

Exploración de las Mutaciones del Gen *TP53* en Células Tumorales de Mama: Evaluación de su Influencia en la Diversidad de Tipos de Cáncer Mamario a través de Análisis Bioinformático con RStudio.



UNIVERSIDAD CATÓLICA DE CÓRDOBA
Facultad de Ingeniería
Licenciatura en Bioinformática

Estudiante

Zuliani Favro Francina:
E-mail: 2002300@ucc.edu.ar

Directores de Trabajo Final

Prof. Dra. Ximena Volpini:
E-mail: ximenavolpini@unc.edu.ar
Bioq. Juan Nahuel Quiroz:
E-mail: juan.quiroz@mi.unc.edu.ar

Docente De Trabajo Final

Prof. Dra. Ximena Volpini:
E-mail: ximenavolpini@unc.edu.ar

Agradecimientos

Quiero expresar mis agradecimientos a mi directora de tesis y de carrera, Dra. Ximena Volpini, por acompañarme en este momento, por guiarme no solo con este trabajo final, sino a lo largo de la facultad y en mi vida personal.

A mi director de tesis Bioq. Juan Nahuel Quiroz por regalarme parte de su tiempo y guiarme en este proceso.

A la Universidad Católica de Córdoba y a la Facultad de Ingeniería por brindar un espacio de estudio de excelencia, tanto en infraestructura y equipamiento como en calidez humana.

A mis papás por brindarme la posibilidad de poder estudiar. También por siempre apoyarme, no solo en la facultad, sino también en mi vida. Sin su apoyo no se si sería posible que hoy me encuentre en el lugar que estoy.

También quiero agradecer a mis hermanas por permitirme tenerme como ejemplo para sus vidas y así como yo les enseñó a ellas, también aprendo de ellas.

Por último pero no menos importante, a mis amigos con los que compartí muchas tardes de estudio, charlas y momentos especiales y son quienes me apoyan en todo momento, y me enseñaron a no rendirme.

ÍNDICE

RESUMEN.	1.....	6
ABSTRACT.	2.....	7
ACRÓNIMOS.....		9
1.0 INTRODUCCIÓN.....		10
MARCO TEÓRICO.....		10
1.1 La Célula.....		10
1.2 ADN y ARN.....		12
1.3 Ciclo Celular.....		15
1.4 Genes.....		16
1.5 Gen Tp53.....		18
1.6 Mutaciones.....		19
1.7 Mutaciones en el gen TP53.....		21
1.8 Cáncer.....		22
1.8.1 Tipos de cáncer.....		23
1.9 Cáncer de mama.....		25
1.10 Tipos de cáncer de mama.....		28
1.11 El gen TP53 en el cáncer de mama.....		30
1.12 Perfil Tumoral Molecular.....		32
1.12.1 Receptor HER2.....		32
1.12.2 Receptor de Estrógeno (ER).....		33
1.12.3 Receptor de Progesterona (PR).....		33
1.12.4 Terapias.....		34
1.13 Perfiles Tumorales Moleculares del Cáncer de mama.....		35
2.0 ESTADO DEL ARTE.....		37
3.0 HIPÓTESIS.....		38
4.0 OBJETIVOS ESPECÍFICOS DE INVESTIGACIÓN.....		38
5.0 MATERIALES Y MÉTODOS.....		39
6.0 RESULTADOS.....		40
6.1 Predominancia en la mutación TP53 y características de los pacientes.....		40
6.2 Mutaciones concurrentes de TP53.....		45
6.3 Mutaciones en CDI y CLI.....		49
6.4 Perfil Tumoral Molecular.....		54
6.4.1 Análisis del receptor HER2.....		54
6.4.2 Análisis del receptor de Estrógeno (ER).....		56
6.4.3 Análisis del receptor de Progesterona (PR).....		58
6.5 Análisis de los Perfiles Tumorales Moleculares del Cáncer de mama.....		60
6.6 Perfiles Tumorales Moleculares en CDI.....		62
6.7 Perfiles Tumorales Moleculares en CLI.....		64
6.8 Perfiles Tumorales Moleculares CDI vs. CLI.....		66
6.8.1 Frecuencia de las Mutaciones en los Distintos Tipos de Perfiles Moleculares.....		68

6.9 Análisis de la Frecuencia de Hotspots en Mutaciones Missense.....	74
7.0 ÍTEMES DESTACADOS DE LA INVESTIGACIÓN.....	78
8.0 CONCLUSIONES.....	79
9.0 DISCUSIÓN.....	83
PALABRAS DE CIERRE.....	87
10.0 GLOSARIO.....	88
11.0 REFERENCIAS BIBLIOGRÁFICAS.....	90

RESUMEN.1

Los tumores pueden originarse debido a mutaciones en el genoma de las células del cuerpo humano. El término cáncer engloba un grupo numeroso de enfermedades que se caracterizan por el desarrollo de células tumorales, que tienen alta capacidad de proliferar e invadir tejidos adyacentes, así como también lejanos al sitio de origen generando metástasis. En condiciones normales, las células tienen un ciclo de vida regulado, donde se dividen y mueren siguiendo un proceso programado. Sin embargo, en el desarrollo del cáncer, este equilibrio se ve alterado. Si bien no todas las células se encuentran en constante división, las mutaciones, inducidas por factores externos como la radiación o infecciones virales, pueden desencadenar la transformación de células no proliferativas en células tumorales, lo que lleva a la formación de tumores y la pérdida de control sobre el ciclo celular. Las células con alta tasa de proliferación y baja de muerte forman “tumores”, que en su expansión pueden destruir y desplazar a los tejidos normales.

Un gen fundamental para la supresión de tumores y regulación de la división celular es el gen *TP53*, el cual es conocido como un gen supresor tumoral. En respuesta a estímulos que indican daño celular, la proteína del gen *TP53* llamada p53, tiene la capacidad de participar en la preservación de la integridad genómica. Estas medidas abarcan desde el proceso de muerte celular programada hasta la detención temporal del ciclo celular para permitir la reparación del daño hasta inducir la apoptosis de la célula dañada.

Cuando el gen *TP53* presenta mutaciones, su función se ve comprometida, lo que puede tener graves consecuencias para la salud. Si el gen *TP53* está mutado, las células pueden perder su capacidad para controlar el crecimiento y la división celular. Como resultado, las células tumorales pueden proliferar de manera descontrolada dando origen al proceso neoplásico. Es decir, puede diseminarse a otros tejidos y órganos del cuerpo. Esta capacidad de multiplicación y diseminación desenfrenada es el acontecimiento que subyace al desarrollo y progresión del cáncer. Las mutaciones en *TP53* se asocian con una variedad de condiciones genéticas y pueden asociarse a un mayor riesgo de desarrollar múltiples tipos de cáncer a lo largo de la vida del paciente. El gen *TP53* es uno de los genes más frecuentemente mutados en personas con cáncer. Más del 50% de los casos de cáncer involucran mutaciones en este gen, lo que puede resultar en la pérdida de su función o daño en su estructura.

El cáncer de mama y el sarcoma de tejidos blandos y óseos son dos de los tipos de cáncer más prevalentes entre los portadores de mutaciones en el gen *TP53*, representando

conjuntamente más del 50% de todos los tumores observados con esta mutación entre los pacientes.

El cáncer de mama, como se mencionó anteriormente, es especialmente relevante en personas con tumores asociados a las mutaciones del gen *TP53*. Dentro de los subtipos de cáncer de mama, se ha observado una mayor incidencia de Carcinoma Ductal Infiltrante (CDI) y Carcinoma Lobulillar Infiltrante (CLI) en portadores de estas mutaciones. Estos dos subtipos también son los más comunes en la población general.

En el análisis que se llevará a cabo en este trabajo, se comparará si las células tumorales de mama pertenecientes a los tipos de cáncer de mama Carcinoma Ductal Infiltrante (CDI) y Carcinoma Lobulillar Infiltrante (CLI), presentan mutaciones similares o diferentes en el gen *TP53*.

PALABRAS CLAVES: Cáncer de mama, Carcinoma Ductal Infiltrante, Carcinoma Lobulillar Infiltrante, *TP53*, Mutaciones.

ABSTRACT.2

Tumors can originate from the malfunction of cells that make up the human body. The term cancer encompasses a large group of diseases characterized by the development of tumor cells, which divide, grow, and spread uncontrollably. Normal cells divide to form new cells and die over a programmed period. However, in cancer, this orderly process is disrupted; the tumor cell loses the ability to die and divides almost without limit. Such multiplication in the number of cells forms "tumors," which can expand to destroy and replace normal tissues.

A gene that is fundamental for tumor suppression and regulation of cell division is the *TP53* gene, which is known as a tumor suppressor gene. In response to stimuli indicating cellular damage, the protein of the *TP53* gene, called p53, has the ability to take measures to preserve genomic integrity. These measures can range from initiating programmed cell death or temporarily halting the cell cycle to allow for damage repair, to inducing apoptosis of the damaged cell.

When the *TP53* gene presents mutations, its function is compromised, which can have serious health consequences. If the *TP53* gene is mutated, cells may lose their ability to control growth and cell division. As a result, tumor cells can proliferate uncontrollably and even metastasize, meaning they can spread to other tissues and organs in the body. This capacity for uncontrolled multiplication and dissemination underlies the development and progression of cancer. Mutations in *TP53* are associated with a variety of genetic conditions and can leave individuals at greater risk of developing multiple types of cancer throughout their lives.

The *TP53* gene is one of the most frequently mutated genes in people with cancer. More than 50% of cancer cases involve mutations in this gene, which can result in loss of function or structural damage. These mutations can contribute to cancer development and progression by allowing uncontrolled cell growth and evading normal mechanisms of repair and apoptosis.

Breast cancer and soft tissue and bone sarcomas are two of the most prevalent types of cancer among carriers of mutations in the *TP53* gene, jointly representing more than 50% of all tumors observed in this group of patients.

Breast cancer, as mentioned earlier, is particularly significant in individuals with tumors associated with *TP53* gene mutations. Among the subtypes of breast cancer, there is a higher incidence of Infiltrating Ductal Carcinoma (IDC) and Infiltrating Lobular Carcinoma (ILC) in carriers of these mutations. These two subtypes are the most common in the general population and also prevail among patients with mutations in the *TP53* gene.

In the analysis that will be conducted in this work, we will determine whether breast tumor cells, belonging to the types of breast cancer Infiltrating Ductal Carcinoma (IDC) and Infiltrating Lobular Carcinoma (ILC), present different mutations in the *TP53* gene.

KEY WORDS: *Breast Cancer, Infiltrating Ductal Carcinoma, Infiltrating Lobular Carcinoma, TP53, Mutations.*

ACRÓNIMOS

- **ADN** - Ácido desoxirribonucleico
- **ARN** - Ácido ribonucleico
- **ARNm** - Ácido ribonucleico mensajero
- **ARNr** - Ácido ribonucleico ribosómico
- **ARNt** - Ácido ribonucleico de transferencia
- **CDI** - Carcinoma Ductal Infiltrante
- **CLI** - Carcinoma Lobular Infiltrante
- **ER (Estrogen Receptor)** - Receptor de Estrógeno
- **FISH (Fluorescence In Situ Hybridization)** - Hibridación Fluorescente In Situ
- **HER2 (Human Epidermal growth factor Receptor 2)** - Receptor 2 del Factor de Crecimiento Epidérmico Humano
- **IHC (Immunohistochemistry)** - Inmunohistoquímica
- **p53** - Proteína 53
- **PR (Progesterone Receptor)** - Receptor de Progesterona
- **TN** - Triple Negativo
- **TP53 (Tumor Protein 53)** - Proteína Tumoral 53

Acrónimos – Nucleótidos

- A** - Adenina
C - Citosina
G - Guanina
T - Timina
U - Uracilo

1.0 INTRODUCCIÓN

MARCO TEÓRICO

1.1 La Célula

La célula constituye la unidad fundamental que se encuentra en todos los seres vivos, dando origen a la totalidad de los organismos y conformando los diversos tejidos presentes en el cuerpo. En el presente trabajo nos enfocaremos en el estudio de las células eucariotas.

Las tres partes principales de la célula son la membrana celular, el núcleo y el citoplasma. La membrana celular o plasmática, se encuentra en todas las células y separa el interior de la célula del ambiente exterior. La misma se compone de una bicapa lipídica que es semipermeable. Su función se basa en regular el transporte de moléculas que entran y salen de la célula.²⁰ El citoplasma es el espacio comprendido entre la membrana plasmática y el núcleo celular cuya función es albergar los orgánulos celulares y contribuir al movimiento de estos.²¹

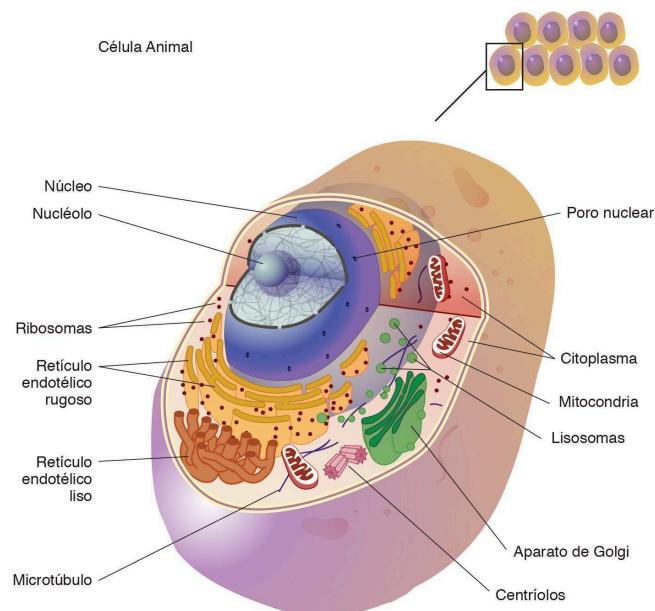


Figura 1: Partes de la célula eucariota.²²

Dentro de la célula se encuentra el núcleo, el cual contiene el nucléolo y la mayor parte del ADN (ácido desoxirribonucleico) celular y es donde se produce la mayor cantidad de ARN. El núcleo es el orgánulo rodeado por membrana en el interior de la célula, que contiene los cromosomas que codifican el material genético. El núcleo contiene una matriz de orificios o

poros en la membrana nuclear, que permite el pasaje selectivo de determinadas moléculas (como las proteínas y los ácidos nucleicos) hacia el interior o el exterior del mismo.²³

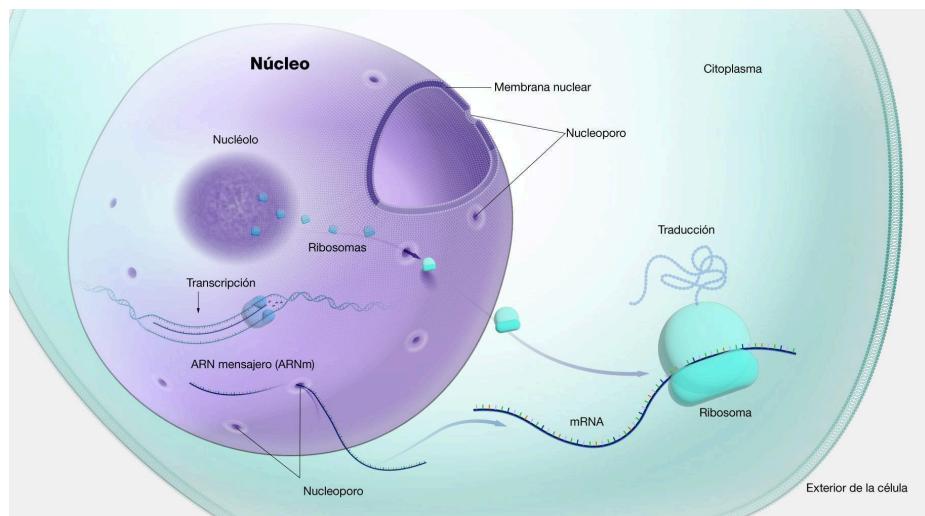


Figura 2: Partes del núcleo, célula eucariota.²³

El ADN de un organismo está organizado en uno o más cromosomas, los cuales son una cadena o un aro de ADN. Un solo cromosoma contiene genes diferentes. En las células eucariotas, los cromosomas son estructuras lineales (cadenas). Cada especie eucariota tiene un número específico de cromosomas en los núcleos de las células de su cuerpo.

Los cromosomas son visibles como estructuras distintivas cuando la célula está preparada para dividirse. Durante las fases de crecimiento y mantenimiento del ciclo celular, los cromosomas adoptan una apariencia de hilos enredados. En este estado, el ADN se encuentra accesible para las enzimas encargadas de su transcripción a ARN, lo que permite la expresión de la información genética contenida.

Ya sea en su forma suelta o compacta, las cadenas de ADN de los cromosomas están unidas a proteínas estructurales, entre ellas una familia de proteínas llamadas histonas.

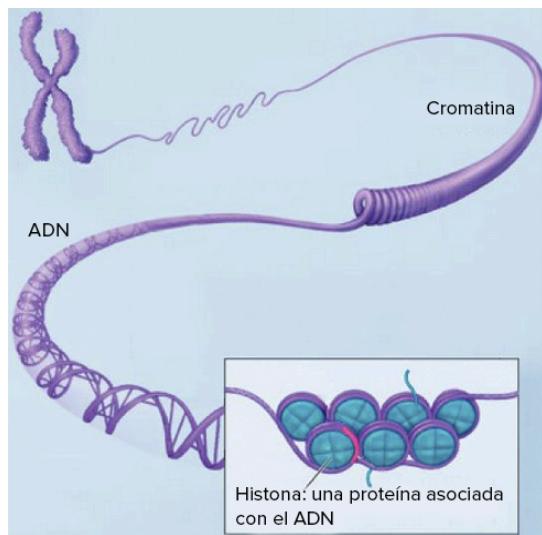


Figura 3: Cromosoma en su estado suelto y compacto.²⁴

Estas proteínas asociadas al ADN lo organizan y compactan para que pueda caber en el núcleo, y también ayudan a determinar qué genes están activos o inactivos. El complejo formado por el ADN y sus proteínas estructurales de soporte se denomina cromatina.²⁵

1.2 ADN y ARN

El ácido desoxirribonucleico (ADN) es la molécula que transporta información genética para el desarrollo y el funcionamiento de un organismo. Está compuesto por dos cadenas complementarias que se enrollan entre sí a la cual se le conoce como doble hélice. Cada hebra tiene una estructura principal compuesta por grupos alternados de azúcar (desoxirribosa) y fosfato. Unida a cada azúcar hay una de cuatro bases: Adenina (A), Citosina (C), Guanina (G) o Timina (T). Las dos hebras del ADN se unen mediante enlaces químicos formados por puentes de hidrógeno, donde la Adenina se empareja con la Timina y la Citosina con la Guanina. La secuencia de las bases a lo largo de la estructura principal del ADN codifica información biológica, por ejemplo, las instrucciones para producir una proteína o molécula de ARN.²⁶

El ácido ribonucleico (ARN) es un ácido presente en todas las células vivas que tiene similitudes estructurales con el ADN. Sin embargo, a diferencia del ADN, es más frecuente que el ARN esté formado por una única cadena. Una molécula de ARN tiene un eje formado por grupos fosfato alternantes y el azúcar ribosa, en lugar de la desoxirribosa del ADN. Unida a cada azúcar hay una de cuatro bases: Adenina (A), Uracilo (U), Citosina (C) o Guanina (G). Existen diferentes tipos principales de ARN en las células: ARN mensajero (ARNm), ARN ribosómico (ARNr) y ARN de transferencia (ARNt). El ARNm transporta la información genética desde el ADN en el núcleo hasta los ribosomas en el citoplasma, donde se utiliza para sintetizar proteínas durante el proceso de traducción. El ARNr forma parte de la estructura de

los ribosomas, donde desempeña un papel fundamental en la síntesis de proteínas al catalizar la unión de aminoácidos durante la traducción del ARNm. Y el ARNt transporta aminoácidos específicos hacia los ribosomas durante la síntesis de proteínas. El ARNt reconoce los codones del ARNm y lleva el aminoácido correspondiente para su incorporación en la cadena polipeptídica en crecimiento.²⁷

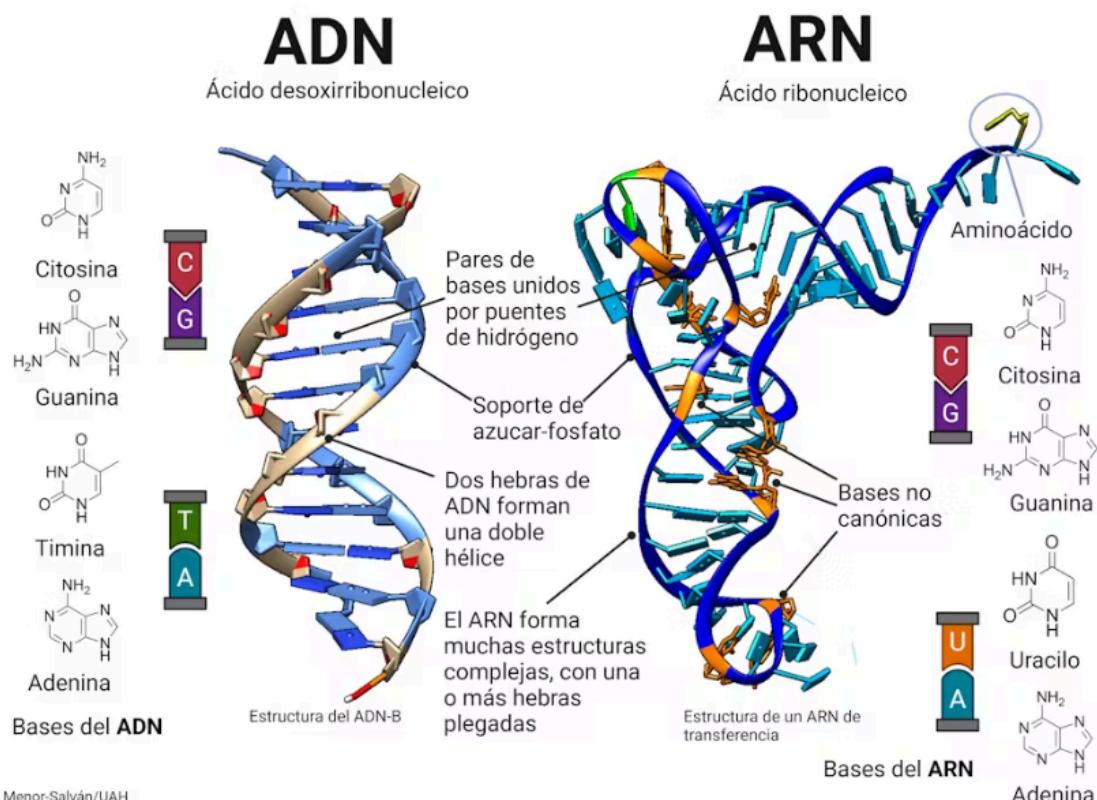


Figura 4: Estructura del ADN y ARN de la célula.²⁸

La transcripción y la traducción son procesos fundamentales en la expresión génica, que permiten a las células sintetizar proteínas a partir de la información contenida en el ADN. La transcripción implica la síntesis de ARN mensajero a partir de una secuencia específica de ADN, mientras que la traducción convierte el ARNm en una secuencia de aminoácidos para formar una proteína funcional. Estos procesos son esenciales para el funcionamiento adecuado de las células y juegan un papel importante en múltiples procesos biológicos.

La transcripción es el proceso mediante el cual una región específica del ADN se copia en una molécula complementaria de ARN. Este proceso es catalizado por la enzima ARN polimerasa, que se une a la región promotora del ADN y desenrolla la doble hélice para exponer la secuencia de nucleótidos. A continuación, la ARN polimerasa sintetiza una cadena de ARN complementaria utilizando una de las hebras de ADN como molde. Durante la transcripción, la secuencia de T en el ADN es reemplazada por Uracilo (U) en el ARN. Una

vez completada la transcripción, el ARN recién sintetizado, conocido como ARNm, se somete a procesamiento, que incluye la eliminación de intrones y la adición de una cola de poliadenina en el extremo 3' para estabilizar la molécula.²⁹

La traducción es el proceso mediante el cual la secuencia de nucleótidos en el ARNm se traduce en una secuencia de aminoácidos para formar una proteína. Este proceso tiene lugar en los ribosomas, que son complejos estructurales compuestos por ARN ribosomal (ARNr) y proteínas. La traducción se lleva a cabo por el ARN de transferencia (ARNt) y los ribosomas, que reconocen y ensamblan los aminoácidos en la secuencia correcta según el código genético. Cada triplete de nucleótidos en el ARNm, llamado codón, especifica un aminoácido particular según el código genético. Los ARNt transportan los aminoácidos correspondientes al codón específico en el ARNm, y los ribosomas catalizan la formación de enlaces peptídicos entre los aminoácidos para formar una cadena polipeptídica. Este proceso continúa hasta que se alcanza un codón de terminación en el ARNm, momento en el cual la síntesis de la proteína se detiene y la cadena polipeptídica se libera del ribosoma.³⁰

La transcripción y la traducción son procesos esenciales para la expresión génica y la síntesis de proteínas, que son fundamentales para el funcionamiento adecuado de las células. Estos procesos son regulados con precisión y están sujetos a una variedad de mecanismos de control que garantizan la producción de proteínas en el momento y lugar adecuados. Las proteínas desempeñan una variedad de funciones en la célula, incluyendo la estructura celular, el transporte de moléculas, la catálisis de reacciones químicas y la transmisión de señales. Por lo tanto, la transcripción y la traducción son procesos críticos que sustentan la vida celular y son objeto de intensa investigación en biología molecular y genética.

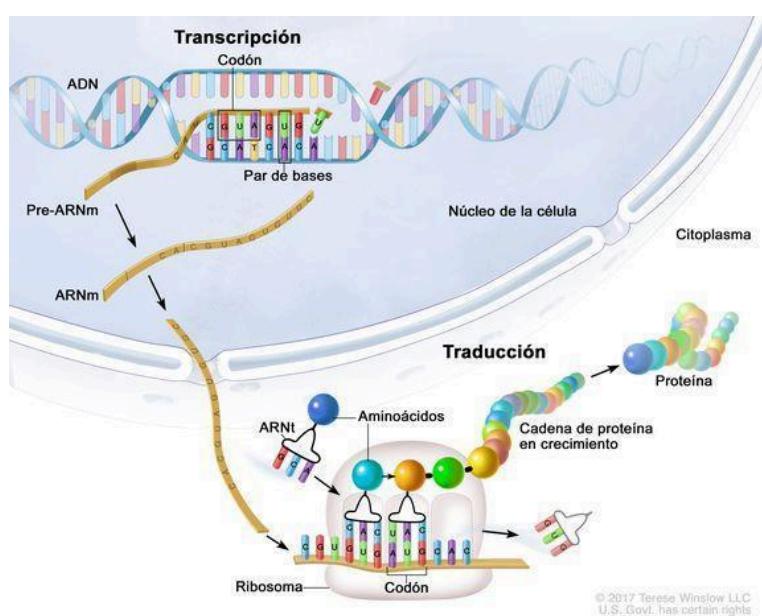


Figura 5: Proceso de transcripción y traducción de una célula.³¹

1.3 Ciclo Celular

El ciclo celular se compone de dos fases, la primera llamada interfase, donde la célula pasa la mayor parte de su tiempo y la segunda fase llamada división celular, donde la célula se reproduce.

La interfase está dividida a su vez en distintas subfases. En la fase G1 la célula comienza a sintetizar proteínas por lo que su tamaño aumenta y se vuelve más funcional. Aquí se producen las proteínas reguladoras del ciclo celular ciclina D y la ciclina E. Luego la célula puede dirigirse a la fase S o desviarse a la fase G0. La fase G0 es un estado de reposo. La fase S es donde la célula sintetiza el doble del ADN, es decir, el ADN se duplica ya que luego se va a dividir en dos células. En esta fase la ciclina E se une con la CDK2 (cinasa dependiente de ciclina 2), y la ciclina D se une con la CDK4 (cinasa dependiente de ciclina 4). La cinasa es una enzima que cataliza la transferencia de grupos fosfato en este caso a las proteínas. Estas dos moléculas posteriormente activarán la fosforilación de la proteína RB, cuya función principal es regular el ciclo celular. Esta proteína, inicialmente, suprime la replicación del ADN hasta que la célula esté en condiciones óptimas, momento en el cual se le adiciona un grupo fosfato para inhibir su actividad.^{32 33}

En la fase G2, la célula se prepara activamente para la división celular mediante la síntesis de proteínas y la duplicación de organelas. La duplicación del ADN no es perfecta, tiene errores, pero se detectan y corrigen en esta fase. Así, en esta etapa ocurre la reparación del ADN. Para controlar la duplicación del ADN participa la enzima CDK1 (cinasa dependiente de ciclina 1) unida a la ciclina B, la cual activa el proceso de mitosis. En esta fase también participa la proteína p53 la cual, a su vez, produce proteína p21 quien inhibe a la CDK1 para bloquear o frenar el proceso de mitosis y la replicación del ADN. Si el ADN no logra recuperarse y la célula se encuentra dañada, se activa el proceso de apoptosis, donde la célula se autoprograma para morir.^{32 33}

Finalmente, si la célula no sufrió ninguna alteración, pasa a la fase de mitosis (M) o fase de división celular. La célula reparte las dos copias de su material genético entre sus dos células hijas. Después de haber completado la fase M, se obtienen dos células y el ciclo celular comienza para cada una de ellas.

