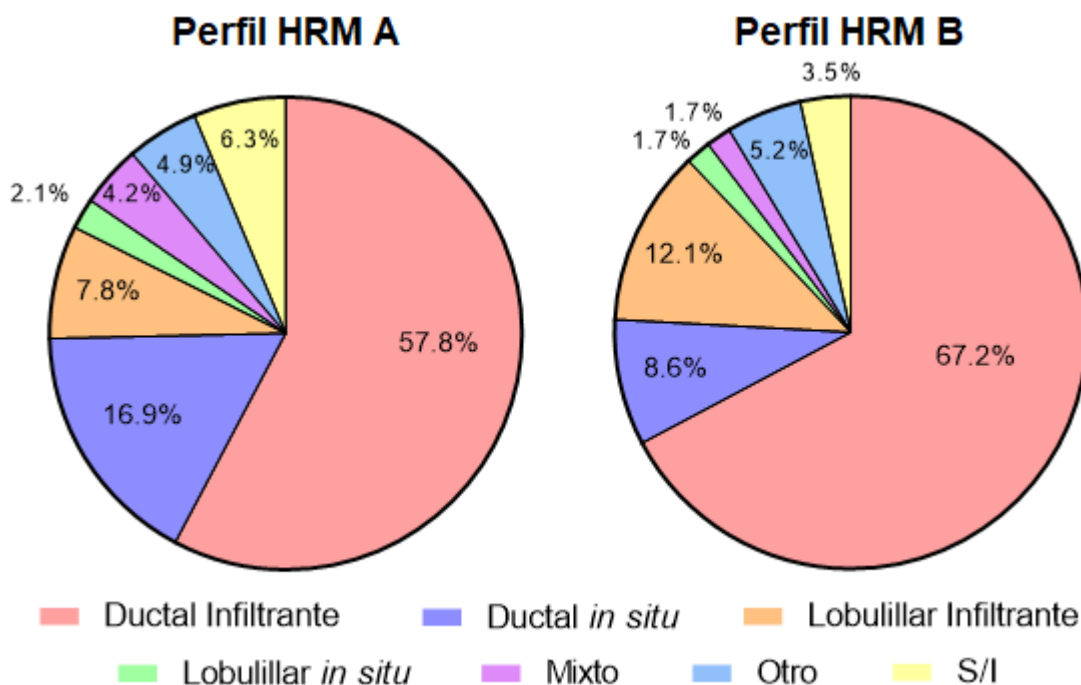


Figura 21. Distribución de los perfiles HRM de *MCPH1* según tipo histológico de casos con CM.

Tabla 16. Comparación de los perfiles de HRM obtenidos entre CM de tipo ductal y CM de tipo lobulillar

Perfil	CM Ductal (%)	CM lobulillar (%)	OR (IC=95%)	p
B	44 (29)	8 (37)		
A	106 (71)	14 (63)	1,3766 (0,5393 – 3,5139)	0,5023

OR, odd ratio; IC, intervalo de confianza.  $p < 0,05$  se considera significativo

### 5.7.3 Distribución por historial familiar de los casos con CM

En la **figura 22** se muestra la disposición de los casos con CM según el tipo de perfil HRM, en donde se pueden observar principalmente los porcentajes en casos sin familiares afectados sube en los casos que no presentan polimorfismos, sin embargo, como lo indica la **tabla 17** al analizar a los casos con CM según si la ausencia o presencia de familiares en cualquier grado presentan diferencias significativas, se puede ver que no existe una asociación al desarrollo del cáncer mama ( $p=0.1336$ ) en los casos que presentan polimorfismos (heterocigotos) en la región estudiada en relación a los casos que no los presentan

