Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Biológicas
Escuela Profesional de Genética y Biotecnología

Análisis transcriptómico *in silico* de genes asociados a interferones en muestras de cáncer de mama triple negativo

TRABAJO DE INVESTIGACIÓN

Para optar el Grado Académico de Bachiller en Genética y Biotecnología

AUTOR

Jherson Oscar CISNEROS GUTIERREZ

ASESORES

Dra. Libertad ALZAMORA GONZALES
Dr. Alexis Germán MURILLO CARRASCO (Coasesor)

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Cisneros, J. (2026). *Análisis transcriptómico in silico de genes asociados a interferones en muestras de cáncer de mama triple negativo*. [Trabajo de investigación de bachiller, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Genética y Biotecnología]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Jherson Oscar Cisneros Gutierrez
Tipo de documento de identidad	DNI
Número de documento de identidad	72971978
URL de ORCID	https://orcid.org/0009-0008-6407-5235
Datos de asesor	
Nombres y apellidos	Libertad Alzamora Gonzales
Tipo de documento de identidad	DNI
Número de documento de identidad	09022430
URL de ORCID	https://orcid.org/0000-0002-7425-7453
Datos de coasesor	
Nombres y apellidos	Alexis Germán Murillo Carrasco
Tipo de documento de identidad	DNI
Número de documento de identidad	75335990
URL de ORCID	https://orcid.org/0000-0002-7372-2608
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Erasmo Honorio Colona Vallejos
Tipo de documento	DNI
Número de documento de identidad	09438212
Miembro del jurado 1	
Nombres y apellidos	Julio Solis Sarmiento
Tipo de documento	DNI

Número de documento de identidad	09671619
Miembro del jurado 2	
Nombres y apellidos	Pedro Eduardo Romero Condori
Tipo de documento	DNI
Número de documento de identidad	44009312
Datos de investigación	
Línea de investigación	B.05.14.16 Bioestadística
Grupo de investigación	Inmunomoduladores y Antitumorales de Origen Natural y Sintético-MODULANS Immunology and Cancer-IMMUCA
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos Latitud: -12.05975 Longitud: -77.08213
Año o rango de años en que se realizó la investigación	Agosto 2024 - Diciembre 2025 2024 - 2025
URL de disciplinas OCDE	Bioinformática https://purl.org/pe-repo/ocde/ford#1.02.03 Oncología https://purl.org/pe-repo/ocde/ford#3.02.21 Inmunología https://purl.org/pe-repo/ocde/ford#3.01.03

Universidad Nacional Mayor de San Marcos
(Universidad del Perú, Decana de América)

FACULTAD DE CIENCIAS BIOLÓGICAS

ACTA DE SESIÓN PARA OPTAR AL GRADO ACADÉMICO DE BACHILLER EN GENÉTICA Y BIOTECNOLOGÍA (MODALIDAD: TRABAJO DE INVESTIGACIÓN)

Siendo las 14:10 horas del 26 de agosto de 2026, en el Salón de Grados de la Facultad de Ciencias Biológicas y en presencia del Jurado formado por los profesores que suscriben, se inició la sesión para optar al **Grado Académico de Bachiller en Genética y Biotecnología (Modalidad: Trabajo de Investigación)** de **JHERSON OSCAR CISNEROS GUTIERREZ.**

Luego de dar lectura y conformidad al expediente **N.º UNMSM-20260029464**, el graduando expuso su Tesis: **ANÁLISIS TRANSCRIPTÓMICO *in silico* DE GENES ASOCIADOS A INTERFERONES EN MUESTRAS DE CÁNCER DE MAMA TRIPLE NEGATIVO** y el Jurado efectuó las preguntas del caso calificando la exposición con la nota 19,calificativo: Aprobado con máximos honores

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología y al Consejo de Facultad para que se apruebe otorgar el **Grado Académico de Bachiller en Genética y Biotecnología (Modalidad: Trabajo de Investigación)** a **JHERSON OSCAR CISNEROS GUTIERREZ** y se eleve lo actuado al Rectorado para conferir el respectivo grado, conforme a ley.

Siendo las 15:17 horas se levantó la sesión.

Ciudad Universitaria, 26 de agosto de 2026.

ERASMO HONORIO COLONA VALLEJOS
PRESIDENTE

LIBERTAD ALZAMORA GONZALES
ASESORA

PEDRO EDUARDO ROMERO CONDORI
MIEMBRO

JULIO SOLIS SARMIENTO
MIEMBRO

UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1
CERTIFICADO DE SIMILITUD

Yo, Libertad Alzamora Gonzales en mi condición de asesora acreditada con el Dictamen de la EP de Genética y Biotecnología **N° 020-2026-EPGBT-FCB** del trabajo de investigación, cuyo título es: "Análisis transcriptómico in silico de genes asociados a interferones en muestras de cáncer de mama triple negativo", presentado por el egresado Sr. Jherson Oscar CISNEROS GUTIERREZ para optar al grado académico de Bachiller en GENÉTICA Y BIOTECNOLOGÍA,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 6% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado académico de bachiller en GENÉTICA Y BIOTECNOLOGÍA.

Lima, 05 de agosto de 2026

Firma de la Asesora UNMSM Firmado digitalmente por ALZAMORA GONZALES Libertad FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 05.08.2026 12:30:06 -05:00

Nombres y Apellidos de la asesora: Dra. Libertad Alzamora Gonzales

DNI: 09022430

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

DEDICATORIA

Este trabajo se lo dedico a mis padres, mi hermano y mi abuela, quienes siempre me motivaron a seguir mis sueños.

AGRADECIMIENTO

A mis padres, Pedro y Marisol, por siempre creer en mí; por quienes me esfuerzo cada día para ser un mejor hijo y profesional.

A mi hermano Omar, por ser un buen hermano mayor, por preocuparse por mí y acompañarme en los eventos importantes.

A mi abuela Bárbara, mi segunda madre, cuyo mayor sueño es verme convertido en un profesional.

A mis profesores, el Dr. Alexis y el Dr. César, por guiarme durante el desarrollo del presente trabajo y permitirme desarrollar mis habilidades analíticas y el uso de herramientas bioinformáticas.

A los profesores, el Dr. Ruy y el Mg. Obert, por permitirme afianzar mi habilidad crítica mediante las discusiones de los resultados.

A la Dra. Libertad Alzamora, por guiar mi camino profesional, formar mis habilidades técnicas y críticas, además de aconsejarme para tomar siempre la mejor decisión.

A mi amiga Jennifer, con quien iniciamos la idea de este trabajo y aprendimos juntos diversas cosas.

A mis grandes amigos y compañeros de universidad: Valery, Sandro y Selene, porque estuvieron ahí en los momentos que más los necesitaba, y con quienes comparto muchas experiencias gratas.

A los grupos de investigación MODULANS e IMMUCA, por abrirme las puertas y permitirme desarrollarme en campos tan apasionantes como la inmunología y la bioinformática; y a todas las personas que directa o indirectamente forman parte de este trabajo, amigos y profesores.

ÍNDICE GENERAL

ÍNDICE DE TABLAS I
ÍNDICE DE FIGURAS II
ÍNDICE DE ABREVIATURAS V
RESUMEN VII
ABSTRACT VIII
1. INTRODUCCIÓN 1
1.1. Objetivos 4
1.1.1. Objetivo General 4
1.1.2. Objetivos Específicos 4
2. ESTADO DEL CONOCIMIENTO 4
3. METODOLOGÍA 7
3.1. Muestras tumorales empleadas en el análisis 7
3.2. Muestra no tumorales empleadas en el análisis 9
3.3. Lista de genes empleados 9
3.3.1. Genes asociados a interferón 9
3.3.2. Genes de referencia 10
3.4. Exploración inicial de datos 10
3.5. Expresión diferencial de genes 11
3.6. Análisis de vías biológicas 12
4. RESULTADOS 15
4.1. Muestras tumorales empleadas en el análisis (CMTN, CMnTN) 15
4.2. Muestras no tumorales empleadas en el análisis 17
4.3. Lista de genes empleados en el análisis. 17
4.3.1. Genes asociados a interferón. 17
4.3.2. Genes de referencia 19
4.4. Análisis de datos inicial 20
4.4.1. Reducción de dimensionalidad 20
4.5. Exploración de los datos de CMTN y CMnTN 28
4.6. Análisis de expresión diferencial de genes 33
4.6.1. CMTN frente a Control 33
4.6.2. CMTN frente a TA 34
4.6.3. CMTN frente a CMnTN 39
5. DISCUSIÓN 44
5.1. Análisis Inicial de Datos 45
5.2. CMTN vs Control 48
5.3. CMTN vs TA 54

5.4. CMTN vs TA (análisis pareado) 55
5.5. CMTN vs CMnTN 57
6. CONCLUSIONES 60
7. RECOMENDACIONES 60
8. REFERENCIAS BIBLIOGRÁFICAS 60

ÍNDICE DE TABLAS

Tabla 1. Características de las cohortes CMTN y CMnTN.

Tabla 2. Distribución de las muestras extraídas del TCGA-BRCA según su origen y clasificación.

Tabla 3. Resumen de las principales características pertenecientes a la cohorte tumoral.

Tabla 4. Distribución de las muestras provenientes de GTEx por sexo.

Tabla 5. Lista de genes asociados a interferón tipo I que fueron incluidos en el análisis de expresión diferencial.

Tabla 6. Lista de genes asociados a interferón tipo II que fueron incluidos en el análisis de expresión diferencial.

Tabla 7. Lista de genes asociados a interferón tipo III que fueron incluidos en el análisis de expresión diferencial.

Tabla 8. Lista de genes de referencia o expresión constitutiva.

Tabla 9. Genes diferencialmente expresados en las vías de interferón según cada grupo de comparación.

ÍNDICE DE FIGURAS

Figura 1. Flujograma general de trabajo incluyendo informaciones solicitadas en el MIABi.

Figura 2. Diagrama de flujo de selección de muestra.

Figura 3. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales utilizando la expresión de 500 genes de mayor varianza.

Figura 4. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales solo de muestras pertenecientes a mujeres.

Figura 5. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales solo de muestras pertenecientes a mujeres utilizando los genes asociados a interferones (n = 85)

Figura 6: Diagrama de cajas de la expresión génica previo a la corrección por efecto lote.

Figura 7: Diagrama de cajas de la expresión génica posterior a la corrección por efecto lote mediante RUVseq.

Figura 8. Histograma del promedio de conteos por gen en una escala logarítmica (Log10).

Figura 9. Análisis de componentes principales utilizando diferentes umbrales de selección de genes más variables entre las muestras.

Figura 10. Histograma de la distribución de valores de coeficiente de variación de cada gen incluido en el análisis.

Figura 11. Análisis de componentes principales utilizando los 1279 genes de mayor variabilidad entre las muestras luego de realizar la corrección por efecto de lote.

Figura 12. Análisis de componentes principales de las muestras utilizando los genes asociados a interferón (n = 85) luego de realizar la corrección por efecto de lote.

Figura 13. Diagrama de Sankey: Representación de la correspondencia entre los grupos formados utilizando información de expresión de receptores y los grupos formados utilizando la información molecular PAM50.

Figura 14. Histograma del promedio de conteos por gen en una escala logarítmica (Log10).

Figura 15. Análisis de componentes principales utilizando diferentes umbrales de selección de genes más variables entre las muestras.

Figura 16. Histograma de la distribución de valores de coeficiente de variación de cada gen incluido en el análisis.

Figura 17. Análisis de componentes principales comparativo utilizando los 1473 genes de mayor variabilidad entre las muestras y que mejor separan a los grupos PAM50.

Figura 18. Diagrama de volcán mostrando genes diferencialmente expresados entre CMTN y Tejido Control.

Figura 19: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. Control).

Figura 20. Diagrama de volcán mostrando genes diferencialmente expresados entre 79 muestras de CMTN y 51 muestras de TA.

Figura 21: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. TA).

Figura 22. Análisis de componentes principales utilizando 500 genes de mayor varianza, mostrando la distribución de 7 muestras de CMTN y sus respectivas muestras de TA.

Figura 23. Diagrama de volcán mostrando genes diferencialmente expresados entre 7 muestras de CMTN y sus respectivos TAs.

Figura 24: Red de interacción proteína-proteína de los genes diferencialmente expresados en análisis pareado (CMTN vs. TA).

Figura 25. Diagrama de volcán mostrando genes diferencialmente expresados entre muestras de CMTN y CMnTN.

Figura 26: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. CMnTN).

Figura 27. Diagramas Euler de la intersección de genes diferencialmente expresados (A. Sobreexpresados y B. Subexpresados) entre el cáncer de mama triple negativo y las cohortes de comparación.

Figura 28. Diagrama de barras mostrando el puntaje de enriquecimiento normalizado (*NES*) de vías bajo el umbral de significancia de 0.071 ($p.adj < 0.071$).

Figura 29. Diagrama de enriquecimiento de las vías de A) interferón tipo I, B) interferón tipo II y C) interferón tipo III. Se muestra la posición de los genes pertenecientes a estas vías dentro de la lista ordenada por la expresión diferencial, representada mediante barras verticales negras.

ÍNDICE DE ABREVIATURAS

ADN: Ácido desoxirribonucleico

ANOVA: Análisis de varianza

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

CTLA-4: Antígeno 4 del linfocito T citotóxico

CXCL10: Ligando 10 de la quimiocina con motivo C-X-C

CXCL9: Ligando 9 de la quimiocina con motivo C-X-C

eIF4F: Complejo factor de iniciación de la traducción eucariota 4F

ELISA: Ensayo por inmunoabsorción ligado a enzimas

GBP1: Proteína 1 de unión a guanilato

IFI6: Proteína 6 inducida por interferón alfa

IFIT2: Proteína 2 con repeticiones tetratricopeptídicas inducida por interferón

IFNAR1: Subunidad 1 del receptor de interferón alfa y beta

IFNAR2: Subunidad 2 del receptor de interferón alfa y beta

IFNB1: Interferón beta 1

IFNGR1: Subunidad 1 del receptor de interferón gamma

IFNGR2: Subunidad 2 del receptor de interferón gamma

IRS1: Sustrato 1 del receptor de insulina

IRS2: Sustrato 2 del receptor de insulina

ISG15: Gen estimulado por interferón de 15 kDa

JAK/STAT: Vía de las quinasas Janus / Transductores de señales y activadores de la transcripción

MHC-I: Complejo principal de histocompatibilidad de clase I

NKT: Célula T Natural Killer

OAS1: 2'-5'-oligoadenilato sintetasa 1

PI3K: Fosfatidilinositol 3-quinasa

PI3K/Akt/mTOR: Vía de señalización celular Fosfatidilinositol 3-quinasa / Proteína quinasa B / Diana de rapamicina en mamíferos

PI3KR1: Subunidad regulatoria 1 de la fosfatidilinositol 3-quinasa

PI3KR2: Subunidad regulatoria 2 de la fosfatidilinositol 3-quinasa

RAS/MAPK: Vía de las proteínas RAS / Quinasas activadas por mitógenos

SOCS1: Supresor de la señalización de citoquinas 1

SOCS3: Supresor de la señalización de citoquinas 3

STAT1: Transductor de señal y activador de la transcripción 1

STAT5: Transductor de señal y activador de la transcripción 5

SVA: Análisis de variables subrogadas

T CD4+: Linfocito T colaborador (positivo para el clúster de diferenciación 4)

T CD8+: Linfocito T citotóxico (positivo para el clúster de diferenciación 8)

TCR: Receptor de células T

Th1: Célula T colaboradora tipo 1

Th2: Célula T colaboradora tipo 2

ZAP70: Proteína quinasa de 70 kDa asociada a la cadena zeta

RESUMEN

Actualmente, el cáncer de mama es la neoplasia con mayor incidencia y mortalidad en mujeres a nivel mundial y nacional. En el Perú, aproximadamente el 18 % de los casos corresponden al subtipo triple negativo, el cual presenta el peor pronóstico debido a su alta agresividad y a la ausencia de terapias dirigidas eficaces. Considerando que el microambiente tumoral influye de manera determinante en el comportamiento del cáncer, el presente estudio tuvo como objetivo analizar la expresión génica de vías relacionadas con interferones en muestras de cáncer de mama triple negativo, comparándolas con tejido no tumoral (Control), tejido adyacente (TA) y con cáncer de mama no triple negativo (CMnTN). Se emplearon herramientas bioinformáticas como RUV-Seq, DESeq2 y clusterProfiler, utilizando datos transcriptómicos provenientes de bases de datos públicas como The Cancer Genome Atlas y Genotype-Tissue Expression, así como conjuntos génicos de vías de interferón obtenidos de la Molecular Signatures Database. Los resultados evidenciaron, principalmente, que los genes asociados a la vía de interferón tipo II presentan una mayor expresión en el cáncer de mama triple negativo en comparación con los grupos Control, TA y CMnTN, lo que sugiere que esta vía podría tener una influencia predominante en las características biológicas de este subtipo tumoral.

Palabras clave: cáncer de mama triple negativo, interferón gamma, interferones, microambiente tumoral, análisis de vías biológicas.

ABSTRACT

Currently, breast cancer is the neoplasm with the highest incidence and mortality in women worldwide and nationally. In Peru, approximately 18% of cases correspond to the triple-negative subtype, which has the worst prognosis due to its high aggressiveness and the lack of effective targeted therapies. Considering that the tumor microenvironment has a determining influence on cancer behavior, the present study aimed to analyze the gene expression of interferon-related pathways in triple-negative breast cancer samples, comparing them with non-tumoral tissue (Control), adjacent tissue (AT) and non-triple-negative breast cancer (non-TNBC). Bioinformatics tools such as RUV-Seq, DESeq2, and clusterProfiler were employed, using transcriptomic data from public databases like The Cancer Genome Atlas and Genotype-Tissue Expression, as well as interferon pathway gene sets obtained from the Molecular Signatures Database. The results mainly evidenced that genes associated with the type-II interferon pathway show higher expression in triple-negative breast cancer compared to the Control, AT, and non-TNBC groups, suggesting that this pathway could have a predominant influence on the biological characteristics of this tumor subtype.

Keywords: triple-negative breast cancer, interferon gamma, interferons, tumor microenvironment, analysis of biological pathways

1. INTRODUCCIÓN

El cáncer de mama es actualmente la neoplasia más frecuente y la principal causa de muerte por cáncer en mujeres a nivel mundial, con 2.3 millones de nuevos casos y 666,000 fallecimientos registrados en 2022 (Bray et al., 2024). En el contexto nacional, representó la primera en incidencia (7,797 nuevos casos) y la tercera en mortalidad (1,951 muertes) en ese mismo año (Bray et al., 2024).