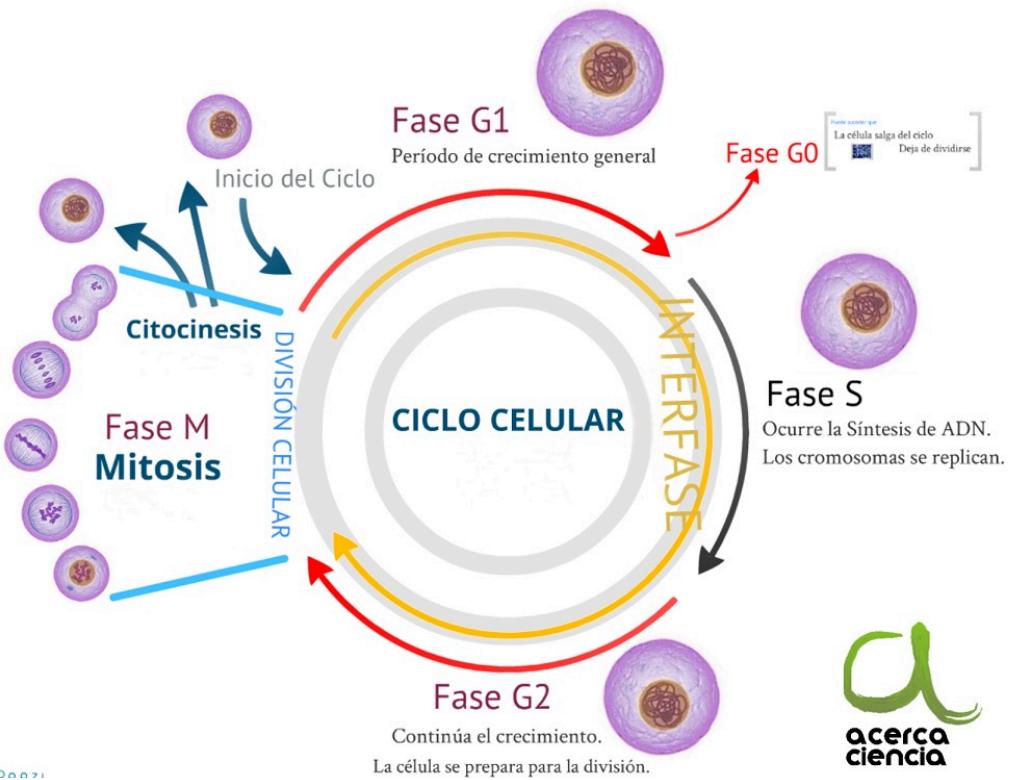


Figura 6: Ciclo celular de una célula.³³

1.4 Genes

Los genes son la unidad básica de la herencia. Se transmiten de los progenitores a la descendencia y contienen la información necesaria para especificar los rasgos físicos y biológicos. La mayoría de los genes codifican para proteínas específicas, o segmentos de proteínas, que tienen diferentes funciones en el cuerpo. Los seres humanos tienen aproximadamente 20.000 genes que codifican para proteínas.³⁴

Cuando un gen se activa, se transcribe en una molécula de ARN mensajero (ARNm), que luego se traduce en una proteína específica durante un proceso llamado síntesis de proteínas. Las proteínas son las moléculas biológicas que realizan la mayoría de las funciones celulares y son esenciales para el crecimiento, desarrollo y mantenimiento de los organismos. Cada gen codifica una característica o función específica en un organismo, como el color de ojos, la producción de enzimas, la respuesta a una señal externa, entre otras.³⁵

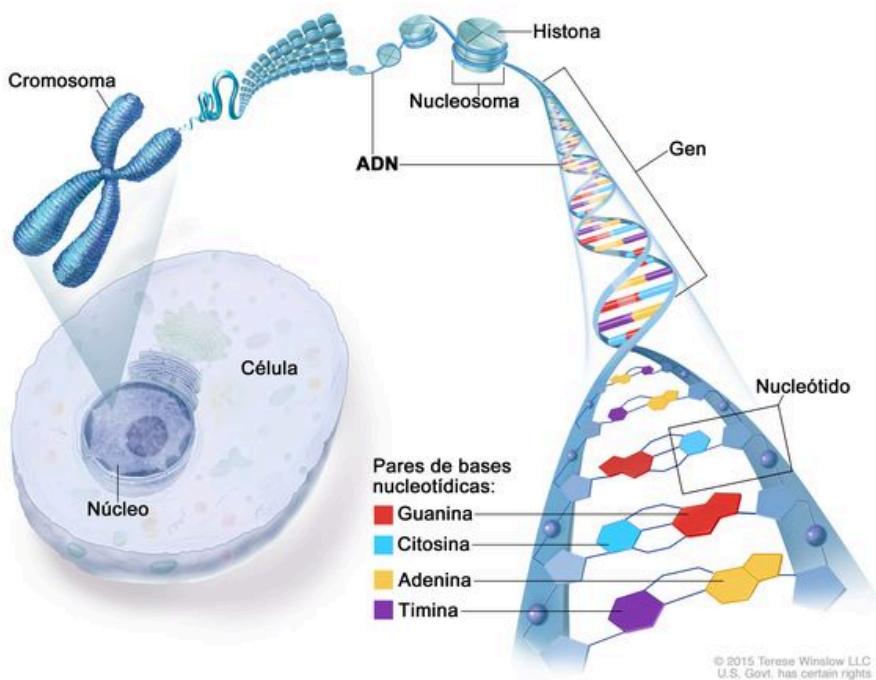


Figura 7: Estructura de un gen.³⁶

Existen dos grupos de genes encargados de regular la proliferación y crecimiento tumoral, los protooncogenes y los genes supresores de tumores. Los protooncogenes son genes que están presentes en el ADN de todos los individuos y su función es regular la proliferación o el crecimiento y la división celular. Cuando en la secuencia del ADN del gen ocurre un cambio en el código genético (independientemente de la cantidad de nucleótidos afectados), esto no permite que el gen sea leído de forma correcta lo que genera un suceso llamado variante genética patogénica o mutación. Si en una célula, un protooncogen sufre una mutación y se convierte en un gen permanentemente activo o sobreexpresado puede provocar una actividad celular anormal, como un crecimiento descontrolado o la proliferación celular sin regulación, lo que promueve el desarrollo tumoral. Cuando los protooncogenes mutan y producen cáncer es cuando se les llama oncogen. La mayoría de las mutaciones que dan origen al proceso neoplástico que involucran oncogenes son adquiridas, no heredadas.³⁷

También existen genes que regulan la división celular y que en condiciones normales la ralentizan. Estos genes son los llamados supresores de tumores. Estos genes tienen una función muy importante y es el control de las mutaciones del ADN. Cuando detectan una mutación en la célula puede que activen sistemas de reparación o activan sistemas para que la célula inicie su muerte y no prolifere.³⁸ De la misma manera que sucede con los oncogenes, si los genes supresores de tumores no funcionan adecuadamente o experimentan alguna alteración, las células pueden crecer sin control, lo que puede desencadenar un proceso tumoral. Cuando se detecta un fallo, como una mutación, los genes supresores de tumor evitan que la célula se divida demasiado rápido. Una diferencia importante entre los oncogenes y los genes supresores de

tumores es que los oncogenes resultan de la activación de protooncogenes, mientras que los genes supresores de tumores favorecen el desarrollo tumoral cuando se inactivan.

1.5 Gen *Tp53*

El gen *TP53* es un gen supresor de tumores ubicado específicamente en el cromosoma 17. La proteína codificada por este gen p53 es una fosfoproteína nuclear.

En las células sanas no estresadas, p53 tiene una vida media corta por su asociación a MDM2, un oncogén que también está regulado. MDM2 regula negativamente al gen *TP53* actuando como una ubiquitina ligasa, la cual reconoce el dominio de transactivación N-terminal de p53, y también como un inhibidor de la activación transcripcional.³⁹

Las señales que activan el gen *TP53* incluyen estímulos genotóxicos, como daño al ADN, estrés celular, hipoxia, desequilibrios en la homeostasis celular y la presencia de oncogenes activados. Cuando las células son expuestas a estímulos genotóxicos, los niveles de p53 se incrementan e interrumpe el ciclo celular (en la fase G1) para dar tiempo a su reparación, pero si el daño es excesivo para ser reparado, entonces se desencadena un proceso de muerte celular o apoptosis mediado por la regulación de la transcripción. La proteína p53 impide la transformación neoplásica siguiendo tres mecanismos entrelazados: activación de la retención temporal del ciclo celular (quiescencia), inducción de la retención permanente del ciclo celular (senescencia) o la activación de la muerte celular programada (apoptosis).

La retención del ciclo celular mediada por p53 puede ser considerada una de las respuestas fundamentales ante el daño del ADN. Se produce al final de la fase G1, principalmente como consecuencia de la transcripción del gen del CDKN1A (p21) dependiente de p53. La proteína p21 inhibe los complejos ciclina-CDK y evita la fosforilación de RB (proteína del retinoblastoma supresora de tumores que se encuentra alterada en muchos tipos de cáncer), deteniendo las células en la fase G1. Esta pausa es esencial para proceder a la reparación del daño del ADN. Además, la proteína p53 también induce la expresión de los genes de reparación del daño del ADN. Si tal daño se repara con éxito, p53 regula positivamente la transcripción de MDM2, lo que provoca la destrucción de p53 y alivia el bloqueo del ciclo celular. Si el daño no se puede reparar, la célula puede entrar en la senescencia inducida por p53 o sufrir apoptosis dirigida por esta proteína. La senescencia inducida por p53 es un arresto permanente del ciclo celular, caracterizado por alteraciones específicas de la morfología y de la expresión genética que la diferencian de la quiescencia y de la detención irreversible del ciclo celular. La senescencia requiere la activación de p53 y/o RB, así como la expresión de sus mediadores (CDKI).^{39 40}

La apoptosis inducida por p53 de las células con lesión irreversible del ADN es el último mecanismo protector que existe contra la transformación neoplásica, y está mediado por varios genes pro-apoptosis, como BAX y PUMA.³⁹

1.6 Mutaciones

Como se mencionó previamente, un gen está compuesto por cuatro bases nitrogenadas: Adenina, Timina, Citosina y Guanina, las cuales se ordenan de distintas maneras para formar distintos codones o tripletes. Los codones son agrupaciones de tres bases nitrogenadas que especifican qué aminoácido se ubicará en cada posición dentro de una proteína. A partir de esto, se define el concepto de marco de lectura.

Una mutación es un cambio o alteración en la secuencia de ADN de un organismo. El tamaño de las mutaciones se puede generar desde un solo nucleótido del ADN hasta un gran segmento de un cromosoma incluyendo varios genes. Las mutaciones pueden producirse a partir de errores en la replicación del ADN durante la división celular, la exposición a mutágenos (agente físico, químico o biológico que altera o cambia la información genética) o una infección viral. Además, existen mutaciones en la línea germinal que pueden transmitirse a la descendencia y también mutaciones somáticas.⁴¹

Al presentarse una mutación, esta puede alterar el marco de lectura de una secuencia, lo que podría modificar la secuencia de aminoácidos y dar lugar a una proteína distinta de la original.

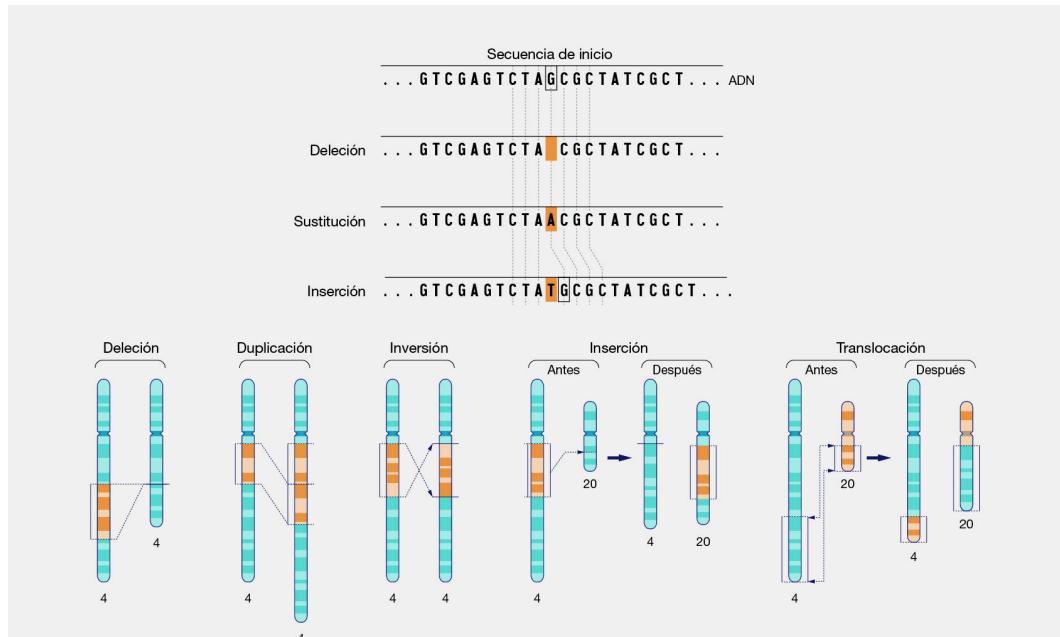


Figura 8: Tipos de mutaciones.⁴¹

Dentro de los distintos tipos de mutación que existen encontramos:

- Deleción: tipo de mutación que implica la pérdida de uno o más nucleótidos de un segmento de ADN.
 - *Frameshift deletion* (Delección con cambio de marco): En una delección con cambio de marco, una o más bases se eliminan del ADN, lo que cambia el marco de lectura del codón y puede llevar a la producción de una proteína completamente diferente o a la pérdida de función de la proteína. En ocasiones las delecciones son tan largas que pueden comprometer un gen entero o varios genes contiguos.
 - *In Frame* (Delección sin cambio de marco): En una delección sin cambio de marco, un número de bases divisible por tres se elimina del ADN, lo que no cambia el marco de lectura del codón y generalmente resulta en la pérdida o ganancia de un pequeño número de aminoácidos.⁴²
- Sustitución: es un tipo de mutación en la que un nucleótido es reemplazado por un nucleótido diferente. Esto puede incluir varios tipos de mutaciones, como las mutaciones de sentido erróneo (*missense mutations*), las mutaciones sin sentido (*nonsense mutations*), las mutaciones silenciosas (*silent mutations*), entre otras.
 - *Missense mutation* (Mutación con cambio de sentido): En una mutación con cambio de sentido, una sola base se cambia en el ADN, lo que resulta en la sustitución de un aminoácido en la proteína codificada por el gen. Dependiendo de la naturaleza del cambio de aminoácido, una mutación de este tipo puede tener efectos variables en la función de la proteína, desde cambios sutiles hasta una pérdida completa de función. Esta sustitución puede cambiar la función de la proteína producida por el gen, ya sea en forma positiva, negativa o neutral, dependiendo de la naturaleza del cambio de aminoácido y su ubicación en la proteína.
 - *Splice Site* (Sitio de Empalme): Una modificación genética en la secuencia de ADN que se produce en la frontera entre un exón y un intrón (sitio de empalme). Esta modificación puede interferir con el proceso de empalme del ARN, lo que resulta en la pérdida de exones o la inclusión de intrones, lo que conduce a una secuencia de codificación de proteínas alterada. También es conocida como variante del sitio de empalme.
 - *Nonsense Mutation* (Mutación Sin Sentido): En una mutación sin sentido, se produce en el ADN cuando un cambio de la secuencia da lugar a un codón de terminación en vez de un codón específico de un aminoácido. La presencia de un nuevo codón de terminación da lugar a la producción de una proteína más corta que probablemente no sea funcional.⁴²
- Inserción: es un tipo de mutación que implica la adición de uno o más nucleótidos en un segmento de ADN. De esta forma se puede alterar el marco de lectura para formar la proteína o insertar aminoácidos extra que son inadecuados.

- *Frame Shift* (Inserción con cambio de marco): En una inserción con cambio de marco, una o más bases se insertan en la secuencia de ADN, lo que desplaza el marco de lectura del codón y puede llevar a la producción de una proteína completamente diferente o a la pérdida de función de la proteína.⁴²

1.7 Mutaciones en el gen *TP53*

Ciertas mutaciones en el gen TP53 causan pérdida de función, evitando la reparación del ADN dañado y bloqueando la apoptosis mediada por p53, lo que permite la supervivencia y proliferación de células con daño genético

Las mutaciones en los genes supresores tumorales, como *TP53*, pueden reducir o desactivar la función de la proteína codificada. Las células que presentan mutaciones en *TP53* son más susceptibles a acumular mutaciones debido a defectos en la reparación del ADN, lo que aumenta el riesgo de cáncer. La importancia de la apoptosis en la prevención del desarrollo de cáncer se ve reflejada por el hecho de que la mutación en *TP53* es la anomalía genética más habitual en cánceres humanos.¹⁴⁰

El gen *TP53* puede presentar mutaciones por diversos factores, entre los cuáles se destaca:

- Exposición a agentes carcinógenos: La exposición a carcinógenos ambientales, como productos químicos, radiación ionizante (como la radiación ultravioleta del sol o la radiación ionizante de fuentes médicas o ambientales) y ciertos virus (como el virus del papiloma humano), puede causar daño al ADN y promover mutaciones en el gen *TP53*.
- Errores durante la replicación del ADN: Durante la replicación del ADN, pueden ocurrir errores que no se reparan adecuadamente, lo que puede llevar a la aparición de mutaciones en el gen *TP53*. Estos errores pueden ser causados por diversos factores, como la exposición a carcinógenos, la presencia de metabolitos tóxicos en las células, o defectos en los mecanismos de reparación del ADN.
- Estrés celular y señales de daño: El gen *TP53* es activado en respuesta a señales de estrés celular y daño en el ADN, como la presencia de radicales libres, especies reactivas de oxígeno, y otros factores que pueden causar daño en el ADN. Cuando estas señales de estrés son persistentes o excesivas, pueden desencadenar la activación continua del gen *TP53*, lo que aumenta el riesgo de mutaciones en el mismo.
- Predisposición genética: Algunas personas pueden tener una predisposición genética a desarrollar mutaciones en el gen *TP53*. Por ejemplo, ciertos síndromes de predisposición al cáncer, como el síndrome de Li-Fraumeni, están asociados con mutaciones hereditarias en el

gen *TP53*, lo que aumenta significativamente el riesgo de desarrollar varios tipos de cáncer a lo largo de la vida.

Cuando el gen *TP53* tiene una mutación que altera su función normal, el oncogén MDM2 puede llevar a cabo la ubiquitinación de *TP53* mutado. La ubiquitina es una pequeña proteína altamente conservada y que se encuentra ampliamente distribuida en todos los organismos eucariotas. Su función principal es la de regular la degradación de proteínas en las células. Este proceso es esencial para mantener la homeostasis celular y está implicado en la regulación de una amplia variedad de funciones biológicas y procesos celulares. Consiste en la adición covalente de moléculas de ubiquitina a otras proteínas, un proceso que es llevado a cabo por un conjunto de enzimas que trabajan en secuencia para activar, transportar y unir la ubiquitina a las proteínas diana. Una vez que una proteína ha sido marcada con ubiquitina, es reconocida y degradada por el proteasoma, un complejo proteico responsable de la degradación de proteínas en las células.³⁹

En condiciones normales, cuando el gen *TP53* detecta daño en el ADN o situaciones estresantes, se activa y desencadena respuestas celulares como el arresto del ciclo celular o la inducción de la apoptosis. Sin embargo, cuando el gen se encuentra mutado, su función puede estar alterada, lo que lleva a que no responda adecuadamente a las señales de daño o estrés. En estos casos, MDM2 puede unirse y ubiquitinizar el gen *TP53* mutado, marcando al mismo para su degradación en el proteosoma. Esta ubiquitinación y degradación del gen *TP53* mutado contribuye a la pérdida de la función supresora de tumores permitiendo el crecimiento celular descontrolado y favoreciendo el desarrollo de cáncer. Es importante tener en cuenta que el sistema MDM2-*TP53* es una de las vías reguladoras críticas en la prevención del cáncer, y las alteraciones en esta vía pueden tener implicaciones en el desarrollo de tumores³⁹.

1.8 Cáncer

El cáncer es una enfermedad que se origina a partir de la proliferación de células tumorales. Las células tumorales suelen derivar de células normales que han adquirido mutaciones en genes que regulan la división celular, la apoptosis, la reparación del ADN y otros procesos celulares críticos. Estas mutaciones pueden ser el resultado de la exposición a carcinógenos, factores genéticos hereditarios, procesos inflamatorios crónicos, radiación u otros factores ambientales. Estas células pierden la capacidad de regular su ciclo celular de manera adecuada y pueden evadir los mecanismos normales de control del crecimiento y muerte celular, lo que resulta en una proliferación descontrolada y la formación de tumores.

El cáncer es una enfermedad de origen genético. Alteraciones en los genes que regulan la función celular, especialmente durante el desarrollo y multiplicación, son responsables de la aparición del cáncer. Estos cambios genéticos que provocan el cáncer pueden originarse por:⁴³

- Errores que ocurren cuando las células se multiplican.
- Daños en el ácido desoxirribonucleico (ADN) por sustancias perjudiciales en el medio ambiente, como las sustancias químicas en el humo del tabaco y los rayos ultravioleta del sol.
- Herencia genética de padres a hijos.

1.8.1 Tipos de cáncer

Una de las formas de clasificar los tipos de cáncer es con el nombre de los órganos o tipo de célula que le dio origen. A continuación, se detallan algunos tipos de cáncer que reciben su nombre según el tipo específico de células donde se origina la enfermedad.⁴⁴

Carcinoma:

Los carcinomas son el tipo más común de cáncer. Se originan a partir de células epiteliales, que son las células que recubren las partes internas y externas del cuerpo. Hay muchos tipos de células epiteliales, por ejemplo:

- Adenocarcinoma: cáncer que se forma en las células epiteliales que producen líquido o moco. El tejido con este tipo de células epiteliales a veces se llama tejido glandular. La mayoría de los cánceres de mama, colon y próstata son adenocarcinomas.
- Carcinoma de células basales: cáncer que se forma en la capa inferior o de base de la epidermis.
- Carcinoma de células escamosas: un cáncer que se forma en las células escamosas, que son células epiteliales que están justo debajo de la superficie externa de la piel. Las células escamosas también recubren muchos otros órganos, como el estómago, los intestinos, el pulmón, la vejiga y los riñones.
- Carcinoma de células transicionales: cáncer que se inicia en un tejido que se llama epitelio transicional o urotelio. Este tejido tiene muchas capas de células epiteliales que se agrandan y achican. Se encuentra en el revestimiento de la vejiga, los uréteres y la pelvis renal (una parte del riñón), y en otros órganos. Algunos cánceres de vejiga, uréter y riñón son carcinomas de células transicionales.⁴⁴

Sarcoma:

Los sarcomas son cánceres que se forman en el hueso y los tejidos blandos, como los músculos, la grasa, los vasos sanguíneos, los vasos linfáticos y el tejido fibroso (como tendones

y ligamentos). El osteosarcoma es el cáncer de hueso más común. Los tipos más comunes de sarcoma de tejido blando son: leiomiosarcoma, sarcoma de Kaposi, histiocitoma fibroso maligno, liposarcoma y dermatofibrosarcoma protuberante.⁴⁴

Leucemia:

Los cánceres que comienzan en el tejido de la médula ósea que se transforma en células sanguíneas se llaman leucemias. Estos cánceres no crean un tumor sólido. En cambio, se acumulan grandes cantidades de glóbulos blancos anormales (células leucémicas y blastocitos leucémicos) en la sangre y la médula ósea, que desplazan a las células sanguíneas normales. La concentración baja de células sanguíneas normales hace que sea más difícil para el cuerpo llevar oxígeno a los tejidos, controlar el sangrado o combatir infecciones. Hay cuatro tipos comunes de leucemia: aguda o crónica (según la velocidad de progresión de la enfermedad), y linfoblástica o mieloide (según el tipo de célula sanguínea donde se inicia). Las leucemias agudas son de crecimiento rápido y las leucemias crónicas son de crecimiento lento.⁴⁴

Linfoma:

El linfoma es un cáncer que se inicia en los linfocitos (células T o células B). Estos glóbulos blancos que luchan contra las enfermedades son parte del sistema inmunitario. En el linfoma, los linfocitos anormales se acumulan en los ganglios y vasos linfáticos, y en otros órganos del cuerpo. Hay dos tipos principales de linfoma:

- Linfoma de Hodgkin: enfermedad en la que las personas tienen linfocitos anormales que se llaman células de Reed-Sternberg. Estas células en general se originan en las células B.
- Linfoma no Hodgkin: Es posible que sean cánceres de crecimiento rápido o lento y que se formen a partir de células B o células T.⁴⁴

Mieloma múltiple:

El mieloma múltiple es un cáncer que se inicia en las células plasmáticas, otro tipo de célula inmunitaria. Las células plasmáticas anormales, llamadas células de mieloma, se acumulan en la médula ósea y forman tumores en los huesos de todo el cuerpo. El mieloma múltiple también se llama mieloma de células plasmáticas o enfermedad de Kahler.⁴⁴

Melanoma:

El melanoma es un cáncer que se inicia en células que se convierten en melanocitos, que son células especializadas que producen melanina (el pigmento que da a la piel su color). La mayoría de los melanomas se forman en la piel, pero a veces aparecen en los tejidos pigmentados, como los del ojo.

Tumores de encéfalo y médula espinal:

Hay distintos tipos de tumores de encéfalo y médula espinal. Estos tumores llevan el nombre de las células en que se forman y el sitio donde aparece el tumor en el sistema nervioso central. Por ejemplo, un astrocitoma se inicia en los astrocitos (células de encéfalo en forma de estrella), que ayudan a mantener sanas las células nerviosas. Los tumores de encéfalo son benignos (no cancerosos) o malignos (cancerosos).

Tumores de células germinativas:

Los tumores de células germinativas son un tipo de tumor que se inicia en las células que luego forman el esperma o los óvulos.⁴⁴

Tumores neuroendócrinos:

Los tumores neuroendócrinos surgen de células que liberan hormonas en la sangre al responder a una señal del sistema nervioso. Estos tumores, que a veces producen concentraciones de hormonas más altas que las normales, causan diversos síntomas. Los tumores neuroendócrinos pueden ser benignos o malignos.⁴⁴

Tumores carcinoides:

Los tumores carcinoides son un tipo de tumor neuroendócrino. Son tumores de crecimiento lento que con frecuencia se encuentran en el aparato digestivo, casi siempre en el recto y el intestino delgado. Es posible que los tumores carcinoides se diseminen al hígado o a otras partes del cuerpo, y que secreten sustancias como la serotonina o las prostaglandinas, que causan un conjunto de afecciones llamadas síndrome carcinoide.⁴⁴

1.9 Cáncer de mama

El cáncer de mama es un tipo de cáncer que se origina en el tejido mamario, manifestándose como una proliferación maligna de las células epiteliales que recubren los conductos o lobulillos mamarios. Es una enfermedad clonal, donde una célula individual producto de una serie de mutaciones somáticas o de línea germinal adquiere la capacidad de dividirse sin control.⁴⁵



ADAM.

Figura 9: Partes del seno.⁴⁶

El tumor resultante, que comienza como anomalía leve, puede pasar a ser grave, invadiendo tejidos vecinos y finalmente, propagándose a otras partes del cuerpo. Es importante comprender la diferencia entre el cáncer "*in situ*" y el cáncer "*invasivo*".