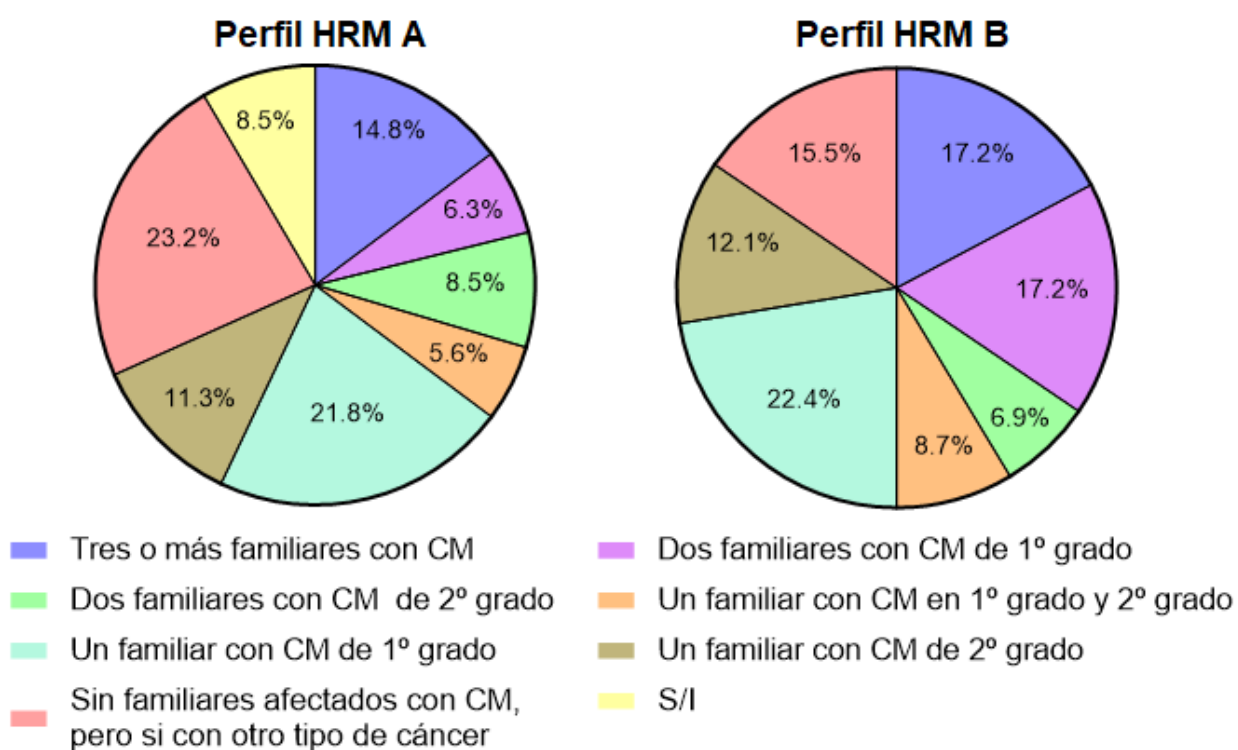


Figura 22. Distribución de los perfiles HRM de *MCPH1* según historial familiar de casos con CM.

Tabla 17. Comparación de los perfiles de HRM obtenidos en casos CMH en ausencia o presencia de familiares con CM

Perfil	Con familiares CM (%)	Sin familiares CM (%)	OR (IC=95%)	p
B	49 (34)	9 (21)		
A	97 (66)	33 (79)	0,5399 (0,2394 – 1,2175)	0,1336

OR, odd ratio; IC, intervalo de confianza.  $p < 0,05$  se considera significativa

## 5.8 Posible impacto de los SNP a nivel de proteínas *in silico*

Se analizó en plataformas bioinformáticas el posible impacto que tendrían los SNPs identificados anteriormente en la proteína MCPH1, esto en cinco plataformas bioinformáticas de uso liberado, disponibles en páginas web que analizan distintas características de la proteína. Las plataformas utilizadas son: 1. SIFT plataforma que evalúa si la sustitución de un aminoácido afecta la función de la proteína; 2. PolyPhen-2 herramienta que predice el posible impacto de una sustitución de aminoácidos en la estructura y función de una proteína mediante consideraciones físicas; 3. PhD-SNP método basado en discriminar mutaciones en proteínas de un solo cambio aminoacídico que pueden ser neutras o relacionadas con enfermedades; 4. DisEMBL herramienta computacional para la predicción de regiones desordenadas o no estructuradas dentro de una secuencia de proteína; y 5. I-Mutant 2.0 programa basado en la predicción de cambios en la estabilidad de proteínas en mutaciones de un solo aminoácido. En la **tabla 17** se describen los resultados de las predicciones resultantes de los programas mencionados.

**Tabla 18. Predicciones resultantes de los cambios en la proteína (UniProt ID: Q8NEM0) según plataformas bioinformáticas**

SNP	Aminoácido que cambia	SIFT	PolyPhen-2	PhD-SNP	DisEMBL	I-Mutant 2.0
<b>c.863C&gt;A (p.P288H)</b>	P [CCT] > H [CAT]	Afecta la función (0.0)	Probablemente dañina (0.998)	Neutro	Desorden por loops	Disminuye estabilidad
<b>c.911G&gt;T (p. R304I)</b>	R [AGA] > I [ATA]	Afecta la función (0.02)	Posiblemente dañina (0.875)	Neutro	Desorden por hot-loops	Disminuye estabilidad
<b>c.940C&gt;G (p.D314H)</b>	D [GAC] > H [CAC]	Afecta la función (0.02)	Probablemente dañina (0.952)	Neutro	Desorden por loops, hot-loops y X-Ray	Disminuye estabilidad

SIFT: probabilidades <0.05 se predicen deletéreas. PolyPhen-2: valores próximos a 1 indican un mayor daño. PhD-SNP; neutro, polimorfismo no relacionado con enfermedad.