Esta enfermedad comprende una amplia diversidad de subtipos que difieren tanto en sus características clínicas como en sus perfiles moleculares. La inmunohistoquímica para los receptores hormonales de estrógeno (RE) y progesterona (RP), así como para el receptor 2 del factor de crecimiento epidérmico humano (HER2), constituye la aproximación clínica estándar para su subclasificación (Orrantia-Borunda et al., 2022).

Entre estos subgrupos destaca el cáncer de mama triple negativo (CMTN), definido por la ausencia de expresión de RE, RP y HER2. Este subtipo se caracteriza por su alta agresividad, elevada tasa proliferativa y mayor propensión a desarrollar metástasis, especialmente en pulmón, hígado y cerebro. Además, la ausencia de dianas terapéuticas específicas limita las opciones de tratamiento sistémico, dejando a la quimioterapia como principal estrategia terapéutica. En conjunto, estas características convierten al CMTN en el subtipo con peor pronóstico (Ensenyat-Mendez et al., 2021).

Estudios recientes en población peruana indican que entre el 16 % y el 18 % del total de casos de cáncer de mama pertenecen a este subtipo. Asimismo, suelen diagnosticarse en estadios más avanzados y con grados histológicos más altos, lo cual agrava el panorama nacional (Zavala et al., 2023; Gómez-Rázuri et al., 2024).

El estudio de las interacciones entre componentes inmunológicos, como citocinas, células del sistema inmunitario y proteínas de membrana, y las células tumorales dentro del

microambiente tumoral resulta fundamental para comprender el comportamiento biológico, la progresión y la respuesta terapéutica en distintos tipos de neoplasias (Zagami & Carey, 2022). Aunque el cáncer de mama se considera, en general, un tumor inmunológicamente "frío", tumor con baja probabilidad de desencadenar una respuesta inmunitaria fuerte debido a la ausencia de los componentes inmunológicos; el CMTN se distingue por poseer un microambiente inmunológico más activos debido a que muestra una mayor infiltración de células inmunitarias (Ensenyat-Mendez et al., 2021).

En este contexto, los interferones conforman una familia de citocinas descritas inicialmente por su capacidad para modular la respuesta inmunitaria e interferir en los procesos de infección viral (Medrano et al., 2017). Desde un punto de vista funcional y estructural, los interferones se agrupan en tres tipos principales. El tipo I (IFN-I) incluye 13 subtipos de interferón alfa (IFNA), además del interferón beta (IFNB), épsilon (IFNE), kappa (IFNK) y omega (IFNW). Todos ellos interactúan con un receptor dimérico compuesto por las subunidades IFNAR1 e IFNAR2; el tipo II (IFN-II) es representado únicamente por el interferón gamma (IFNG), que se une al receptor formado por las subunidades IFNGR1 e IFNGR2; por último, el tipo III (IFN-III) comprende los interferones lambda (IFNL1-4), cuya señalización ocurre a través del receptor heterodimérico compuesto por el receptor 1 de interferón lambda (IFNLR1) y la subunidad beta del receptor de interleucina 10 (IL10RB) (Platanias, 2005; Marín Sánchez et al., 2019).

Además de su reconocida función antiviral, los interferones desempeñan un papel crucial en la inmunovigilancia y en la respuesta inmunitaria innata y adaptativa frente a tumores. Entre sus múltiples efectos se encuentran la inducción de la presentación de antígenos tumorales, la activación y maduración de células inmunitarias como células dendríticas (DCs, del inglés "*dendritic cells*"), macrófagos, linfocitos T y células *Natural Killers* (NK), la producción de

citocinas inflamatorias, la promoción de la apoptosis, la inhibición de la angiogénesis, la supresión de la proliferación celular y la restricción del proceso metastásico (Geoffroy & Bourgeois-Daigneault, 2020; L. Zhou et al., 2020).

No obstante, a pesar de su papel antitumoral, diversos estudios han demostrado que los interferones también pueden tener efectos pro-tumorales en ciertos contextos neoplásicos. Su expresión crónica o desregulada en el microambiente tumoral puede activar rutas inmunosupresoras que favorecen la progresión tumoral. Por ejemplo, pueden inducir la expresión de moléculas como la interleucina 10 (IL-10) y el ligando de muerte programada 1 (PD-L1) en células presentadoras de antígeno y linfocitos T reguladores (Treg), así como inhibir la activación, proliferación y polarización de linfocitos T CD4+ hacia el fenotipo Th1. También se ha observado que promueven la expansión de células supresoras derivadas de linaje mieloide (del inglés, *myeloid derived suppressor cells*, MDSC) y la producción de enzimas inmunorreguladoras como la indolamina 2,3-dioxigenasa (IDO), involucrada en la tolerancia inmunológica. Además, su acción sostenida puede facilitar la diseminación metastásica y provocar el agotamiento funcional de células inmunitarias, como los linfocitos T citotóxicos infiltrados, reduciendo así su eficacia en el control del crecimiento tumoral (Pingwara et al., 2017; Provance & Lewis-Wambi, 2019; Boukhaled et al., 2021; Larsen et al., 2023). Además de ello, se ha evidenciado que la expresión de interferones influye en la respuesta del CMTN a terapias, como la inmunoterapia dirigidas contra PD-L1 y CTLA-4 (Sceneay et al., 2019) o a la quimioterapia (Legrier et al., 2016)

Dado que hasta la fecha no existe un estudio que compare de manera conjunta la expresión de genes asociados a los tres tipos de interferones, y considerando la importancia de estas moléculas en el microambiente tumoral, este trabajo se enfocó en analizar las diferencias en la expresión de dichos genes entre distintos tipos de muestras. Para ello, se evaluó el transcriptoma

de CMTN, cáncer de mama no triple negativo (CMnTN) y tejido mamario libre de cáncer, utilizando datos públicos provenientes de *The Cancer Genome Atlas* (TCGA) y *Genotype-Tissue Expression* (GTEx).

1.1. Objetivos

1.1.1. Objetivo General

Evaluar y caracterizar la expresión de genes asociados a la señalización del interferón en el cáncer de mama triple negativo, en comparación con tejido sano y tumores de mama no triple negativos (CMnTN).

1.1.2. Objetivos Específicos

- ➢ Comparar la expresión de genes relacionados con interferones entre muestras tumorales de CMTN y tejido mamario libre de cáncer.
- ➢ Comparar la expresión de genes relacionados con interferones entre muestras tumorales de CMTN y CMnTN.

2. ESTADO DEL CONOCIMIENTO

Diversos estudios han aportado, directa o indirectamente, evidencia sobre la expresión de genes relacionados con interferones en cáncer de mama y entre sus distintos subtipos clínicos.

Uno de los primeros trabajos en este tema fue desarrollado por Perou et al. (1999) utilizando microarreglos de cDNA. En este estudio, los autores identificaron en células epiteliales mamarias no cancerosas un grupo de genes inducidos por interferones tipo I y II (IFNA e IFNG). Además, observaron que estos genes también respondían a estímulos como confluencia y senescencia celular. Al evaluar este mismo conjunto en 13 tumores de mama, encontraron una expresión marcadamente heterogénea (tumores con alta, intermedia y baja expresión) probablemente mediada por la vía de STAT1. Los autores propusieron que la variabilidad

observada en estos genes, junto con un conjunto adicional relacionado a proliferación, podría contribuir a la clasificación molecular de los 13 tumores estudiados. Esta idea fue reforzada posteriormente por Perou et al. (2000), quienes, con una cohorte más amplia (40 tumores), confirmaron que la expresión del módulo de genes inducidos por interferón era altamente variable y aportaba a la subclasificación del cáncer de mama.

Existen estudios que no mencionan subtipos tumorales pero evalúan la expresión de los genes en relación con la presencia de los receptores hormonales y HER2. En este sentido, podemos mencionar el estudio de Chavey et al., (2007) que en un ensayo empleando ELISA en extractos tumorales, encontraron que IFNG se expresaba en mayor cantidad en tejido tumoral que en tejido no tumoral, y que sus niveles eran más altos en tumores RE negativos y RP negativos, evaluados independientemente. De manera complementaria, Todorović-Raković et al. (2017), en un seguimiento de 73 pacientes tratadas únicamente con cirugía o radioterapia, reportaron una correlación negativa entre la expresión del ARNm de *IFNG* y la expresión del receptor de progesterona. En la misma línea, Walter et al. (2017) demostraron que el receptor de progesterona puede suprimir la expresión de genes inducidos tanto por IFN-I como por IFN-II, incluso de manera independiente de ligando. Por otro lado, Yaghoobi et al., (2018) mediante reacción en cadena de la polimerasa (PCR) en tiempo real, evaluaron la expresión de genes de IFNG y el ARN antisentido del IFNG (IFNG-AS1), un ARN largo no codificante que regula su transcripción, en pacientes con cáncer de mama. Los autores encontraron mayores niveles de ambos en tejido tumoral respecto al tejido adyacente, aunque la diferencia en IFNG no alcanzó significancia estadística ($p = 0.07$). Además reportaron mayor expresión de IFNG en tumores HER2 negativos y en grado 1 respecto a grado 2, sugiriendo que la disminución de IFNG podría ser un mecanismo asociado a inmunoevasión tumoral. En estudios realizados en líneas celulares de cáncer de mama

inflamatorio, Ogony et al. (2016) observaron que IFNA presentaba mayor expresión en la línea de CMTN SUM149 en comparación con SUM190, representativa de un subtipo HER2 positivo.

Estudios recientes han manifestado un creciente interés en esta área de investigación centrada en los interferones en CMTN. Los resultados del trabajo de Heimes et al. (2020), utilizando datos de expresión génica por microarreglos, mostraron una mayor expresión individual de *IFNG* y de genes asociados (*IFNGR1, IFNGR2, IRF1, JAK1, JAK2, STAT1*) en tumores del subtipo basal, además de estar relacionado a un mejor pronóstico de supervivencia libre de metástasis.

Utilizando datos de expresión génica obtenidos por secuenciación, Weng et al. (2025) observaron una mayor expresión de *IFNG* en tumores CMTN en comparación a los otros subtipos tumorales mamarios. Además; dividiendo al CMTN en 3 grupos basado en análisis de componentes principales, se reportó una heterogeneidad intrínseca en su expresión. Asimismo, su expresión tuvo un impacto significativo positivo en la supervivencia en tumores del subtipo basal y estaba correlacionado positivamente con la expresión de genes de proteína de puntos de control inmunitarias como *CD274*, *PD-1*, *CTLA-4* y *TIGHT,* sugiriendo su papel como biomarcador de aplicación de inmunoterapia.

Así, también, Liu et al. (2025) reportaron la expresión diferencial de 139 genes asociados a interferón en CMTN frente a tejido normal adyacente. De los cuales *LAMP3*, *CD276*, *POLR2F* y *STXBP1,* este último subexpresado en CMTN, se utilizaron para constituir una firma pronóstica de riesgo de supervivencia. El grupo CMTN de bajo riesgo tuvo una mayor carga mutacional y un menor nivel de disfunción o escape inmunológico tumoral, sugiriendo la mejor respuesta a inmunoterapia. Además, se predijo que la sobreexpresión de estos genes aumenta la resistencia a 126 fármacos y la sensibilidad a 84, resaltando el impacto en la medicina de precisión.

A diferencia de todo lo anteriormente mencionado, esta investigación se centra en los genes asociados a la vías de señalización de los tres tipos de interferones, realizando un análisis comparativo de la expresión génica en muestras de CMTN frente a tejido mamario de donante libre de cáncer, tejido adyacente y tejido tumoral CMnTN, este último permitiendo además aumentar el conocimiento sobre las diferencias existentes entre los subtipos de cáncer de mama.

3. METODOLOGÍA

Este trabajo sigue los estándares estipulados dentro del *Minimum Information about a Bioinformatics Investigation* (MIABi) el cuál establece la información mínima que debe reportarse para permitir la reproducibilidad de un análisis bioinformático como las metas y objetivos del estudio; la base de datos utilizada, la fecha de descarga, el tipo y tamaño de los datos; el tipo de análisis empleado, con especificación de los métodos, software y algoritmos; y las técnicas de transformación y normalización de datos aplicadas (Tan et al., 2010)

3.1. Muestras tumorales empleadas en el análisis

Se descargaron los conteos crudos (en inglés, *raw counts*) de expresión génica correspondientes al transcriptoma global de muestras tumorales de pacientes con carcinoma de mama invasivo (*Breast Invasive Carcinoma - BRCA*), pertenecientes al proyecto *The Cancer Genome Atlas* (TCGA), a través de la plataforma XenaBrowser (fecha de descarga: 16/08/2024). Estos datos fueron generados utilizando la tecnología de secuenciación masiva de ARN Illumina RNA-seq. Asimismo, se obtuvieron los datos clínicos o metadatos asociados a cada muestra.

Las muestras tumorales fueron agrupadas en dos cohortes principales: CMTN y CMnTN, de acuerdo con la información de la expresión de los receptores hormonales (RE, RP) y del receptor HER2, siguiendo los criterios especificados en la **Tabla 1**. Las muestras que carecían de esta información fueron excluidas del análisis. En el caso de los receptores de estrógeno y progesterona, la información fue obtenida por inmunohistoquímica, mientras que en el caso del

marcador HER2, se contó con resultados provenientes de dos técnicas distintas: inmunohistoquímica e hibridación *in situ*. En aquellas muestras en las que existía discrepancia entre ambos métodos, se priorizó el resultado de la hibridación *in situ*, conforme a lo reportado por Bogdanovska-Todorovska et al. (2018).

Tabla 1. *Características de las cohortes CMTN y CMnTN. Adaptación realizada por el autor a partir de los datos publicados por Gao & Swain (2018).*

Cohortes	**Subtipo**	**HER2**	**RE**	**RP**
CMTN	Triple negativo	-	-	-
CMnTN	HER2+	+	-	-
	Luminal	+/-	+	+/-
			+/-	+

Nota: (-) negativo, (+) positivo, (+/-) positivo o negativo; HER2: Receptor 2 del factor de crecimiento epidérmico humano; RE: Receptor de estrógeno; RP: Receptor de progesterona; CMTN: Cáncer de mama triple negativo. CMnTN: Cáncer de mama no triple negativo.

Para asegurar la homogeneidad del contenido de cada cohorte, se utilizó la información del sistema de codificación de muestras establecido por el TCGA (Heath et al., 2021), con el fin de incluir únicamente muestras provenientes de tumor primario y excluir aquellas de origen metastásico u otro tipo de tejido.

Para comparar las diferencias en proporción de los datos clínicos (variables categóricas) más relevantes entre nuestras cohortes se aplicó la prueba de chi-cuadrado de Pearson. Sin embargo, se aplicó la prueba exacta de Fisher para los casos donde la frecuencia esperada sea menor a 5 o la aproximación de Fisher por simulación de Monte-Carlo, si además la variable posee múltiples categorías (Crawley, 2013).

3.2. Muestra no tumorales empleadas en el análisis

Se descargaron los conteos crudos de expresión génica correspondientes al transcriptoma global de muestras de tejido mamario no tumoral o de referencia provenientes de donantes sin cáncer del *The Adult Genotype Tissue Expression* (GTEx, versión 10 del proyecto), los cuales formarán la cohorte Control de este estudio. Estos datos fueron generados utilizando la tecnología de secuenciación masiva de ARN (Illumina RNA-seq). Así mismo, se descargaron los metadatos asociados a cada muestra (GTEx Consortium, 2013).

3.3. Lista de genes empleados

3.3.1. Genes asociados a interferón

La categoría de "genes asociados" a cada tipo de interferón comprende no solo los genes que codifican para las proteínas de interferón, sino también aquellos que participan en sus vías de señalización y algunos que son inducidos por su actividad o que regulan su expresión y producción. El conjunto principal de genes se obtuvo de la base de datos *Molecular Signatures Database (MSigDB)*, específicamente de la colección curada de WikiPathways (c2.cp.wikipathways.v2025.1.Hs.entrez.gmt). Esta colección incorpora las rutas actualizadas según la versión más reciente de WikiPathways, liberada el 10 de julio de 2024. Se utilizaron las siguientes colecciones: WP_INTERFERON_TYPE_I_SIGNALING, para los genes de la vía de señalización del interferón tipo I (IFN-I) ; WP_TYPE_II_INTERFERON_SIGNALING, para los del tipo II (IFN-II) y WP_TYPE_III_INTERFERON_SIGNALING, para los del tipo III (IFN-III) (Agrawal et al., 2024). Esta recopilación fue complementada con algunos genes adicionales seleccionados mediante una revisión manual de literatura científica, con el fin de incorporar elementos relevantes no representados en los repositorios anteriores (Platanias, 2005).

3.3.2. Genes de referencia

Los genes de referencia, también llamados genes de expresión constitutiva en tejidos mamarios (tumores o no tumorales), se emplearon para mitigar el "efecto de lote" o las variaciones debidas a la distinta procedencia de los datos de expresión génica de tejido mamario (TCGA y GTEx) que pueden influir en los análisis posteriores (Leek et al., 2010). Para tal fin, estos genes se caracterizan por presentar una expresión media o alta en tejido mamario y por mantener niveles relativamente estables entre distintas condiciones biológicas. Dichas condiciones incluyen fenotipos cancerosos (en diversos estadios y subtipos tumorales de cáncer de mama), tejido normal adyacente al tumor, líneas celulares sometidas a diferentes tratamientos y tejido mamario sano (Maltseva et al., 2013). La selección de estos genes se realizó cuidadosamente a partir de una revisión exhaustiva de la literatura científica (ver **Tabla 8**).

3.4. Exploración inicial de datos

Antes de realizar el análisis de expresión diferencial de genes, se exploró la distribución espacial de las muestras según su expresión génica mediante herramientas de reducción de dimensionalidad como el análisis de componentes principales (PCA, del inglés, *principal component analysis*), el cual se realizó utilizando la función plotPCA del paquete DESeq2, en caso de realizar un PCA solo con los 500 genes de mayor varianza, y *prcomp* del paquete base de R, en caso de realizar un PCA mediante filtros sucesivos por expresión génica y coeficiente de variación.