En el tipo de cáncer "*in situ*", las células tumorales se encuentran dentro del tejido mamario y no han invadido los tejidos circundantes ni se han diseminado hacia otras partes del cuerpo. Este tipo de cáncer se considera en etapa temprana y a menudo se detecta mediante mamografías de detección. Aunque el cáncer "*in situ*" no ha invadido otros tejidos, aún requiere tratamiento para prevenir su progresión a una forma invasiva.^{47 48}

El tipo de cáncer invasivo o también conocido como cáncer infiltrante, implica la propagación de las células tumorales hacia los tejidos vecinos. Si el cáncer se disemina a través de los conductos o los lobulillos mamarios y penetra en los tejidos vecinos, se considera invasivo. Este carcinoma invasivo puede hacer una metástasis o diseminarse a otras partes del cuerpo por el torrente sanguíneo o sistema linfático.^{47 48}

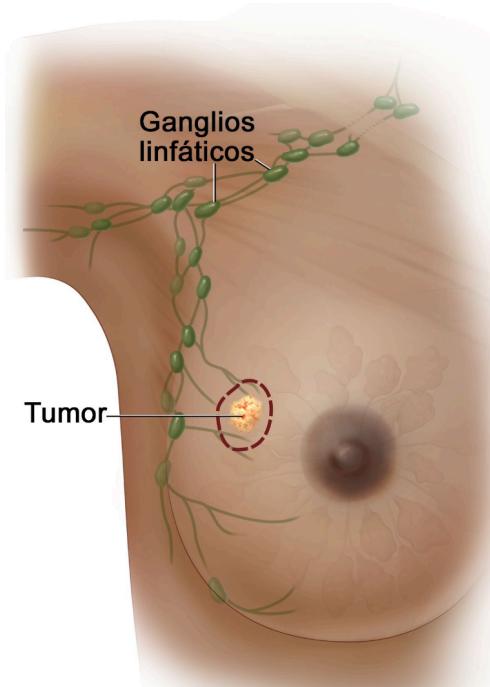


Figura 10: Tumor de mama.⁴⁹

Los factores que contribuyen al desarrollo del cáncer de mama incluyen tanto aspectos hereditarios como ambientales. Entre los factores hereditarios, destacan las mutaciones en los genes BRCA1 y BRCA2, que son algunas de las más comunes y están asociadas con un mayor riesgo de cáncer de mama y ovario. Además, factores externos como el estilo de vida y el entorno también pueden inducir cambios genéticos que aumentan el riesgo de desarrollar esta enfermedad.⁵⁰

Algunos de los factores que favorecen el desarrollo de cáncer de mama son:⁵⁰

- Edad avanzada
- Historial de cáncer de mama o enfermedad benigna de la mama
- Riesgo heredado de cáncer de mama, incluyendo la aparición de cambios en los genes BRCA1 y BRCA2
- Tejido mamario denso
- Un historial reproductivo que lleva a una mayor exposición a la hormona estrógeno, incluyendo:
 - Menstruar a una edad temprana
 - Tener mayor edad al dar a luz por primera vez o nunca haber dado a luz
 - Comenzar la menopausia a una edad más avanzada
- Recibir terapia hormonal para los síntomas de la menopausia
- Radioterapia en el pecho
- Obesidad
- Beber alcohol

El cáncer de mama es uno de los principales causantes de muertes por cáncer a nivel global y en nuestro país. Sin embargo, se ha demostrado que la detección temprana de este tipo de tumor tiene un impacto significativo en la prevención y tratamiento de la enfermedad. Por esta razón, el reconocimiento de signos y la realización de controles periódicos son fundamentales en el proceso de detección. A continuación, se presentan los signos y síntomas del cáncer de mama que deben ser monitoreados.⁵⁰

- Un bulto nuevo o engrosamiento en el seno, o cerca de éste, o en la axila
- Un cambio en el tamaño o la forma de la mama
- Un hoyuelo o arrugas en la piel del seno. Puede parecer la piel de una naranja
- Un pezón vuelto hacia adentro en el seno
- Secreción del pezón que no es leche materna.
- Piel escamosa, enrojecida o inflamada en el área del pezón o en la mama
- Dolor en cualquier parte de la mama

1.10 Tipos de cáncer de mama

Existen dos tipos principales de cáncer de mama que serán objeto de estudio en esta investigación, el carcinoma ductal infiltrante (CDI) y el carcinoma lobular infiltrante (CLI). El carcinoma ductal infiltrante (CDI) y el carcinoma lobulillar infiltrante (CLI) son los tipos más prevalentes de cáncer de mama, representando más del 90% de los casos. Dada su relevancia, hemos decidido centrarnos en estos subtipos para el estudio en esta tesis.⁴⁸

El CDI es el tipo más frecuente de cáncer de mama representando aproximadamente el 80 % de los casos.^{48 51} Se origina en el revestimiento de los conductos galactóforos, que son tubos delgados que transportan la leche desde los lobulillos de la mama hasta el pezón. A medida que avanza, el carcinoma se extiende desde los conductos hacia el tejido circundante. Además, tiene la posibilidad de diseminarse a otras partes del cuerpo a través de la sangre y el sistema linfático. En muchos casos, el carcinoma ductal infiltrante no produce síntomas y se detecta cuando el médico ve un área sospechosa que aparece en una mamografía.

Cáncer ductal invasivo de mama

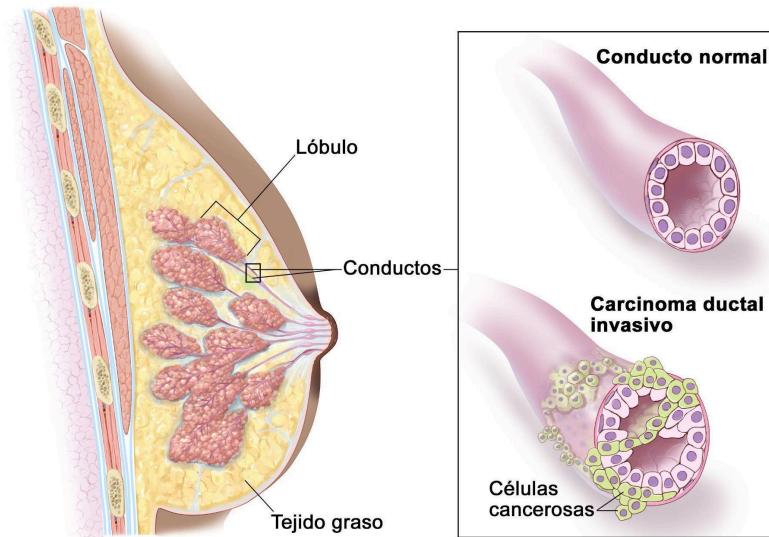


Figura 11: Células tumorales correspondientes al tipo de cáncer de mama carcinoma ductal invasivo de mama.⁵²

En segundo lugar se encuentra el CLI, el cual abarca entre el 10 % y el 12 % de los casos.⁴⁸ Este tipo de cáncer se origina en la glándula mamaria, específicamente en los lobulillos, que son las estructuras responsables de la producción de leche, y se extiende al tejido mamario circundante. En algunos casos, puede diseminarse a otras partes del cuerpo a través de los sistemas sanguíneo y linfático. El carcinoma lobular invasivo no presenta síntomas y se detecta únicamente cuando una mamografía revela un área sospechosa. A diferencia de otros tipos de cáncer de mama, el CLI es menos propenso a formar un bulto duro, ya que las células tumorales lobulares tienden a crecer en patrones lineales y a formar una capa en lugar de un nódulo sólido.

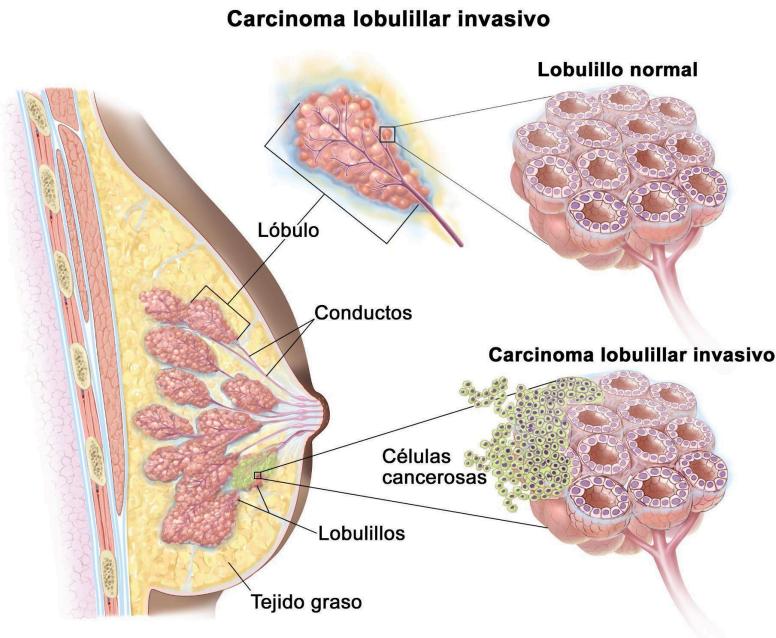


Figura 12: Células tumorales correspondientes al tipo de cáncer de mama carcinoma lobular invasivo de mama.⁵³

1.11 El gen *TP53* en el cáncer de mama

El gen con la mutación más frecuente en personas con cáncer es *TP53*. La mayoría de las mutaciones del gen *TP53* son adquiridas, lo que significa que se desarrollan a lo largo de la vida de una persona debido a la exposición a factores ambientales o a errores durante la replicación del ADN. Si bien las mutaciones de la línea germinal en el gen *TP53* son raras, los pacientes portadores de estas mutaciones están en mayor riesgo de desarrollar muchos tipos diferentes de cáncer.¹

En la carcinogénesis, más del 70% de los cánceres humanos presentan un defecto en este gen, y las neoplasias malignas restantes tienen defectos en los genes que interactúan con *TP53*. Se ha observado una pérdida de los dos alelos del gen *TP53* en prácticamente todos los tipos de cáncer, incluyendo carcinomas de pulmón, colon y mama, que son las tres principales causas de muerte por cáncer. Una de las principales razones por las que el estudio de *TP53* adquiere gran relevancia reside en que en numerosos casos de cáncer esta proteína aparece mutada. Las mutaciones de *TP53* que más frecuentemente se han descrito provocan un cambio de un aminoácido por otro diferente (mutaciones *missense* o de cambio de sentido) y se han localizado principalmente en el dominio de unión a ADN. El cambio de aminoácido impide el reconocimiento específico del ADN por el Dominio de Unión al ADN (DBD), lo que altera la funcionalidad de la proteína p53. Esto significa que la mutación afecta la capacidad de la proteína p53 para unirse y regular la actividad de los genes específicos mediante la interacción

con secuencias de ADN específicas. Esta alteración en la función del DBD puede llevar a una pérdida parcial o total de la capacidad de la proteína p53 para controlar la proliferación celular, reparar el ADN dañado y promover la apoptosis, lo que contribuye al desarrollo y progresión del cáncer.¹

Tomando como referencia los gráficos publicados en el artículo *TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use* publicado en *Cold Spring Harbor Perspectives in Biology*, en la **Figura 13** se indica la proporción de tumores con mutaciones somáticas del gen *TP53*, donde se puede observar que aproximadamente el 25% de los casos de cáncer de mama presentan mutaciones en el gen *TP53*.¹

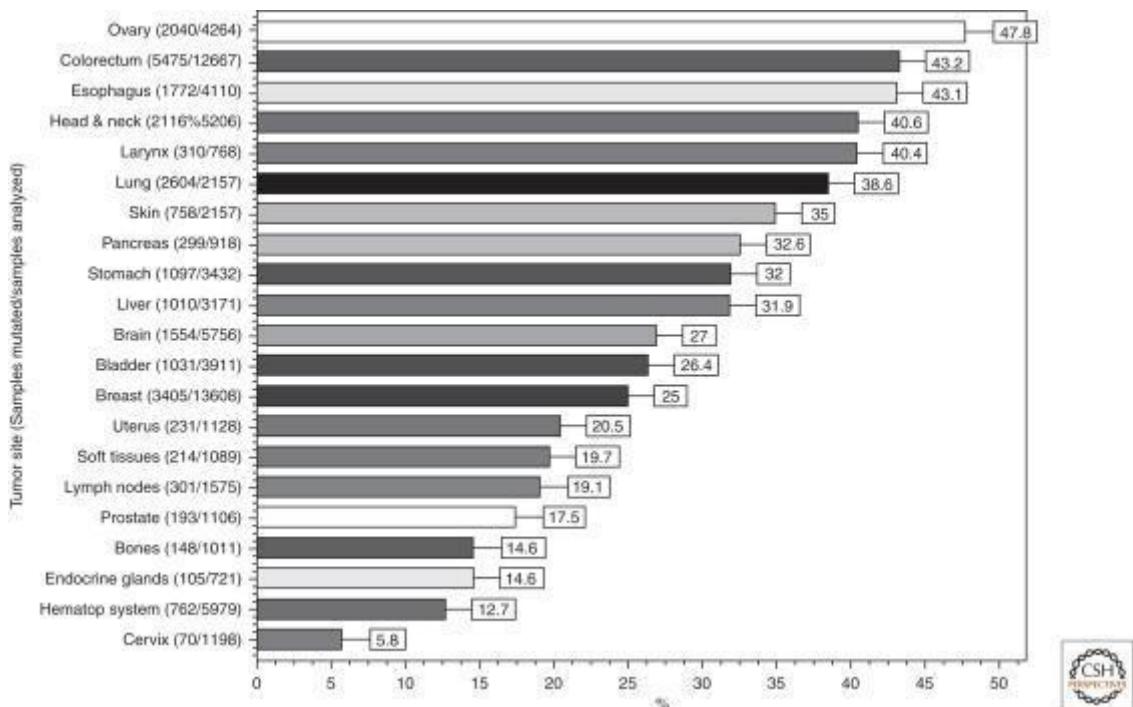


Figura 13: Prevalencia de mutaciones en *TP53* en cánceres esporádicos. Se indica la proporción de tumores con mutaciones somáticas en *TP53*. Datos de la Base de Datos *TP53* de la IARC.¹

1.12 Perfil Tumoral Molecular

Los perfiles tumorales moleculares del cáncer de mama se definen principalmente por la expresión de ciertos receptores en las células tumorales, específicamente los receptores de estrógeno (ER), de progesterona (PR) y el receptor 2 del factor de crecimiento epidérmico humano (HER2). A continuación, abordaremos cada uno de estos receptores y su importancia en la clasificación y el tratamiento del cáncer de mama.

1.12.1 Receptor HER2

El receptor 2 del factor de crecimiento epidérmico humano HER2 (receptor tirosina-proteína quinasa), también conocido como ErbB2, es un protooncogén localizado en el brazo largo del cromosoma 17. Funciona como un receptor para el factor de crecimiento epidérmico humano, desempeñando un papel crucial en el crecimiento y la división normal de las células. La sobreexpresión está asociada a una mayor actividad biológica del tumor, lo que se traduce en una tasa de división celular acelerada debido a la mayor respuesta de las células a los niveles normales de factores de crecimiento⁹. En el cáncer de mama, entre el 15% y el 20% de los pacientes tienen amplificación del HER2. Este subgrupo de pacientes con amplificación de HER2 (HER2+) puede beneficiarse de la terapia dirigida a HER2. Fármacos con anticuerpos anti-HER2 como *trastuzumab* y *pertuzumab*, inhibidores de la tirosina quinasa (ITC), inhibidores de moléculas pequeñas como *lapatinib*, *neratinib* y *pirotinib*, y conjugados de anticuerpos y fármacos como *trastuzumab deruxtecan (DS-8201)* y *trastuzumab emtansina (T-DM1)* también muestran buenos efectos curativos en estos pacientes⁹.

El estado de HER2 en los tumores de mama se evalúa mediante pruebas específicas para determinar si las células tumorales tienen un nivel alto de proteína HER2. Las principales pruebas utilizadas son primero la inmunohistoquímica (IHC) y luego la hibridación fluorescente *in situ* (FISH). Los resultados de las pruebas de HER2 son fundamentales para guiar tanto al paciente como al equipo de salud en la toma de decisiones sobre el tratamiento más adecuado. Cuando la prueba IHC para HER2 es negativa, estos cánceres no responden a tratamientos con medicamentos dirigidos a HER2. Cuando la prueba IHC para HER2 es positiva, estos cánceres suelen tratarse con medicamentos que tienen a la proteína HER2 como blanco. Hay un tercer resultado que se clasifica como "ambiguo" y, en este caso, es necesario realizar una prueba adicional de FISH para clarificar el resultado.

1.12.2 Receptor de Estrógeno (ER)

El receptor de estrógeno (ER) hace referencia a un grupo de receptores celulares que son activados por la hormona estrógeno. Las hormonas son sustancias que funcionan como mensajeros químicos en el cuerpo, afectando la actividad de células y tejidos en distintos lugares. Estas hormonas suelen llegar a su destino a través de la sangre. Los estrógenos, que son hormonas sexuales esteroideas de tipo femenino, son producidos principalmente por los ovarios, la placenta durante el embarazo, y en menores cantidades por las glándulas suprarrenales. La función principal del receptor de estrógeno es actuar como un factor de transcripción que se une al ADN para regular la expresión génica. Sin embargo, el receptor de estrógeno también tiene funciones adicionales que son independientes de la unión al ADN.

Los receptores de estrógeno son proteínas que, cuando se unen al estrógeno, pueden estimular el crecimiento de células de cáncer de mama. La presencia o ausencia de estos receptores influye en el comportamiento del cáncer y en las opciones de tratamiento.

Para determinar si las células del cáncer de mama tienen receptores hormonales, se analizan muestras de tejido tumoral que se extraen mediante cirugía. Si las células tumorales contienen receptores de estrógeno, es un cáncer positivo para el ER. También se describe como cáncer sensible al estrógeno o receptivo al estrógeno. Se estima que entre el 67 % y el 80 % de los cánceres de mama en las mujeres dan positivo para el receptor de estrógeno^{10 11}.

1.12.3 Receptor de Progesterona (PR)

La progesterona es una hormona esteroidea producida principalmente por los ovarios y, en menores cantidades, por las glándulas suprarrenales. Desempeña un papel crucial en el ciclo menstrual y en el embarazo. También tiene efectos sobre otros tejidos, incluyendo el tejido mamario, donde promueve el desarrollo glandular. No todas las células expresan su receptor (PR), las que lo hacen se encuentran normalmente en la mama, los ovarios, el útero y el cuello uterino. Algunos tipos de cáncer, incluidos los de mama, ovario, útero y cervix, también producen este receptor. El PR es una proteína ubicada dentro de las células, activada por la hormona esteroidea progesterona. Este receptor permite que una célula responda a las acciones de la hormona progesterona. Una vez producida, la proteína PR permanece en el núcleo de la célula. La hormona progesterona que ingresa a la célula puede alterar su comportamiento al activar los PR.

Los patólogos analizan los cánceres de mama para determinar si las células tumorales están produciendo PR. Esta prueba se realiza porque las células tumorales que producen este receptor requieren la acción de la hormona progesterona para crecer. Un paciente con un cáncer que produce PR tiene más probabilidades de beneficiarse de medicamentos bloqueadores de hormonas. La respuesta al medicamento depende de la cantidad de receptor de progesterona producido. Los tumores con niveles más altos de PR dependen más de la progesterona para crecer y, por lo tanto, responden mejor al tratamiento. Cuando las células tumorales tienen receptores de progesterona, el cáncer se considera positivo para el receptor de progesterona (PR positivo o PgR positivo). En hombres, alrededor de 90 % de los cánceres de mama dan positivo para el receptor de estrógeno y alrededor de 80 % dan positivo para el receptor de progesterona¹².

1.12.4 Terapias

Terapia hormonal:

La terapia hormonal se emplea en cánceres de mama receptor hormonal positivo (HR+), es decir, aquellos que dependen de hormonas como el estrógeno o la progesterona para su crecimiento. Esta terapia actúa disminuyendo o impidiendo el crecimiento de tumores sensibles a las hormonas, ya sea bloqueando la producción de estas hormonas o interfiriendo en sus efectos sobre las células tumorales⁵⁶.

- Tamoxifeno: Es un medicamento utilizado en la terapia hormonal para el cáncer de mama en células tumorales con receptores de estrógeno positivos (HR+). Bloquea el efecto del estrógeno en las células mamarias, frenando el crecimiento tumoral, aunque en otros tejidos, como el útero y los huesos, actúa como estrógeno. Es eficaz para reducir el riesgo de recurrencia en mujeres con alto riesgo, aquellas tratadas por carcinoma ductal *in situ*, y en cánceres de mama metastásicos. En mujeres premenopáusicas se usa ampliamente, mientras que en mujeres postmenopáusicas suelen emplearse inhibidores de la aromatasa.
56 57
- Inhibidores de la aromatasa (anastrozol, letrozol y exemestano): bloquean la producción de estrógeno en mujeres postmenopáusicas al inhibir la enzima aromatasa, que convierte hormonas de tejido adiposo en estrógeno. Esto reduce los niveles hormonales que podrían estimular el crecimiento de células cancerosas en tumores que dependen del estrógeno. En mujeres premenopáusicas, se pueden usar con supresión ovárica para mayor efectividad.⁵⁷

Terapia Dirigida a HER2:

La terapia dirigida a HER2 es específica para tumores que sobreexpresan la proteína HER2, asociada con un crecimiento celular acelerado en el cáncer de mama. Incluye medicamentos como trastuzumab (Herceptin), pertuzumab y lapatinib, que se unen a la proteína HER2 para bloquear señales que estimulan el crecimiento tumoral, frenando la proliferación de las células cancerosas. Los anticuerpos monoclonales, como el trastuzumab, son diseñados para adherirse a la proteína HER2 en las células malignas y detener su crecimiento. Este tratamiento puede administrarse junto con quimioterapia y suele prolongarse de seis meses a un año en etapas iniciales, o de forma continua en cáncer avanzado mientras sea efectivo.⁵⁸

Quimioterapia:

La quimioterapia se utiliza para destruir células tumorales y, debido a que afecta tanto a células cancerosas como a células normales de rápida división, se emplea frecuentemente en combinación con otras terapias. Su aplicación puede ser neoadyuvante para reducir el tamaño del tumor antes de la cirugía, o adyuvante para reducir el riesgo de recurrencia posterior a la operación. Los medicamentos de quimioterapia se administran generalmente de forma intravenosa o oral, y viajan a través del cuerpo para atacar las células cancerígenas. En casos donde el cáncer se ha propagado al líquido de la médula espinal, la quimioterapia también puede administrarse intratecal, directamente sobre esta área para tratar la propagación.⁵⁹

1.13 Perfiles Tumorales Moleculares del Cáncer de mama

Los “perfiles de cáncer de mama” se asignan con el fin de clasificar el tumor en subtipos con diferentes características biológicas y respuestas a los tratamientos disponibles. Estos perfiles son fundamentales para personalizar el tratamiento y predecir el pronóstico de cada paciente.

En este trabajo, se analizaron 4 perfiles:

- *Lumina A*: este perfil se caracteriza por tener ER y PR positivos, y HER2 negativo. Es el tipo de perfil menos agresivo y se asocia con un mejor pronóstico. Los tumores con este perfil responden favorablemente a la terapia hormonal, como el tamoxifeno y los inhibidores de la aromatasa, y generalmente tienen una baja necesidad de quimioterapia.
- *Luminal B*: este perfil se caracteriza por tener ER positivos, PR que pueden ser positivos o negativos, y HER2 positivo, o bien, ER positivo, PR negativo, y HER2 negativo. Este tipo de perfil es más agresivo y tiene un peor pronóstico que el perfil Luminal A. Los

tumores con este perfil pueden requerir terapia hormonal y, si HER2 es positivo, también terapia dirigida a HER2. Además, en este perfil, la necesidad de quimioterapia es mayor.

- *HER2-enriquecido*: este perfil se caracteriza por tener ER negativos, PR negativos, y HER2 positivo. Este tipo de cáncer tiene una alta proliferación celular y es considerado agresivo, con un mal pronóstico si no se administra un tratamiento específico. No obstante, responde bien a terapias dirigidas contra HER2, como *trastuzumab* y *pertuzumab*, y suele requerir quimioterapia.
- Triple Negativo, en este perfil, los ER, PR, y HER2 son negativos. Este tipo de cáncer es muy agresivo y tiene el peor pronóstico entre todos los perfiles. No responde ni a la terapia hormonal ni a la dirigida contra HER2 y es tratado principalmente con quimioterapia.

2.0 ESTADO DEL ARTE

El gen supresor tumoral *TP53* ha sido el gen tumoral secuenciado más extensamente antes de la era de la secuenciación a gran escala, lo que generó una gran cantidad de datos para identificar vínculos entre carcinógenos, huellas dactilares de mutaciones y tumorigénesis.

Las mutaciones *TP53* han demostrado ser biomarcadores extremadamente complejos. A pesar de los avances en la comprensión mecanicista de la estructura y función de *TP53*, la investigación de *TP53* aún no ha generado aplicaciones de gran impacto en el tratamiento y la terapia del cáncer. De alguna manera, la complejidad del *TP53* puede haber actuado como un elemento disuasorio para las aplicaciones clínicas. El hecho de que las mutaciones en *TP53* sean diversas en sus efectos biológicos y que puedan ocurrir en muchas etapas diferentes antes o durante el desarrollo del tumor, hace imposible derivar mensajes simples que sean uniformemente aplicables a todos los contextos clínicos. Traducir la mutación *TP53* a las clínicas requerirá ensayos clínicos grandes y estructurados en los que los pacientes con un estado definido de *TP53* sean reclutados sobre la base de criterios de inclusión específicos, aleatorizados para el tratamiento de acuerdo con regímenes determinados y seguidos para determinar criterios de valoración terapéuticos y clínicos a largo plazo. Las bases de datos, como las actuales bases de datos de mutaciones de *TP53*, tendrán un papel fundamental en la recopilación, estructuración y anotación de estos datos, lo que permitirá la interpretación de la mutación de *TP53* y su uso informado en la práctica estándar de patología molecular.¹

Las mutaciones somáticas del gen *TP53* son frecuentes en la mayoría de los cánceres humanos.⁵⁴ También se consideran posibles marcadores pronósticos y predictivos, así como objetivos para la intervención farmacológica.¹

Es importante destacar que *TP53* es el gen mutado con mayor frecuencia en el cáncer de mama invasivo. Aunque está mutado en alrededor del 35% de todos los casos de cáncer de mama, se observa una mutación en aproximadamente el 80% de los tumores triple negativos (TN), caracterizados por la ausencia de expresión de los receptores de estrógeno (ER), progesterona (PR) y HER2. Dada su alta prevalencia, la mutación en *TP53* representa un biomarcador potencial y un posible objetivo terapéutico en pacientes con cáncer de mama, particularmente en aquellos con el subtipo triple negativo.²

En este trabajo de investigación se analiza cómo las células tumorales de mama, correspondientes a los tipos de cáncer de mama Carcinoma Lobular Infiltrante (CLI) y Carcinoma Ductal Infiltrante (CDI), presentan diferentes mutaciones en el gen *TP53*. Además,

se evaluará si existe una relación entre estos tipos de cáncer de mama y sus perfiles moleculares cuando el gen *TP53* se encuentra mutado.

Como objetivo general nos proponemos detectar si predomina un tipo específico de mutación del gen *TP53* según el tipo de cáncer de mama comparando estos dos subtipos. También determinar si los perfiles moleculares de las células tumorales se relacionan con los distintos subtipos de cáncer de mama, especialmente en aquellos con mutaciones en *TP53*. Además, se investigará el gen *TP53* para ampliar el conocimiento sobre sus mutaciones y su impacto en el desarrollo y progresión del cáncer de mama. El análisis se llevará a cabo utilizando principalmente la herramienta *RStudio*.

3.0 HIPÓTESIS

Existen diferencias en los perfiles tumorales moleculares entre los tumores de mama CDI y CLI, lo que podría asociarse con la respuesta a terapias antitumorales.

4.0 OBJETIVOS ESPECÍFICOS DE INVESTIGACIÓN

- Identificar si las células tumorales de mama, que corresponden a los tipos de cáncer de mama Carcinoma Lobular Infiltrante (CLI) y Carcinoma Ductal Infiltrante (CDI), comparten las mutaciones en el gen *TP53* empleando información recopilada de bases de datos y realizando análisis estadísticos con R.
- Determinar si predomina un tipo de mutación del gen *TP53* según el tipo de cáncer de mama.
- Evaluar si los perfiles moleculares de las células tumorales se relacionan con los distintos subtipos de cáncer de mama, especialmente en aquellos con mutaciones en *TP53*.
- Investigar el gen *TP53* para ampliar el conocimiento sobre sus mutaciones y funciones en relación con el cáncer de mama, reflexionando sobre el foco del tratamiento dependiendo de los factores específicos del tipo de mutación.

5.0 MATERIALES Y MÉTODOS

Base de datos. Para llevar a cabo la investigación se tomaron 200 informes de distintos pacientes desde las bases de datos *TCGA Cell*³, *TCGA Firehose Legacy*⁴, *TCGA Nature*⁵ y *TCGA PanCancer Atlas*⁶, proporcionadas por la plataforma *cBioPortal*⁷. El *cBioPortal* para la genómica del cáncer, se desarrolló originalmente en el Memorial Sloan Kettering Cancer Center (MSK). El sitio público *cBioPortal* está alojado en el Centro de Oncología Molecular de MSK.⁸

RStudio software. Para llevar a cabo los análisis, se utilizó la herramienta *RStudio*. La información obtenida de la base de datos se organizó en diferentes carpetas, separando los datos por pacientes y asignando identificadores correspondientes a cada uno. Posteriormente, se procedió a realizar una limpieza de los datos para asegurar su precisión y consistencia. Una vez completada esta etapa, se procedió al análisis estadístico de la información, permitiendo obtener resultados precisos y relevantes.

ChatGPT4.0. Se utilizó la herramienta de inteligencia artificial para consultar y mejorar la redacción. También se utilizó para consultas de error del código.