## 6. DISCUSIÓN

El cáncer de mama es considerado uno de los mayores problemas de salud pública en el mundo, el más frecuente en la mujer en países desarrollados y en vías de desarrollo, además la primera causa oncológica de muerte en mujeres en población adulta.

En la actualidad, está en constante aumento la realización de investigaciones orientadas a descubrir genes responsables de la considerable fracción restante de casos de CMH, en los cuales los genes de alta penetrancia más estudiados, como *BRCA1/2* o *P53*, no presentan mutaciones. Gracias a la aplicación de nuevas técnicas de secuenciación y al avance del conocimiento en relación a los genes involucrados en la replicación y reparación del ADN, se han creado paneles de genes que exhiben vínculos con un mayor riesgo de desarrollar CM. Entre estos paneles destinados al estudio del cáncer hereditario, se han incluido los genes más estudiados relacionados al tema, así como también genes incipientes, en los que se catalogan los genes estudiados en este trabajo *RECQL* y *MCPH1* (Cybulski & cols., 2015)(Mantere & cols., 2016). Diversos estudios han abordado este tema en diferentes zonas geográficas y, por ende, en poblaciones diversas, obteniendo resultados coherentes en distintas partes del mundo en lo que respecta al aumento del riesgo de CM.

El objetivo principal de este estudio consistió en investigar la asociación entre las posibles variantes encontradas en los genes *RECQL* y *MCPH1* con un mayor grado de susceptibilidad al desarrollo de CM dentro de familias chilenas. Para lograr este propósito, se adoptó un enfoque metodológico basado en el análisis de casos con CM y controles buscando o realizando un “screening” de variaciones en un sector de cada gen. Es importante destacar que, hasta la fecha, no se han realizado investigaciones acerca de estos genes en poblaciones latinoamericanas en el contexto específico del CM de origen hereditario (Zavala, V. & cols., 2019).

En esta investigación, se incluyeron individuos con antecedentes familiares de CM, debido a que en la literatura se sugiere que hay una mayor probabilidad de identificar genes propensos a variantes relacionadas con el CM en este grupo. Asimismo, resultó fundamental recopilar datos demográficos de estos casos para poder analizar cómo la exposición a un factor de riesgo varía entre diferentes segmentos poblacionales. Esto abarcó consideraciones como la edad aproximada en la que se diagnostica el CM que puede tener factores hereditarios, dividiéndola en dos grupos: igual o menor a 45 años, y



mayor a 45 años; el tipo histológico del tumor, ya que se ha visto que dependiendo del tejido de origen y la propagación del CM este puede ser más agresivo, distinguiendo entre ductal infiltrante e *in situ* y lobulillar infiltrante e *in situ*; y finalmente, la historia y grado de parentesco familiar entre las personas que han desarrollado cáncer dentro del núcleo familiar y sobre todo CM, permitiendo una evaluación de varios factores que pueden contribuir al riesgo a CMH.

La metodología empleada demuestra innovación, ya que la técnica PCR-HRM, tradicionalmente utilizada para secuencias cortas de 70 a 150 pares de bases, con el fin de analizar una mutación puntual, se empleó de manera poco tradicional al amplificar fragmentos de mayor longitud, es decir, 390 pares de bases para *RECQL* y 445 pares de bases para *MCPH1*. Este enfoque permitió la detección de mutaciones en un contexto más amplio en lugar de limitarse al análisis de una sola mutación puntual. En el caso del gen *MCPH1*, se lograron identificar tres variantes distintas. Este resultado resalta el potencial de esta técnica para explorar otros genes, especialmente en regiones exónicas que comparten similitudes en términos de longitud de secuencia. Dando como resultado una vista completa de las posibles variantes presentes en el exón, lo que podría ser de gran importancia en futuras investigaciones genéticas. La posible asociación de un perfil HRM con variaciones genéticas específicas podría en un futuro también utilizarse como una manera de realizar un screening y ser utilizada como una herramienta diagnóstica más rápida y menos costosa que una secuenciación. Esta utilidad clínica no ha sido descrita anteriormente. Aunque es una proyección positiva de esta metodología, la asociación clara y no ambigua de un perfil HRM con un riesgo determinado debe ser evaluado con estudios con un gran tamaño muestral y que no varíe entre poblaciones (universalidad como marcador).