Posteriormente, se utilizó el paquete RUVseq y su función interna RUVg, el cuál utiliza la información de expresión génica de genes de referencia para mitigar el efecto de lote y eliminar otras variaciones no deseadas que pueden influenciar en los resultados (Risso et al., 2014). Para evaluar estratégicamente el impacto de la corrección, se evaluó la expresión de los siguientes genes: homeobox A5 (*HOXA5*) debido a su baja expresión en cáncer de mama en comparación a

tejido normal (Fan et al., 2022), receptor del factor de crecimiento epidérmico (*EGFR*) como marcador de la cohorte CMTN (Shetty et al., 2025), proteína de unión a GATA 3 (*GATA3*) como marcador de la cohorte CMnTN debido a su mayor expresión en tumores mamario que mantienen la expresión de sus receptores hormonales y HER2 (Chen et al., 2024), genes de receptores hormonales (*PGR, ESR1*), el gen del receptor HER2 (*ERBB2*) y genes asociados a interferones (*CXCL10, IRS1*). La significancia estadística de la diferencia en la expresión génica entre las cohortes se determinó utilizando la prueba de Kruskal-Wallis.

3.5. Expresión diferencial de genes

El análisis de expresión diferencial de genes (DEG, *differential expression gene*) se llevó a cabo utilizando la herramienta bioinformática DESeq2. Para obtener estimaciones más estables del logaritmo de la razón de cambio ($Log_2(Fold\ Change)$, LFC), se aplicó el método de contracción bayesiana empírica mediante el paquete apeglm (*Approximate Posterior Estimation for GLM*) (Zhu et al., 2019).

Se consideraron como genes sobreexpresados a aquellos que tuvieron un $LFC > 1$ y subexpresados a los que tuvieron $LFC < -1$. Este límite evitará enfocarnos en los genes que podrían poseer cambios estadísticamente significativos pero que, debido al pequeño efecto o cambio en su expresión, no poseen relevancia biológica; además de permitir la comparación y reproducibilidad con otras tecnologías e investigaciones (MAQC Consortium et al., 2006; Love et al., 2014). En ambos casos, se tomó como diferencia significativa solo aquellos que poseían un valor p reajustados para pruebas múltiples utilizando el procedimiento de Benjamini y Hochberg menor a 0.05 ($p.adj < 0.05$). El algoritmo DESeq2, apeglm y la elección del método de reajuste de valor p fueron realizados siguiendo las recomendaciones del manual oficial (Love et al., 2014).

La visualización de los resultados se realizó mediante gráficos de volcán *(volcano plots*), generados con el paquete EnhancedVolcano (Blighe, 2018).

3.6. Análisis de vías biológicas

El análisis de vías biológicas se llevó a cabo mediante el método computacional de Análisis de Enriquecimiento de Conjunto de Genes (GSEA, del inglés *Gene Set Enrichment Analysis*). Esta metodología se implementó utilizando el paquete R ClusterProfiler. GSEA es un método estadístico que permite identificar conjuntos de genes (vías) que están significativamente enriquecidos en una condición fenotípica dada, proporcionando un contexto biológico a los resultados de la expresión diferencial (Nguyen et al., 2019). Para la ejecución del GSEA prerankeado (*preranked GSEA*), fue necesario generar una lista de genes clasificada. Como métrica de clasificación (*ranking metric*), se seleccionó el estadístico de Wald obtenido de los resultados del análisis DEG entre CMTN y CMnTN. La elección del estadístico de Wald es debido a que esta métrica combina la significancia estadística con la dirección del cambio en la expresión génica, es fundamental y recomendada para el preranking en GSEA (Subramanian et al., 2005; Love et al., 2014). Los valores p nominales fueron reajustados para pruebas múltiples utilizando el procedimiento de Benjamini y Hochberg, siguiendo las recomendaciones del manual oficial Subramanian et al., (2005). Específicamente, este análisis se realizó utilizando únicamente las vías contenidas en WikiPathways, las cuales fueron previamente descargadas de la colección MSigDB.

La manipulación, filtración y representación de datos se realizaron en el entorno estadístico R versión 4.4.1. El flujo de trabajo, los softwares y bases de datos utilizados son resumidos en la **Figura 1**.

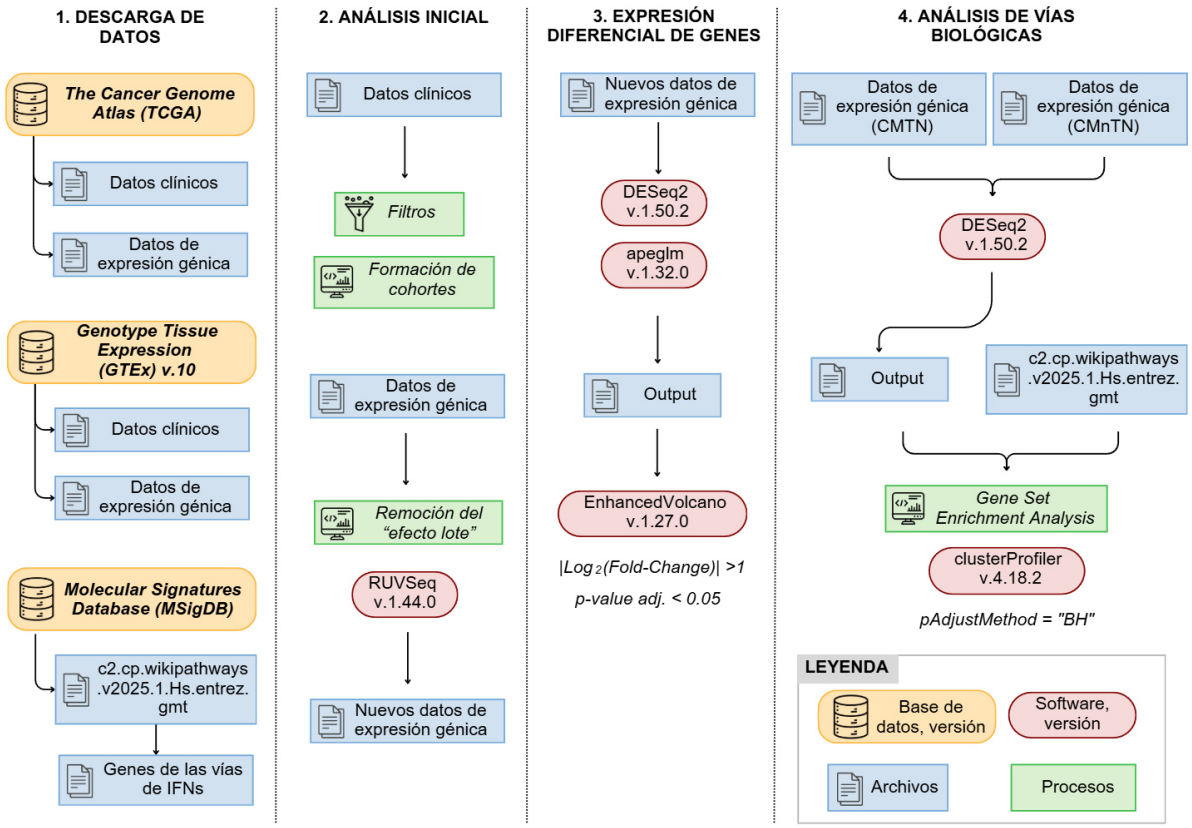


Figura 1. Flujograma general de trabajo incluyendo informaciones solicitadas en el MIABi. *Nota*: Se incluyen las versiones de las bases de datos y softwares utilizados. *p-value adj*: valor p ajustado por el método de Benjamini y Hochberg.

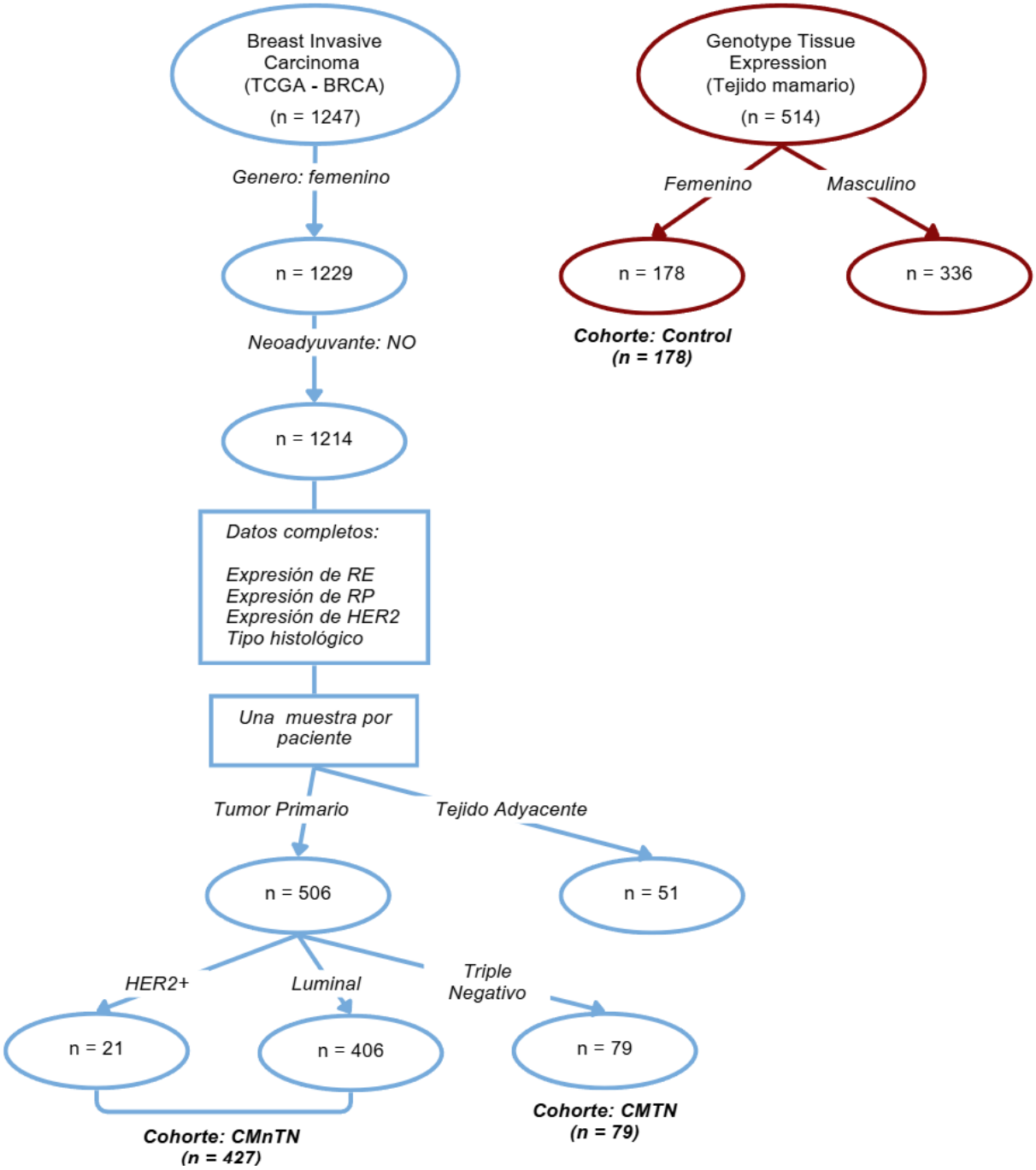


Figura 2. Diagrama de flujo de selección de muestra. Se detallan los criterios de inclusión aplicados a los datos provenientes de TCGA-BRCA (celeste) y GTEx (marrón). Se reporta el tamaño muestral retenido (n) tras aplicar cada criterio. *Nota*: CMTN: cohorte cáncer de mama triple negativo; CMnTN: cohorte cáncer de mama no triple negativo; RE: receptor de estrógeno; RP: receptor de progesterona; HER2: receptor 2 del factor de crecimiento epidérmico humano.

4. RESULTADOS

4.1. Muestras tumorales empleadas en el análisis (CMTN, CMnTN)

Se obtuvo un total de 557 muestras provenientes del TCGA luego del proceso de selección y filtrado según sus datos clínicos, las cuales fueron divididas según su código de barras en tumor primario y tejido sólido normal, del cual hablaremos más adelante. Las 506 muestras provenientes de tumor primario fueron divididas en CMTN y CMnTN (**Tabla 2**). Los datos clínicos de mayor relevancia pertenecientes a las muestras se encuentran resumidos en la **Tabla 3**.

Tabla 2. *Distribución de las muestras extraídas del TCGA-BRCA según su origen y clasificación.*

Clasificación	Código de barras del TCGA-BRCA	Origen	Número de muestras
CMTN	TCGA-XX-XXXX-01	Tumor Primario	79
CMnTN			427
Tejido Adyacente	TCGA-XX-XXXX-11	Tejido sólido normal	51

Nota. TCGA-BRCA: Cohorte de cáncer de mama del proyecto *The Cancer Genome Atlas*; CMTN: Cáncer de mama triple negativo; CMnTN: Cáncer de mama no triple negativo.

Tabla 3. *Resumen de las principales características pertenecientes a la cohorte tumoral.*

Característica	CMTN (n = 79)	CMnTN (n = 427)	valor p [2*]
Clasificación IHQ			**<0.001**
Her2 +	0 (0%)	21 (4.9%)	
Luminal	0 (0%)	406 (95%)	
Triple Negativo	79 (100%)	0 (0%)	
Tipo de carcinoma			**<0.001**
Ductal infiltrante	70 (89%)	326 (76%)	

Lobulillar infiltrante	2 (2.5%)	66 (15%)	
Carcinoma medular	1 (1.3%)	0 (0%)	
Carcinoma metaplásico	2 (2.5%)	0 (0%)	
Carcinoma mucinoso	0 (0%)	2 (0.5%)	
Otro	4 (5.1%)	33 (7.7%)	
Reaparición pos tratamiento inicial			**0.3**
Sí	10 (14%)	36 (9.6%)	
No	62 (78.5%)	338 (79.2%)	
Margen			**0.3**
Cercano	5 (6.3%)	15 (3.6%)	
Negativo	71 (90%)	374 (89%)	
Positivo	3 (3.8%)	32 (7.6%)	
Estadio Patológico			**0.13**
I	17 (21.5%)	65 (15.2%)	
II	51 (64.6%)	256 (60%)	
III	8 (10.1%)	95 (22.2%)	
IV	1 (1.3%)	6 (1.4%)	
X	0 (0%)	2 (0.5%)	

Nota: CMTN: Cáncer de mama triple negativo; CMnTN: Cáncer de mama no triple negativo; IHQ: inmunohistoquímica. *2**: Prueba de chi-cuadrado de Pearson; prueba exacta de Fisher; prueba exacta de Fisher para datos de recuento con valores p simulado (basado en 2000 réplicas).

4.2. Muestras no tumorales empleadas en el análisis

Durante el análisis de los identificadores o código de barras de las muestras provenientes del TCGA se identificaron 51 muestras clasificadas como “tejido sólido normal”, es decir, tejido

peritumoral extirpado de pacientes con cáncer, que también contaban con datos transcriptómicos. Estas muestras fueron aprovechadas y agrupadas en una nueva cohorte denominada "Tejido Adyacente" (TA) (**Tabla 2**) (Hang et al., 2022). De estas, 7 muestras proceden de individuos dentro de la cohorte CMTN, 42 de Luminal y 2 de Her2 +.

Se obtuvieron 514 muestras de tejido mamario provenientes de la base de datos GTEx. Debido a que estos datos provienen de donantes sin tumor, no se tiene una tabla de datos clínicos, solo se conoce el sexo del donante (**Tabla 4**).

Tabla 4. *Distribución de las muestras provenientes de GTEx por sexo*.

Sexo	**GTEx - Control (n = 514)**
Femenino	178 (35%)
Masculino	336 (65%)

Nota: GTEx: Base de datos Genotype Tissue Expression.

4.3. Lista de genes empleados en el análisis.

4.3.1. Genes asociados a interferón.

Se obtuvo una lista total de 104 genes asociados a las tres clases de interferones, los cuales son resumidos en tres tablas (Tablas 5 a 7) según el tipo de interferón al que están asociados.

Tabla 5. *Lista de genes asociados a interferón tipo I que fueron incluidos en el análisis de expresión diferencial (n = 71).*

Símbolo	**Entrez ID**	**Símbolo**	**Entrez ID**	**Símbolo**	**Entrez ID**
CBL	867	IFNA8	3445	PIK3R1	5295
CREB1	1385	IFNAR1	3454	PIK3R2	5296
CRK	1398	IFNAR2	3455	PRMT1	3276
CRKL	1399	IFNB1	3456	PTPN11	5781
EIF4A1	1973	IFNE	338376	PTPN6	5777
EIF4B	1975	IFNK	56832	PTPRC	5788

EIF4E	1977	IFNW1	3467	RAC1	5879
EIF4EBP1	1978	IRF9	10379	RAP1A	5906
FYN	2534	IRS1	3667	RAPGEF1	2889
GAB2	9846	IRS2	8660	REL	5966
GNB2L1	10399	JAK1	3716	RPS6	6194
IFNA1	3439	LCK	3932	RPS6KA4	8986
IFNA10	3446	MAP2K3	5606	RPS6KA5	9252
IFNA13	3447	MAP2K6	5608	RPS6KB1	6198
IFNA14	3448	MAP3K1	4214	RPTOR	57521
IFNA16	3449	MAPK14	1432	SOCS1	8651
IFNA17	3451	MAPKAP1	79109	SOCS3	9021
IFNA2	3440	MLST8	64223	STAT1	6772
IFNA21	3452	MTOR	2475	STAT2	6773
IFNA4	3441	PDCD4	27250	STAT3	6774
IFNA5	3442	PIAS1	8554	STAT4	6775
IFNA6	3443	PIAS3	10401	STAT5A	6776
IFNA7	3444	PIK3CD	5293	TYK2	7297
				VAV1	7409
				ZAP70	7535

Tabla 6. *Lista de genes asociados a interferón tipo II que fueron incluidos en el análisis de expresión diferencial (n = 37).*

Símbolo	**Entrez ID**	**Símbolo**	**Entrez ID**	**Símbolo**	**Entrez ID**
CIITA	4261	IFNG	3458	OAS1	4938
EIF2AK2	5610	IFNGR1	3459	PRKCD	5580
CXCL10	3627	IFNGR2	3460	PSMB9	5698
CXCL9	4283	IL1B	3553	PTPN11	5781
CYBB	1536	IRF1	3659	REG1A	5967
GBP1	2633	IRF2	3660	SOCS1	8651
HIST2H4A	8370	IRF4	3662	SOCS3	9021
HLA-B	3106	IRF8	3394	SPI1	6688

ICAM1	3383	IRF9	10379	STAT1	6772
IFI6	2537	ISG15	9636	STAT2	6773
IFIT2	3433	JAK1	3716	TAP1	6890
IFNA2	3440	JAK2	3717	NOS2	4843
				IFNB1	3456

Tabla 7. *Lista de genes asociados a interferón tipo III que fueron incluidos en el análisis de expresión diferencial (n = 10).*

Símbolo	Entrez ID	Símbolo	Entrez ID	Símbolo	Entrez ID
IRF9	10379	IFNL2	282616	IL10RB	3588
JAK1	3716	IFNL3	282617	STAT1	6772
IFNL1	282618	IFNLR1	163702	STAT2	6773
				TYK2	7297

4.3.2. Genes de referencia

Se recopiló una lista de 17 genes que cumplían con los requisitos para ser utilizados como genes de expresión constitutiva en el tejido mamario. Es decir, su expresión no está influenciada por las categorías analizadas en el presente trabajo (tumoral y no tumoral). Solo 16 de ellos fueron utilizados para mitigar el efecto de lote, dado que el gen *CCSER2* no contó con datos de expresión génica en todas las muestras.