6.0 RESULTADOS

6.1 Predominancia en la mutación *TP53* y características de los pacientes

El gen *TP53* es un gen supresor de tumores, el cual se ubica específicamente en el cromosoma 17. La proteína codificada por el gen se llama p53, la misma es una fosfoproteína que se localiza en el núcleo celular. El gen *TP53* es uno de los genes supresores de tumores más frecuentemente mutados en personas con cáncer. Como se menciona en el apartado 1.11 del Marco Teórico, más del 70% de los casos de cáncer y aproximadamente el 25% de los pacientes con cáncer de mama poseen mutaciones en este gen, lo que resulta en la pérdida de su función o daño en su estructura.

Para llevar a cabo el estudio se analizaron en total 200 mutaciones del gen *TP53* de distintos pacientes con cáncer de mama de tipo CDI y CLI, obtenidas de la base de datos *CBioPortal*. Para analizar estos datos, se utilizó la herramienta *RStudio*, la cual permitió filtrar específicamente las mutaciones del gen *TP53* y extraer estadísticas relevantes de los pacientes. Previamente, se realizó una limpieza de datos para asegurar la precisión y calidad de la información antes del análisis.

Cada paciente fue asociado a dos tablas: una que contenía su información personal y otra que incluía los datos de la muestra. Posteriormente se realizó una limpieza de datos, en la cual se conservó información únicamente del gen *TP53* junto con los parámetros necesarios para el análisis. Como resultado, se obtuvo una tabla consolidada que contiene toda la información relevante del paciente.

```

#Cargo las bases de datos (información del paciente - información de la mutación)
#----Primer trabajo sobre la base de datos de mutación
#install.packages("readr")
library(readr)
datos126<- read_tsv("/Users/francinazuliani/Desktop/TF/Database/TCGA-B6-A0IE/table.tsv")
#View(datos126)
genes_TP53_126<- subset(datos126, Gene == "TP53")
#View(genes_TP53_12)
TP53_data_126 <- genes_TP53_126[, c("Gene", "Protein Change", "Annotation", "Functional Impact", "Chromosome", "HGVSc", "MS",
#View(TP53_data_12)

#----Segundo trabajo sobre la base de datos de información
clinical_data_126 <- read_tsv("/Users/francinazuliani/Desktop/TF/Database/TCGA-B6-A0IE/clinical_data.tsv")
#View(clinical_data_12)
cl_data_126 <- subset(clinical_data_126,
                        Attribute == "Diagnosis Age" |
                        Attribute == "Neoplasm Histologic Type Name" |
                        Attribute == "Patient Primary Tumor Site" |
                        Attribute == "Sex" |
                        Attribute == "IHC-HER2" |
                        Attribute == "ER Status By IHC" |
                        Attribute == "PR status by ihc" |
                        Attribute == "Tumor Tissue Site")
#View(cl_data_12)
#-----
#install.packages("tidyverse")
library(tidyverse)
cl_data_transposed_126 <- pivot_wider(cl_data_126, names_from = Attribute, values_from = Value)
#View(cl_data_transposed_12)

#Agrego la columna ID a TP53_data y cl_data_transposed
cl_data_transposed_126<- cl_data_transposed_126 %>%
  mutate(ID = "TCGA-B6-A0IE")
TP53_data_126 <- TP53_data_126 %>% mutate(ID = "TCGA-B6-A0IE")
#-----
# Unir los subsets
#install.packages("dplyr")
library(dplyr)
c_data_126 <- merge(cl_data_transposed_126, TP53_data_126, by = "ID", all = TRUE)
View(c_data_126)

```

Código utilizado para realizar la unión de las tablas y la limpieza de datos.

ID	Diagnosis Age	ER Status By IHC	IHC-HER2	Neoplasm Histologic Type Name	Patient Primary Tumor Site	PR status by ihc	Sex	Tumor Tissue Site	Gene	Protein Change
TCGA-B6-A0IE	38	Negative	Negative	Infiltrating Lobular Carcinoma	Left Upper Outer Quadrant Left	Negative	Female	Breast	TP53	X307_splice

Annotation	Functional Impact	Chromosome	HGVSc	MS	Mutation Type	Variant Type	Allele Freq	Variant Reads
OncoKB: Likely Oncogenic, level NA, resistance NA;CI...	MutationAssessor: NA;SIFT: NA;Polyphen-2: NA	17	ENST0000269305.4:c.919+1G>A	Somatic	Splice_Site	SNP	0.456869	143

Tabla resultante con información completa del paciente.

Luego de realizar este proceso para cada paciente, se procedió a unificar las tablas y a filtrar según los tipos de tumor a analizar, es decir CDI y CLI.

```

datos_combinados <- rbind(datos1, datos2, datos3, datos4, datos5, datos6, datos7, datos9, datos10,
                           datos11, datos13, datos14, datos15, datos16, datos17, datos18, datos19,
                           datos20, datos21, datos22, datos23, datos24, datos25, datos27, datos28,
                           datos29, datos30, datos31, datos32, datos33, datos34, datos35, datos36, datos37,
                           datos38, datos39, datos40, datos41, datos42, datos43, datos44, datos45, datos46,
                           datos47, datos48, datos49, datos50, datos51, datos52, datos53, datos54, datos55,
                           datos56, datos57, datos58, datos59, datos60, datos61, datos62, datos63, datos64,
                           datos65, datos66, datos67, datos68, datos69, datos70, datos71, datos72, datos73,
                           datos74, datos75, datos76, datos77, datos78, datos79, datos80, datos81, datos82,
                           datos83, datos84, datos86, datos87, datos88, datos89, datos90, datos91,
                           datos92, datos93, datos94, datos95, datos96, datos97, datos98, datos99, datos100,
                           datos101, datos102, datos103, datos104, datos106, datos107, datos108,
                           datos110, datos111, datos113, datos114, datos115, datos116,
                           datos118, datos119, datos120, datos121, datos122, datos123, datos124,
                           datos125, datos126, datos127, datos128, datos129, datos130, datos131, datos132,
                           datos133, datos134, datos135, datos136, datos137, datos138, datos139, datos140.

View(datos_combinados)
ncol(datos185)

setwd("C:\\\\Users\\\\francina.zuliani\\\\Desktop\\\\Extras\\\\TF\\\\Database")
write.table(datos_combinados, file = "datos_combinados.tsv", sep = "\\t", quote = FALSE, row.names = FALSE)

#-----FILTRO POR TIPOS QUE NECESITO-----
df <- subset(datos_combinados, `Neoplasm Histologic Type Name` == 'Infiltrating Lobular Carcinoma' | `Neoplasm Histologic Type Name` == 'Infiltrating Ductal Carcinoma')
View(df)
setwd("C:\\\\Users\\\\francina.zuliani\\\\Desktop\\\\Extras\\\\TF\\\\Database")
write.table(df, file = "df.tsv", sep = "\\t", quote = FALSE, row.names = FALSE)

```

Código para unificar tablas y filtrar por tipo de cáncer CLI y CDI.

Además de las estadísticas sobre mutaciones, también se analizaron los datos demográficos y clínicos de los pacientes, como edad, género y la expresión de las proteínas HER2, ER y PR, para proporcionar un contexto más completo al estudio. Las características clínicas de los pacientes se muestran en la **Tabla 1** y en las **Figuras 14 y 15**.

La información obtenida fue cargada en la **Tabla 1**, y a continuación se presenta el código utilizado para obtener dichos resultados.

```

#2
# Histograma de la edad de diagnóstico

library(ggplot2)
ggplot(df, aes(x = `Diagnosis Age`)) +
  geom_histogram(binwidth = 5, fill = "blue", color = "black", alpha = 0.7) +
  labs(title = "Distribución de Edad de Pacientes con Cáncer de Mama", x = "Edad de Diagnóstico", y = "Frecuencia") +
  theme_minimal()

```

De las muestras evaluadas, se identificó que el 100% de las muestras procedían de pacientes de género femenino. El estudio abarcó pacientes con edades comprendidas entre los 30 y 90 años. El rango de edad más común fue de 50 a 59 años, representando el 40.5% de los casos, seguido por el grupo de edad de 40 a 49 años, que constituyó el 22.5% de los casos como se presenta en la **Tabla 1**.

Rango de Edad	Cantidad	%
50 - 59	81	40.5
40 - 49	45	22.5
60 - 69	34	17
70 - 79	17	8.5
30 - 39	12	6

Tabla 1: Distribución de Edades en Cáncer de Mama con Mutación en *TP53*. Esta tabla muestra los cinco rangos de edades más frecuentes en pacientes con cáncer de mama que presentan una mutación en el gen *TP53*, junto con el porcentaje de casos correspondiente a cada rango de edad.

Posteriormente, se realizó un análisis de distribución de edades según tipo de cáncer de mama. Para esto, se realizaron los siguientes códigos, donde el primero divide las mutaciones del gen *TP53* de los pacientes por tipo de cáncer y los dos siguientes códigos generan los gráficos de barra correspondientes.

```
library(readr)

# Crear un nuevo conjunto de datos con solo las columnas necesarias para hacer el dot plot
dp <- subset(datos_combinados, select = c(`Neoplasm Histologic Type Name`, `Mutation Type`))
#View(dp)

# Filtrar el conjunto de datos

#TIPO 1: Infiltrating Ductal Carcinoma
dp1 <- subset(df, `Neoplasm Histologic Type Name` == 'Infiltrating Ductal Carcinoma')
View(dp1)

#TIPO 3: Infiltrating Lobular Carcinoma
dp3 <- subset(df, `Neoplasm Histologic Type Name` == 'Infiltrating Lobular Carcinoma')
View(dp3)
```

Clasificación de casos por tipo de cáncer de mama.

```
library(dplyr)
library(ggplot2)

# Crear una nueva columna con los rangos de edades
dp3 <- dp3 %>%
  mutate(Rango_Edad = cut(`Diagnosis Age`, breaks = seq(20, 100, by = 10), labels = c("20-29", "30-39", "40-49", "50-59", "60-69", "70-79", "80-89", "90-99")))

# Crear el bar plot
ggplot(dp3, aes(x = Rango_Edad, fill = Rango_Edad)) +
  geom_bar() +
  labs(title = "Edades de Diagnóstico por Rango del Tipo de Cáncer Infiltrating Lobular Carcinoma",
       x = "Rango de Edad",
       y = "Frecuencia") +
  theme_minimal()
```

Código perteneciente a la Figura 14.

```

library(dplyr)
library(ggplot2)

# Crear una nueva columna con los rangos de edades
dp1 <- dp1 %>%
  mutate(Rango_Edad = cut(`Diagnosis Age`, breaks = seq(20, 100, by = 10), labels = c("20-29", "30-39", "40-49", "50-59", "60-69", "70-79", "80-89", "90-99")))

View(dp1)

# Crear el bar plot
ggplot(dp1, aes(x = Rango_Edad, fill = Rango_Edad)) +
  geom_bar() +
  labs(title = "Edades de Diagnóstico por Rango del Tipo de Cáncer Infiltrating Ductal Carcinoma",
       x = "Rango de Edad",
       y = "Frecuencia") +
  theme_minimal()

```

Código perteneciente a la Figura 15.

Al analizar los grupos de edad según el tipo de cáncer, se observó que el CLI fue más frecuente en el grupo de edad de 50 a 59 años, representando el 45% como se muestra en la **Figura 14**. En cuanto al CDI, el grupo de edad de 50 a 59 años también predominó, aunque con una frecuencia menor del 33% como se muestra en la **Figura 15**. Además, se observa que entre los pacientes estudiados, para el rango de 40-49 años hay una mayor frecuencia de casos de CLI que de CDI.

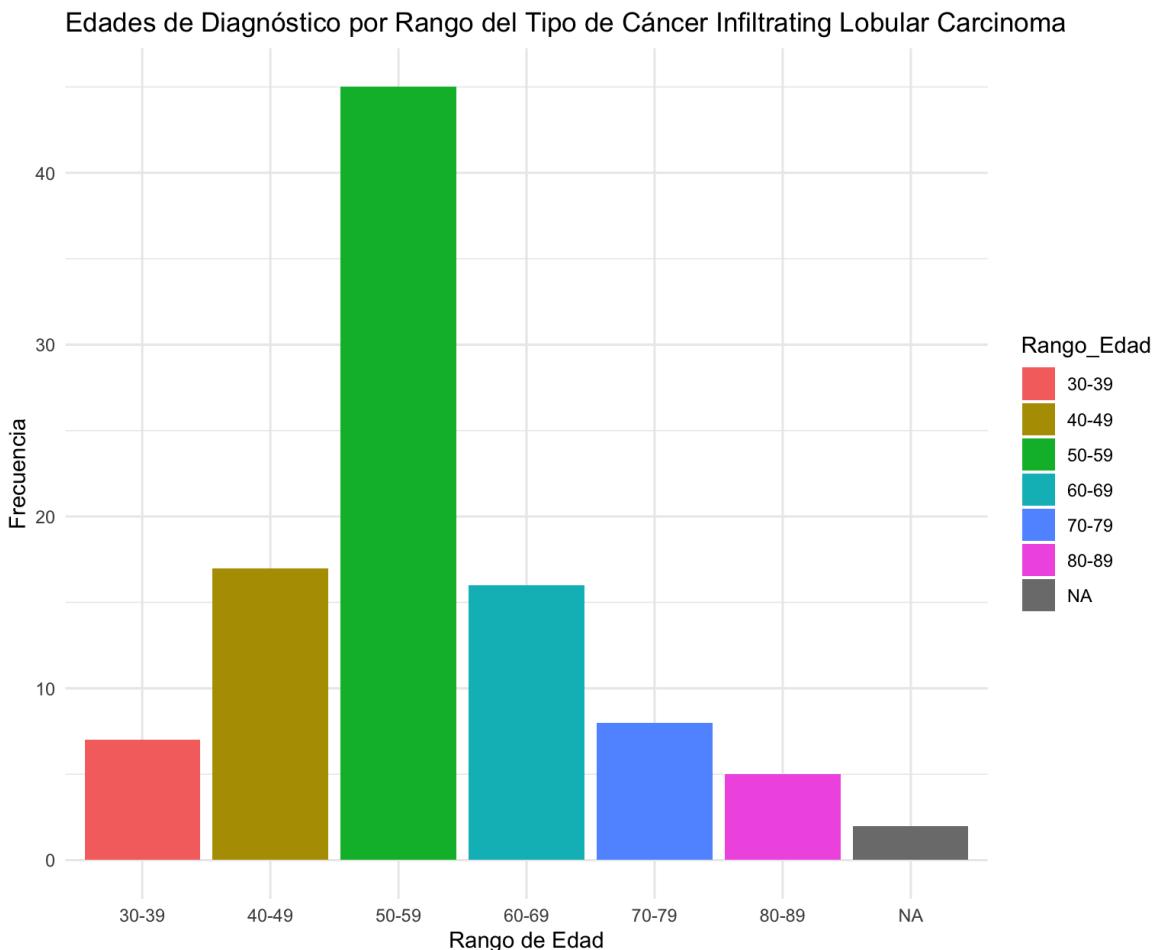


Figura 14: Histograma donde se muestra la distribución de edades de los pacientes con cáncer del tipo CLI.

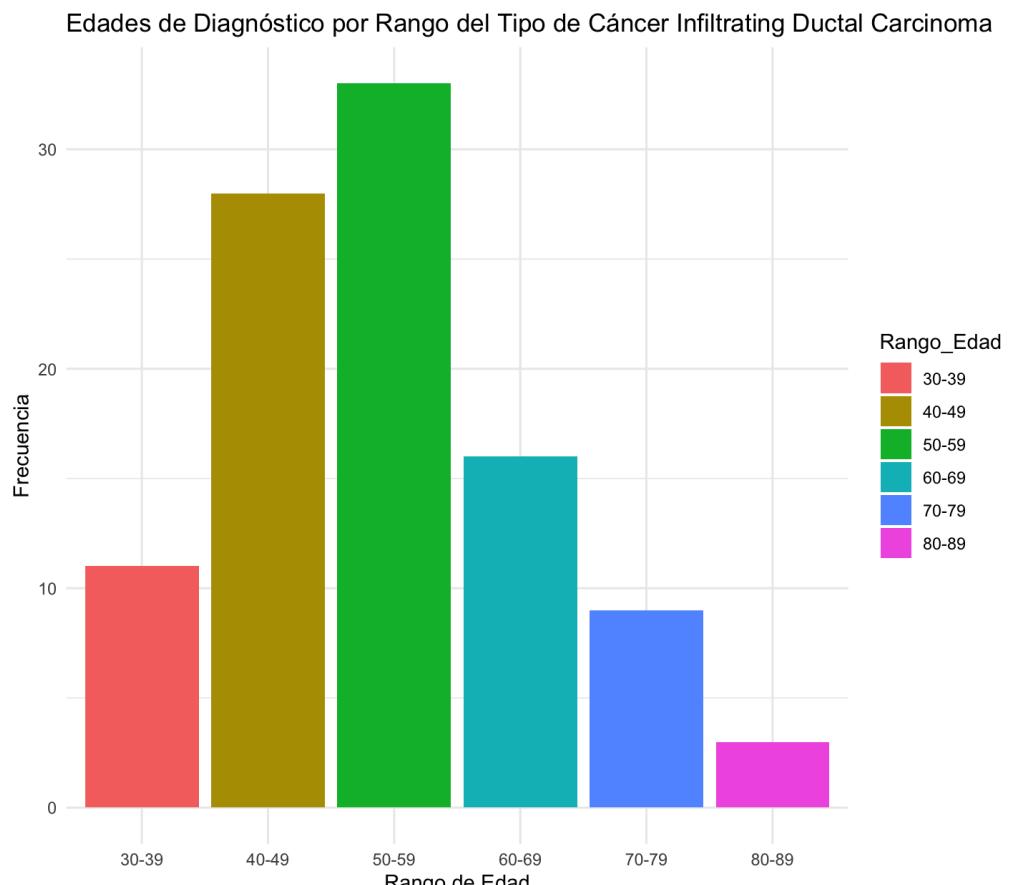


Figura 15: Histograma donde se muestra la distribución de edades de los pacientes con cáncer del tipo CDI.

6.2 Mutaciones concurrentes de *TP53*

A continuación, se evaluaron las mutaciones somáticas de los pacientes. Entre las muestras analizadas, se identificaron 200 mutaciones somáticas de *TP53* en 143 sitios mutantes diferentes. La **Figura 16** presenta un gráfico de dispersión que ilustra la distribución de estos sitios. Los resultados se obtuvieron de los siguientes códigos, donde se calculó la cantidad de sitios mutantes únicos y la cantidad de mutaciones por sitio mutante. Los resultados fueron luego representados en el gráfico de dispersión (**Figura 16**).

```

# Cantidad de sitios mutantes diferentes

library(dplyr)

# Calcular la cantidad de sitios mutantes diferentes
cantidad_sitios_mutantes <- df %>%
  distinct(HGVSc) %>% # Extraer los valores únicos del campo HGVSc
  nrow()

print(cantidad_sitios_mutantes)

#-
library(dplyr)
library(ggplot2)

# Calcular la cantidad de mutaciones en cada sitio mutante
mutaciones_por_sitio <- df %>%
  group_by(HGVSc) %>%
  summarize(Cantidad_Mutaciones = n())

```

Cálculo de la cantidad de sitios mutantes únicos y la cantidad de mutaciones por sitio mutante.

```

# Crear el gráfico de dispersión
ggplot(mutaciones_por_sitio, aes(x = HGVSc, y = Cantidad_Mutaciones)) +
  geom_point(color = "blue", size = 2) +
  labs(title = "Mutaciones somáticas de TP53 en diferentes sitios mutantes",
       x = "Sitio Mutante",
       y = "Cantidad de Mutaciones") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))

```

Código utilizado para el diseño del gráfico de dispersión.

Mutaciones somáticas de TP53 en diferentes sitios mutantes

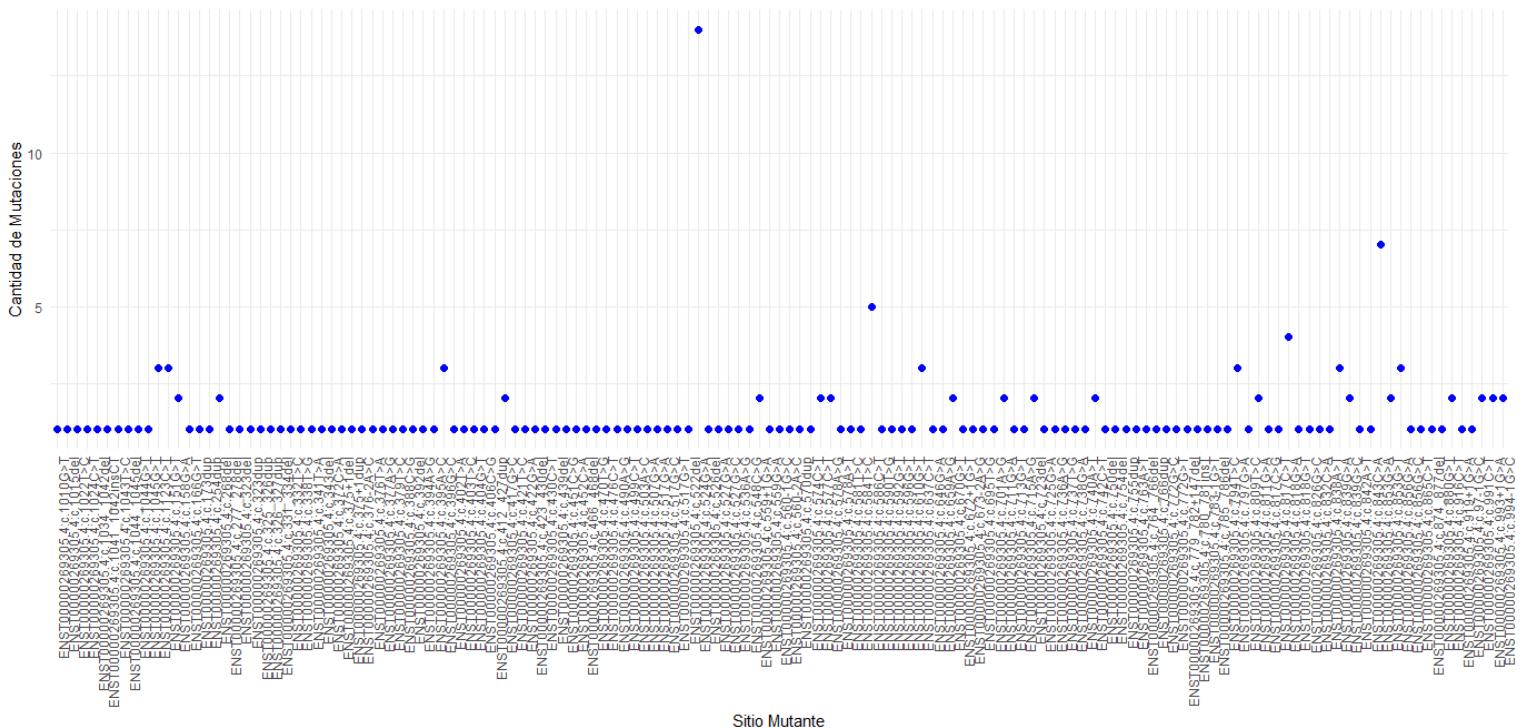


Figura 16: Gráfico de dispersión donde se muestra la distribución de los diferentes sitios mutantes del gen **TP53** en el cáncer de mama.

Posteriormente, se calculó la cantidad de tipos de mutaciones y su frecuencia. Los resultados se representaron mediante un gráfico de torta, tal como se muestra en la **Figura 17**.

```
library(dplyr)

# Calcular los conteos de cada tipo de mutación
conteos_mutaciones <- df %>%
  count(`Mutation Type`)

# Calcular los porcentajes de cada tipo de mutación
porcentajes_mutaciones <- conteos_mutaciones %>%
  mutate(Porcentaje = n / sum(n) * 100)

# Gráfico de torta
library(ggplot2)

colores <- c("blue", "red", "green", "orange", "purple", "pink")

ggplot(porcentajes_mutaciones, aes(x = "", y = Porcentaje, fill = `Mutation Type`, label = paste0(round(Porcentaje, 1), "%"))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  geom_text(position = position_stack(vjust = 0.5), color = "white") + # Agregar etiquetas de porcentaje
  labs(title = "Porcentaje de Mutaciones",
       fill = "Tipo de Mutación",
       y = "Porcentaje") +
  theme_void() +
  theme(legend.position = "right") +
  scale_fill_manual(values = colores)
```

Cálculo de cantidad de tipos de mutaciones, cálculo de frecuencias y diseño de gráfico de torta.

Se detectaron seis tipos de mutaciones distintas. Las mutaciones con cambio de sentido (*Missense*) fueron el tipo más común y abarcaron el 56% de todas las mutaciones observadas. Las mutaciones sin sentido (*Nonsense*) mostraron comprende el 19.5%, seguidas por las delecciones con cambio de marco (*Frameshift deletion*) con un 9%, las mutaciones en el sitio de empalme (*Splice Site*) con un 8%, las inserciones con cambio de marco (*Frameshift insertion*) con un 6%, y, en menor porcentaje, las delecciones sin cambio de marco (*In frame deletion*) con un 1.5%. Estos datos se pueden observar en la **Figura 17**.

Porcentaje de Mutaciones

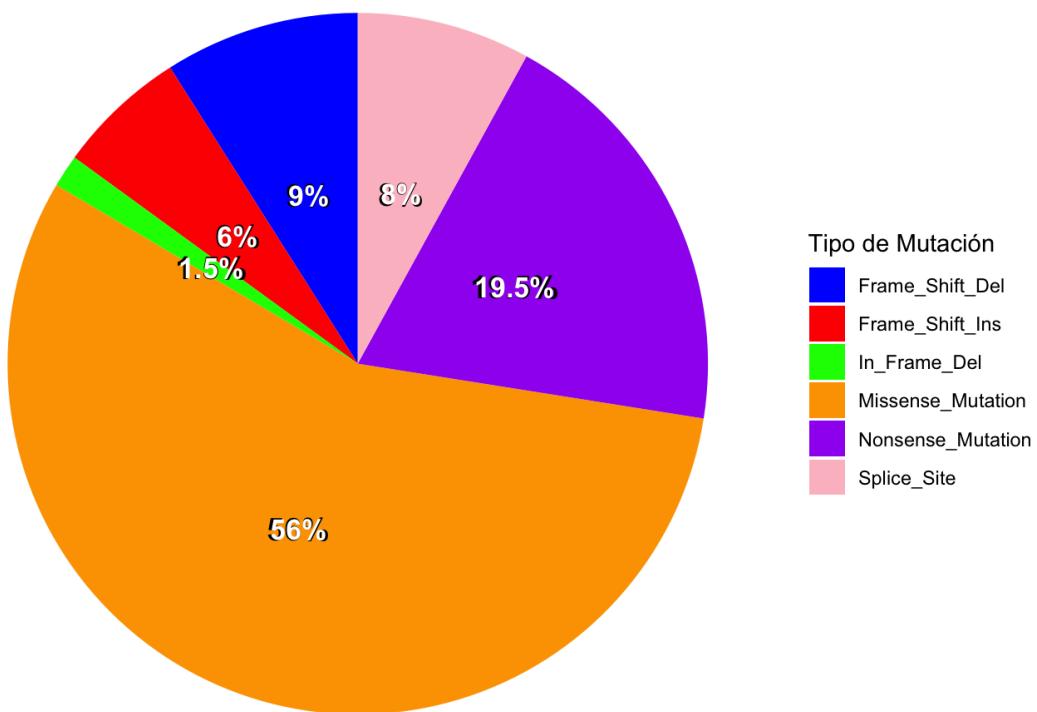


Figura 17: Gráfico de torta donde se muestran la proporción de los diferentes tipos de mutaciones somáticas de *TP53* que se encuentran en el cáncer de mama de los pacientes a analizar.

6.3 Mutaciones en CDI y CLI

Con el fin de analizar los tipos de mutaciones en cada subtipo de cáncer de mama, se realizó un histograma general (**Figura 18**) , utilizando para ello el siguiente código.

#1

```
library(ggplot2)

colores <- c("blue", "red", "green", "orange", "purple", "pink")

ggplot(df, aes(x = `Neoplasm Histologic Type Name`, fill = `Mutation Type`)) +
  geom_bar(position = "dodge") +
  labs(title = "Relación entre tipo de mutaciones somáticas y tipo de cáncer de mama",
       x = "Neoplasm Histologic Type Name",
       y = "Frecuencia") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Rotar etiquetas del eje x
  scale_fill_manual(values = colores) # Asignar colores personalizados
```

Código perteneciente al histograma de la Figura 18.

Los resultados indican que en los tipos más comunes de cáncer de mama, CDI y CL. Las células tumorales presentan patrones de mutación similares en el gen *TP53*, siendo la mutación Missense (cambio de sentido) la más frecuente en ambos casos. Sin embargo, se observa que el tipo CLI muestra una mayor proporción de esta mutación en comparación con el CDI, como se observa en la **Figura 18**. Además, ambos tipos de carcinomas comparten un perfil de frecuencia similar en cuanto a las mutaciones somáticas analizadas.