Los resultados de este estudio consisten en la asociación de los perfiles PCR-HRM, los cuales se diferencian por cambios en la temperatura de “melting”, debido a posibles variantes nucleotídicas que afecta esta temperatura en la secuencia estudiada. En el caso de *RECQL* no se revelaron la presencia de variantes de cualquier tipo como delección, inserción, duplicación o polimorfismos de un solo nucleótidos heterocigotos u homocigotos para alelos de baja frecuencia en la región analizada del gen *RECQL*. A través de la metodología de PCR-HRM, solo se encontró un perfil en el total de muestras tanto en casos con CM como en controles. Esto fue corroborado posteriormente al comparar algunas de las muestras secuenciadas con la secuencia de referencia informada en el GenBank. Sin embargo, existen otras zonas de interés en este gen descritas en la **tabla 10** por la cantidad

y frecuencia de variantes, como pueden ser los exones 5 o 13. En este último se ubica la variante “c.1667\_1667+3delAGTA”, que ha sido altamente estudiada en relación al CM (Cybulski & cols., 2015), por lo cual sería relevante analizarla en la población chilena. Debido a lo anterior el resto de los resultados se enfocan en el gen *MCPH1*.

En el gen *MCPH1* se encontraron dos perfiles de HRM, clasificados de forma arbitraria como A y B. El perfil A se encontró en el 71% de los casos con CM y en un 72% en los controles. El perfil B se detectó en el 29% de los casos con CM, en tanto que en los controles en un 28%. Por otro lado, al realizar la secuenciación de algunas de las muestras según el perfil HRM se detectaron tres polimorfismos: c.863C>A (rs35590577), c.911G>T (rs2083914) y c.940C>G (rs930557). La frecuencia descrita a nivel mundial para c.863C>A es para el alelo C de 97,6% y para el alelo A de 2,4%, mientras que a nivel latinoamericano para el alelo C de 98,4% y para el alelo A de 1,6%. En el caso de c.911G>T la frecuencia descrita a nivel mundial para el alelo G es 85% y para el alelo T de 15%, mientras que a nivel latinoamericano para el alelo G de 90% y para el alelo T de 10 %. Por el lado de c.940C>G la frecuencia descrita a nivel mundial para el alelo C es 78,5% y para el alelo G de 21,5%, mientras que a nivel latinoamericano para el alelo C es 82,7% y para el alelo G de 17,3% (Phan, L & cols., 2020), se puede ver que en latinoamérica el alelo de mayor frecuencia presenta un porcentaje más elevado que en todo el mundo.

Se puede inferir por medio de la secuenciación, que el perfil de HRM de tipo B corresponde en un 100% a muestras heterocigotas para el polimorfismo c.940C>G, no obstante los perfiles HRM dependen principalmente del contenido G/C y largo de la secuencia. Este SNP no produce un cambio en la Tm que se puede discriminar y por lo tanto no sería detectable por esta metodología; sin embargo, este siempre se detecta en pareja con alguno de los otros dos SNP encontrados, los cuales si provocan un cambio de Tm. De todas formas cabe la posibilidad que en muestras que no fueron secuenciadas se encuentren heterocigotos para c.863C>A y c.911G>T, donde no se detecte la formas heterocigota de c.940C>G. Resumiendo los datos obtenidos en esta memoria nos da como resultado tres posibles haplotipos para esta región (**tabla 14**). Sin embargo, tanto en la literatura como en este trabajo el SNP c.863C>A presenta un baja frecuencia para el alelo A, por lo cual se puede asumir que el cambio del perfil HRM se debe a heterocigotos para el polimorfismo c.911G>T. Mientras tanto que el perfil HRM de tipo A, la secuenciación entrega resultados que indican solamente la presencia de homocigotos para los alelos de mayores frecuencias descritos en la literatura para los tres polimorfismos, por lo que se comprende la forma más habitual detectada como perfil HRM en estas muestras.

En relación al análisis caso-control de los heterocigotos para las variantes identificadas en *MCPH1*, no se pudo establecer una asociación significativa que indique que los casos de cáncer de mama tienen una predisposición mayor para desarrollar esta enfermedad en comparación con los controles (OR = 0,9521 [IC95% 0,6308 – 1,4371]). Considerando la cercanía del valor de OR a 1, se puede llegar a proponer que no existe efecto protector ni de riesgo en estos tres polimorfismos en relación al CM. Sin embargo, estudios con una mayor cantidad de casos con CM pueden entregar nuevos resultados.