Tabla 8. *Lista de genes de referencia o expresión constitutiva (n = 17)*.

Genes	Referencia bibliográfica
CCSER2, SYMPK, ANKRD17, PUM1	Tilli et al. (2016)
ACTB, RPS23, HUWE1, EEF1A1, SF3A1, MRPL19	Maltseva et al. (2013)
RPL13A	Kılıç et al.(2014)
SDHA	Gur-Dedeoglu et al. (2009)

TBP, RPLP0	Lyng et al. (2008)
HNRNPL, PCBP1, RER1	Jo et al. (2019)

4.4. Análisis de datos inicial

4.4.1. Reducción de dimensionalidad

Antes de realizar análisis estadísticos como la expresión diferencial de genes, es fundamental comprender las características globales del conjunto de datos de expresión génica (Woo & Jung, 2023). Por este motivo, se efectuó un primer PCA no supervisado utilizando los 500 genes con mayor varianza en su expresión entre los grupos o cohortes (**Figura 3**). Los valores de expresión fueron previamente transformados mediante la función *varianceStabilizingTransformation* del paquete DESeq2, siguiendo las recomendaciones de Love et al. (2014).

En este primer análisis, los grupos tumorales (CMTN y CMnTN), las muestras de tejido adyacente y el tejido Control tienden a formar tres agrupaciones, ubicando al tejido adyacente como un punto intermedio entre los controles y los tumores. Así mismo, se observa una tendencia del grupo Control a formar dos agrupaciones internas, lo cual evidencia la presencia de variabilidad no deseada incluida dentro de este grupo (**Figura 3**).

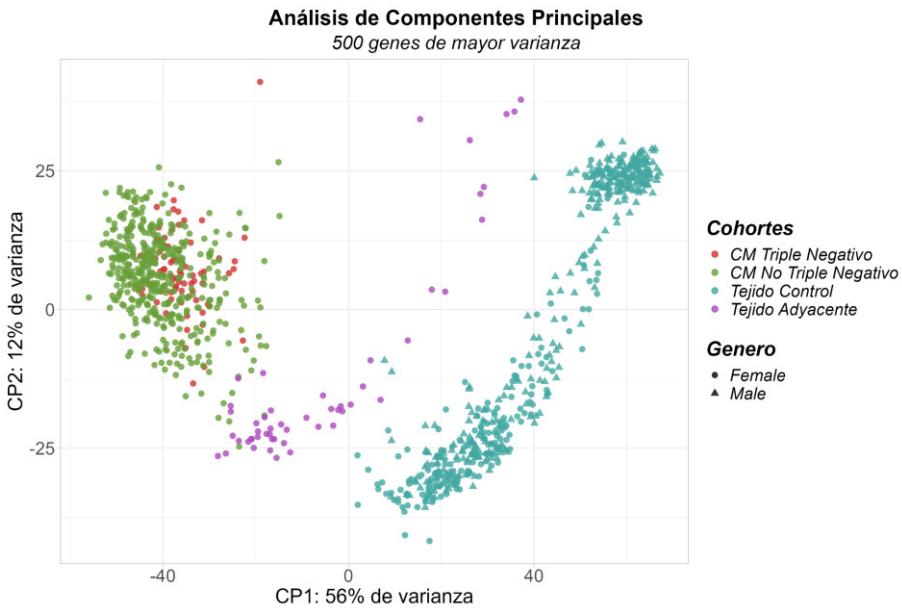


Figura 3. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales utilizando la expresión de 500 genes de mayor varianza. *Nota*: CP1: Componente principal 1, CP2: Componente principal 2.

Para minimizar esta variabilidad, y debido a que la totalidad de las muestras tumorales provienen de mujeres, se optó por eliminar del análisis las muestras Control proveniente de donantes masculinos (n = 336), quedando en este grupo un total de 178 muestras y se repitió el PCA (**Figura 4**). Como resultado se obtuvo una mejor separación de las cohortes tumorales y del tejido adyacente, al mismo tiempo que la cohorte Control estuvo menos dispersa. Posteriormente, se evaluó la dispersión utilizando exclusivamente genes relacionados con interferón. En este caso, las muestras se agruparon en dos grandes grupos: controles por un lado y tumores junto con tejidos adyacentes por el otro (**Figura 5**).

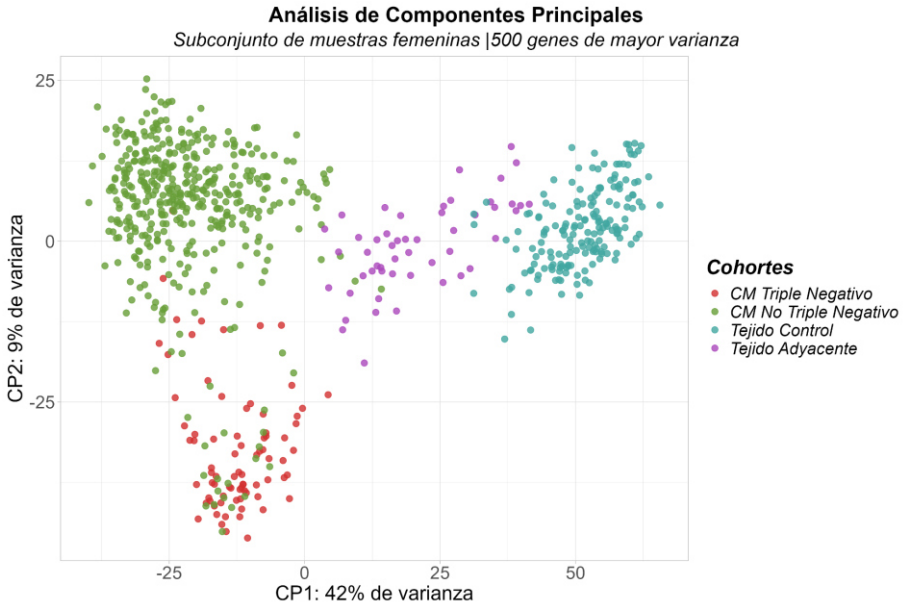


Figura 4. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales solo de muestras pertenecientes a mujeres. Se utilizan los 500 genes de mayor varianza. *Nota*: CP1: componente principal 1, CP2: Componente principal 2

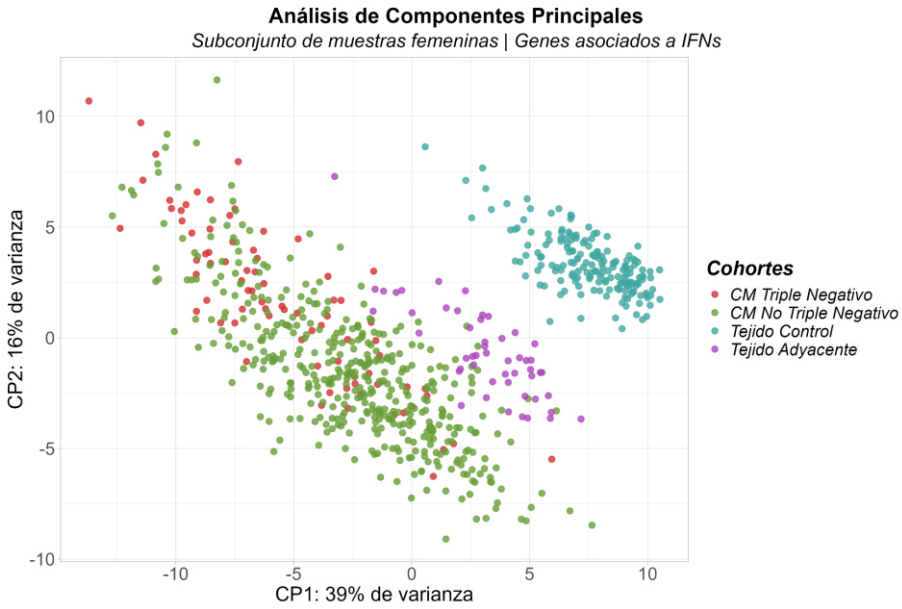


Figura 5. Análisis de componentes principales: Distribución de las cohortes tumorales y no tumorales solo de muestras pertenecientes a mujeres utilizando los genes asociados a interferones (n = 85). *Nota*: CP1: componente principal 1, CP2: Componente principal 2

Dado que los datos transcriptómicos de las muestras tumorales y de tejido adyacente provienen de TCGA y las muestras Control de GTEx, es esperable la presencia de diferencias

sistemáticas en la preparación de librerías, protocolo y plataformas de secuenciación, reconocidas como fuentes comunes de efectos de lote.

Para mitigar el sesgo técnico y preservar las diferencias biológicas relevantes, se aplicó la corrección por lote mediante el algoritmo *Remove Unwanted Variation (RUV)* implementado en el paquete RUVSeq empleando como referencia el conjunto de genes constitutivos de la **Tabla 8**. Al comparar las **Figuras 6 y 7** se puede apreciar que el algoritmo estabiliza la expresión de los genes constitutivos, conservando las variación originada por las diferencias biológicas entre las cohortes: mayor expresión de *HOXA5* en cohortes no tumorales (Control y TA), mayor expresión de *EGFR* en la cohorte CMTN, mayor expresión de GATA3 en la cohorte CMnTN, la baja expresión de genes de los receptores hormonales y HER2 en la cohorte CMTN. Se evidencia también que la variabilidad en la expresión de los genes *CXCL10* e *IRS1* asociados a interferones se preserva tras la corrección.

Con el fin de verificar la efectividad de la corrección, se generó un nuevo PCA utilizando la matriz de expresión ajustada. Previamente, se filtraron los genes de baja expresión (promedio > 99; log10(expresión + 1) > 2) (**Figura 8**). Los datos resultantes fueron nuevamente transformados con DESeq2, a partir del cual se calculó el coeficiente de variación para cada gen. Mediante un análisis exploratorio se evaluaron umbrales de variabilidad de 0.05, 0.075, 0.1, 0.125, 0.15 y 0.2 (**Figura 9**). Tras una inspección visual de los datos, se estableció como corte óptimo el valor de 0.1, el cual conservó los 1279 genes de mayor variabilidad (**Figura 10**) y proporcionó la mejor discriminación entre las cohortes (**Figuras 11**).

Tras mitigar el efecto de lote, las muestras CMTN se agrupan de manera más clara y tienden a separarse de CMnTN, TA y controles. Sin embargo, algunas muestras de CMnTN se solapan con CMTN y viceversa, lo que justifica un análisis más profundo. Cabe destacar que, al

comparar los PCA de genes asociados a interferón antes (**Figura 5**) y después (**Figura 12**) de la corrección por RUVSeq, se observa que la separación entre cohortes desaparece tras la normalización, formando una sola agrupación.

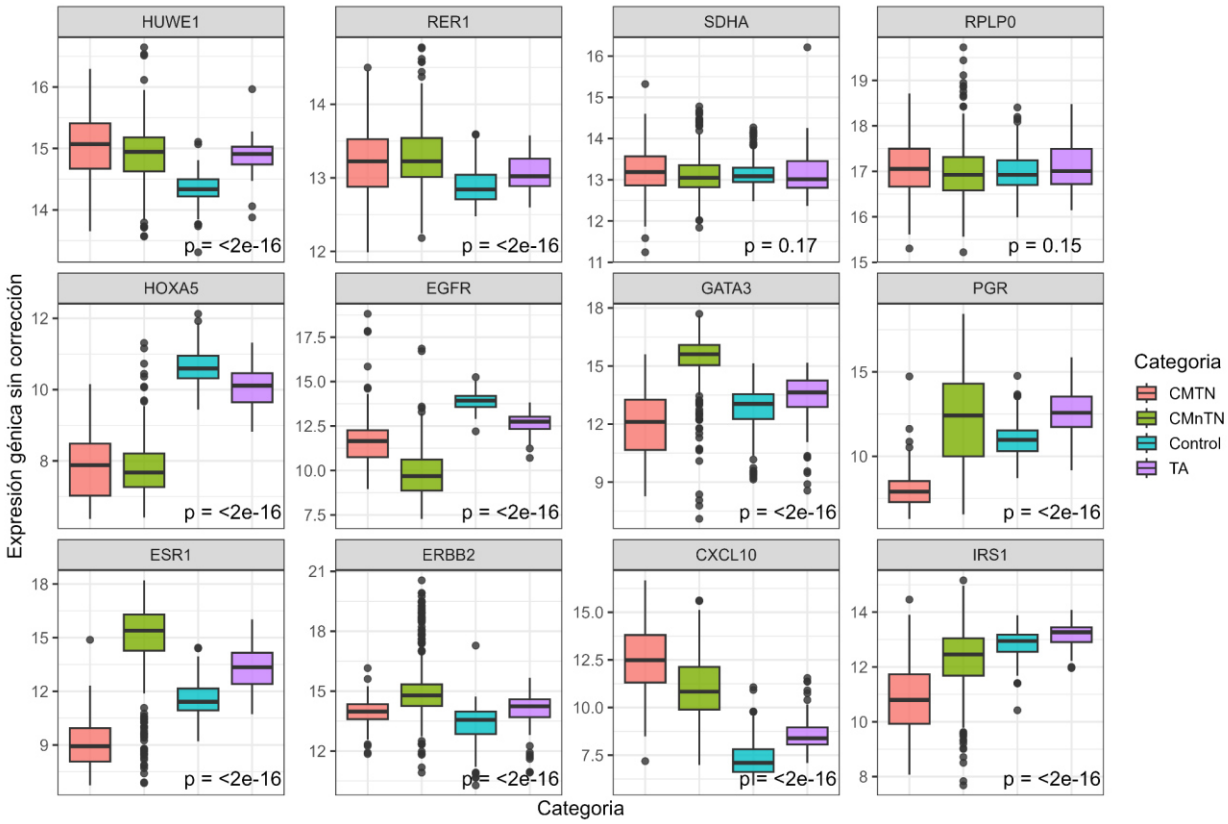


Figura 6: Diagrama de cajas de la expresión génica previo a la corrección por efecto lote. Se representan la expresión de genes de referencia (*HUWE1, RER1, SDHA, RPLP0*), un gen supresor de tumores (*HOXA5*), un gen marcador de CMTN (*EGFR*), un gen marcador de CMnTN (*GATA3*), genes de receptores hormonales (*PGR, ESR1*), el gen del receptor HER2 (*ERBB2*) y genes asociados a interferones (*CXCL10, IRS1*). *Nota*: La significancia estadística (*p*) se determinó a partir de la prueba de Kruskal-Wallis.

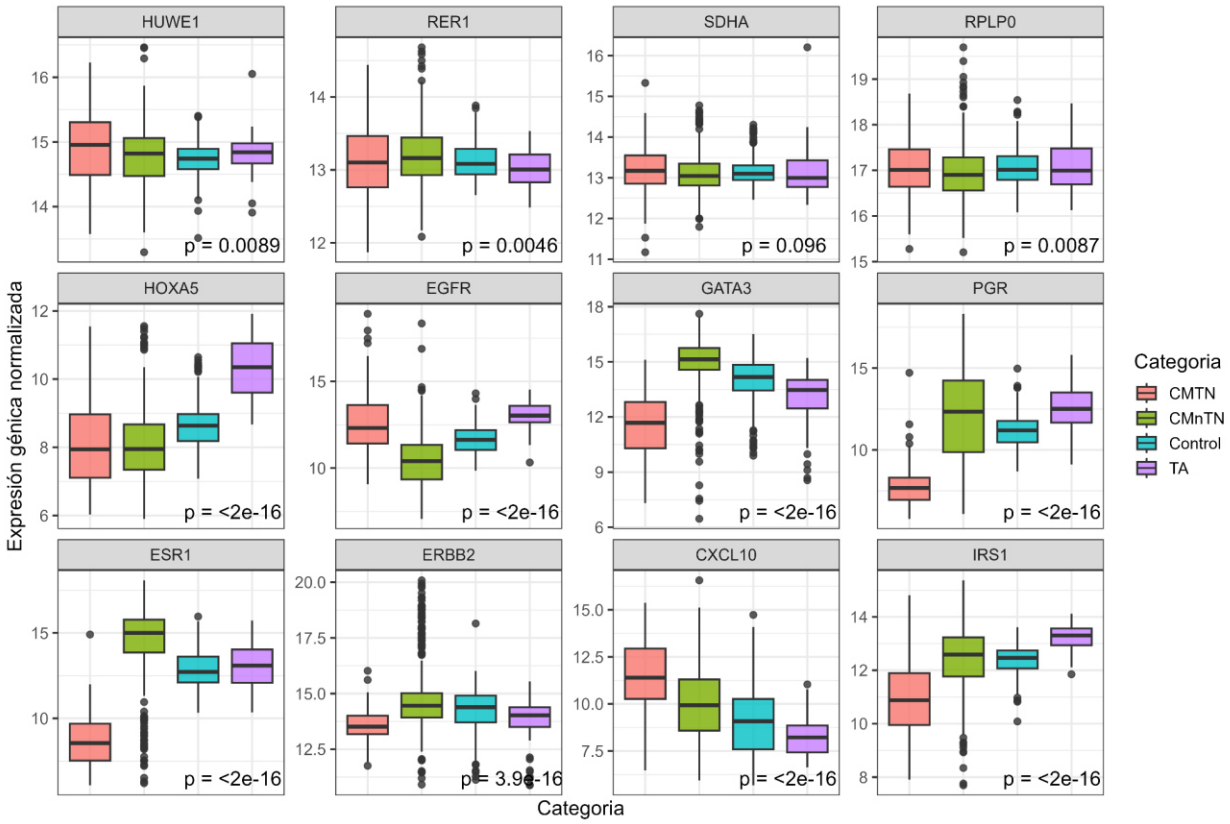


Figura 7: Diagrama de cajas de la expresión génica posterior a la corrección por efecto lote mediante RUVseq. Se representan la expresión de genes de referencia (*HUWE1, RER1, SDHA, RPLP0*), un gen supresor de tumores (*HOXA5*), un gen marcador de CMTN (*EGFR*), un gen marcador de CMnTN (*GATA3*), genes de receptores hormonales (*PGR, ESR1*), el gen del receptor HER2 (*ERBB2*) y genes asociados a interferones (*CXCL10, IRS1*). *Nota*: La significancia estadística (*p*) se determinó a partir de la prueba de Kruskal-Wallis.