Relación entre tipo de mutaciones somáticas y tipo de cáncer de mama

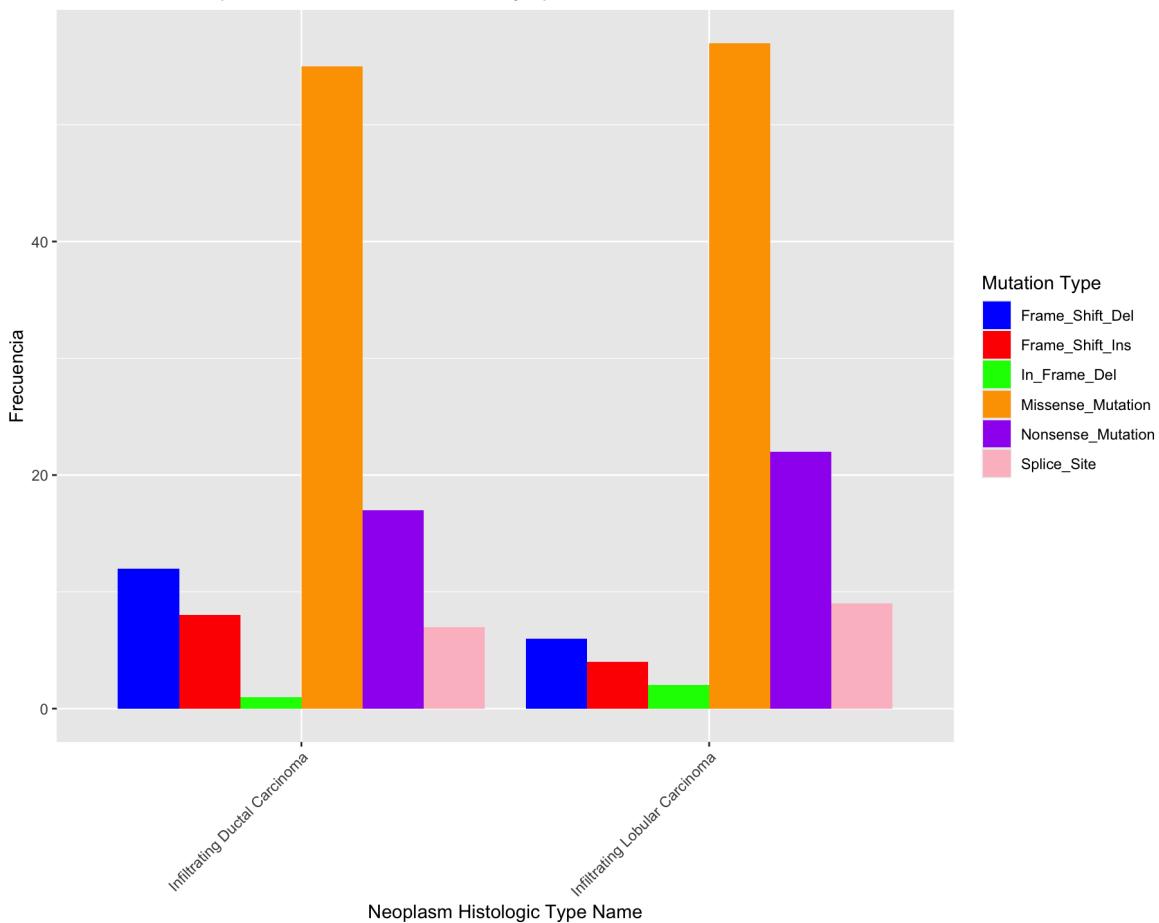


Figura 18: Histograma que compara los distintos tipos de mutaciones somáticas entre los dos subtipos de cáncer de mama: Carcinoma Lobular Infiltrante (CLI) y Carcinoma Ductal Infiltrante (CDI).

A los fines de mostrar por separado las mutaciones presentes, se realizó un análisis específico de los subtipos de cáncer de mama CLI y CDI de forma individual. Para ello, se emplearon los siguientes códigos: el primero corresponde a la **Figura 19**, y el segundo, a la **Figura 20**.

```

----- GRÁFICO DE TORTA
library(ggplot2)
#Frecuencias de cada tipo de mutación
frecuencias <- table(dp3$`Mutation Type`)
# Porcentajes
porcentajes <- prop.table(frecuencias) * 100
# Crear un dataframe con las frecuencias y los porcentajes
df_frecuencias2 <- data.frame(Mutation_Type = names(frecuencias),
                               Frecuencia = as.vector(frecuencias),
                               porcentajes = porcentajes)

colores <- c("blue", "red", "green", "orange", "purple", "pink") # Ejemplo de colores

# Crear el gráfico de torta
ggplot(df_frecuencias2, aes(x = "", y = Frecuencia, fill = Mutation_Type)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste0(round(porcentajes, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black",      # Borde negro
            fontface = "bold",
            size = 5,             # Tamaño mayor para el borde
            label.size = 0.8,
            #nudge_x = 0.1,        # Desplazamiento para simular el borde
            #nudge_y = 0.1,        # Desplazamiento para simular el borde
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(porcentajes, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white",     # Texto blanco
            fontface = "bold",
            size = 4.5,           # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Infiltrating Lobular Carcinoma",
       fill = "Mutation Type") +
  theme_void() +
  scale_fill_manual(values = colores)

```

Código perteneciente al gráfico de torta de la Figura 19.

```

#----- GRÁFICO DE TORTA
library(ggplot2)
# Calcular las frecuencias de cada tipo de mutación
frecuencias <- table(dp1$`Mutation Type`)
# Porcentajes
porcentajes <- prop.table(frecuencias) * 100
# Crear dataframe con las frecuencias y los porcentajes
df_frecuencias1 <- data.frame(Mutation_Type = names(frecuencias),
                                Frecuencia = as.vector(frecuencias),
                                porcentajes = porcentajes)

# Definir los colores deseados
colores <- c("blue", "red", "green", "orange", "purple", "pink") # Ejemplo de colores
# Crear el gráfico de pastel con etiquetas de porcentaje y colores personalizados
ggplot(df_frecuencias1, aes(x = "", y = Frecuencia, fill = Mutation_Type)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste0(round(porcentajes, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black", # Borde negro
            fontface = "bold",
            size = 5, # Tamaño mayor para el borde
            label.size = 0.8,
            #nudge_x = 0.1, # Desplazamiento para simular el borde
            #nudge_y = 0.1, # Desplazamiento para simular el borde
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(porcentajes, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white", # Texto blanco
            fontface = "bold",
            size = 4.5, # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Infiltrating Ductal Carcinoma",
       fill = "Mutation Type") +
  theme_void() +
  scale_fill_manual(values = colores)

```

Código perteneciente al gráfico de torta de la Figura 20.

El análisis de las mutaciones en el gen *TP53* en células tumorales de CLI presenta una distribución variada de los tipos de mutación. El gráfico muestra que predominan las mutaciones de tipo “mutación con cambio de sentido” (*Missense Mutation*), la que representa el 57% de las mutaciones identificadas. Estas mutaciones implican un cambio en un solo nucleótido que resulta en la sustitución de un aminoácido en la proteína p53. Por otro lado, las mutaciones de tipo la “mutaciones sin sentido” (*Nonsense Mutation*) representan el 22% de las mutaciones identificadas, lo que indica la generación de un codón de parada prematuro que interrumpe la síntesis de la proteína p53. Las mutaciones de tipo Sitio de Empalme (*Splice Site*), “Deleción con cambio de marco” (*Frameshift Deletion*), Inserción con cambio de marco (*Frameshift Insertion*) y “Deleción sin cambio de marco” (*In Frame Deletion*) se observan en frecuencias más bajas, 21% en conjunto.

Infiltrating Lobular Carcinoma

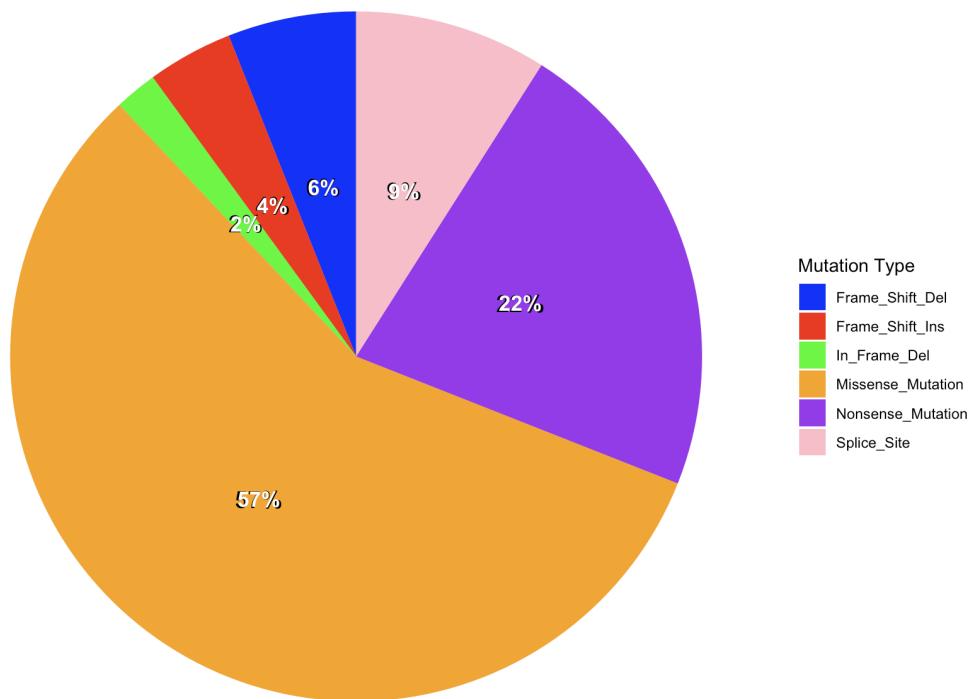


Figura 19: Gráfico de torta que muestran la proporción de los diferentes tipos de mutaciones somáticas de *TP53* que se encuentran en el tipo de cáncer de mama Carcinoma Lobular Infiltrante (CLI).

El análisis de las mutaciones en el gen *TP53* en células tumorales de CDI también presenta una distribución variada de los tipos de mutaciones somáticas. El gráfico muestra que las mutaciones de tipo *Missense Mutation* son las que predominan, representando el 55% de las mutaciones identificadas. Estas mutaciones implican un cambio en un solo nucleótido que resulta en la sustitución de un aminoácido en la proteína p53, como se mencionó previamente. Por otro lado, las mutaciones de tipo la “mutaciones sin sentido” (*Nonsense Mutation*) representan el 17% de las mutaciones identificadas, lo que indica la generación de un codón de parada prematuro que interrumpe la síntesis de la proteína p53. Las mutaciones de tipo Sitio de Empalme (*Splice Site*), “Delección con cambio de marco” (*Frameshift Deletion*), Inserción con cambio de marco (*Frameshift Insertion*) y “Delección sin cambio de marco” (*In Frame Deletion*) presentan porcentajes más bajos, representando en conjunto el 28%.

Infiltrating Ductal Carcinoma

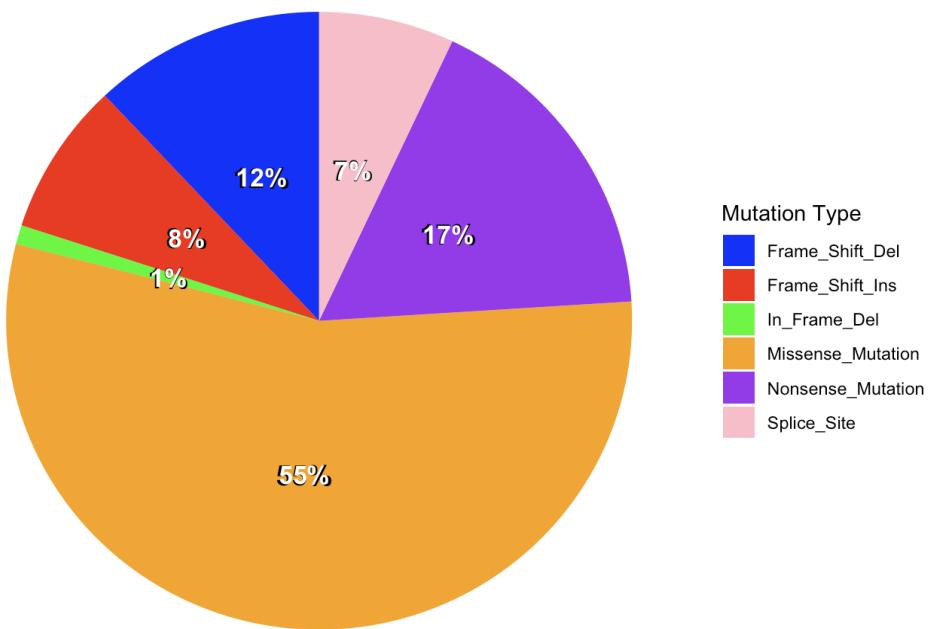


Figura 20: Gráfico de torta que muestran la proporción de los diferentes tipos de mutaciones somáticas de *TP53* que se encuentran en el tipo de cáncer de mama Carcinaoma Ductal Infiltrante (CDI).

6.4 Perfil Tumoral Molecular

Los perfiles tumorales moleculares del cáncer de mama hacen referencia a la expresión de los receptores de estrógeno (ER), receptores de progesterona (PR), y el receptor 2 del factor de crecimiento epidérmico humano (HER2). En este apartado, se presenta un análisis detallado del perfil tumoral molecular de los pacientes abarcados en este estudio.

6.4.1 Análisis del receptor HER2

Para continuar los análisis, se evaluó la distribución porcentual de la expresión de HER2 en las muestras de pacientes con cáncer de mama, calculando los porcentajes de casos con resultados positivos y negativos. Posteriormente, los resultados fueron representados mediante un histograma (**Figura 21**).

```

#----HER2-----
library(dplyr)
library(ggplot2)

# Calcular los porcentajes de cada categoría en IHC-HER2
porcentajes_IHC_HER2 <- df %>%
  group_by(`IHC-HER2`) %>%
  summarize(Conteo = n()) %>%
  mutate(Porcentaje = Conteo / sum(Conteo) * 100)

# Gráfico de barras con porcentajes
ggplot(porcentajes_IHC_HER2, aes(x = `IHC-HER2`, y = Porcentaje, fill = `IHC-HER2`)) +
  geom_bar(stat = "identity", width = 0.7) +
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")), vjust = -0.5) +
  labs(title = "Porcentajes de HER2",
       x = "HER2 Status",
       y = "Porcentaje") +
  scale_fill_manual(values = c("Positive" = "skyblue", "Negative" = "salmon")) +
  theme_minimal()

```

Código utilizado para el cálculo y la representación en histograma de los casos positivos y negativos de expresión de HER2.

Durante el análisis de los datos, se identificó que el 78.5% de los casos exhiben un resultado negativo para HER2, lo que indica que las células tumorales no expresan niveles elevados de la proteína HER2. Además, se observó que el 16% de los casos muestran un resultado positivo para HER2, lo que significa que las células tumorales sí expresan la proteína HER2. Estos resultados destacan una diferencia evidente, con una predominancia de casos HER2 negativos, como se muestra en la **Figura 21**. Es importante señalar que cuando los resultados son equívocos o ambiguos, se requiere el uso de métodos adicionales para definir su clasificación como positiva o negativa.

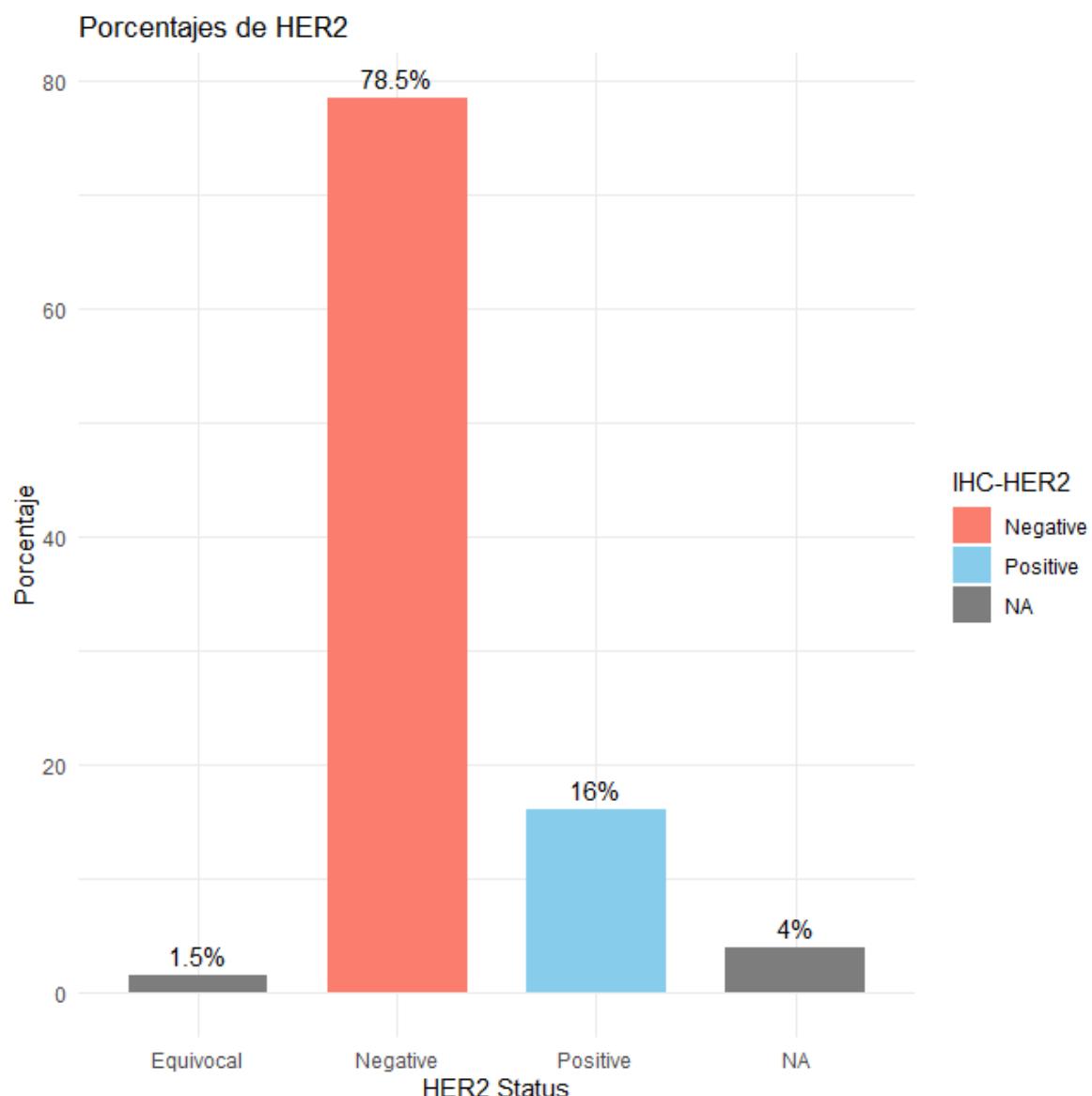


Figura 21: Histograma donde se muestra la distribución porcentual de la expresión de HER en las muestras de pacientes con cáncer de mama.

6.4.2 Análisis del receptor de Estrógeno (ER)

Aquí, se evaluó la distribución porcentual de la expresión de ER en las muestras de pacientes con cáncer de mama estudiadas, calculando los porcentajes de casos con resultados positivos y negativos. Posteriormente, los resultados fueron representados mediante un histograma (**Figura 22**).

```

#----ER-----
library(dplyr)
library(ggplot2)

# Calcular los porcentajes de cada categoría en IHC-ER
porcentajes_IHC_ER <- df %>%
  group_by(`ER Status By IHC`) %>%
  summarize(Conteo = n()) %>%
  mutate(Porcentaje_ER = Conteo / sum(Conteo) * 100)

# Crear el gráfico de barras con porcentajes
ggplot(porcentajes_IHC_ER, aes(x = `ER Status By IHC`, y = Porcentaje_ER, fill = `ER Status By IHC`)) +
  geom_bar(stat = "identity", width = 0.7) +
  geom_text(aes(label = paste0(round(Porcentaje_ER, 1), "%")), vjust = -0.5) +
  labs(title = "Porcentajes de ER",
       x = "ER Status",
       y = "Porcentaje") +
  scale_fill_manual(values = c("Positive" = "skyblue", "Negative" = "salmon")) +
  theme_minimal()

```

Código utilizado para el cálculo y la representación en histograma de los casos positivos y negativos de expresión de ER.

Durante el análisis de los datos, se identificó que el 58% de los casos exhiben un resultado positivo para ER, lo que indica que las células tumorales expresan niveles elevados del receptor ER. Además, se observó que el 40.5% de los casos muestran un resultado negativo para ER, lo que significa que las células tumorales no expresan al receptor ER. Estos resultados destacan una diferencia leve, con una predominancia de casos ER positivos, como se muestra en la **Figura 22**. Esto se condice con los datos bibliográficos mencionados previamente, que reportan que la mayoría de casos de cáncer de mama, son ER+.

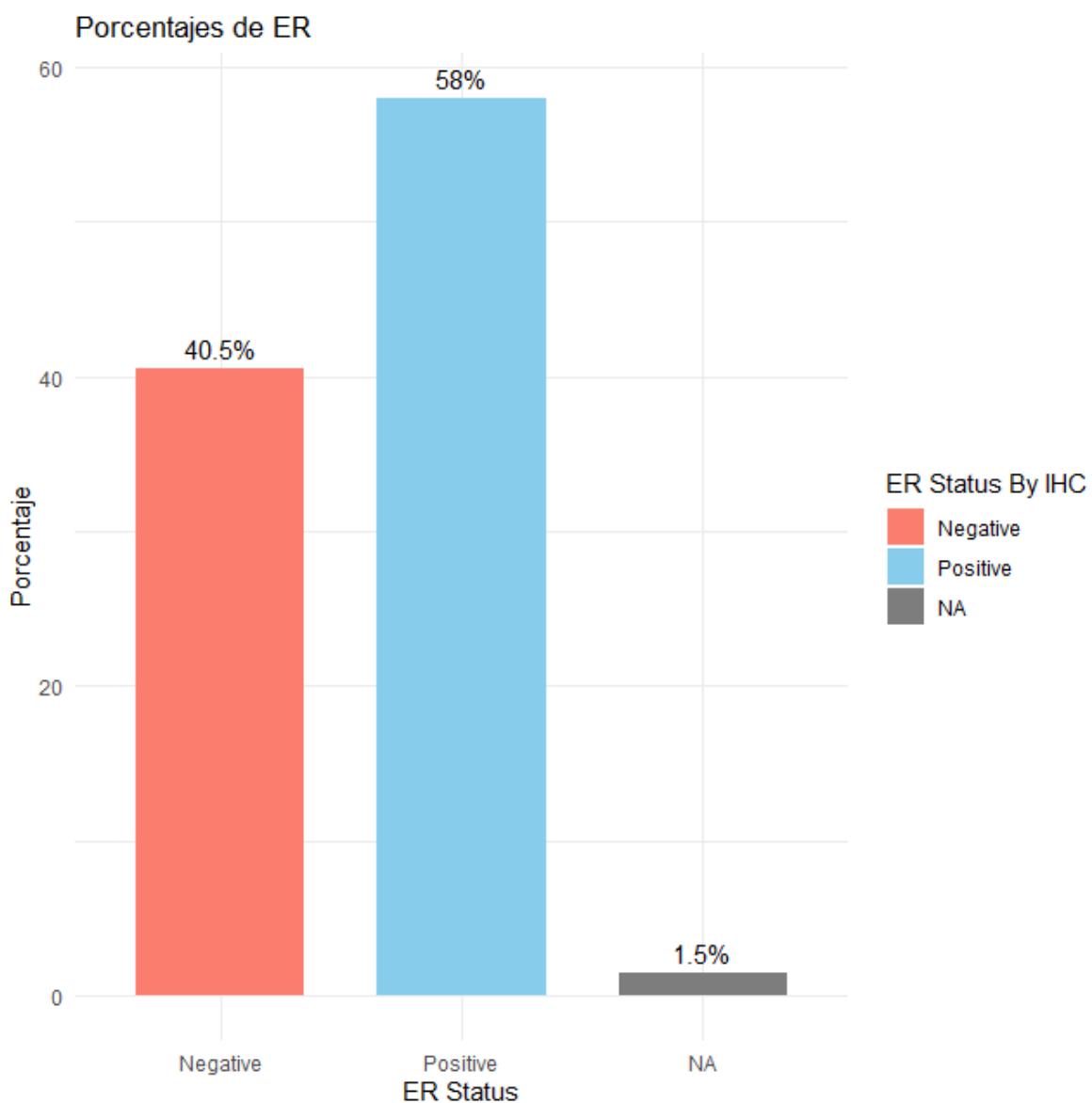


Figura 22: Histograma donde se muestra la distribución porcentual de la expresión de ER en las muestras de pacientes con cáncer de mama.

6.4.3 Análisis del receptor de Progesterona (PR)

A continuación, se evaluó la distribución porcentual de la expresión de PR en las muestras de pacientes con cáncer de mama, calculando los porcentajes de casos con resultados positivos y negativos. Posteriormente, los resultados fueron representados mediante un histograma (**Figura 23**).

```

#----PR-----
library(dplyr)
library(ggplot2)

# Porcentajes de cada categoría en IHC-ER
porcentajes_IHC_PR <- df %>%
  group_by(`PR status by ihc`) %>%
  summarize(Conteo = n()) %>%
  mutate(Porcentaje_PR = Conteo / sum(Conteo) * 100)

# Gráfico de barras con porcentajes
ggplot(porcentajes_IHC_PR, aes(x = `PR status by ihc`, y = Porcentaje_PR, fill = `PR status by ihc`)) +
  geom_bar(stat = "identity", width = 0.7) +
  geom_text(aes(label = paste0(round(Porcentaje_PR, 1), "%")), vjust = -0.5) +
  labs(title = "Porcentajes de PR",
       x = "PR Status",
       y = "Porcentaje") +
  scale_fill_manual(values = c("Positive" = "skyblue", "Negative" = "salmon")) +
  theme_minimal()

```

Código utilizado para el cálculo y la representación en histograma de los casos positivos y negativos de expresión de PR.

Durante el análisis de los datos, se identificó que el 46.5% de los casos analizados exhiben un resultado positivo para el receptor de progesterona (PR), lo que indica que las células tumorales expresan niveles elevados de este receptor. Además, se observó que el 46% de los casos muestran un resultado negativo para el PR, lo que significa que las células tumorales no expresan este receptor. Estos resultados no muestran una diferencia importante, como se muestra en la **Figura 23**.

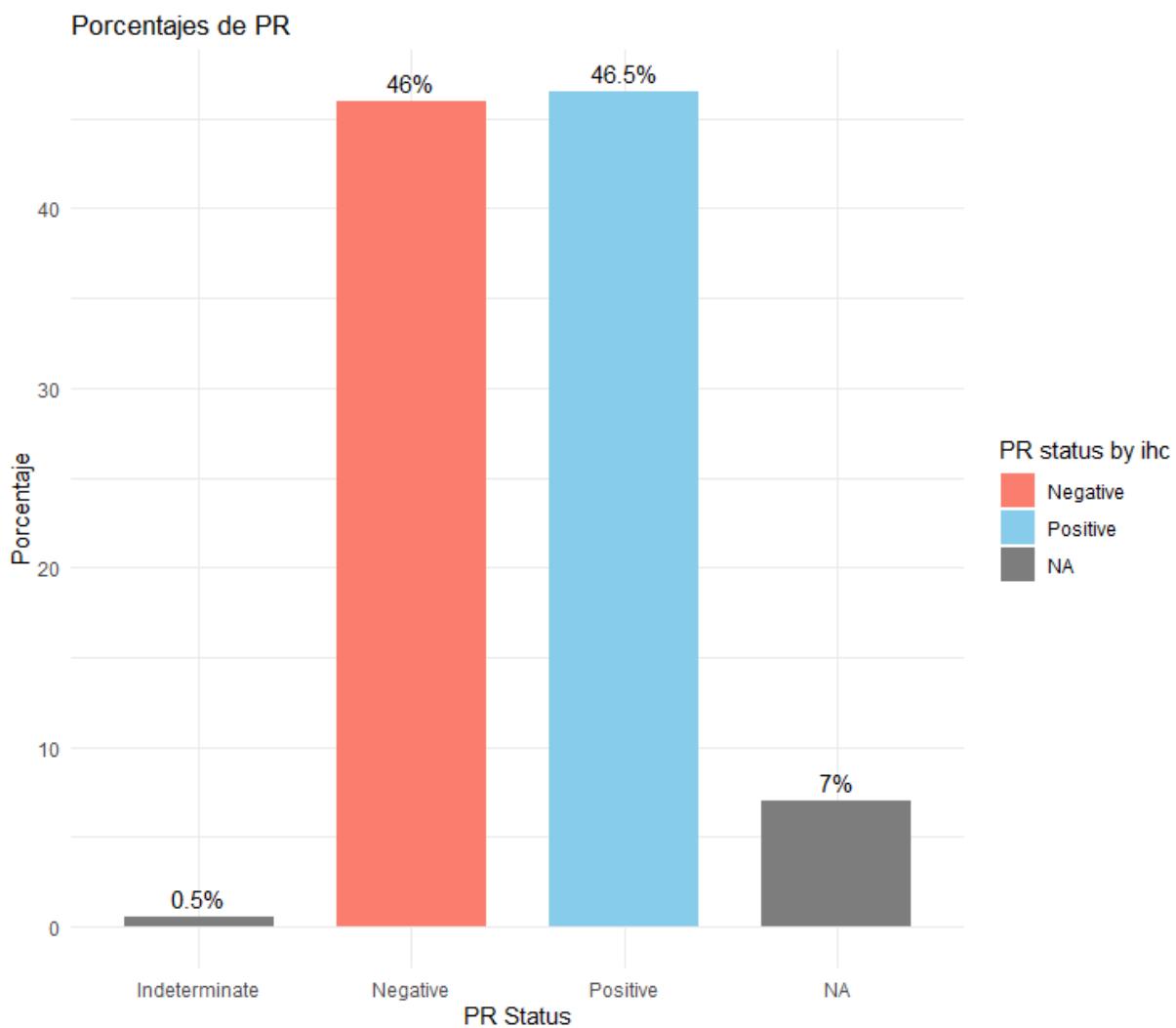


Figura 23: Histograma donde se muestra la distribución porcentual de la expresión de PR en las muestras de pacientes con cáncer de mama.