Si se analizan las frecuencias alélica a nivel mundial de las variantes encontradas en el gen *MCPH1*, según el proyecto 1000 genomas, nos señala que el alelo C en c.863C>A, alelo G en c.911G>T y alelo C en c.940C>G, son los alelos de mayor frecuencia. En la actualidad no existen publicaciones que asocien los SNPs detectados en este trabajo con la predisposición a desarrollar CM. Los resultados de esta memoria no permiten plantear que las variantes encontradas en la región estudiada del gen *MCPH1* se asocien a un mayor riesgo de desarrollar CM entre casos con CM agrupados por edades, tipo histológico e historial familiar de CM. Principalmente los heterocigotos para las tres variantes encontradas no se asocian significativamente con el aumento del riesgo o efecto protector en los tres grupos analizados, edad igual o menor a 45 años, y mayor a 45 años (**tabla 15**) presentando un OR = 1,4172 [IC95% 0,7673 – 2,6179], tipo histológico ductal o lobulillar (**tabla 16**) presentando un OR = 1,3766 [IC95% 0,5393 – 3,5139] y casos con CM con familiares afectados de CM o sin familiares afectados de CM pero si con otros tipos de cánceres (**tabla 17**) presentando un OR= 0,5399 [IC95% 0,2394 – 1,2175].

Como se mencionó anteriormente, no hay literatura que asocie los SNPs detectados del gen *MCPH1* con un mayor riesgo a desarrollar CM. Sin embargo, al realizar el análisis *in silico* de los cambios aminoacídicos en las distintas plataformas bioinformáticas, se indica posibles efectos funcionales de estos SNPs (**tabla 18**). En el caso de la variante c.863C>A se reemplaza el aminoácido Prolina (apolar e hidrofílico) por Histidina (básico e hidrofílico) alterando su función, disminuyendo su estabilidad, presentando desorden estructural y en conclusión resultando en un cambio probablemente dañino. En tanto la variante c.911G>T se reemplaza al aminoácido Arginina (básico e hidrofílico) por Isoleucina (apolar e hidrofóbico) alterando su función, disminuyendo su estabilidad, presentando desorden de la estructura secundaria en “hot-spots” y en conclusión resultando en un cambio probablemente dañino. En la variante c.940C>G se reemplaza al aminoácido Aspartato (ácido e hidrofílico) por Histidina (básico e hidrofílico) alterando su función, disminuyendo su estabilidad, presentando desorden de la estructura secundaria en “hot-spots”, cambios

en la estructura 3D por medio de cristalografía de rayos X y en conclusión resultando en un cambio probablemente dañino. Aunque estos análisis indican posibles cambios funcionales, es importante indicar que una proporción de las muestras eran heterocigotas para los SNPs, y con la metodología empleada no se puede determinar cuál de los dos alelos finalmente se expresa y cuál será el aminoácido que esté presente finalmente en la proteína. No se detectó ninguna asociación con otros factores de riesgo como edad menor de desarrollo de CMH, tipo histológico más agresivo, etc. Debido a la carencia de los antecedentes sobre el tipo molecular del CMH de las muestras analizadas no se puede hacer una asociación con respuesta a terapias hormonales y tampoco se tiene antecedentes de la sobrevivencia de los casos. En este estudio, el análisis de las variantes detectadas sobre funcionalidad de la proteína se realizó en plataformas bioinformáticas con un resultado de alteración de la función; sin embargo, en un estudio sobre su asociación con microcefalia, se consideraron alteraciones benignas a diferencia de nuestro análisis (Richards & cols., 2015).

Existieron limitaciones al realizar este trabajo de tesis, una de ellas fue la cantidad limitada de material genético de casos con CM en buen estado disponibles para llevar a cabo el estudio. Debido al tiempo desde su recolección posiblemente la integridad del material y el tamaño del fragmento grande que se pretendía amplificar no permitió amplificar más muestras. En relación a la cantidad de casos con CM esta es reducida, por lo cual no se logra una potencia estadística del 80%, debido a esto los resultados pueden no detectar cambios significativos. Además, otra limitante se originó en la elección del sector específico que se amplifica mediante PCR-HRM. Específicamente, en el caso del gen RECQL, donde no se encontraron variantes discernibles para su análisis. En relación a la metodología empleada para detectar variantes mediante PCR-HRM, surgió un problema cuando se identificaron dos mutaciones similares que generan un cambio idéntico en la  $T_m$ , por lo que, la presencia de estos cambios o variantes no permite una correcta discriminación con esta metodología.