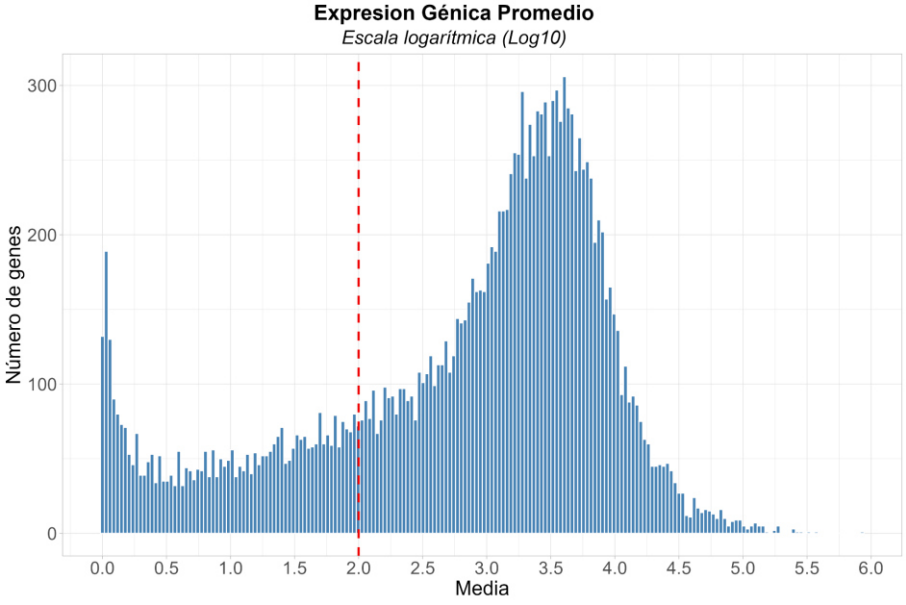


Figura 8. Histograma del promedio de conteos por gen en una escala logarítmica (Log10). *Nota*: La línea punteada roja representa el punto de corte que se utiliza como filtro (expresión promedio igual a 100).

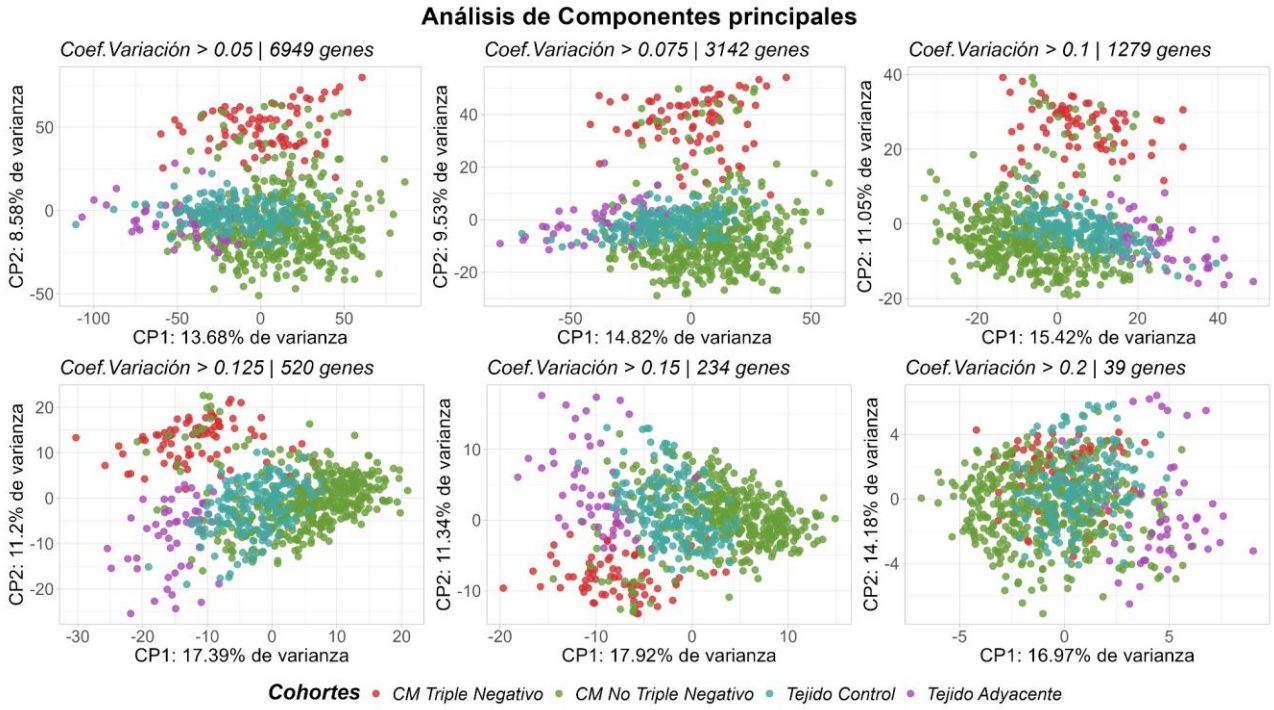


Figura 9. Análisis de componentes principales utilizando diferentes umbrales de selección de genes más variables entre las muestras. De izquierda a derecha y de arriba a abajo, se incluyen los siguientes umbrales: 0.05, 0.075, 0.1, 0.125, 0.15 y 0.2. *Nota*: CP1: Componente principal 1, CP2: Componente principal 2.

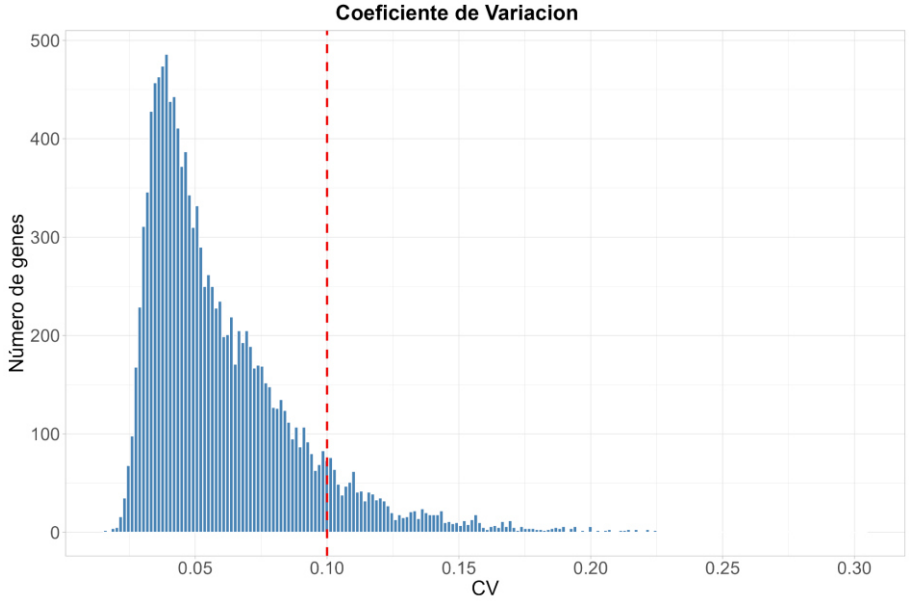


Figura 10. Histograma de la distribución de valores de coeficiente de variación de cada gen incluido en el análisis. La línea punteada roja representa el punto de corte (coeficiente de variación igual a 0.1). *Nota*: CV: Coeficiente de variación; CP1: componente principal 1; CP2: Componente principal 2.

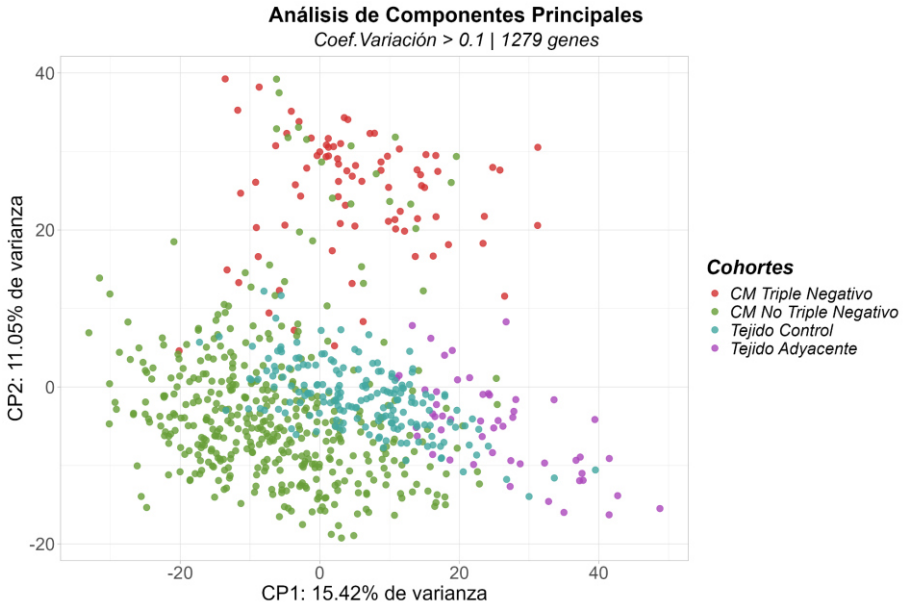


Figura 11. Análisis de componentes principales utilizando los 1279 genes de mayor variabilidad entre las muestras luego de realizar la corrección por efecto de lote. *Nota*: CP1: Componente principal 1; CP2: Componente principal 2.

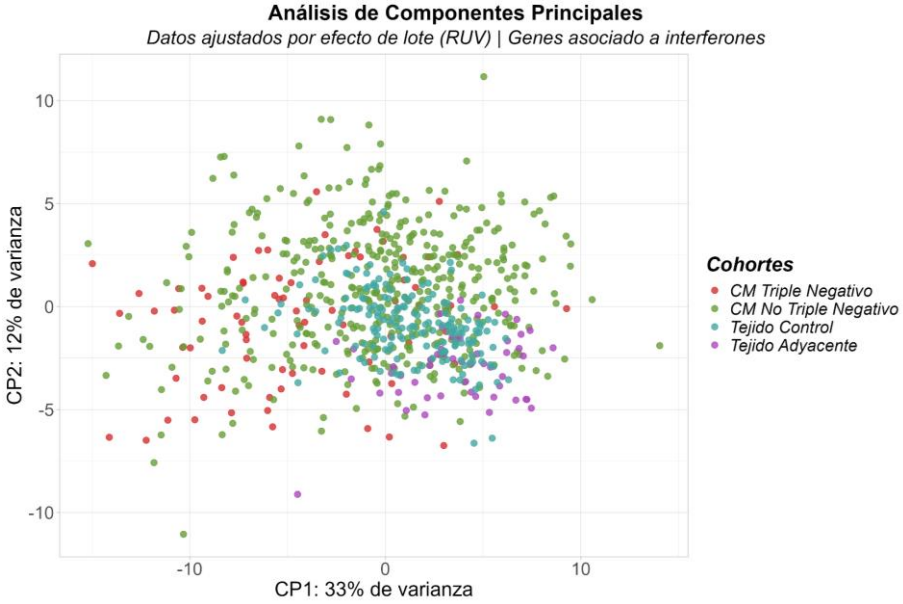


Figura 12. Análisis de componentes principales de las muestras utilizando los genes asociados a interferón (n = 85) luego de realizar la corrección por efecto de lote. *Nota*: CP1: componente principal 1, CP2: Componente principal 2.

4.5. Exploración de los datos de CMTN y CMnTN

Dadas las inconsistencias observadas en la agrupación de CMTN y CMnTN (**Figura 11**), se exploró con mayor detalle la composición de estas cohortes. Se empleó la anotación molecular PAM50 incluida en los datos clínicos para evaluar la correspondencia entre las categorías moleculares y las definidas según la expresión inmunohistoquímica de sus receptores de membrana (**Figura 13**).

Ambos grupos inmunohistoquímicos resultaron ser heterogéneos, con todos los subtipos moleculares presentes en diferentes proporciones: CMTN mostró mayor contribución de muestras "Similar a Basal" (82.278% de casos) y HER2-enriquecido (10.127%), mientras que "Similar a Normal" (5.063%) y luminales (2.532%) fueron menos frecuentes. Por el contrario, CMnTN presentó mayor proporción de muestras Luminal A (59.016%) y B (27.635%), y menor contribución de "Similar a Basal" (4.215%), "Similar a Normal" (3.044%) y HER2-enriquecido (6.089%).

Se realizó un PCA adicional utilizando únicamente muestras tumorales ajustadas por efecto de lote. Tras filtrar los genes de baja expresión (promedio > 99; log10 (expresión + 1) > 2) (**Figura 14**) y aplicar la transformación de varianza de DESeq2, se evaluaron distintos umbrales de variabilidad (0.05 - 0.2) (**Figura 15**). El umbral óptimo de 0.1 retuvo 1473 genes (**Figura 16**), logrando la discriminación más clara entre grupos (**Figuras 17**).

El PCA anotado por subtipo molecular evidencia una tendencia de los datos a tener dos agrupaciones definidas: una correspondiente a muestras "Similar a Basal" y otra que integra Luminal A y B. En contraste, los subtipos HER2-enriquecido y "Similar a Normal" presentan menor separación o diferenciación. La comparación revela que las muestras de CMnTN que se agrupan con CMTN pertenecen principalmente al subtipo "Similar a Basal", mientras que el caso atípico de CMTN agrupado con CMnTN corresponde a Luminal B. En la zona intermedia, donde

ambos grupos se superponen, predominan muestras clasificadas como HER2-enriquecido según PAM50.

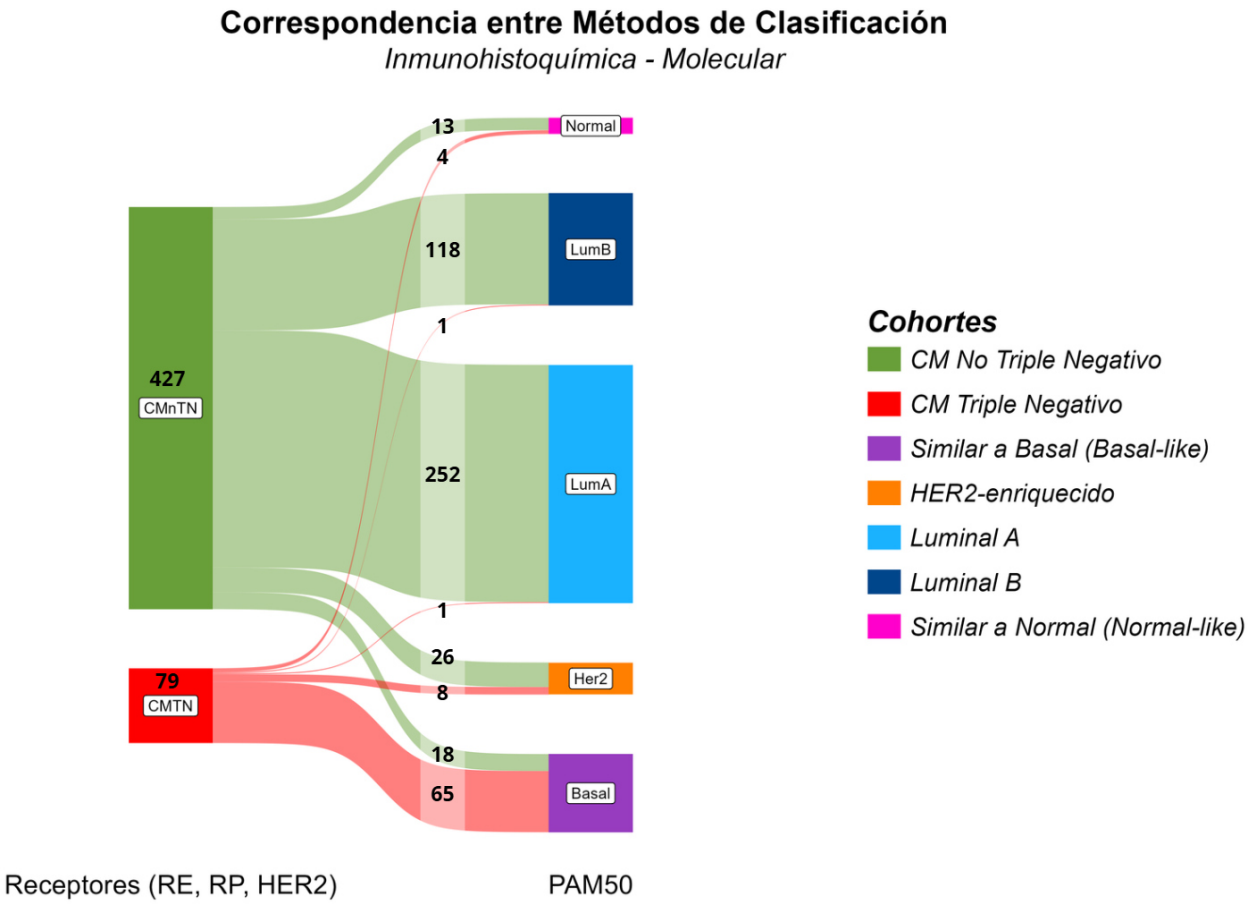


Figura 13. Diagrama de Sankey: Representación de la correspondencia entre los grupos formados utilizando información de expresión de receptores y los grupos formados utilizando la información molecular PAM50. Se indica el número de muestras por cada categoría. *Nota*: RE:Receptor de Estrógenos; RP: Receptor de Progesterona; HER2: Receptor 2 del factor de crecimiento epidérmico humano

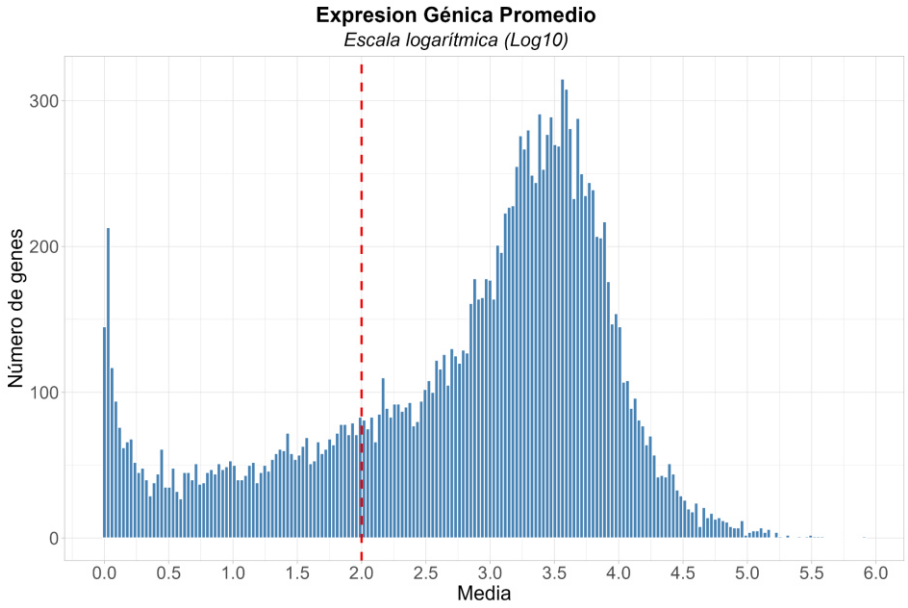


Figura 14. Histograma del promedio de conteos por gen en una escala logarítmica (Log10). La línea punteada roja representa el punto de corte que se utiliza como filtro (expresión promedio igual a 100). *Nota*: Solo se consideran los conteos de genes de las muestras que pertenecen a CMTN y CMnTN.