6.5 Análisis de los Perfiles Tumorales Moleculares del Cáncer de mama

Para comprender mejor los distintos tipos de perfiles moleculares en el cáncer de mama, se establecieron parámetros específicos que permitieran clasificar y analizar las muestras de cada paciente. Estos parámetros facilitan el registro de las características particulares de cada perfil molecular, permitiendo una evaluación detallada. A continuación, los resultados de este análisis se presentan en un histograma (**Figura 24**).

```

# Filtrar los datos para cada grupo
grupo_LA <- df %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Positive", `PR status by ihc` == "Negative")

grupo_LB <- df %>%
  filter(
    `ER Status By IHC` == "Positive" & `IHC-HER2` == "Positive" | 
    `ER Status By IHC` == "Positive" & `IHC-HER2` == "Negative" & `PR status by ihc` == "Negative"
  )

grupo_HE <- df %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Negative", `PR status by ihc` == "Negative")

grupo_TN <- df %>%
  filter(`IHC-HER2` == "Negative", `ER Status By IHC` == "Negative", `PR status by ihc` == "Negative")

# Porcentajes para cada grupo
porcentajes_LA <- nrow(grupo_LA) / nrow(df) * 100
porcentajes_LB <- nrow(grupo_LB) / nrow(df) * 100
porcentajes_HE <- nrow(grupo_HE) / nrow(df) * 100
porcentajes_TN <- nrow(grupo_TN) / nrow(df) * 100

# Crear un dataframe con los porcentajes
porcentajes_df <- data.frame(
  Grupo = c("LA", "LB", "HE", "TN"),
  Porcentaje = c(porcentajes_LA, porcentajes_LB, porcentajes_HE, porcentajes_TN)
)

# Gráfico de barras
ggplot(porcentajes_df, aes(x = Grupo, y = Porcentaje, fill = Grupo)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")), vjust = -0.5) +
  labs(title = "Distribución Perfiles Moleculares",
       x = "Grupo",
       y = "Porcentaje") +
  scale_fill_manual(values = c("LA" = "skyblue", "LB" = "salmon", "HE" = "yellow", "TN" = "green")) +
  theme_minimal()

```

Código que muestra la definición de los nuevos parámetros, el cálculo de los perfiles moleculares y su representación en un histograma.

A continuación, se analizaron los distintos perfiles tumorales y moleculares en el total de los datos. Como se observa en la **Figura 24**, entre las muestras analizadas, predomina el perfil Triple Negativo con un 27.5%. Le sigue el perfil Luminal B con un 9.5%. Los tumores Luminal B, son menos agresivos que los Triple Negativos. El perfil HER2-enriquecido se presenta en el 6% de los casos. Por último, el menos predominante es el perfil Luminal A con un porcentaje de 3%.

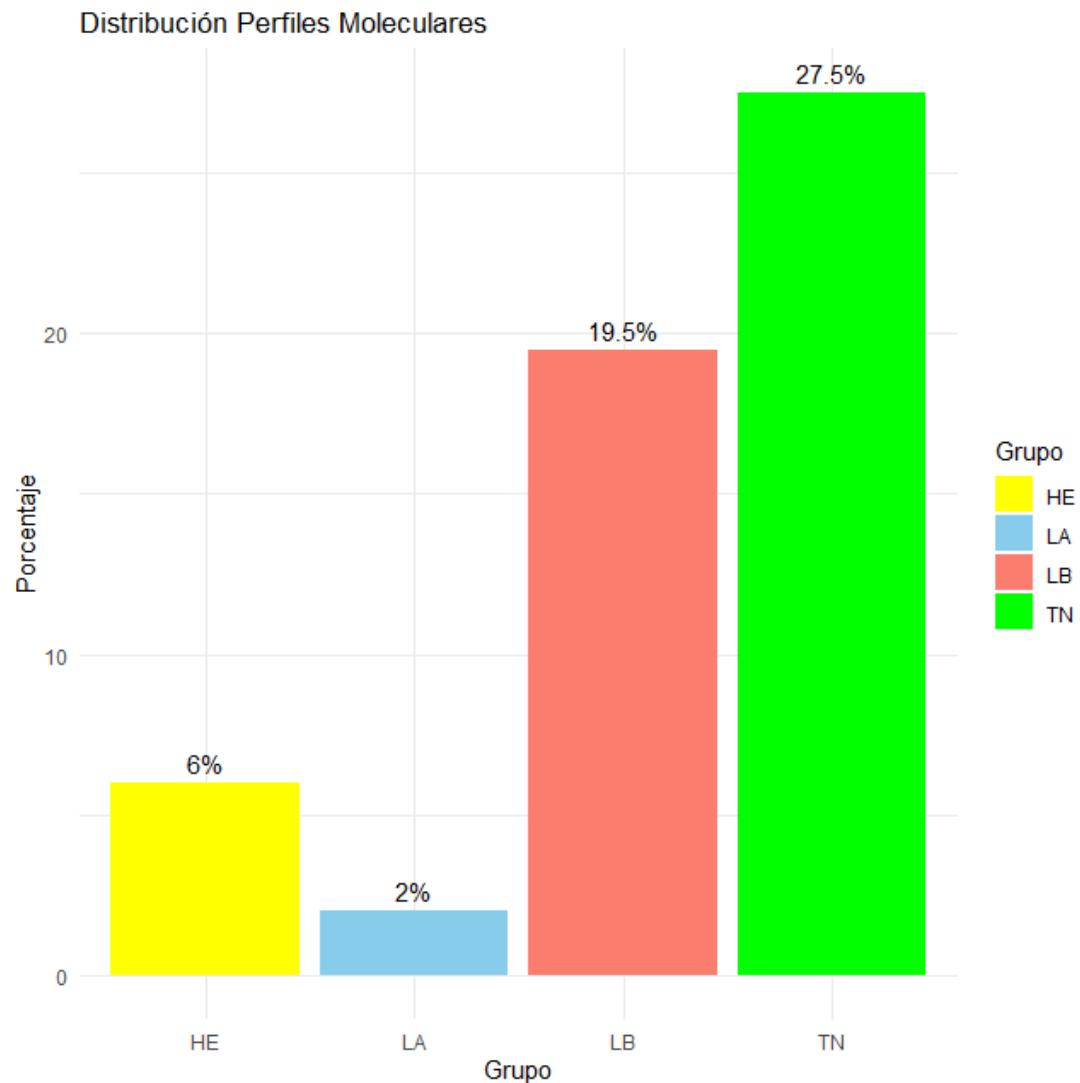


Figura 24: Histograma donde se muestra la distribución porcentual de la expresión de los distintos perfiles tumorales y moleculares en las muestras de pacientes con cáncer de mama. LA: Lumina A, LB: Lumina B, HE: HER2-enriquecido y TN: TRIPLE negativo.

6.6 Perfiles Tumorales Moleculares en CDI

Para determinar qué perfil molecular predomina en el CDI, se aplicó el código que define los nuevos parámetros, registrando las condiciones de cada perfil molecular y separando las muestras de los pacientes según su tipo de perfil y cáncer CDI. Los resultados se representaron en el histograma que se muestra en la **Figura 25**.

```

#-----er y her2 y pr
library(dplyr)
library(ggplot2)

# Filtrar los datos para cada grupo
grupo_LA_1 <- dp1 %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Positive", `PR status by ihm` == "Negative")
grupo_LB_1 <- dp1 %>%
  filter(
    (`ER Status By IHC` == "Positive" & `IHC-HER2` == "Positive") |
    (`ER Status By IHC` == "Positive" & `IHC-HER2` == "Negative" & `PR status by ihm` == "Negative")
  )
grupo_HE_1 <- dp1 %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Negative", `PR status by ihm` == "Negative")
grupo_TN_1 <- dp1 %>%
  filter(`IHC-HER2` == "Negative", `ER Status By IHC` == "Negative", `PR status by ihm` == "Negative")

# Calcular los porcentajes para cada grupo
porcentajes_LA_1 <- nrow(grupo_LA_1) / nrow(dp1) * 100
porcentajes_LB_1 <- nrow(grupo_LB_1) / nrow(dp1) * 100
porcentajes_HE_1 <- nrow(grupo_HE_1) / nrow(dp1) * 100
porcentajes_TN_1 <- nrow(grupo_TN_1) / nrow(dp1) * 100

# Crear dataset con porcentajes
porcentajes_dp1 <- data.frame(
  Grupo = c("LA", "LB", "HE", "TN"),
  Porcentaje_1 = c(porcentajes_LA_1, porcentajes_LB_1, porcentajes_HE_1, porcentajes_TN_1)
)

# Gráfico de barras
ggplot(porcentajes_dp1, aes(x = Grupo, y = Porcentaje_1, fill = Grupo)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = paste0(round(Porcentaje_1, 1), "%")), vjust = -0.5) +
  labs(title = "Distribución Perfiles Moleculares en CDI",
       x = "Grupo",
       y = "Porcentaje") +
  scale_fill_manual(values = c("LA" = "skyblue", "LB" = "salmon", "HE" = "yellow", "TN" = "green")) +
  theme_minimal()

```

Código que muestra la definición de los nuevos parámetros, el cálculo de los perfiles moleculares en el tipo de cáncer CDI y su representación en un histograma.

El perfil que predomina entre las muestras estudiadas de CDI es el perfil Triple Negativo, abarcando un 40% de los casos. Este tipo de cáncer es conocido por su agresividad y la falta de receptores hormonales y HER2, lo que limita las opciones de tratamiento a la quimioterapia. Le sigue el perfil Luminal B con un porcentaje de 18%. Los tumores Luminal B tienden a tener una menor respuesta a la terapia hormonal en comparación con Luminal A, pero pueden beneficiarse de tratamientos adicionales como la quimioterapia. El perfil HER2-enriquecido representa un 7% de los casos. Este subtipo se caracteriza por la sobreexpresión del receptor HER2 y, aunque es agresivo, responde bien a terapias dirigidas como *trastuzumab*. En menor proporción, con un 3%, se encuentra el perfil Luminal A, conocido por ser menos agresivo y responder bien a la terapia hormonal, lo que generalmente se traduce en un mejor pronóstico.

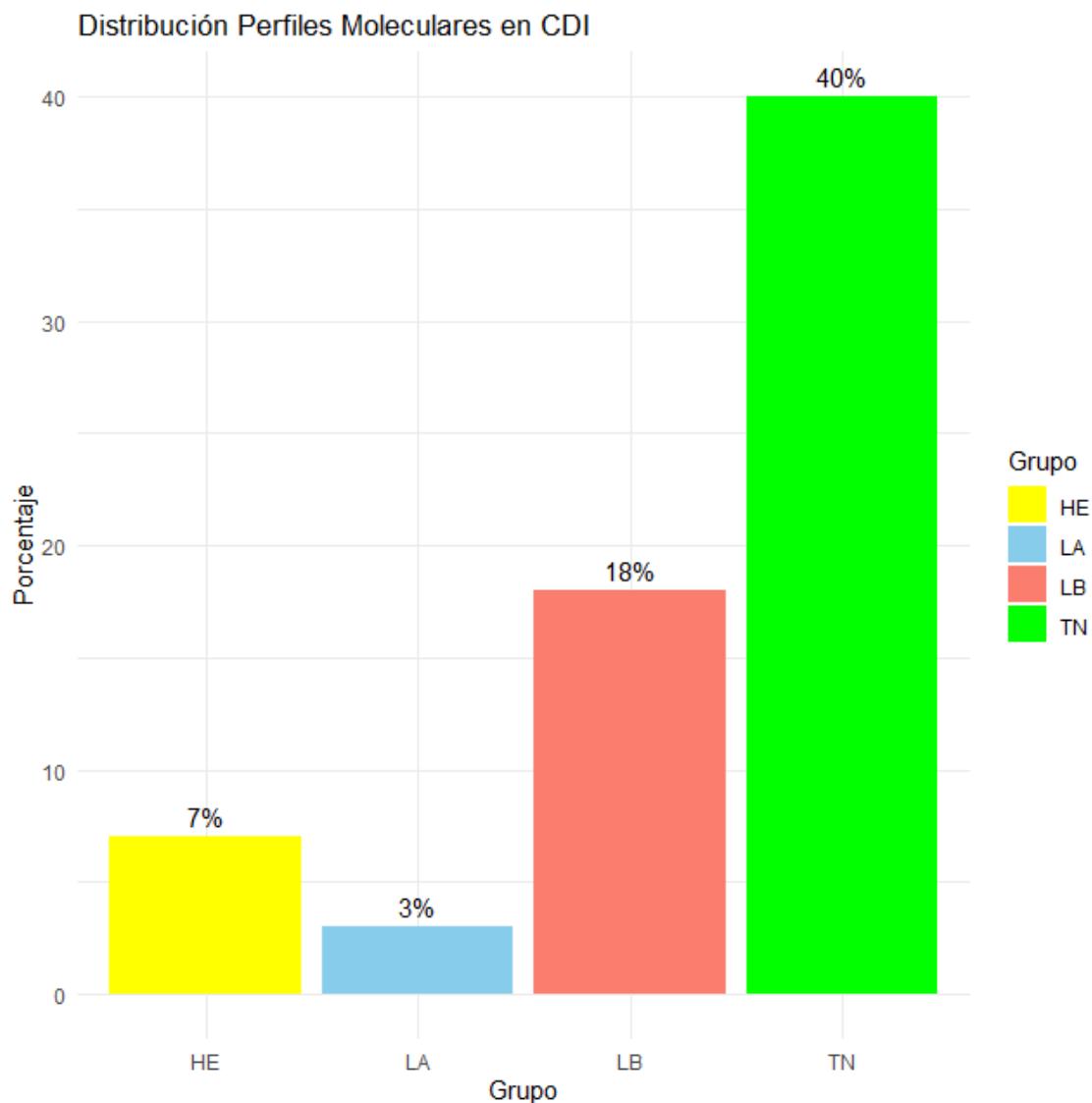


Figura 25: Histograma donde se muestra la distribución porcentual de la expresión de los distintos perfiles tumorales y moleculares en CDI. LA: Lumina A, LB: Lumina B, HE: HER2-enriquecido y TN: TRIPLE negativo.

6.7 Perfiles Tumorales Moleculares en CLI

El mismo análisis se llevó a cabo para determinar qué perfil molecular predomina en CLI. Para ello se realizó un proceso similar a CDI donde se aplicó el código que define nuevos parámetros registrando las condiciones de cada perfil molecular y separando las muestras de los pacientes según su tipo de perfil y cáncer CLI. Los resultados se representaron en un histograma, que se muestra en la **Figura 26**.

```

#-----er y her2 y pr
library(dplyr)
library(ggplot2)

# Filtrar los datos para cada grupo
grupo_LA_3 <- dp3 %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Positive", `PR status by ihm` == "Negative")
grupo_LB_3 <- dp3 %>%
  filter(
    `ER Status By IHC` == "Positive" & `IHC-HER2` == "Positive") |
    `ER Status By IHC` == "Positive" & `IHC-HER2` == "Negative" & `PR status by ihm` == "Negative"
  )
grupo_HE_3 <- dp3 %>%
  filter(`IHC-HER2` == "Positive", `ER Status By IHC` == "Negative", `PR status by ihm` == "Negative")
grupo_TN_3 <- dp3 %>%
  filter(`IHC-HER2` == "Negative", `ER Status By IHC` == "Negative", `PR status by ihm` == "Negative")

# Calcular los porcentajes para cada grupo
porcentajes_LA_3 <- nrow(grupo_LA_3) / nrow(dp3) * 100
porcentajes_LB_3 <- nrow(grupo_LB_3) / nrow(dp3) * 100
porcentajes_HE_3 <- nrow(grupo_HE_3) / nrow(dp3) * 100
porcentajes_TN_3 <- nrow(grupo_TN_3) / nrow(dp3) * 100

# Crear un dataframe con los porcentajes
porcentajes_dp3 <- data.frame(
  Grupo = c("LA", "LB", "HE", "TN"),
  Porcentaje_3 = c(porcentajes_LA_3, porcentajes_LB_3, porcentajes_HE_3, porcentajes_TN_3)
)

# Crear el gráfico de barras
ggplot(porcentajes_dp3, aes(x = Grupo, y = Porcentaje_3, fill = Grupo)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = paste0(round(Porcentaje_3, 1), "%")), vjust = -0.5) +
  labs(title = "Distribución Perfiles Moleculares en CLI",
       x = "Grupo",
       y = "Porcentaje") +
  scale_fill_manual(values = c("LA" = "skyblue", "LB" = "salmon", "HE" = "yellow", "TN" = "green")) +
  theme_minimal()

```

Código que muestra la definición de los nuevos parámetros, el cálculo de los perfiles moleculares en el tipo de cáncer CLI y su representación en un histograma.

Los resultados obtenidos se pueden observar en la **Figura 26**. El perfil que predomina en CLI es el perfil Luminal B con un 21%. Este perfil se caracteriza por una combinación de receptores hormonales positivos y una mayor proliferación celular, lo que puede implicar un tratamiento que combine terapia hormonal y quimioterapia. El siguiente perfil en frecuencia es el Triple Negativo, con un 15%. El perfil HER2-enriquecido se presenta en un 5% de los casos. Finalmente, el perfil Luminal A es el menos común dentro de los tumores CLI, representando solo el 1%.

Distribución Perfiles Moleculares en CLI

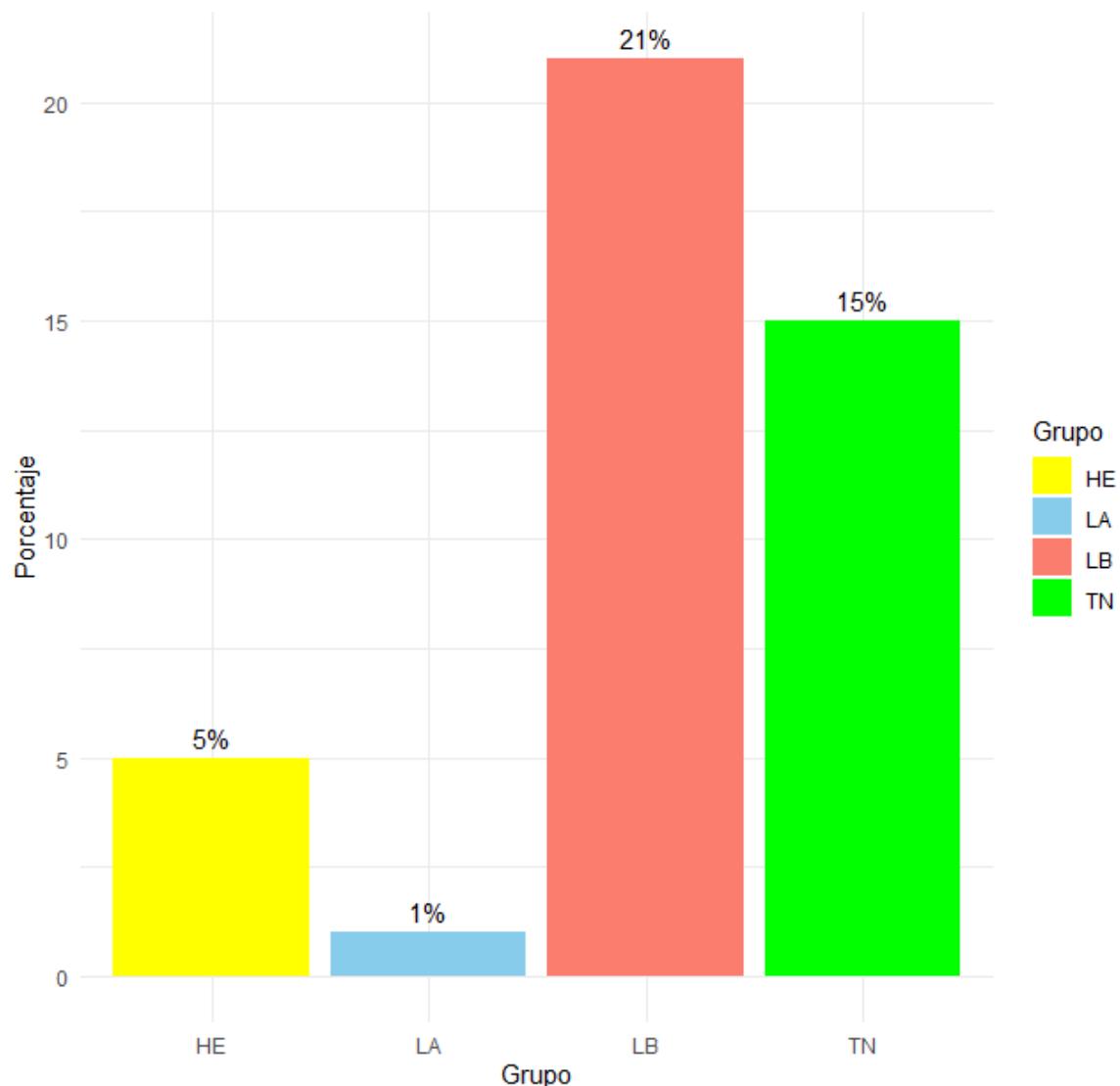


Figura 26: Histograma donde se muestra la distribución porcentual de la expresión de los distintos perfiles tumorales y moleculares en CLI. LA: Lumina A, LB: Lumina B, HE: HER2-enriquecido y TN: TRIPLE negativo.

6.8 Perfiles Tumorales Moleculares CDI vs. CLI

Para poder mostrar un análisis más comparativo, se realizó un histograma donde se muestra la distribución porcentual de la expresión de los distintos perfiles tumorales diferenciados por tipo de cáncer, según sea CDI y CLI.

```

#----- Gráfico Barras Acumulado ----|
library(ggplot2)

data <- data.frame(
  Tipo_Cancer = rep(c("IDC", "ILC"), each = 4),
  Estado = rep(c("LA", "LB", "HE", "TN"), times = 2),
  Porcentaje = c(3, 18, 7, 40, 1, 21, 5, 15)
)

# Creo el gráfico de barras apiladas
ggplot(data, aes(x = Estado, y = Porcentaje, fill = Tipo_Cancer)) +
  geom_bar(stat = "identity", position = "stack") +
  geom_text(aes(label = paste0(Porcentaje, "%")),
            position = position_stack(vjust = 0.5),
            size = 3) +
  labs(title = "Distribución acumulada de perfiles moleculares por tipo de cáncer",
       x = "Estado",
       y = "Porcentaje") +
  scale_fill_manual(values = c("IDC" = "skyblue", "ILC" = "salmon")) +
  theme_minimal()

```

Código que genera un histograma acumulado mostrando los porcentajes de los perfiles moleculares en Carcinoma Ductal Infiltrante (CDI) y Carcinoma Lobular Infiltrante (CLI).

A partir de la **Figura 27**, que muestra los porcentajes de expresión de los perfiles moleculares *Luminal A*, *Luminal B*, *HER2-enriquecido* y *Triple Negativo* diferenciados por tipos de cáncer, CDI y CLI, se puede observar que en el CDI predomina el perfil molecular *Triple Negativo*, representando alrededor del 40% de los casos; mientras que en el CLI este perfil se presenta en el 15%. En el CLI predomina el perfil *Luminal B* en un 21%, en comparación con el CDI donde este perfil representa el 18% de los casos. Además, se observa que el perfil menos frecuente en ambos tipos de cáncer es el *Luminal A*, con un 3% en CDI y un 1% en CLI.

Es importante destacar que se excluyeron del análisis los estados ambiguos y vacíos, ya que no son imprescindibles para la interpretación de los resultados. Estos hallazgos sugieren una diferencia significativa en la distribución de los perfiles moleculares entre los tipos de cáncer estudiados, lo que refleja las importantes implicancias de los perfiles moleculares no sólo para el diagnóstico, sino también para el tratamiento idóneo para los pacientes.

Distribución acumulada de perfiles moleculares por tipo de cáncer

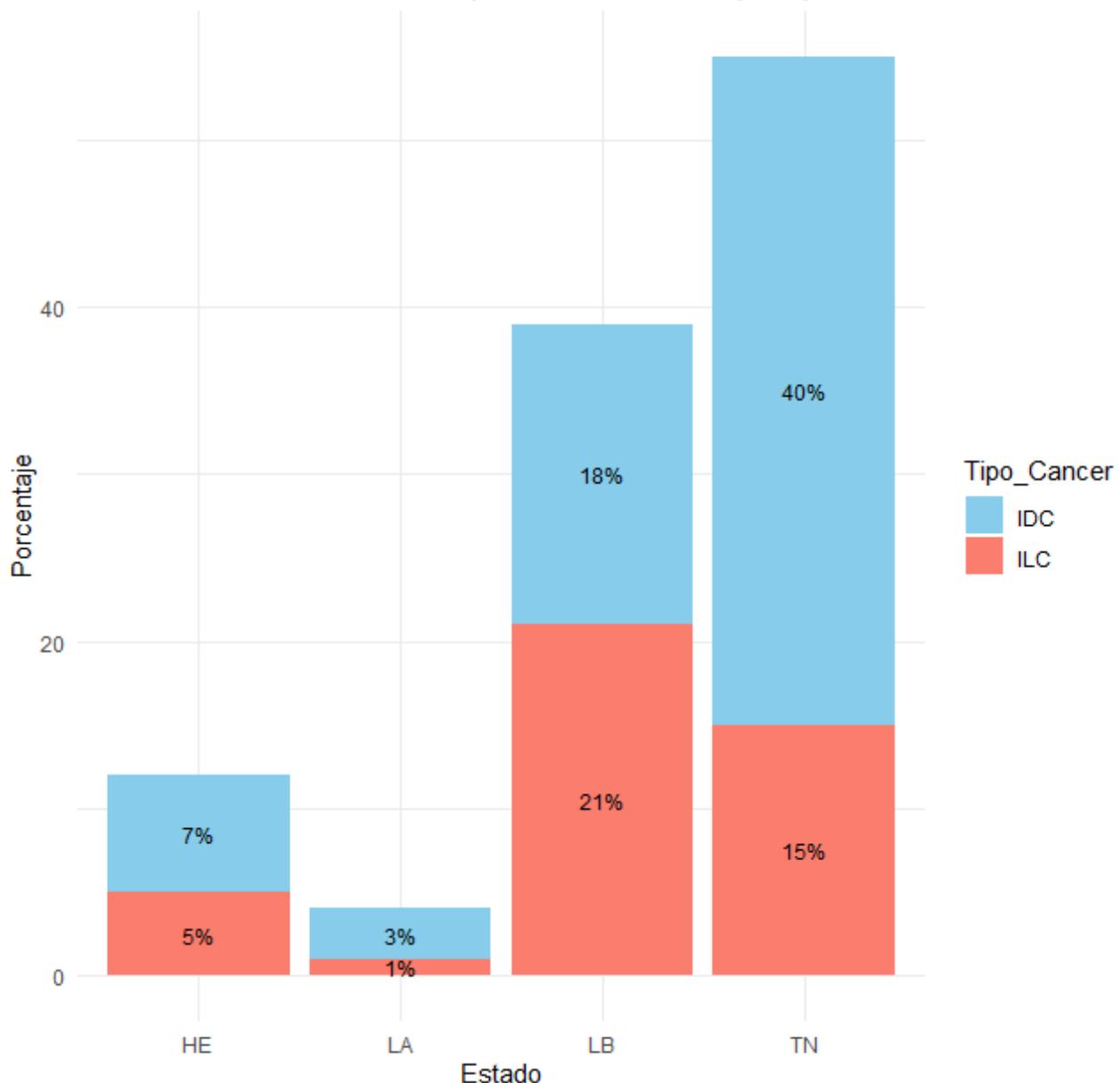


Figura 27: Histograma donde se muestra la distribución porcentual de la expresión de los distintos perfiles tumorales diferenciados por tipo de cáncer IDC y ILC. LA: Lumina A, LB: Lumina B, HE: HER2-enriquecido y TN: TRIPLE negativo.