Posibles mejoras en una investigación posterior, en caso de analizar regiones distintas de estos u otros genes relacionados con la reparación y replicación del ADN, en donde se encuentren variantes significativas por medio de esta metodología, son en caso de encontrar nuevamente en la secuenciación dos variantes que tengan el mismo cambio de  $T_m$  realizar una búsqueda acotada mediante la elección de nuevos partidores en ambas variantes para comprobar o descartar la presencia de ellas. Otra posible mejora es complementar los análisis *in silico* de los cambios provocados en la proteína, se sugiere realizar experimentos *in vitro* utilizando líneas celulares para evaluar de manera directa los

efectos de las variantes e identificar sus niveles por medio de Western Blot.

En este trabajo se evidencia que en las regiones de ambos genes estudiadas por PCR-HRM no se lograron encontrar variantes que se asocien con el aumento del riesgo a desarrollar CM en la población chilena analizada.

## 7. CONCLUSIONES

- En el exón 6 del gen *RECQL* se detectó un único perfil de HRM mediante la técnica PCR-HRM, tanto en casos como en controles.
- En el exón 8 del gen *MCPH1* se detectaron dos perfiles de HRM mediante la técnica de PCR-HRM, tanto en casos como en controles.
- El análisis de perfiles HRM de *MCPH1* no reveló diferencias significativas en el estudio caso-control.
- No se observaron diferencias significativas en los SNPs de *MCPH1* en la relación con la edad, el tipo histológico del CM y el historial familiar de CM, basándose en los perfiles HRM.
- Se detectaron tres SNPs mediante secuenciación en el gen *MCPH1*: c.863C>A (rs35590577), c.911G>T (rs2083914) y c.940C>G (rs930557).
- Los tres SNPs encontrados **no** han sido relacionados con el cáncer de mama.
- Según los datos de secuenciación las frecuencias alélicas obtenidas no difieren de lo reportado en 1000 genomas u en otras poblaciones.
- El análisis *in silico* del impacto en la proteína MCPH1 de las variantes encontradas en el gen *MCPH1* entregaron resultados que indican alteraciones en la estructura, función y estabilidad de la proteína.
- Los resultados cumplen con el objetivo general de determinar si las variaciones en la región seleccionada de los genes *RECQL* y *MCPH1* en la población chilena con cáncer de mama hereditario están asociadas con un mayor riesgo de desarrollar esta enfermedad.
- La hipótesis planteada fue refutada, ya que no se encontraron variaciones en *RECQL* y *MCPH1* que estuvieran asociadas con un mayor riesgo de desarrollar CM en la población estudiada.

## 8. BIBLIOGRAFÍA

1. Alsolami, M., Aboalola, D., Malibari, D., Alghamdi, T., Alshekhi, W., Jad, H., Rumbold-Hall, R., Altowairqi, A. S., Bell, S. M., & Alsiary, R. A. (2023). The emerging role of MCPH1/BRIT1 in carcinogenesis. *Frontiers in oncology*, 13, 1047588.
2. American Cancer Society (2021). *Cancer Statistics Center*.
3. Beaber EF, Buist DSM, Barlow WE, Malone KE, Reed SD, Li CI. (2014). Recent oral contraceptive use by formulation and breast cancer risk among women 20 to 49 years of age. *Cancer Res.*, 74(15):4078–4089.
4. Benedict, B., van Bueren, M. A., van Gemert, F. P., Liefink, C., Guerrero Llobet, S., & TeRiele, H. (2020). The RECQL helicase prevents replication fork collapse during replication stress. *Life science alliance*, 3(10).
5. Bravi F, Decarli A, Russo AG. (2018). Risk factors for breast cancer in a cohort of mammographic screening program: a nested case-control study within the FRiCaM study. *Cancer Med.*, 7(5):2145–2152
6. Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical biochemistry*, 162(1), 156-159.
7. Cicconi, A., Rai, R., Xiong, X., Broton, C., Al-Hiyasat, A., Hu, C., Dong, S., Sun, W., Garbarino, J., Bindra, R. S., Schildkraut, C., Chen, Y., & Chang, S. (2020). Microcephalin 1/BRIT1-TRF2 interaction promotes telomere replication and repair, linking telomere dysfunction to primary microcephaly. *Nature communications*, 11(1), 5861
8. Cobain EF, Milliron KJ, Merajver SD. (2016). Updates on breast cancer genetics: Clinical implications of detecting syndromes of inherited increased susceptibility to breast cancer. *Semin Oncol.*, 43(5):528–535.
9. Cybulski, C., Carrot-Zhang, J., Kluźniak, W., Rivera, B., Kashyap, A., Akbari, M. R. (2015). Germline RECQL mutations are associated with breast cancer susceptibility. *Nature*

*genetics*, 47(6), 643–646.