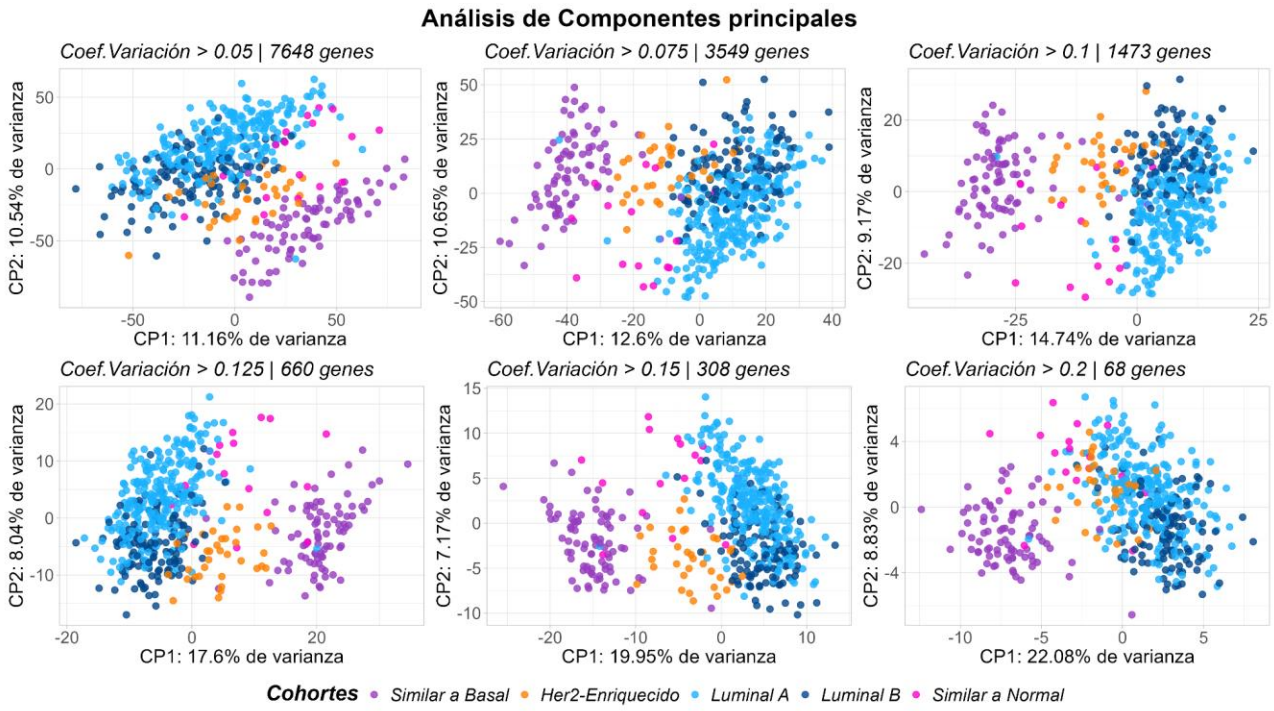


Figura 15. Análisis de componentes principales utilizando diferentes umbrales de selección de genes más variables entre las muestras. De izquierda a derecha y de arriba a abajo, se incluyen los siguientes umbrales: 0.05, 0.075, 0.1, 0.125, 0.15 y 0.2. *Nota*: Solo se consideran los conteos de

las muestras que pertenecen a CMTN y CMnTN. CP1: Componente principal 1, CP2: Componente principal 2.

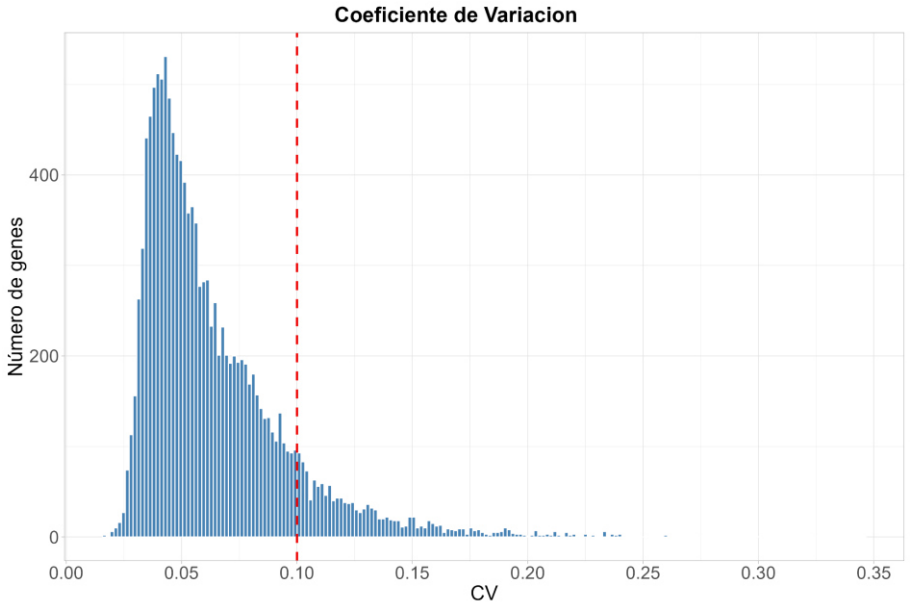


Figura 16: Histograma de la distribución de valores de coeficiente de variación de cada gen incluido en el análisis. La línea punteada roja representa el punto de corte (coeficiente de variación igual a 0.1).

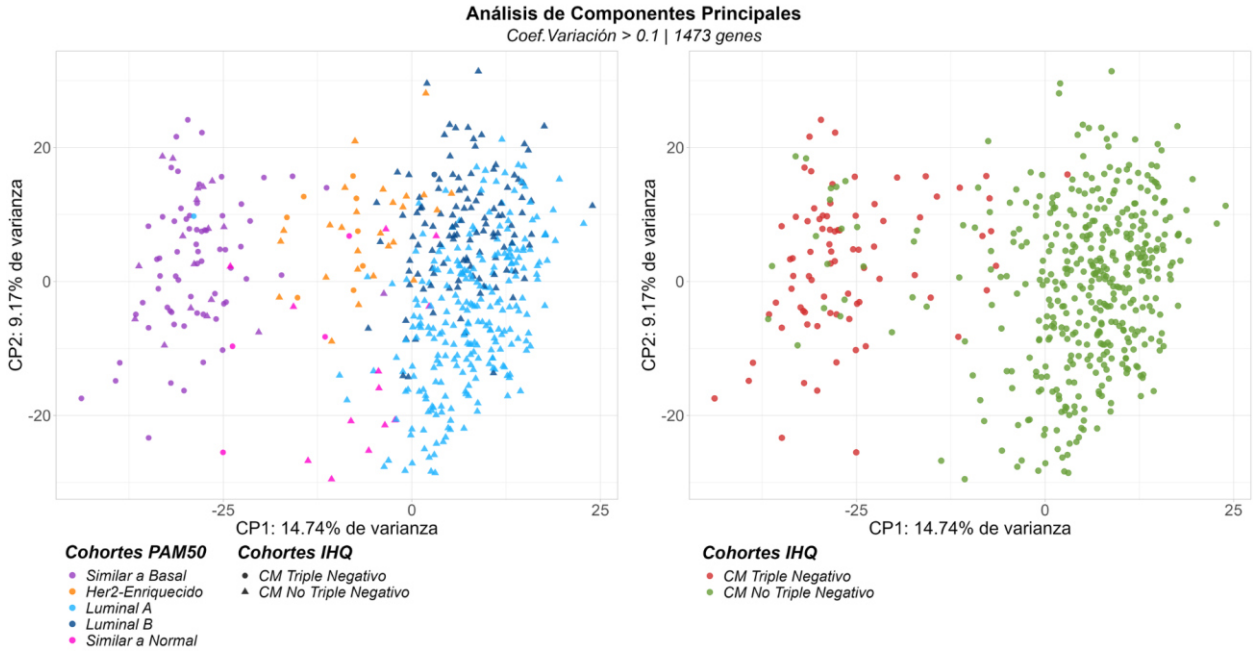


Figura 17. Análisis de componentes principales comparativo utilizando los 1473 genes de mayor variabilidad entre las muestras y que mejor separan a los grupos PAM50. *Nota*: A la izquierda los colores del PCA representan la anotación por PAM50 y las figuras la anotación por IHQ. A la

derecha, los colores representan la anotación por IHQ. CP1: Componente principal 1, CP2: Componente principal 2.

4.6. Análisis de expresión diferencial de genes

Una vez consolidadas las cohortes y corregido el efecto de lote, se procedió al análisis de expresión diferencial utilizando DESeq2.

4.6.1. CMTN frente a Control

La comparación entre el grupo CMTN y las muestras Control permitió identificar 28 genes sobreexpresados y 7 genes subexpresados (**Figura 18 y 19**). En la vía del IFN-I se observó la sobreexpresión de ocho genes (*EIF4E, IFNAR2, LCK, PIK3R2, PTPRC, STAT1, VAV1* e *IFNB1*) y la subexpresión de siete genes (*IRS1, IRS2, MAP2K6, PIK3R1, RPS6KA5, STAT5A* y *TYK2*). Respecto a la vía del IFN-II, se detectó la sobreexpresión de 18 genes (*STAT1, CXCL10, CXCL9, CYBB, GBP1, HLA-B, ICAM1, IFI6, IFIT2, IFNB1, IFNG, IFNGR2, IL1B, IRF1, ISG15, OAS1, PSMB9* y *TAP1*). En la vía del IFN-III se evidenció la sobreexpresión de cinco genes (*STAT1, IFNL1, IFNL2, IFNL3* e *IL10RB*) y la subexpresión de *TYK2*.

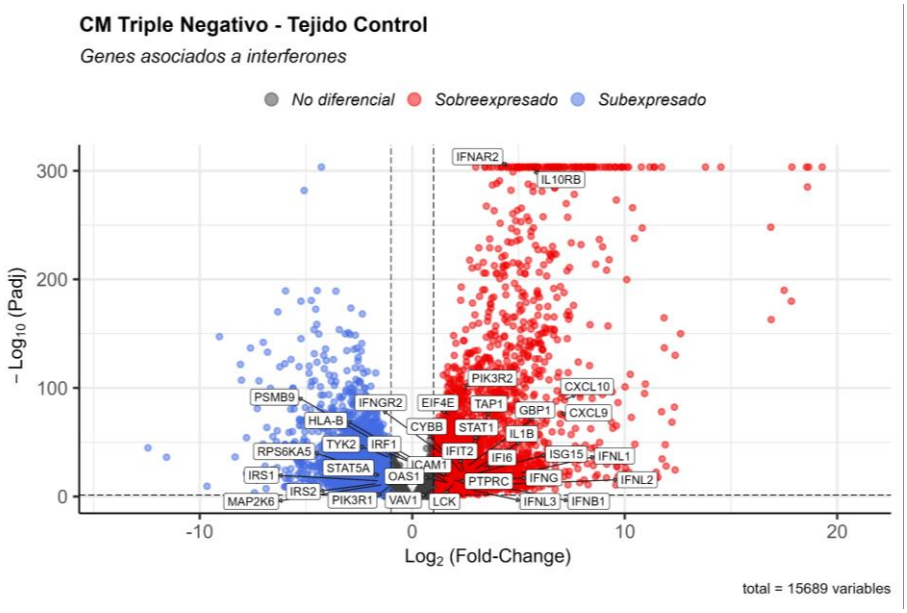


Figura 18. Diagrama de volcán mostrando genes diferencialmente expresados entre CMTN y Tejido Control. Cada gen es representado por un punto; los puntos rojos simbolizan genes sobreexpresados (*Fold-Change* > 2) y azules simbolizan genes subexpresados (*Fold-Change* < -2). Se considera que la diferencia es significativa si el valor p ajustado (*p.adj*) es menor que 0.05. Se incluyeron etiquetas para genes asociados a las vías de interferones.

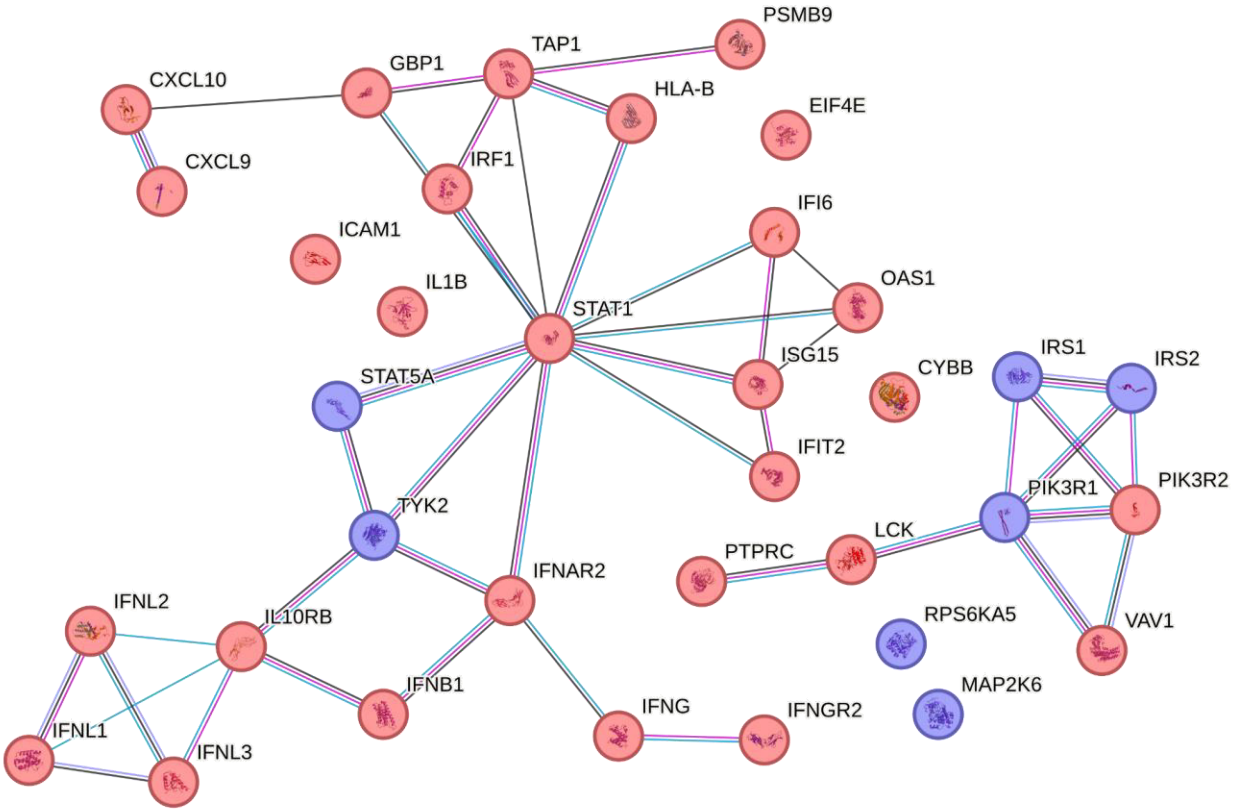


Figura 19: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. Control). *Nota.* Las proteínas en los nodos en rojo corresponden a genes sobreexpresados y los nodos en azul a genes subexpresados en la cohorte de CMTN. Las líneas representan el tipo de evidencia que respalda la asociación: bases de datos curadas (línea celeste), determinación experimental (línea rosada), homología proteica (línea morada) y coexpresión (línea negra). Figura generada mediante la plataforma STRING (*Search Tool for Retrieval of Interacting Genes/Proteins*).

4.6.2. CMTN frente a TA

La comparación entre el grupo CMTN y la totalidad de las muestras de tejido adyacente (TA) presentó la sobreexpresión de 28 genes y la subexpresión de 7 genes en comparación con el grupo TA (**Figura 20 y 21**). En la vía del interferón tipo I (IFN-I) se identificaron nueve genes

sobreexpresados (*EIF4EBP1, IRF9, LCK, PIK3R2, PTPRC, SOCS1, STAT1, ZAP70* e *IFNB1*) y siete genes subexpresados (*EIF4B, IRS1, IRS2, PIK3R1, RPS6KA5, SOCS3* y *STAT5A*). Respecto a la vía del interferón tipo II (IFN-II), se detectó la sobreexpresión de 20 genes (*IRF9, SOCS1, STAT1, CXCL10, CXCL9, EIF2AK2, GBP1, HLA-B, ICAM1, IFI6, IFIT2, IFNB1, IFNG, IRF1, IRF4, ISG15, OAS1, PSMB9, REG1A* y *TAP1*), mientras que solo *SOCS3* se encontró subexpresado. En cuanto a la vía del interferón tipo III (IFN-III), se evidenció la sobreexpresión de cinco genes (*IRF9, STAT1, IFNL1, IFNL2* e *IFNL3*).