6.8.1 Frecuencia de las Mutaciones en los Distintos Tipos de Perfiles Moleculares

A continuación, se realizó un análisis para determinar si existe alguna mutación predominante en cada perfil molecular (HE, LA, LB, TN). Para ello, se reutilizó el código que agrupa las muestras según sus respectivos perfiles moleculares, y se calcularon las frecuencias de mutación en cada uno de ellos. Como resultado, se generaron cuatro gráficos circulares que ilustran dichas frecuencias (**Figura 28**, **Figura 29**, **Figura 30**, **Figura 31**).

```

library(ggplot2)

colores <- c("red", "green", "orange", "purple", "pink")

ggplot(porcentajes_mutaciones_HE, aes(x = "", y = Porcentaje, fill = `Mutation Type`, label = paste0(round(Porcentaje, 1), "%"))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  # Capa para el borde negro
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black",      # Borde negro
            fontface = "bold",
            size = 7,             # Tamaño mayor para el borde
            label.size = 1,
            #nudge_x = 0.1,       # Desplazamiento para simular el borde
            #nudge_y = 0.1,       # Desplazamiento para simular el borde
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white",     # Texto blanco
            fontface = "bold",
            size = 6.5,           # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Frecuencia de Mutaciones Del Perfil Molecular HE",
       fill = "Tipo de Mutación",
       y = "Porcentaje") +
  theme_void() +
  theme(legend.position = "right") +
  scale_fill_manual(values = colores)

```

Código perteneciente a la Figura 28.

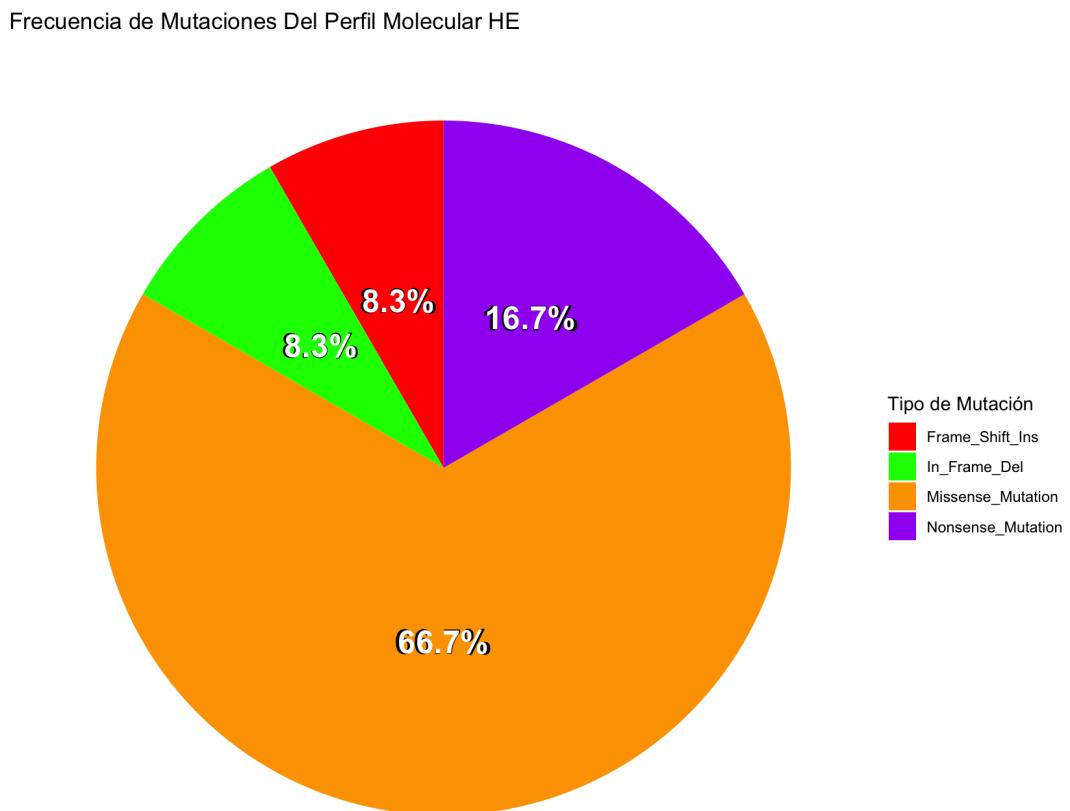


Figura 28: Frecuencia de mutaciones del perfil molecular HE.

Podemos observar que en el perfil molecular HE predomina la “mutación con cambio de sentido” (*Missense Mutation*) con un 66.7%. Asimismo, se puede apreciar que se encontró 4 tipos de mutaciones distintas.

```

-----LA-----
library(dplyr)
# Calcular los conteos de cada tipo de mutación
conteos_mutaciones_LA <- grupo_LA %>%
  count(`Mutation Type`)
# Calcular los porcentajes de cada tipo de mutación
porcentajes_mutaciones_LA <- conteos_mutaciones_LA %>%
  mutate(Porcentaje = n / sum(n) * 100)
# Gráfico de torta
#Figura 29
library(ggplot2)

colores <- c("orange", "purple", "blue", "red", "green", "pink")
ggplot(porcentajes_mutaciones_LA, aes(x = "", y = Porcentaje, fill = `Mutation Type`, label = paste0(round(Porcentaje, 1), "%"))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black",      # Borde negro
            fontface = "bold",
            size = 7,             # Tamaño mayor para el borde
            label.size = 1,
            #nudge_x = 0.1,       # Desplazamiento para simular el borde
            #nudge_y = 0.1,       # Desplazamiento para simular el borde
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white",     # Texto blanco
            fontface = "bold",
            size = 6.5,           # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Frecuencia de Mutaciones Del Perfil Molecular LA",
       fill = "Tipo de Mutación",
       y = "Porcentaje") +
  theme_void() +
  theme(legend.position = "right") +
  scale_fill_manual(values = colores)

```

Código perteneciente a la Figura 29.

Frecuencia de Mutaciones Del Perfil Molecular LA

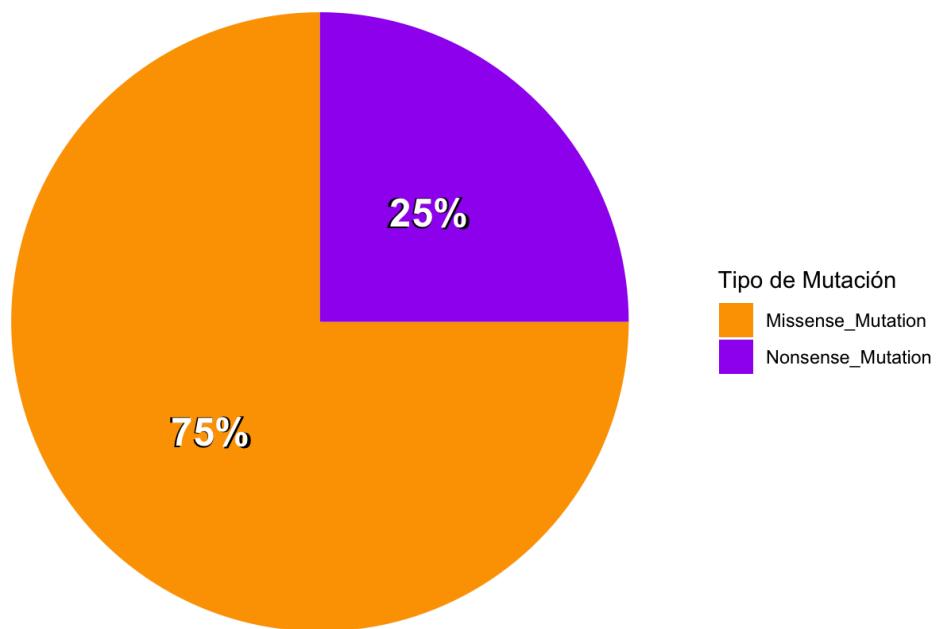


Figura 29: Frecuencia de mutaciones del perfil molecular LA.

Así mismo, en el perfil molecular LA predomina la “mutación con cambio de sentido” (*Missense Mutation*) con un 75%. En este caso, se puede apreciar que se encontraron sólo 2 tipos de mutaciones distintas (*Missense y Nonsense Mutations*).

```

#----LB-----
library(dplyr)
# Calcular los conteos de cada tipo de mutación
conteos_mutaciones_LB <- grupo_LB %>%
  count(`Mutation Type`)
# Calcular los porcentajes de cada tipo de mutación
porcentajes_mutaciones_LB <- conteos_mutaciones_LB %>%
  mutate(Porcentaje = n / sum(n) * 100)
# Gráfico de torta
library(ggplot2)

colores <- c("blue", "red", "orange", "purple", "pink", "green")

ggplot(porcentajes_mutaciones_LB, aes(x = "", y = Porcentaje, fill = `Mutation Type`, label = paste0(round(Porcentaje, 1), "%"))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black",      # Borte negro
            fontface = "bold",
            size = 7,             # Tamaño mayor para el borte
            label.size = 1,
            #nudge_x = 0.1,       # Desplazamiento para simular el borte
            #nudge_y = 0.1,       # Desplazamiento para simular el borte
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white",     # Texto blanco
            fontface = "bold",
            size = 6.5,          # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Frecuencia de Mutaciones Del Perfil Molecular LB",
       fill = "Tipo de Mutación",
       y = "Porcentaje") +
  theme_void() +
  theme(legend.position = "right") +
  scale_fill_manual(values = colores)

```

Código perteneciente a la Figura 30.

Frecuencia de Mutaciones Del Perfil Molecular LB

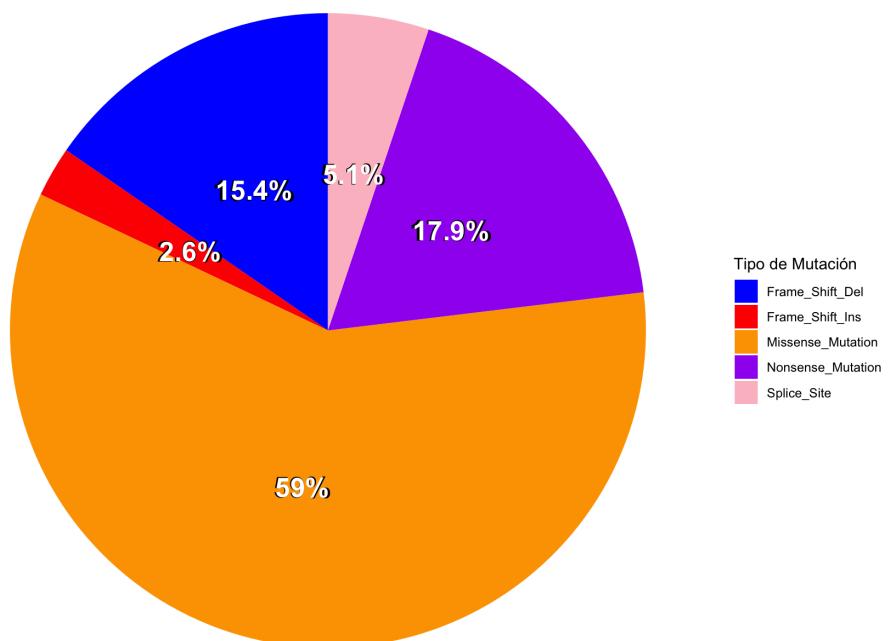


Figura 30: Frecuencia de mutaciones del perfil molecular LB.

En el perfil molecular LB predomina la “mutación con cambio de sentido” (*Missense Mutation*) con un 59%. En este caso, se puede apreciar que se encontró 5 tipos de mutaciones distintas.

```

#----TN-----
library(dplyr)
# Calcular los conteos de cada tipo de mutación
conteos_mutaciones_TN <- grupo_TN %>%
  count(`Mutation Type`)
# Calcular los porcentajes de cada tipo de mutación
porcentajes_mutaciones_TN <- conteos_mutaciones_TN %>%
  mutate(Porcentaje = n / sum(n) * 100)
# Gráfico de torta
library(ggplot2)

colores <- c("blue", "red", "green", "orange", "purple", "pink")

ggplot(porcentajes_mutaciones_TN, aes(x = "", y = Porcentaje, fill = `Mutation Type`, label = paste0(round(Porcentaje, 1), "%"))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "black",           # Borde negro
            fontface = "bold",
            size = 7,                 # Tamaño mayor para el borde
            label.size = 1,
            #nudge_x = 0.1,           # Desplazamiento para simular el borde
            #nudge_y = 0.1,           # Desplazamiento para simular el borde
            check_overlap = TRUE) +
  # Capa para el texto blanco (contenido principal)
  geom_text(aes(label = paste0(round(Porcentaje, 1), "%")),
            position = position_stack(vjust = 0.5),
            color = "white",          # Texto blanco
            fontface = "bold",
            size = 6.5,               # Tamaño del texto
            check_overlap = TRUE) +
  labs(title = "Frecuencia de Mutaciones Del Perfil Molecular TN",
       fill = "Tipo de Mutación",
       y = "Porcentaje") +
  theme_void() +
  theme(legend.position = "right") +
  scale_fill_manual(values = colores)

```

Código perteneciente a la Figura 31.

Frecuencia de Mutaciones Del Perfil Molecular TN

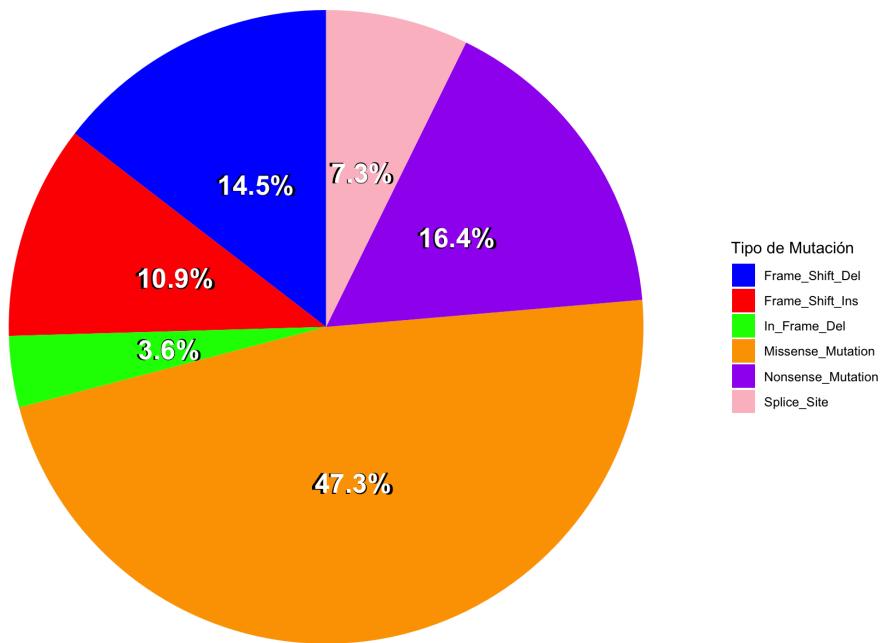


Figura 31: Frecuencia de mutaciones del perfil molecular TN.

También en el perfil molecular TN predomina la “mutación con cambio de sentido” (*Missense Mutation*) con un 47.3%. De manera interesante, en este caso, se puede apreciar que se encontraron 6 tipos de mutaciones distintas.

Teniendo en cuenta los análisis previamente realizados, se puede concluir que, en los cuatro perfiles moleculares estudiados, la mutación más frecuente corresponde a la “mutación con cambio de sentido” (*Missense Mutation*). Además, estos datos descartarían que algún tipo de mutación específica esté particularmente asociada a un perfil molecular.

6.9 Análisis de la Frecuencia de *Hotspots* en Mutaciones *Missense*

Dado que las mutaciones *Missense* son las más frecuentes en las muestras de pacientes con el gen *TP53* mutado, se llevó a cabo un análisis de la frecuencia de los puntos críticos o *hotspots*, que son las regiones donde estas mutaciones ocurren con mayor frecuencia. Este análisis se realizó con el propósito de identificar patrones específicos de mutación que podrían tener implicaciones en la función de la proteína p53 y su relación con la carcinogénesis. Para llevar a cabo el mismo, se generó un código que filtró las muestras, seleccionando únicamente aquellas que contenían mutaciones del tipo *Missense*. A continuación, se empleó la variable *Protein Change* como referencia principal para evaluar la distribución y frecuencia de las mutaciones. Los resultados obtenidos se presentan en la **Tabla 2**, donde se muestran los 15 *hotspots* más frecuentes.

```

library(dplyr)

# Filtrar Mutation Type = "Missense_Mutation"
df_filtered <- df %>%
  filter(`Mutation Type` == "Missense_Mutation") %>%
  group_by(`Protein Change`) %>%
  summarise(count = n()) %>%
  arrange(desc(count))

# Calcular el porcentaje
total_mutations <- sum(df_filtered$count)
df_filtered <- df_filtered %>%
  mutate(percentage = (count / total_mutations) * 100) %>%
  mutate(percentage = paste0(round(percentage, 2), "%"))

#top_10 <- df_filtered %>%
#  head(10)

#print(top_10)

View(df_filtered)

```

Código perteneciente a la Tabla 2.

	Protein Change	count	percentage
1	R175H	14	12.5%
2	E285K	7	6.25%
3	R273H	4	3.57%
4	E286K	3	2.68%
5	G266E	3	2.68%
6	K132N	3	2.68%
7	R280K	3	2.68%
8	E271K	2	1.79%
9	H193R	2	1.79%
10	H193Y	2	1.79%
11	K139N	2	1.79%
12	M237I	2	1.79%
13	R248W	2	1.79%
14	R280T	2	1.79%
15	V173L	2	1.79%

Tabla 2: Frecuencia de *hotspots* en muestras de pacientes con Cáncer de Mama con Mutación Missense en *TP53*. Esta tabla muestra los 15 *hotspots* más frecuentes en pacientes con cáncer de mama que presentan una mutación en el gen *TP53*, junto a su porcentaje.

A partir de este análisis, podemos observar que el *hotspot* más frecuente es R175H con una frecuencia de 12.5%. Los hotspots son regiones específicas dentro de un gen donde las mutaciones ocurren con mayor frecuencia. En este caso, el *hotspot* R175H corresponde a una mutación recurrente en el gen *TP53*, particularmente en una zona crítica de la proteína p53. Las mutaciones en estos hotspots tienen un impacto significativo en la función de la proteína, lo cual puede contribuir al desarrollo y progresión de ciertos tipos de cáncer, como el cáncer de mama. Identificar estos *hotspots* es esencial para entender mejor cómo las mutaciones en estas áreas pueden influir en la carcinogénesis.

En resumen, al analizar las características de los dos tipos de cáncer, CDI y CLI, se observa una predominancia de mutaciones puntuales en el gen *TP53*, especialmente las denominadas "Mutación con cambio de sentido" (*Missense*) y "Mutaciones sin sentido" (*Nonsense Mutation*). Estas alteraciones genéticas son conocidas por su implicación en la progresión tumoral y la resistencia al tratamiento.

Además, entre los casos con mutaciones en el gen *TP53*, se destaca una mayor frecuencia de tumores con expresión negativa de HER2 y positiva de receptores de estrógeno (ER), mientras que no se observa una diferencia significativa en la expresión de receptores de progesterona (PR).

En cuanto al perfil molecular, se identificó que el tipo CDI presenta una prevalencia del perfil Triple Negativo, caracterizado por la ausencia de receptores hormonales y HER2, lo que sugiere un fenotipo más agresivo y con menor número de opciones terapéuticas convencionales. Por otro lado, en el tipo CLI, el perfil predominante es Luminal B, que exhibe una combinación de receptores hormonales positivos y una mayor proliferación celular, lo que puede influir en la elección del tratamiento.

Adicionalmente, se destaca que el rango etario más común en ambos tipos de cáncer es de 50 a 60 años, lo que resalta la importancia de la detección temprana y el monitoreo en esta población de pacientes.

Cabe destacar que no se identificó un perfil molecular que se asocie exclusivamente a una mutación específica. Todos los perfiles moleculares analizados presentan una alta frecuencia de "mutación con cambio de sentido" (*Missense Mutation*), lo que indica que este tipo de alteración genética es común en los diferentes subtipos, sin estar vinculada de manera única a un perfil molecular en particular.

En conclusión, estos hallazgos subrayan la diversidad molecular y clínica entre los subtipos de cáncer de mama, lo que enfatiza la necesidad de enfoques terapéuticos individualizados y estrategias de manejo específicas para cada paciente.

7.0 ÍTEMES DESTACADOS DE LA INVESTIGACIÓN

- Se estudiaron los perfiles y características de los dos principales tipos de cáncer de mama CDI y CLI.
- Se analizó el rango etario donde predominan los tipos de cáncer de mama.
- Se analizaron las mutaciones predominantes en cada tipo de cáncer de mama.
- Se analizaron la frecuencia de los receptores HER2, ER y PR en el cáncer de mama.
- Se analizaron los perfiles moleculares en cada tipo de cáncer de mama.
- Se analizaron las frecuencias de *hotspots* en el perfil de mutaciones predominantes.
- Todos estos resultados fueron abordados empleando la herramienta RStudio.

8.0 CONCLUSIONES

A partir de estudios previos, se sabe que las mutaciones somáticas en el gen *TP53* son una de las alteraciones más frecuentes en los cánceres humanos. También que el 25% de los casos de cáncer de mama presentan mutaciones en el gen *TP53*. Por lo que a partir de estos estudios, se llevó a cabo un análisis para identificar las características predominantes en dos de los tipos más comunes de cáncer de mama: el Carcinoma Ductal Infiltrante (CDI) y el Carcinoma Lobulillar Infiltrante (CLI). Se analizó si las células tumorales de mama diferenciadas por tipo de cáncer, presentan distintas mutaciones en el gen *TP53*, si predomina un tipo de mutación del gen *TP53* y un perfil molecular, según el tipo de cáncer de mama. El objetivo de este análisis fue comprender mejor la diversidad de esta enfermedad y abordar la importancia de la implementación de terapias más personalizadas y efectivas para los pacientes.

Para lograr realizar este análisis, a partir de datos tomados de bases de datos de la plataforma *cBioPortal*, se ejecutó un procesamiento y estudio en *RStudio*. Donde se realizó una recolección de datos y limpieza de los mismos, para luego obtener los resultados que se analizaron posteriormente. Tras examinar los resultados obtenidos, se constata que el cáncer de mama, tanto del tipo CDI como CLI, predominan principalmente en el rango de edad entre los 50 y 60 años. Asimismo, se destaca una notable incidencia en el rango de edad comprendido entre los 40 y 50 años en el caso del CDI. Estos hallazgos subrayan la importancia crítica de la detección temprana del cáncer de mama y del seguimiento de esta población de pacientes. La comprensión de estas tendencias de edad puede facilitar la implementación de estrategias efectivas de detección y tratamiento.

Para verificar el cumplimiento de la hipótesis “*Existen diferencias en los perfiles tumorales moleculares entre los tumores de mama CDI y CLI, lo que podría asociarse con la respuesta a terapias antitumorales*”, se estudiaron los diferentes tipos de mutaciones presentes en las células tumorales de los pacientes seleccionados. Se detectaron seis tipos de mutaciones distintas, siendo el tipo más frecuente las mutaciones con cambio de sentido (*Missense*), que abarcan un 56%. Las mutaciones sin sentido (*Nonsense*) comprendieron un 19.5%, seguidas por las delecciones con cambio de marco (*Frameshift deletion*) con un 9%, las mutaciones en el sitio de empalme (*Splice Site*) con un 8%, las inserciones con cambio de marco (*Frameshift insertion*) con un 6%, y, en menor porcentaje, las delecciones sin cambio de marco (*In frame deletion*) con un 1.5%.

A partir de estos resultados, se procedió a realizar un análisis con énfasis en los tipos de cáncer de mama CDI y CLI. Los resultados del tipo de cáncer de mama CLI mostraron que hay

una predominancia del tipo de mutación “mutación con cambio de sentido” (*Missense Mutation*), representando el 57% de las mutaciones identificadas. Por otro lado, las mutaciones de tipo “mutaciones sin sentido” (*Nonsense Mutation*) representan el 22% de las mutaciones identificadas. Las mutaciones de tipo Sitio de Empalme (*Splice Site*) se encuentran en el 9% de las células tumorales, siguiéndole el tipo de mutación “Delección con cambio de marco” (*Frameshift Deletion*) representando un 6%, Inserción con cambio de marco (*Frameshift Insertion*) con un 4% y por último la mutación menos frecuente es la “Delección sin cambio de marco” (*In Frame Deletion*) con un 2%. Luego se analizaron particularmente los resultados del tipo de cáncer de mama CDI, los cuales muestran que hay una predominancia del tipo de mutación “mutación con cambio de sentido” (*Missense Mutation*), representando el 55% de las mutaciones identificadas. Por otro lado, las mutaciones de tipo “mutaciones sin sentido” (*Nonsense Mutation*) representan el 17% de las mutaciones. Las mutaciones de tipo “Delección con cambio de marco” (*Frameshift Deletion*) representaron un 11%, las mutaciones Sitio de Empalme (*Splice Site*) 8% de las células tumorales al igual que las mutaciones de “Inserción con cambio de marco” (*Frameshift Insertion*) y por último la mutación menos frecuente fue la “Delección sin cambio de marco” (*In Frame Deletion*) con un 1%.

En respuesta a la hipótesis planteada, los tipos de cáncer de mama CDI y CLI presentan una variedad de mutaciones, y en ambos predomina la misma. Además, se descartó que predomine un tipo de mutación del gen *TP53*, según el tipo de cáncer de mama, al menos en la población de estudio.

A partir de estos resultados, se procedió a analizar si existen otros factores o características en los pacientes que presentan células tumorales con el gen *TP53* mutado, y que se destaquen en cada tipo de cáncer de mama, CDI y CLI. Para ello, se evaluaron los estados de los tres receptores que predominan en el cáncer de mama: receptores de estrógeno (ER), progesterona (PR) y el receptor del factor de crecimiento epidérmico humano 2 (HER2). Estos receptores son indicadores importantes en la biología tumoral para determinar los perfiles moleculares de los pacientes. También tienen implicaciones significativas en la selección y respuesta al tratamiento.

A partir del análisis realizado, se pudo observar en los resultados que hay una gran predominancia del receptor HER2 negativo, esto significa que las células tumorales expresan poca o ninguna proteína HER2. Este tipo de cáncer puede desarrollarse más lentamente que el cáncer con el receptor HER2 positivo. También es menos probable que las células tumorales se propaguen a otras partes del cuerpo, es decir que es menor la probabilidad de generar metástasis. La terapia dirigida contra HER2 no funcionará para este tipo de cáncer.

El análisis de ER mostró que predomina el estado positivo, es decir que la mayoría de los tumores de mama analizados en el estudio tienen células que expresan receptores para la hormona estrógeno. La presencia de receptores de estrógeno en las células tumorales indica que el crecimiento del tumor puede estar influenciado por las hormonas.

Por último, se analizó el PR. Los resultados mostraron que los niveles de PR no presentan una gran diferencia entre PR positivo y negativo, lo que indica una distribución más homogénea de tumores que expresan este receptor y aquellos que no lo hacen. Este hallazgo sugiere que la presencia o ausencia del receptor de progesterona no es un factor distintivo significativo en la mayoría de los casos de cáncer de mama analizados.

Una vez analizados los receptores, se procedió a identificar el perfil molecular y tumoral de los pacientes. Inicialmente, se realizó un estudio general sin discriminar por tipo de cáncer, con el objetivo de determinar cuáles perfiles moleculares eran más comunes y cuáles eran menos frecuentes. Los resultados demostraron que el perfil molecular prevalente en el cáncer de mama es el Triple Negativo, seguido del perfil Luminal B. Estos hallazgos son significativos, ya que los diferentes perfiles moleculares implican variaciones en el comportamiento del cáncer y, por lo tanto, en la respuesta a los tratamientos o terapias.

Para continuar con la línea del estudio se procedió a identificar el perfil molecular y tumoral de los pacientes separando a los mismos por el tipo de cáncer que presentaban, es decir CDI y CLI. Relacionar los perfiles moleculares con los tipos histológicos de cáncer de mama permite un enfoque más detallado y personalizado en el tratamiento y manejo del cáncer. Además, proporciona una visión más completa de la biología del tumor, lo cual es crucial para desarrollar estrategias terapéuticas eficaces.

Cuando se realizó el análisis en CDI los resultados demostraron que el perfil que predomina en este tipo de cáncer es el Triple Negativo. El conocer que en el tipo de cáncer de mama CDI, en un paciente que tiene el gen *TP53* mutado predomine el Triple Negativo, donde no expresa el ER, PR O HER-2, tiene características clínicas que incluyen alta invasividad, alto potencial metastásico, propensión a recaídas y mal pronóstico. Debido a que los tumores Triple Negativos carecen de expresión de ER, PR y HER2, no son sensibles a la terapia endocrina ni al tratamiento con HER2, y todavía faltan regímenes de tratamiento estandarizados de cáncer de mama triple negativo¹⁴.