10. Debnath, S., & Sharma, S. (2020). RECQ1 Helicase in Genomic Stability and Cancer. *Genes*, 11(6), 622.

11. Deng, C.-X. (2006). BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Research*, 34(5), 1416–1426.

12. Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., ... & Futreal, P. A. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic acids research*, 39, D945–D950.

13. Frank C, Fallah M, Ji J & cols. (2014) The population impact of familial cancer, a major cause of cancer. *Int J Cancer* 1348:1899–1906

14. Giordano SH, Buzdar AU, Hortobagyi GN. (2002). Breast Cancer in Men. *Ann Intern Med.*, 137(8):678–687.

15. Global Cancer Observatory. New Global Cancer Data: GLOBOCAN 2018 | UICC. 2018.

16. Gracia-Aznarez, F. J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., ... & Benitez, J. (2013). Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. *PloSone*, 8(2), e55681.

17. Jara, L. & cols. 2017. Mutations in BRCA1, BRCA2 and other breast and ovarian cancer susceptibility genes in Central and South American populations. *Biological research*, 50.

18. Jo, Y. H., Kim, H. O., Lee, J., Lee, S. S., Cho, C. H., Kang, I. S., Choe, W. J., Baik, H. H., & Yoon, K. S. (2013). MCPH1 protein expression and polymorphisms are associated with risk of breast cancer. *Gene*, 517(2), 184–190.

19. Kim Y, Yoo K-Y, Goodman MT. (2015). Differences in Incidence, Mortality and Survival of Breast Cancer by Regions and Countries in Asia and Contributing Factors. *Asian Pac J Cancer Prev.*, 16(7):2857–2870

20. Kwong, A., Shin, V. Y., Cheuk, I. W., Chen, J., Au, C. H., Ho, D. N., ... & Narod, S. A.

(2016). Germline RECQL mutations in high risk Chinese breast cancer patients. *Breast cancer research and treatment*, 157(2), 211-215.

**21.** Lin, S. Y., Rai, R., Li, K., Xu, Z. X., & Elledge, S. J. (2005). BRIT1/MCPH1 is a DNA damage responsive protein that regulates the Brca1–Chk1 pathway, implicating checkpoint dysfunction in microcephaly. *Proceedings of the National Academy of Sciences*, 102(42), 15105-15109.

**22.** Marta Prieto, M. (2011). Epidemiología del cáncer de mama en Chile. *Revista Médica Clínica Las Condes*, 22(4), 428–435.

**23.** Mantere, T., Winqvist, R., Kauppila, S., Grip, M., Jukkola-Vuorinen, A., & Pylkäs, K. (2016). Targeted Next-Generation Sequencing Identifies a Recurrent Mutation in MCPH1 Associating with Hereditary Breast Cancer Susceptibility. *PLoS genetics*, 12(1).

**24.** McDonald, E. S., Clark, A. S., Tchou, J., Zhang, P., & Freedman, G. M. (2016). Clinical Diagnosis and Management of Breast Cancer. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine*, 57 Suppl 1, 9S–16S.

**25.** Meza-Junco, J., Montaña-Loza, A., & Aguayo-González, A. (2006). Bases moleculares del cáncer. *Revista de investigación clínica; órgano del Hospital de Enfermedades de la Nutrición*, 58(1), 56–70

**26.** Peng, G., Yim, E. K., Dai, H., Jackson, A. P., Burgt, I.v, Pan, M. R., Hu, R., Li, K., & Lin, S. Y. (2009). BRIT1/MCPH1 links chromatin remodelling to DNA damage response. *Nature cell biology*, 11(7), 865–872.

**27.** Phan, L., Jin, Y., Zhang, H., Qiang, W., Shekhtman, E., Shao, D., ... & Kattman, B. L. (2020). ALFA: allele frequency aggregator. *National Center for Biotechnology Information, US National Library of Medicine*, 10.

**28.** Pike, A. C., Shrestha, B., Popuri, V., Burgess-Brown, N., Muzzolini, L., Costantini, S., ... & Gileadi, O. (2009). Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proceedings of the National Academy of Sciences*, 106(4), 1039-1044. 30.