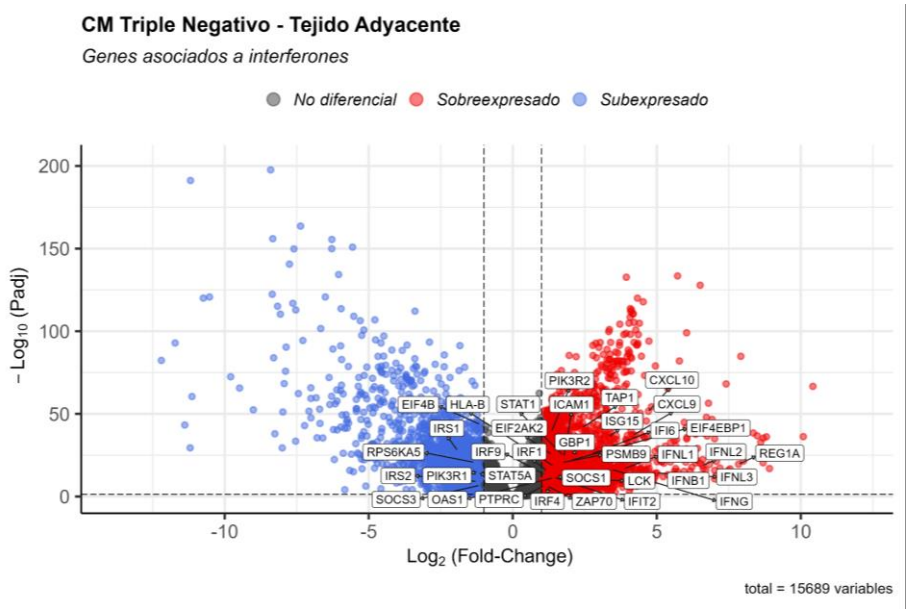


Figura 20. Diagrama de volcán mostrando genes diferencialmente expresados entre 79 muestras de CMTN y 51 muestras de TA. Cada gen es representado por un punto; los puntos rojos simbolizan genes sobreexpresados (*Fold-Change* > 2) y azules simbolizan genes subexpresados (*Fold-Change* < -2). Se considera que la diferencia es significativa si el valor p ajustado es menor que 0.05. Se incluyeron etiquetas para genes asociados a las vías de interferones.

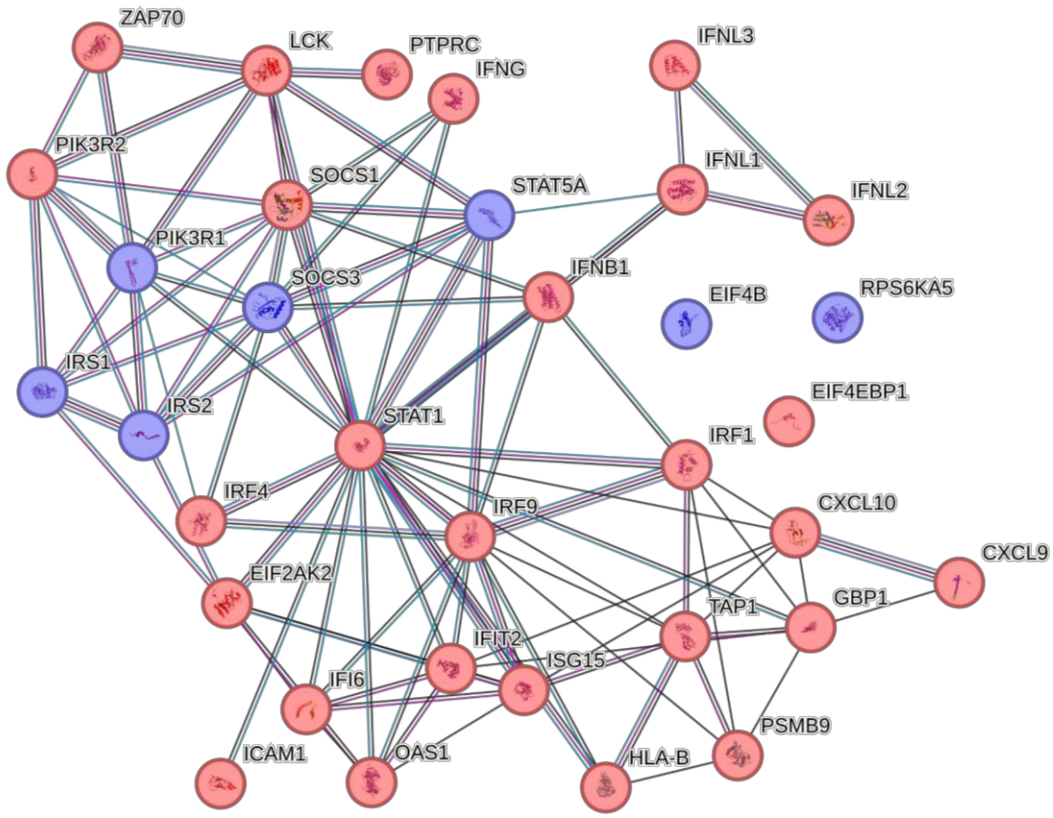


Figura 21: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. TA). *Nota.* Las proteínas en los nodos en rojo corresponden a genes sobreexpresados y los nodos en azul a genes subexpresados en la cohorte de CMTN. Las líneas representan el tipo de evidencia que respalda la asociación: bases de datos curadas (línea celeste), determinación experimental (línea rosada), homología proteica (línea morada) y coexpresión (línea negra). Figura generada mediante la plataforma STRING (*Search Tool for Retrieval of Interacting Genes/Proteins*).

De manera complementaria, se realizó un análisis pareado en un subconjunto de catorce muestras, correspondientes a siete pacientes con tumores CMTN y sus respectivos tejidos adyacentes (TA). El PCA mostró una clara separación entre las muestras CMTN y sus respectivos tejidos adyacentes, destacando además dos muestras de TA con un perfil de expresión génica divergente respecto al resto del grupo (**Figura 22**).

El análisis diferencial de genes en este subconjunto reveló la sobreexpresión de cinco genes y la subexpresión de cinco genes en CMTN frente a sus tejidos pareados (**Figura 23 y 24**). En la

vía del IFN-I, *PIK3R2* se encontró sobreexpresado, mientras que *EIF4B, IRS1, IRS2, PIK3R1 y STAT5A* se encuentran subexpresados. En la vía del IFN-II, se observaron cuatro genes sobreexpresados (*CXCL10, CXCL9, GBP1* e *IFI6*), y no se detectaron cambios en la expresión de genes asociados a la vía del IFN-III.

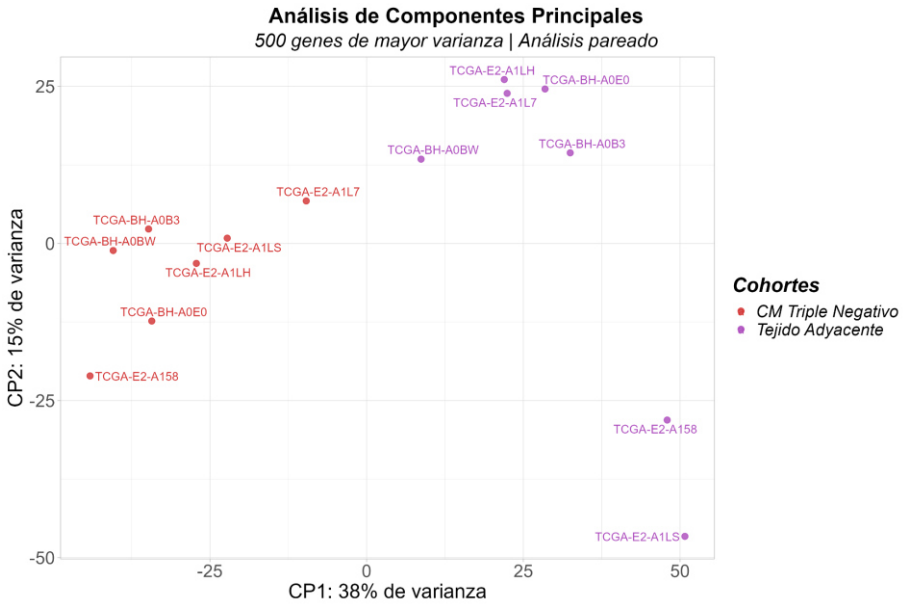


Figura 22. Análisis de componentes principales utilizando 500 genes de mayor varianza, mostrando la distribución de 7 muestras de CMTN y sus respectivas muestras de TA. Se evidencia una clara separación entre el grupo de CMTN y el TA. Se adicionaron los identificadores de pacientes.

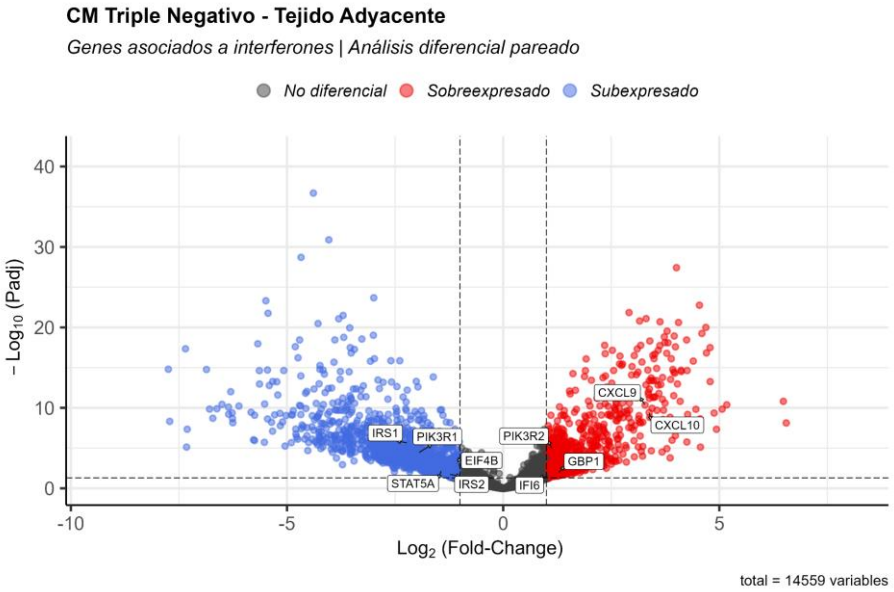


Figura 23. Diagrama de volcán mostrando genes diferencialmente expresados entre 7 muestras de CMTN y sus respectivos TAs. Cada gen es representado por un punto; los puntos rojos simbolizan genes sobreexpresados (*Fold-Change* > 2) y azules simbolizan genes subexpresados (*Fold-Change* < -2). Se considera que la diferencia es significativa si el valor p ajustado es menor que 0.05. Se incluyeron etiquetas para genes asociados a las vías de interferones.

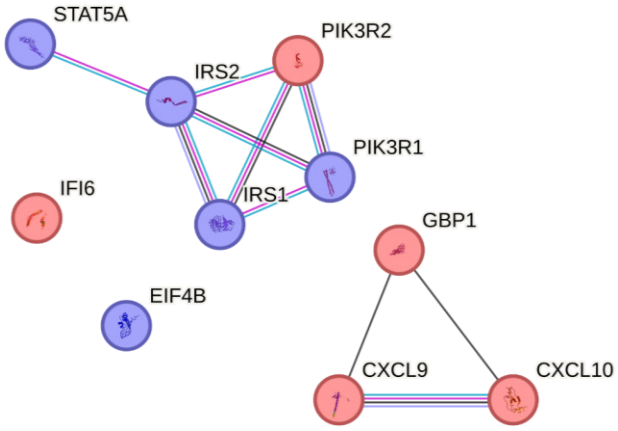


Figura 24: Red de interacción proteína-proteína de los genes diferencialmente expresados en análisis pareado (CMTN vs. TA). *Nota.* Las proteínas en los nodos en rojo corresponden a genes sobreexpresados y los nodos en azul a genes subexpresados en la cohorte de CMTN. Las líneas representan el tipo de evidencia que respalda la asociación: bases de datos curadas (línea celeste), determinación experimental (línea rosada), homología proteica (línea morada) y coexpresión (línea negra). Figura generada mediante la plataforma STRING (*Search Tool for Retrieval of Interacting Genes/Proteins*).

4.6.3. CMTN frente a CMnTN

Con el objetivo de identificar las características de expresión específicas del subtipo CMTN en comparación con otros subtipos tumorales, se realizó un DEG frente al grupo CMnTN. En este contraste, las muestras del subtipo CMTN mostraron la sobreexpresión de 12 genes y la subexpresión de 2 genes.

En la vía del IFN-I, se identificó la sobreexpresión de *IFNAR2, SOCS1, ZAP70* e *IFNE*, mientras que *IRS1* y *MAP3K1* se encontraron subexpresados. En la vía del interferón tipo II (IFN-II), se detectó la sobreexpresión de nueve genes: *SOCS1, CXCL9, CXCL10, GBP1, ICAM1, IFNG, IRF4, REG1A y TAP1*. Por otro lado, en la vía del interferón tipo III (IFN-III) no se observaron cambios significativos en la expresión de los genes analizados (**Figura 25 y 26**). Los resultados encontrados mediante análisis de expresión diferencial de genes del grupo CMTN frente a los cuatro grupos de referencia mencionados anteriormente son resumidos mediante un diagrama de Euler en la **Figura 27** y en la **Tabla 9**.

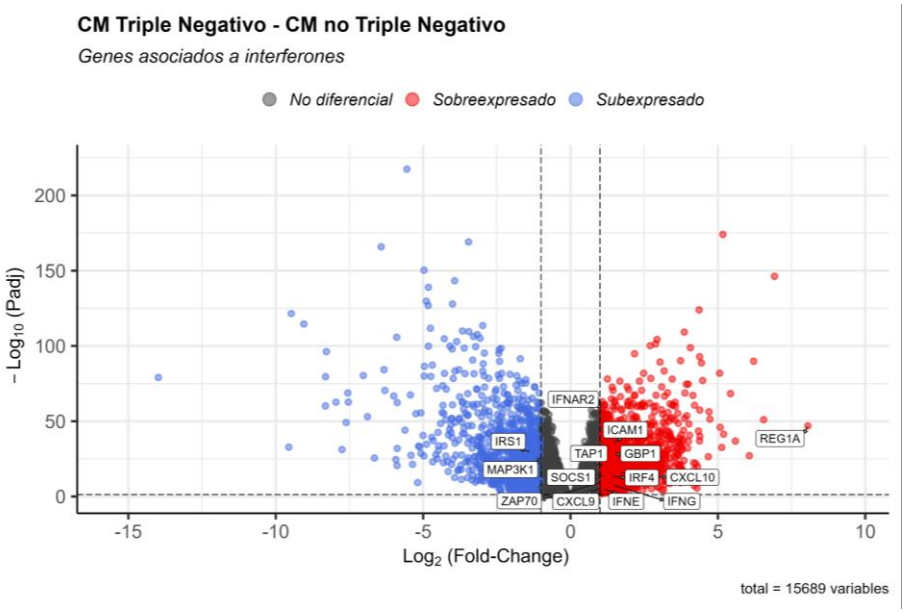


Figura 25. Diagrama de volcán mostrando genes diferencialmente expresados entre muestras de CMTN y CMnTN. Cada gen es representado por un punto; los puntos rojos simbolizan genes sobreexpresados (*Fold-Change* > 2) y azules simbolizan genes subexpresados (*Fold-Change* < -

2). Se considera que la diferencia es significativa si el valor p ajustado es menor que 0.01. Se incluyeron etiquetas para genes asociados a la vía de interferones.

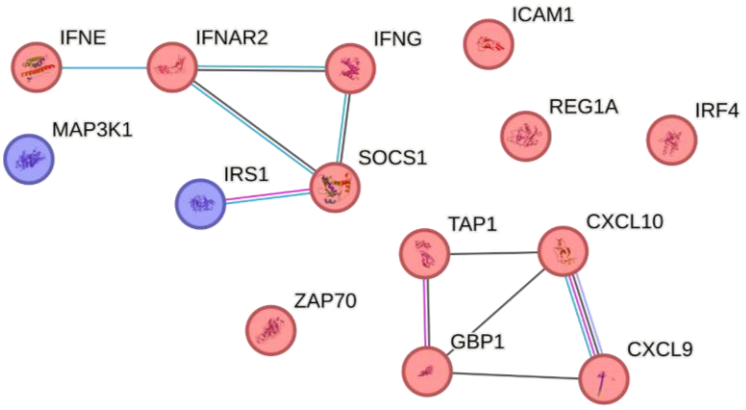


Figura 26: Red de interacción proteína-proteína de los genes diferencialmente expresados (CMTN vs. CMnTN). *Nota.* Las proteínas en los nodos en rojo corresponden a genes sobreexpresados y los nodos en azul a genes subexpresados en la cohorte de CMTN. Las líneas representan el tipo de evidencia que respalda la asociación: bases de datos curadas (línea celeste), determinación experimental (línea rosada), homología proteica (línea morada) y coexpresión (línea negra). Figura generada mediante la plataforma STRING (*Search Tool for Retrieval of Interacting Genes/Proteins*).