Por otro lado, en el tipo de cáncer CLI el perfil que predomina es el Luminal B. Conocer que en el tipo de cáncer de mama CLI, en un paciente que tiene el gen *TP53* mutado predomine el perfil Luminal B, nos indica que los tumores son de mayor grado y peor pronóstico en comparación con los Luminal A. Estos tumores pueden beneficiarse de la terapia

hormonal junto con la quimioterapia. Este perfil constituye entre el 10 y el 20% de los tumores luminales. Tiene una expresión moderadamente baja de receptores de estrógeno y una mayor expresión de genes de proliferación y ciclo celular. Representa el grupo de tumores luminales con peor pronóstico. Se benefician de la terapia hormonal y en mayor porcentaje de la quimioterapia en comparación con el grupo anterior (Luminal A)¹⁵.

Se llevó a cabo un análisis para determinar si existe una mutación predominante según el tipo de perfil molecular. Los resultados muestran que la "mutación con cambio de sentido" (*Missense Mutation*) es la que aparece con mayor frecuencia en todos los perfiles moleculares. Por lo tanto, se puede concluir que ninguna mutación específica se asocia exclusivamente a un perfil molecular particular, pero este tipo de mutación en particular está presente de manera generalizada en todos los perfiles moleculares.

Finalmente, tras el análisis de los *hotspots* en las muestras de pacientes con cáncer de mama que presentan mutaciones en el gen *TP53*, se observa que el *hotspot* más frecuente corresponde a la mutación R175H, la cual presenta una frecuencia del 12.5%. Las proteínas p53 mutadas, codificadas por estas mutaciones *hotspot*, pierden funciones respecto de su forma p53 no mutada y ganan funciones que, por el contrario, promueven el desarrollo del cáncer, incluida la promoción de la proliferación de células cancerosas, la migración, la invasión, la iniciación, la reprogramación metabólica, la angiogénesis y la conferencia de resistencia a los medicamentos a las células cancerosas. Cada vez hay más evidencia que ha demostrado que estas mutaciones *hotspot* pierden la función en su forma no mutada y adquieren funciones oncogénicas que promueven la progresión del cáncer. Entre estas mutaciones *hotspots*, p53-R175H tiene la mayor incidencia en diversos cánceres, incluidos los cánceres de mama, pulmón, colorrectal y pancreático⁵⁵.

En este trabajo se pudo llevar a cabo el análisis de los tipos de cáncer de mama CDI y CLI, y entre las muestras analizadas se pudo observar los distintos perfiles moleculares según tipo de tumor. Se constató que los tipos de cáncer de mama CDI y CLI de pacientes que presentan células tumorales con el gen *TP53* mutado presentan la misma mutación predominante y que el tipo de mutación específica no tiene relación exclusiva con el tipo de cáncer. Al evaluar los perfiles moleculares y tumorales, se observó una relación significativa entre el tipo de cáncer de mama y el perfil molecular. Es decir, se pudo asociar que el CDI presenta una prevalencia del perfil Triple Negativo mientras que en CLI el perfil predominante es Luminal B.

9.0 DISCUSIÓN

En este apartado se plantean algunos puntos sin ampliar en otras secciones del trabajo. Algunos están relacionados con los resultados obtenidos del análisis en *Rstudio*, mientras que otros intentan ir más allá de lo expuesto en las secciones de resultados. También es importante señalar algunas consideraciones y limitaciones relacionadas con los datos y el análisis realizado.

En la interpretación de los resultados del estudio realizado, observamos numerosas tendencias y patrones claros entre las mutaciones del gen *TP53*, los perfiles moleculares y tumorales y los subtipos de cáncer de mama CDI y CLI. *TP53* (p53) es el gen mutado con mayor frecuencia en el cáncer de mama invasivo. Aunque está mutado en 30 a 35% de todos los casos, p53 está mutado en aproximadamente 80% de los tumores triple negativos (TN) (es decir, tumores negativos para ER, PR y HER2). Debido a esta alta prevalencia, la p53 mutada es tanto un biomarcador potencial como un objetivo terapéutico para pacientes con cáncer de mama, especialmente para aquellos con el subtipo TN². Este motivo es uno de los principales por el cuál se decidió inclinar este estudio a la investigación del gen *TP53*. También se tuvo en cuenta cómo influye el gen *TP53* en el cáncer de mama. *TP53* es uno de los genes mutados con mayor frecuencia en el cáncer de mama y se ha informado que hasta entre el 30% y el 40% de los pacientes con cáncer de mama portan mutaciones en *TP53*².

Se ha observado que hay una variedad de mutaciones presentes en ambos tipos de cáncer de mama y que su distribución es variada. Los resultados muestran patrones similares de mutaciones en el gen *TP53* entre el CLI y el CDI, los dos principales tipos de cáncer de mama estudiados. La mutación de tipo “Mutación con cambio de sentido” (*Missense Mutation*) como la predominante, aunque el CLI presenta una diferencia poco significativa en la proporción de esta mutación en comparación con el CDI. Una mutación de este tipo puede tener distintos efectos en la función de la proteína, que van desde cambios sutiles hasta una pérdida completa de función. Mientras que la mutación que menos predomina es la “Deleción sin cambio de marco” (*In Frame Deletion*), en ambos tipos de cáncer de mama.

Cuando el gen *TP53* experimenta una mutación y pérdida de función, se ve comprometida su capacidad para corregir o eliminar células tumorales dañadas. Como consecuencia, estas células pueden proliferar sin control, lo que contribuye al desarrollo y crecimiento de tumores en el tejido mamario. Esta disfunción en el gen *TP53* puede ser un factor crucial en la progresión del cáncer de mama y en la resistencia a la terapia. Por lo tanto, comprender las características de un cáncer en el cual el gen *TP53* está mutado es esencial para desarrollar estrategias de tratamiento efectivas y dirigidas.

Un punto importante para discutir es la posibilidad de que los resultados y análisis de este estudio puedan contribuir a la selección de terapias específicas dirigidas a contrarrestar los efectos de la mutación en el gen *TP53*. El gen *TP53* ha sido el gen tumoral secuenciado más extensamente antes de la era de la secuenciación a gran escala, generando una gran cantidad de datos para identificar vínculos entre carcinógenos, huellas dactilares de mutaciones y tumorigénesis. En términos de aplicaciones clínicas, las mutaciones *TP53* han demostrado ser biomarcadores extremadamente complejos. A pesar de los impresionantes avances en la comprensión mecanicista de la estructura y función de p53, la investigación de p53 aún no ha generado aplicaciones de gran impacto en el tratamiento y la terapia del cáncer. De alguna manera, la complejidad del campo p53 puede haber actuado como un elemento disuasivo para las aplicaciones clínicas¹.

Otro punto importante para discutir es la asociación de perfiles moleculares y tumorales con los tipos de cáncer de mama CDI y CLI de pacientes con células tumorales el gen *TP53* presenta mutaciones. Múltiples estudios han confirmado que la frecuencia de mutación de *TP53* difiere entre los subtipos moleculares¹⁶. En este estudio también se mencionó que estas diferencias en los perfiles de mutación podrían tener importantes implicaciones biológicas y clínicas. Además, el conocimiento de estas diferencias podría permitir el desarrollo de terapias más dirigidas y personalizadas para cada subtipo de cáncer de mama, lo que podría mejorar significativamente los resultados clínicos y la calidad de vida de los pacientes. Esto se confirmó que es posible en artículos publicados. Conocer el subtipo de cáncer de mama puede ayudar a la práctica clínica a establecer el mejor tratamiento. Los estudios de expresión genética han demostrado la heterogeneidad del cáncer de mama, y esta heterogeneidad tiene una base molecular; sin embargo, hasta la fecha no se ha determinado si estas características moleculares influirían de manera inequívoca en el tratamiento clínico del cáncer de mama¹⁷.

Aunque los resultados no coincidieron completamente con la hipótesis inicial, se lograron cumplir los objetivos propuestos. Sin embargo, este análisis presenta algunas limitaciones. La muestra utilizada podría no ser representativa de la población total de pacientes con cáncer de mama, y sería necesario un mayor tamaño de muestra para validar los hallazgos. Además, factores como la región geográfica, junto con otros factores genéticos y ambientales, podrían influir en la variabilidad de los resultados. Por ejemplo, la interacción de *TP53* con otros genes relacionados con la reparación del ADN, el ciclo celular o la respuesta inmune podría modificar su efecto en el cáncer. Asimismo, factores ambientales como la exposición a carcinógenos, la dieta y el estilo de vida también podrían afectar tanto el desarrollo del cáncer de mama como la respuesta al tratamiento. Incorporar estos datos haría los resultados más específicos.

Entre las dificultades técnicas, uno de los principales retos fue la localización de bases de datos adecuadas. No obstante, una vez obtenidas, el trabajo con los datos resultó menos complejo. El análisis en RStudio presentó complicaciones iniciales debido a que el aprendizaje del lenguaje fue desde cero a causa del trayecto formativo del Plan de estudios primario. Aquí, las materias de Programación y Bases de Datos fueron de gran ayuda. La manipulación y

visualización de datos, así como el uso de paquetes especializados para analizar mutaciones genéticas y crear gráficos, exigieron un manejo avanzado de la herramienta, pero se tornaron más fluidos al disponer de los datos correctos.

Para comprender completamente la contribución de las mutaciones en *TP53* al cáncer de mama y para desarrollar estrategias de prevención y tratamiento más efectivas, es importante considerar la interacción compleja entre los factores genéticos y ambientales. Se propone para estudios futuros, que se integren datos genéticos y ambientales más completos ya que podrían ayudar a explicar estas interacciones y a identificar nuevas estrategias para la prevención y el tratamiento del cáncer de mama. Esta comprensión puede informar decisiones clínicas, como la selección de terapias específicas dirigidas a contrarrestar los efectos de la mutación en el gen *TP53* y restaurar su función normal. Además, el conocimiento de estas características puede ayudar a identificar subgrupos de pacientes que puedan beneficiarse de enfoques terapéuticos personalizados, lo que podría mejorar significativamente los resultados clínicos y la calidad de vida de los pacientes con cáncer de mama.

La aplicación de análisis genéticos en la investigación y tratamiento del cáncer de mama plantea diversas cuestiones éticas que debieron ser cuidadosamente consideradas. En este contexto, el manejo de datos clínicos y genéticos sensibles exige asegurar la confidencialidad, el consentimiento informado y el respeto por los derechos de los pacientes. Además, es crucial evaluar el uso justo y equitativo de los resultados en la práctica clínica, garantizando que los hallazgos contribuyan de manera ética al tratamiento y diagnóstico de la enfermedad. A continuación, se analizarán las principales cuestiones éticas derivadas de este estudio.

En primer lugar, la recopilación y el análisis de datos genéticos plantean preocupaciones sobre la privacidad y la confidencialidad de la información genética de los pacientes. Es fundamental garantizar que se implementen medidas adecuadas para proteger la privacidad de los pacientes y que se obtenga su consentimiento informado antes de realizar cualquier análisis genético.

Los conjuntos de datos heredados de TCGA *firehose* se importan directamente desde el Centro de coordinación de datos de TCGA original a través de *Broad Firehose*. Además, se seleccionan activamente conjuntos de datos de la literatura, incluyendo aquellos publicados junto con manuscritos. En ocasiones, se contacta a los investigadores para obtener información adicional, como atributos clínicos. Todas las llamadas de mutación (en formato VCF o MAF) se procesan mediante una canalización interna para anotar las variantes de manera consistente en todos los estudios¹⁸. Es recomendable que los/as investigadores/as aseguren que las personas participantes puedan controlar durante todo el proceso de investigación el tipo y nivel de información que quieren proporcionar sobre sí mismas y que la información proporcionada se resguarde de modo tal que garantice la privacidad y respete la intimidad de las personas¹⁹.

También se sugiere que los/as investigadores/as tomen todas las precauciones necesarias para proteger la privacidad y la confidencialidad de la información de las personas

participantes en el estudio, conforme a la legislación sobre protección de datos personales (por ejemplo, omitiendo los datos que pudiesen identificar a los individuos o limitando su acceso solo a personas autorizadas). Es preferible no utilizar datos personales identificables si el estudio puede realizarse sin ellos. En los casos donde sea necesario registrar datos identificables, los investigadores deberían justificar esta necesidad ante el CEI y detallar las estrategias para proteger la confidencialidad, como la disociación, codificación o anonimización¹⁹.

En investigaciones que utilicen entornos virtuales o herramientas digitales se recomienda tener especial cuidado con las medidas de protección de la privacidad, asegurando que no se revele ni infiera información personal al compartir, vincular o publicar los datos. Es importante que los investigadores evalúen los riesgos para la privacidad, los mitiguen en lo posible y describan los riesgos restantes en su proyecto. Asimismo, se sugiere monitorear y examinar todas las interacciones con los datos en cada etapa de la investigación¹⁹.

Se aconseja a los/as investigadores/as tomar todas las precauciones necesarias para almacenar y custodiar los datos de las investigaciones de manera segura, conforme a la legislación sobre protección de datos personales. Los/as investigadores/as deben explicitar ante el CEI las estrategias diseñadas para ello¹⁹.

Finalmente, tanto los/as investigadores/as como los CEI deberían procurar que la investigación beneficie, de manera directa o indirecta, a las personas, grupos o comunidades participantes, minimizando los posibles riesgos asociados a su participación¹⁹.

Respecto de la publicación de datos de pacientes en las bases de datos públicas, desde mi perspectiva, los datos científicos y tecnológicos deberían estar disponibles para todos, siempre que se respeten la confidencialidad, identidad y consentimiento de los pacientes. Al resguardar esta información exclusivamente, se pierden oportunidades valiosas de evolución y avance. Cada investigador/a puede observar y analizar los mismos datos desde diferentes perspectivas, lo que aumenta las posibilidades de descubrir algo nuevo y significativo. Por lo tanto, compartir los datos no sólo enriquece a la comunidad científica, sino que acelera el progreso hacia soluciones que mejoren la calidad de vida de las personas. Sin embargo, el debate sobre si estos datos deberían ser accesibles de manera gratuita o no es un tema que considero que podría explorarse en otro momento. Lo esencial aquí es reconocer que la ciencia avanza colectivamente, y al compartir datos, se abre la puerta a contribuciones más diversas e innovadoras.

PALABRAS DE CIERRE

La comprensión más profunda de las mutaciones en el gen *TP53* y su impacto en el cáncer de mama podría conducir a avances significativos en el desarrollo de terapias más efectivas y personalizadas.

Este trabajo no solo me ha permitido analizar el conocimiento en este campo, sino que también ha sido un cierre significativo en mi carrera universitaria. Me brindó la oportunidad de integrar y aplicar los conocimientos adquiridos a lo largo de estos cuatro años, eligiendo un tema de interés personal e integrando conceptos clave de genética con las herramientas y metodologías bioinformáticas. Esta experiencia ha sido importante para consolidar mi formación y abrir nuevas perspectivas en mi desarrollo profesional.

10.0 GLOSARIO

ADN: El ácido desoxirribonucleico (ADN) codifica la información que la célula necesita para fabricar proteínas. Se conforma de cuatro nucleótidos o bases: adenina (A), citosina (C), guanina (G) y timina (T).

Aminoácido: Unidad base que actúa como estructura fundamental de las proteínas. Hay 20 aminoácidos distintos que se combinan para formar proteínas.

Apoptosis: Proceso de muerte celular programada que elimina células dañadas o innecesarias del organismo.

ARN: El ácido ribonucleico (ARN) es un ácido presente en todas las células vivas que tiene similitudes estructurales con el ADN. Se conforma de cuatro nucleótidos o bases: adenina (A), citosina (C), guanina (G) y uracilo (U).

Cáncer: Tipo de cáncer que se desarrolla a partir de células epiteliales, es decir, que empieza en la piel o en los tejidos que revisten o cubren los órganos internos.

Células tumorales: Células que han experimentado cambios genéticos y moleculares que les permiten proliferar sin control, formando tumores.

Hibridación Fluorescente In Situ (FISH): Técnica de laboratorio

utilizada para detectar y localizar la presencia o ausencia de secuencias específicas de ADN en cromosomas.

Inmunohistoquímica (IHC): Técnica de laboratorio utilizada para visualizar la expresión de proteínas específicas en tejidos mediante el uso de anticuerpos.

Marco de Lectura: es una de las posibles formas en que se puede dividir una secuencia de nucleótidos de ADN o ARN para formar un grupo de tripletes consecutivos no solapados.

Metástasis: Proceso mediante el cual las células cancerosas se diseminan desde el tumor original a otras partes del cuerpo.

Mutación: Una mutación es un cambio en la secuencia de ADN de un organismo.

Nucleótido: Estructura fundamental básica de los ácidos nucleicos (ARN y ADN). Consta de una molécula de azúcar (ya sea ribosa en el ARN o desoxirribosa en el ADN) unida a un grupo fosfato y a una base nitrogenada.

Proteína: Molécula biológica compuesta por una o más cadenas de aminoácidos.

Receptores: Proteínas que se encuentran en la superficie celular o dentro de las células y que se unen a moléculas específicas (ligandos), como hormonas, neurotransmisores o factores de crecimiento, desencadenando una respuesta celular.

Traducción: Proceso biológico en el cual el ARNm es decodificado para sintetizar proteínas en los ribosomas.

Supresor Tumoral: Gen que protege a una célula de convertirse en cancerosa, generalmente mediante la regulación del ciclo celular o la inducción de apoptosis.

Transcripción: Proceso biológico en el cual una secuencia de ADN es copiada en una molécula de ARN mensajero

11.0 REFERENCIAS BIBLIOGRÁFICAS

1. Olivier, M., Hollstein, M., & Hainaut, P. (2010). TP53 mutations in human cancers: Origins, consequences, and clinical use. *Cold Spring Harbor Perspectives in Biology*, 2(1), a001008–a001008. <https://doi.org/10.1101/cshperspect.a001008>
2. Duffy, M. J., Synnott, N. C., & Crown, J. (2018). Mutant p53 in breast cancer: potential as a therapeutic target and biomarker. *Breast Cancer Research and Treatment*, 170(2), 213–219. <https://doi.org/10.1007/s10549-018-4753-7>
3. cBioPortal for Cancer Genomics. (s/f). Cbioportal.org. https://www.cbioportal.org/study?id=brca_tcga_pub2015
4. cBioPortal for Cancer Genomics. (s/f-b). Cbioportal.org. https://www.cbioportal.org/study?id=brca_tcga
5. cBioPortal for Cancer Genomics. (s/f-c). Cbioportal.org. https://www.cbioportal.org/study?id=brca_tcga_pub
6. cBioPortal for Cancer Genomics. (s/f-d). Cbioportal.org. https://www.cbioportal.org/study?id=brca_tcga_pan_can_atlas_2018
7. cBioPortal for Cancer Genomics. (s/f-e). Cbioportal.org. <https://www.cbioportal.org/datasets>
8. About us. (s/f). Cbioportal.org. <https://docs.cbioportal.org/about-us/>
9. Liu, B., Yi, Z., Guan, Y., Ouyang, Q., Li, C., Guan, X., Lv, D., Li, L., Zhai, J., Qian, H., Xu, B., Ma, F., & Zeng, Y. (2022). Molecular landscape of TP53 mutations in breast cancer and their utility for predicting the response to HER-targeted therapy in HER2 amplification-positive and HER2 mutation-positive amplification-negative patients. *Cancer Medicine*, 11(14), 2767–2778. <https://doi.org/10.1002/cam4.4652>
10. Kohler BA, Sherman RL, Howlader N, et al. Annual Report to the Nation on the Status of Cancer, 1975–2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state. *Journal of the National Cancer Institute* 2015; 107(6):djv048.
11. Joe BN. Clinical features, diagnosis, and staging of newly diagnosed breast cancer. In: UpToDate. Burstein HJ, Vora SR (eds.). Waltham, MA: UpToDate, 2019.
12. Rudlowski C. Male breast cancer. *Breast Care (Basel)* 2008; 3(3):183–189.
13. Estado del receptor hormonal del cáncer de seno. (s/f). Cancer.org. <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/comprehension-de-un-diagnostico-de-cancer-de-seno/estado-del-receptor-hormonal-del-cancer-de-seno.html>
14. Yin, L., Duan, J.-J., Bian, X.-W., & Yu, S.-C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Research: BCR*, 22(1). <https://doi.org/10.1186/s13058-020-01296-5>

15. Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L. E., Gómez-Valles, F. O., & Ramírez-Valdespino, C. A. (2022). Subtypes of Breast Cancer. En *Breast Cancer* (pp. 31–42). Exon Publications.
16. Fedorova O, Daks A, Shuvalov O, et al. "Atenuación del mutante p53 como enfoque para el tratamiento del cáncer Her2 positivo". *Descubrimiento de muerte celular*. 2020; 6 :100.
17. Sonnenblick A, Fumagalli D, Sotiriou C, Piccart M. ¿Es importante la diferenciación en subtipos moleculares de cáncer de mama para la estadificación, la terapia local y sistémica y el seguimiento? *Tratamiento del cáncer Rev.* 2014; 40 (9): 1089–1095.
18. Architecture overview. (s/f). Cbioportal.org. <https://docs.cbioportal.org/architecture-overview/>
19. de Investigación en Salud, D. (s/f). Guía para el diseño, desarrollo y evaluación ética de investigaciones sociales en Salud. Gob.ar. Recuperado el 24 de mayo de 2024, de https://www.argentina.gob.ar/sites/default/files/2020/07/guia_etica_2023-op3.pdf
20. *Membrana plasmática* | NHGRI. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Membrana-plasmatica#:~:text=%E2%80%8BMembrana%20plasm%C3%A1tica&text=La%20membrana%20plasm%C3%A1tica%20se%20componen,y%20salen%20de%20la%20c%C3%A9lula>
21. colaboradores de Wikipedia. (2024, 19 agosto). *Citoplasma*. Wikipedia, la Enciclopedia Libre. <https://es.wikipedia.org/wiki/Citoplasma#:~:text=Su%20funci%C3%B3n%20es%20albergar%20los,se%20dan%20en%20las%20c%C3%A9lulas>
22. Célula. (s/f). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Celula>
23. Núcleo celular. (s/f). Genome.gov. <https://www.genome.gov/es/genetics-glossary/N%C3%BAcleo-celular>
24. (S/f). Kastatic.org. <https://cdn.kastatic.org/ka-perseus-images/2f9d3f30c7c756be527ee310fb7c73229c2998b.png>
25. *Diccionario de cáncer del NCI*. (s. f.). Cancer.gov. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/cromatina#>
26. *Ácido desoxirribonucleico (ADN)*. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/%C3%81cido-desoxirribonucleico#:~:text=El%20ADN%20consta%20de%20c%C3%A1didos%20que%20están%20enlazados%20entre%20s%C3%ADntesis%20de%20ADN>
27. *Ácido ribonucleico (ARN)*. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/ARN#:~:text=El%20ARN%20consta%20de%20c%C3%A1didos%20que%20est%C3%A1n%20enlazados%20entre%20s%C3%ADntesis%20de%20ARN>
28. Menor-Salván, C. (2022, junio 2). Nuevas evidencias sobre el mundo ARN: un poco más cerca de entender el origen de la vida. The Conversation. <http://theconversation.com/nuevas-evidencias-sobre-el-mundo-arn-un-poco-mas-cerca-de-entender-el-origen-de-la-vida-183294>
29. *Transcripción*. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Transcripcion>
30. *Traducción*. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Traducion>
31. (S/f-b). Cancer.gov. <https://nci-media.cancer.gov/pdq/media/images/791096-571.jpg>

32. Ciclo celular | NHGRI. (s. f.-b). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Ciclo-cellular#:~:text=El%20ciclo%20celular%20es%20el,c%C3%A9lula%20se%20prepara%20para%20dividirse>.
33. Ciclo celular. (2012, octubre 15). Acerca Ciencia. <https://www.acercaciencia.com/amp/2012/10/15/ciclo-cellular/>
34. Gen. (s. f.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Gen#:~:text=La%20mayor%C3%ADa%20de%20los%20genes,genes%20que%20codifican%20para%20prote%C3%ADnas>.
35. Gen. (s. f.-b.). Genome.gov. <https://www.genome.gov/es/genetics-glossary/Gen#:~:text=Los%20genes%20se%20transmite,n%20de,diferentes%20funciones%20en%20el%20cuerpo>.
36. Diccionario de cáncer del NCI. (s. f.-b.). Cancer.gov. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/gen>
37. Pauta, M. (2024, 12 agosto). *Genes, protooncogenes y oncogenes. ¿Qué son?* MatchTrial. <https://matchtrial.health/genes-protooncogenes-y-oncogenes-que-son/>
38. Definición de gen supresor de tumoresNational Cancer Institute (.gov)<https://www.cancer.gov/espanol/diccionario-cancer>
39. Mauricio Lema. (2014, 4 enero). *MolBiol: p53 - función normal* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=JSgHcwC22IM>
40. Rangel-López, A., Piña-Sánchez, P., & Salcedo, M. (2006). Variaciones genéticas del gen supresor de tumores TP53: relevancia y estrategias de análisis. *Revista de investigacion clinica; organo del Hospital de Enfermedades de la Nutricion*, 58(3), 254–264. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0034-83762006000300010
41. Mutación. (s/f). Genome.gov. <https://genome.gov/es/genetics-glossary/Mutacion>
42. Tipos de mutaciones. (2021, 14 junio). Guía Metabólica. <https://metabolicas.sjdhospitalbarcelona.org/noticia/tipos-mutaciones>
43. Genética del cáncer. (s. f.-c.). Cancer.gov. <https://www.cancer.gov/espanol/cancer/causas-prevencion/genetica>
44. ¿Qué es el cáncer? (s. f.). Cancer.gov. <https://www.cancer.gov/espanol/cancer/naturaleza/que-es#:~:text=Los%20carcinomas%20son%20el%20tipo,al%20microscopio%2C%20parecen%20peque%C3%B1as%20columnas>.
45. colaboradores de Wikipedia. (2024a, agosto 14). *Cáncer de mama*. Wikipedia, la Enciclopedia Libre. https://es.wikipedia.org/wiki/C%C3%A1ncer_de_mama
46. Enciclopedia médica. (s. f.). <https://ssl.adam.com/content.aspx?productid=118&isarticlelink=false&pid=5&gid=000913&sitename=nuestrasalud.adam.com&login=NUES7600>
47. Tipos de cáncer de mama. (s. f.-b.). <https://www.mammotome.com/es/es/patient-resources-category/tipos-de-cancer-de-mama>
48. Invasive Breast Cancer (IDC/ILC). (s. f.). American Cancer Society. <https://www.cancer.org/cancer/types/breast-cancer/about/types-of-breast-cancer/invasive-breast-cancer.html>

49. Instituto Nacional del Cáncer. (n.d.). *Tratamiento del cáncer de seno durante el embarazo (PDQ®)-Versión para pacientes.* <https://www.cancer.gov/espanol/tipos/seno/paciente/tratamiento-seno-embarazo-pdq>
50. National Library of Medicine. (s. f.). *Cáncer de seno.* Cáncer de Mama | MedlinePlus En Español. <https://medlineplus.gov/spanish/breastcancer.html#:~:text=El%20c%C3%A1ncer%20de%20mama%20es,c%C3%A9stas%20suelen%20formar%20un%20tumor>.
51. *Cancer of the Breast (Female) - Cancer Stat Facts.* (s. f.). SEER. <https://seer.cancer.gov/statfacts/html/breast.html>
52. (S/f-d). Cancer.gov. <https://nci-media.cancer.gov/pdq/media/images/807074.jpg>
53. National Cancer Institute. (n.d.). Carcinoma lobulillar infiltrante. Instituto Nacional del Cáncer. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/carcinoma-lobulillar-infiltrante>
54. *Cancer of the Breast (Female) - Cancer Stat Facts.* (s. f.). SEER. <https://seer.cancer.gov/statfacts/html/breast.html>
55. Chiang YT, Chien YC, Lin YH, Wu HH, Lee DF, Yu YL. La función del mutante p53-R175H en el cáncer. Cánceres (Basilea). 13 de agosto de 2021;13(16):4088. doi: 10.3390/cancers13164088. PMID: 34439241; PMCID: PMC8391618.
56. Instituto Nacional del Cáncer. (s.f.). Terapia hormonal para el cáncer de seno. <https://www.cancer.gov/espanol/tipos/seno/hoja-informativa-terapia-hormonal-seno>
57. American Cancer Society. (s.f.). Terapia hormonal para el cáncer de seno. <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/tratamiento/terapia-hormonal-para-el-cancer-de-seno.html>
58. American Cancer Society. (s.f.). Terapia dirigida para el cáncer de seno. <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/tratamiento/terapia-dirigida-para-el-cancer-de-seno.html>
59. American Cancer Society. (s.f.). Quimioterapia para el cáncer de seno. <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/tratamiento/quimioterapia-para-el-cancer-de-seno.html>