- 29.** Popuri, V., Croteau, D. L., Brosh, Jr, R. M., & Bohr, V. A. (2012). RECQ1 is required for cellular resistance to replication stress and catalyzes strand exchange on stalled replication fork structures. *Cell Cycle*, 11(22), 4252-4265.
- 30.** Pulvers, J. N., Journiac, N., Arai, Y., & Nardelli, J. (2015). MCPH1: a window into brain development and evolution. *Frontiers in cellular neuroscience*, 9, 92.
- 31.** Puranam, K. L., & Blackshear, P. J. (1994). Cloning and characterization of RECQL, a potential human homologue of the Escherichia coli DNA helicase RecQ. *Journal of Biological Chemistry*, 269(47), 29838-29845.
- 32.** Rai, R., Phadnis, A., Haralkar, S., Badwe, R. A., Dai, H., Li, K., & Lin, S. Y. (2008). Differential regulation of centrosome integrity by DNA damage response proteins. *Cell cycle*, 7(14), 2225–2233.
- 33.** 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.
- 34.** Sánchez, H., Paul, M. W., Grosbart, M., van Rossum-Fikkert, S. E., Lebbink, J. H., Kanaar, R., & Wyman, C. (2017). Architectural plasticity of human BRCA2–RAD51 complexes in DNA break repair. *Nucleic acids research*, 45(8), 4507-4518.
- 35.** Sami, F., & Sharma, S. (2013). Probing genome maintenance functions of human RECQ1. *Computational and structural biotechnology journal*, 6(7), e201303014.
- 36.** Sharma, S., Stumpo, D. J., Balajee, A. S., Bock, C. B., Lansdorp, P. M., Brosh Jr, R. M., & Blackshear, P. J. (2007). RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability. *Molecular and cellular biology*, 27(5), 1784-1794.
- 37.** Sharma, S., & Brosh Jr, R. M. (2007). Human RECQ1 is a DNA damage responsive protein required for genotoxic stress resistance and suppression of sister chromatid exchanges. *PloS one*, 2(12), e1297.

- 38.** Singh, N., Basnet, H., Wiltshire, T. D., Mohammad, D. H., Thompson, J. R., Héroux, A., & Mer, G. (2012). Dual recognition of phosphoserine and phosphotyrosine in histone variant H2A. X by DNA damage response protein MCPH1. *Proceedings of the National Academy of Sciences*, *109*(36), 14381-14386.
- 39.** Singh, N., Wiltshire, T. D., Thompson, J. R., Mer, G., & Couch, F. J. (2012). Molecular basis for the association of microcephalin (MCPH1) protein with the cell division cycle protein 27 (Cdc27) subunit of the anaphase-promoting complex. *Journal of Biological Chemistry*, *287*(4), 2854-2862.
- 40.** Singletary, S. E. (2003). Rating the Risk Factors for Breast Cancer. *Annals of Surgery*, *237*(4), 474–482
- 41.** Stratton, M. R., & Rahman, N. (2008). The emerging landscape of breast cancer susceptibility. *Nature genetics*, *40*(1), 17-22.
- 42.** Sun, J., Wang, Y., Xia, Y., Xu, Y., Ouyang, T., Li, J., ... & Xie, Y. (2015). Mutations in RECQL gene are associated with predisposition to breast cancer. *PLoS genetics*, *11*(5), e1005228.
- 43.** Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians*, *71*(3), 209–249.
- 44.** Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular biology and evolution*, *38*(7), 3022–3027.
- 45.** Torabi Dalivandan, S., Plummer, J., & Gayther, S. A. (2021). Risks and Function of Breast Cancer Susceptibility Alleles. *Cancers*, *13*(16), 3953
- 46.** Thakur P, Seam RK, Gupta MK, Gupta M, Sharma M, Fotedar V. (2017). Breast cancer risk factor evaluation in a Western Himalayan state: A case–control study and comparison with the Western World. *South Asian J Cancer.*, *6*(3):106–109.

47. Venkatesh, T., Nagashri, M. N., Swamy, S. S., Mohiyuddin, S. M. A., Gopinath, K. S., & Kumar, A. (2013). Primary Microcephaly Gene MCPH1 Shows Signatures of Tumor Suppressors and Is Regulated by miR-27a in Oral Squamous Cell Carcinoma. *PLoS ONE*, 8(3).
48. Xu, X., Lee, J., & Stern, D. F. (2004). Microcephalin Is a DNA Damage Response Protein Involved in Regulation of CHK1 and BRCA1\*♦. *Journal of Biological Chemistry*, 279(33), 34091-34094.
49. Zaqout, S., Morris-Rosendahl, D., & Kaindl, A. M. (2017). Autosomal recessive primary microcephaly (MCPH): an update. *Neuropediatrics*, 48(03), 135-142.
50. Zavala, V. A., Serrano-Gomez, S. J., Dutil, J., & Fejerman, L. (2019). Genetic epidemiology of breast cancer in Latin America. *Genes*, 10(2), 153.
51. Zhou, Z. W., Tapias, A., Bruhn, C., Gruber, R., Sukchev, M., & Wang, Z. Q. (2013). DNA damage response in microcephaly development of MCPH1 mouse model. *DNA repair*, 12(8), 645-655.