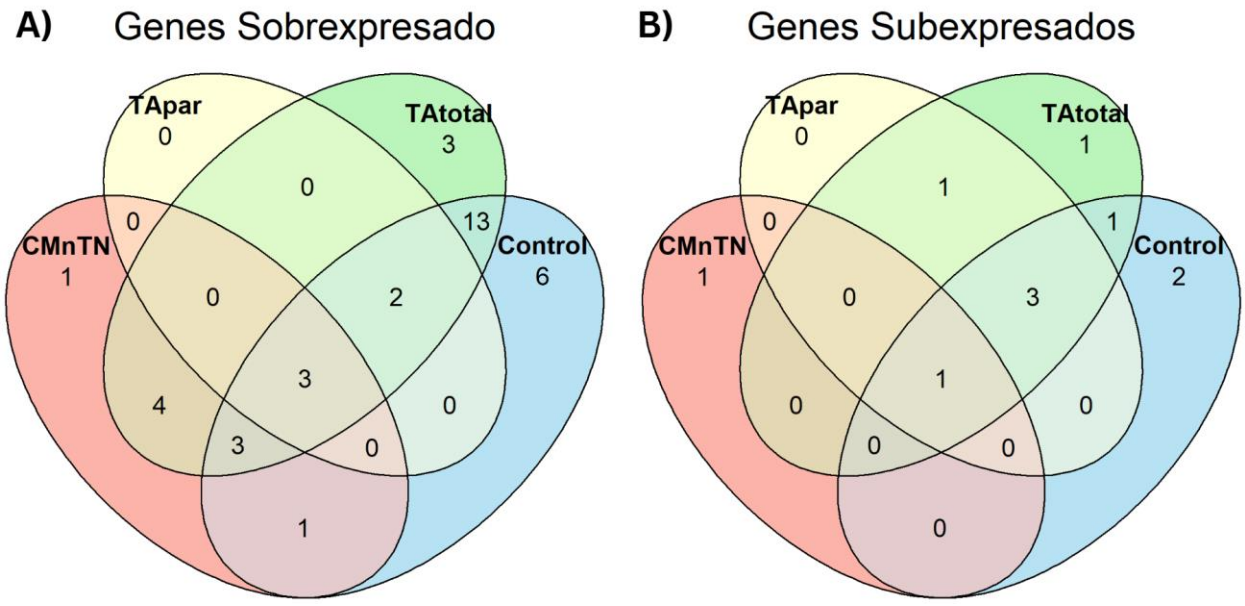


Figura 27. Diagramas Euler de la intersección de genes diferencialmente expresados (A. Sobreexpresados y B. Subexpresados) entre el cáncer de mama triple negativo y las cohortes de comparación. *Nota*: CMnTN: Cáncer de mama no triple negativo; TApar: Tejido adyacente solo de muestras pareadas; TAtotal: Total de muestras de tejido adyacente.

Tabla 9: Genes diferencialmente expresados en las vías de interferón según cada grupo de comparación.

Vía	Regulación	CMTN - Control	CMTN - TAtotal	CMTN - TApar	CMTN - CMnTN
IFN-I	**Sobreexpresados**	*EIF4E, IFNAR2, LCK, PIK3R2, PTPRC, STAT1, VAV1, IFNB1*	*EIF4EBP1, IRF9, LCK, PIK3R2, PTPRC, SOCS1, STAT1, ZAP70, IFNB1*	*PIK3R2*	*IFNAR2, SOCS1, ZAP70, IFNE*
	Subexpresados	*IRS1, IRS2, MAP2K6, PIK3R1, RPS6KA5, STAT5A, TYK2*	*EIF4B, IRS1, IRS2, PIK3R1, RPS6KA5, SOCS3, STAT5A*	*EIF4B, IRS1, IRS2, PIK3R1, STAT5A*	*IRS1, MAP3K1*
IFN-II	**Sobreexpresados**	*STAT1, CXCL10, CXCL9, CYBB, GBP1, HLA-B, ICAM1, IFI6, IFIT2, IFNB1, IFNG, IFNGR2, IL1B, IRF1, ISG15, OAS1, PSMB9, TAP1*	*IRF9, SOCS1, STAT1, CXCL10, CXCL9, EIF2AK2, GBP1, HLA-B, ICAM1, IFI6, IFIT2, IFNB1, IFNG, IRF1, IRF4, ISG15, OAS1, PSMB9, REG1A, TAP1*	*CXCL10, CXCL9, GBP1, IFI6*	*SOCS1, CXCL9, CXCL10, GBP1, ICAM1, IFNG, IRF4, REG1A, TAP1*
	Subexpresados	—	*SOCS3*	—	—
IFN-III	**Sobreexpresados**	*STAT1, IFNL1, IFNL2, IFNL3, IL10RB*	*IRF9, STAT1, IFNL1, IFNL2, IFNL3*	—	—
	Subexpresados	*TYK2*	—	—	—

Nota: CMTN: Cáncer de mama triple negativo; CMnTN: Cáncer de mama no triple negativo; TAtotal: Total de muestras de tejido adyacente; TApar: Tejido adyacente solo de muestras pareadas; IFN: Interferón. El símbolo (—) indica la ausencia de genes.

Con el fin de evaluar de manera integral la activación de las vías de interferón y no solo los genes individuales, se realizó un análisis de enriquecimiento de conjuntos génicos (*gene set enrichment analysis*, GSEA) empleando las rutas definidas en WikiPathways (**Figura 28**). Para cada vía se obtuvieron la puntuación de enriquecimiento (ES), la puntuación de enriquecimiento normalizada (NES), el valor p nominal (*p*) y el valor p ajustado por Benjamini–Hochberg (*p.adj*).

El análisis reveló que la vía del IFN-II presenta el mayor puntuación de enriquecimiento y una significancia estadística de *p.adj* = 0.0708 (NES = 1.76; ES = 0.49; *p* = 0.004), superando a las vías del IFN-I (NES = 1.12; ES = 0.286; *p* = 0.27; *p.adj* = 0.601) e IFN-III (NES = 1.19; ES = 0.458; p = 0.256; *p.adj* = 0.588) (**Figura 29**), las cuales resultaron ser menos significativas estadísticamente.

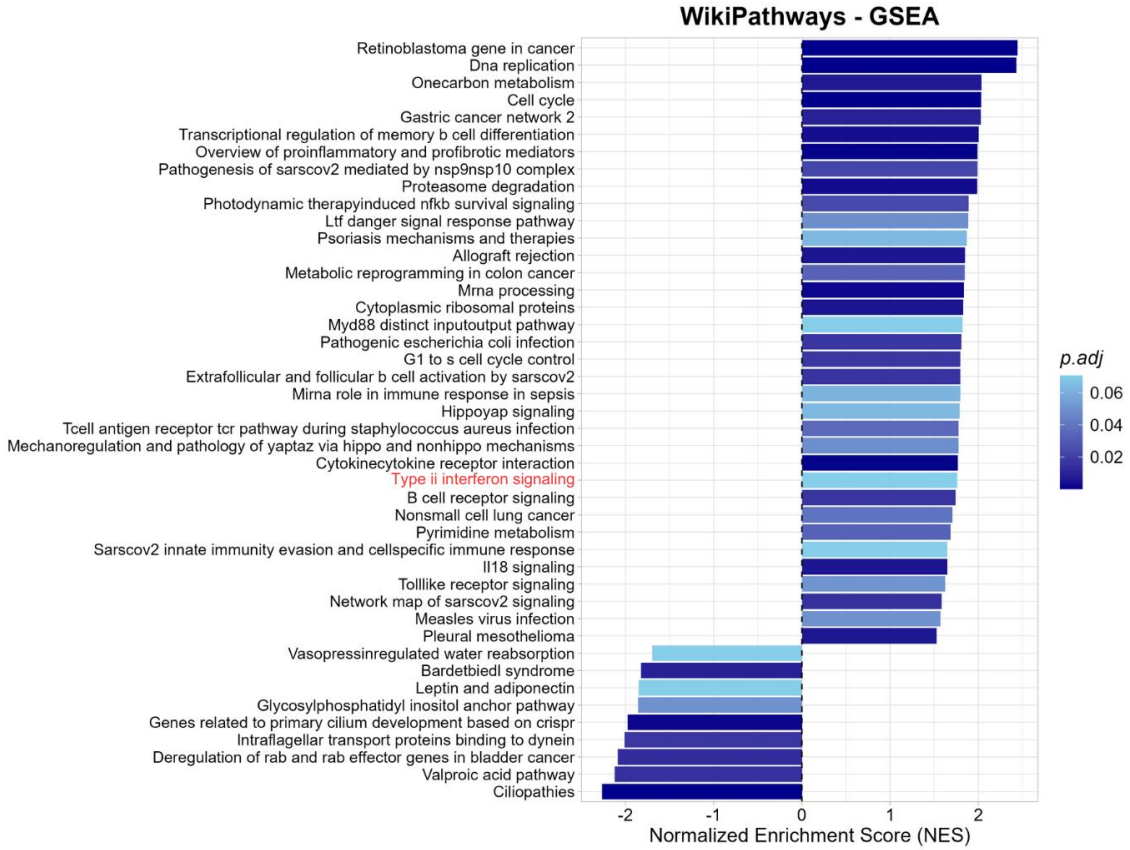


Figura 28. Diagrama de barras mostrando el puntaje de enriquecimiento normalizado (*NES*) de vías bajo el umbral de significancia de 0.071 (*p.adj* < 0.071). Se destaca la posición de la vía relacionada con interferón tipo 2 (*Type 2 interferon signaling*) dentro del conjunto de vías de mayor NES. *Nota*: NES, puntuación de enriquecimiento normalizada; *p.adj*, valor p nominal reajustado por el método Benjamini–Hochberg.

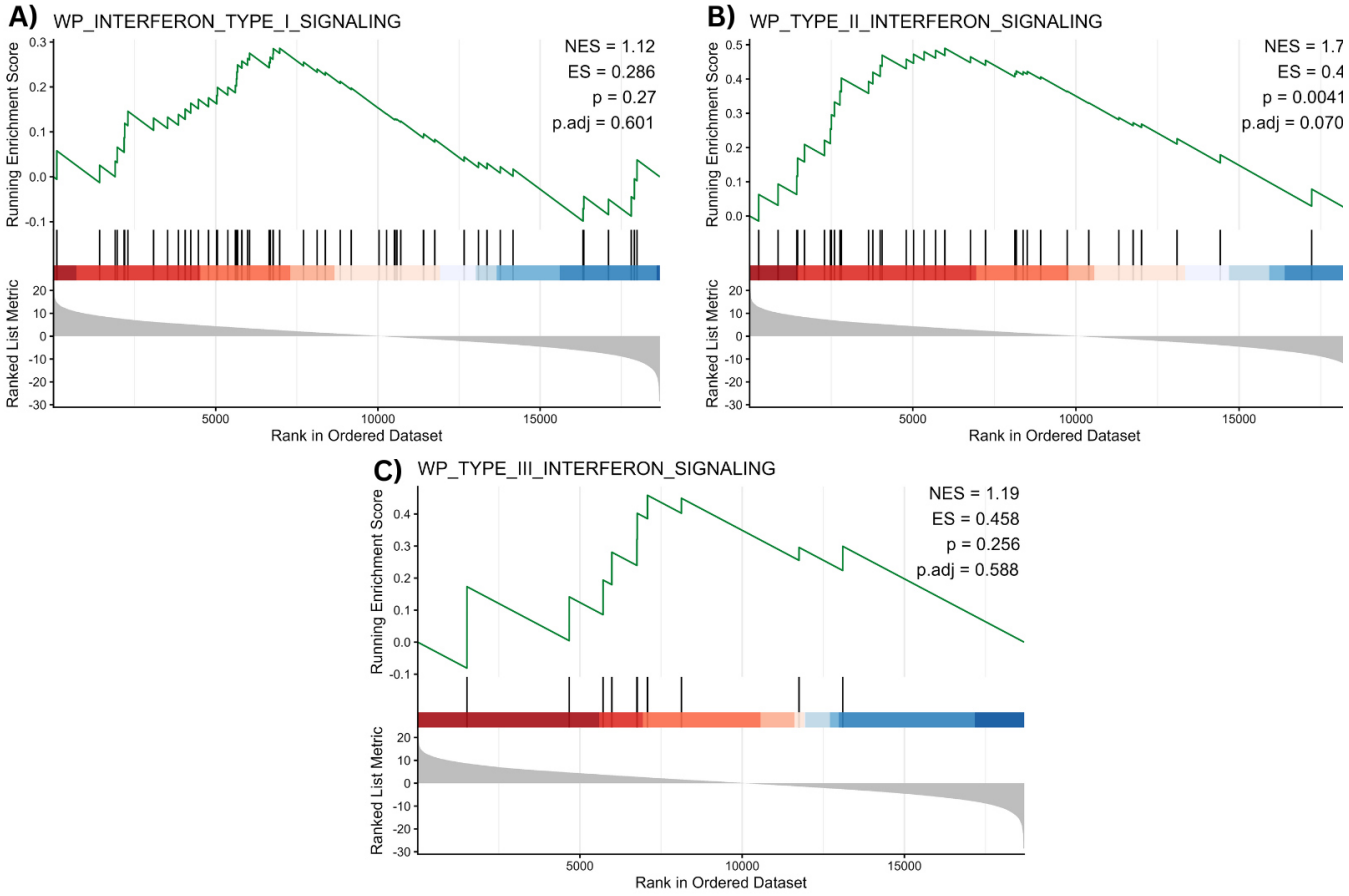


Figura 29. Diagrama de enriquecimiento de las vías de A) interferón tipo I, B) interferón tipo II y C) interferón tipo III. Se muestra la posición de los genes pertenecientes a estas vías dentro de la lista ordenada por la expresión diferencial, representada mediante barras verticales negras. *Nota*: ES, puntuación de enriquecimiento; NES, puntuación de enriquecimiento normalizada; *p*, valor p nominal; *p.adj*, valor p nominal reajustado por el método Benjamini–Hochberg.

5. DISCUSIÓN

El cáncer de mama triple negativo (CMTN) es el subtipo tumoral más agresivo y con peor pronóstico dentro del espectro del cáncer de mama(Zagami & Carey, 2022). Comprender la interacción entre el sistema inmunológico, incluyendo moléculas y células inmunitarias, y las

células cancerígenas dentro del microambiente tumoral resulta esencial para entender el comportamiento del tumor y su respuesta a los tratamientos (Zagami & Carey, 2022).

Actualmente, las tecnologías de secuenciación y análisis de alto rendimiento han permitido la generación masiva de datos biológicos, muchos de los cuales se encuentran disponibles en repositorios públicos gracias a los esfuerzos de grandes consorcios como *The Cancer Genome Atlas* (TCGA) y el proyecto *Genotype-Tissue Expression* (GTEx). Estas bases de datos constituyen una fuente valiosa para formular y responder nuevas preguntas biológicas (Oza et al., 2023).

5.1. Análisis Inicial de Datos

Un paso fundamental para realizar un análisis estadístico apropiado y reproducible es garantizar la integridad y transparencia de los datos mediante una exploración inicial. En otras palabras, es necesario verificar que los datos observados concuerden con las expectativas sobre su comportamiento (Baillie et al., 2022). Con esta finalidad, el PCA permite representar la dispersión de las muestras según sus características de expresión génica.

La **Figura 3** muestra los datos tras aplicar los criterios de filtrado descritos en la metodología. En este punto se observa que las muestras correspondientes al grupo Control presentan una distribución no homogénea, con tendencia a agruparse en dos subconjuntos. Lopes-Ramos et al. (2020), en su análisis de muestras provenientes del GTEx, reportaron diferencias de expresión génica influenciadas por el sexo en diversos tejidos. En particular, el tejido mamario mostró el mayor número de genes diferencialmente expresados entre muestras masculinas y femeninas (4181 genes, de los cuales 4009 son autosómicos). Por ello, resultó pertinente eliminar del análisis los datos de muestras masculinas del grupo Control para evitar sesgos comparativos. Como consecuencia, la **Figura 4** muestra una mejor coherencia en el patrón de agrupamiento.

Las **Figuras 4 y 5** reflejan el tejido adyacente como un fenotipo intermedio entre el tejido tumoral y el sano. Sin embargo, esta distribución puede deberse tanto al efecto biológico (tejido tumoral, tejido normal adyacente, tejido proveniente de individuos sin cáncer o control) como a un posible efecto de lote, dado que los tejidos tumorales y los normales adyacentes provienen de una base de datos distinta a la del tejido Control (Goh et al., 2017). Dado el alto grado de colinealidad entre ambos efectos, se optó por utilizar la metodología RUV, que presenta un desempeño robusto incluso bajo estas condiciones (**Figura 6, 7 y 11**), a diferencia de métodos como la ANOVA de dos vías, SVA o ComBat, ya que emplea genes control para la normalización (Risso et al., 2014).

En la **Figura 11** se observa el resultado tras aplicar RUV, mostrando un patrón de distribución similar al reportado por Aran et al. (2017), donde también se aplicó esta metodología. En ambos casos se evidencia la formación de dos subconjuntos de muestras tumorales: uno más relacionado con el tejido adyacente y control, y otro más alejado. En nuestro estudio, dichos subconjuntos corresponden a las cohortes CMTN y CMnTN, siendo esta última la más cercana al tejido adyacente y Control.

Las diferencias respecto al trabajo de Aran et al. (2017) pueden atribuirse a la selección de los genes de control. Dichos autores emplearon un conjunto de genes housekeeping, cuyos niveles de expresión se mantienen relativamente constantes entre tejidos; sin embargo, se han señalado que su uso no siempre puede ser óptimo, dado que la expresión de algunos genes de esta categoría puede variar en función del estado fisiológico o patológico del tejido específico (Gagnon-Bartsch & Speed, 2012; Jo et al., 2019). En este trabajo, se seleccionaron como genes control solo aquellos con variación relativamente estable específicamente en el tejido mamario, independientemente del fenotipo normal o canceroso (**Tabla 8**). El patrón de distribución también se correlaciona con el

trabajo de Andreou et al. (2024), donde se aprecia la cercanía entre tejido Control y tejido adyacente, además de separarse del tejido tumoral.

El patrón de dispersión obtenido concuerda con las diferencias biológicas conocidas entre las cohortes. El tejido adyacente o peritumoral, aunque presenta un fenotipo similar al tejido sano, muestra algunas alteraciones en su perfil de expresión génica debido a la influencia del tumor, tales como cambios en la remodelación de la matriz extracelular y la activación de genes asociados a la transición epitelio-mesénquima (EMT), por lo que suele considerarse un fenotipo intermedio (Aran et al., 2017; Gadaleta et al., 2022).

Respecto a la cohorte tumoral, las **Figuras 13 y 17** muestran que el CMnTN está compuesto principalmente por los subtipos Luminal A y Luminal B, caracterizados por un fenotipo celular más diferenciado y una expresión génica más semejante a la del tejido normal, debido a la presencia de genes asociados a RE, RP y, en algunos casos, HER2. Por el contrario, el CMTN, conformado principalmente por tumores del subtipo similar a basal, presenta un fenotipo menos diferenciado y una muy baja o nula expresión de los receptores mencionados, lo que se asocia con su mayor agresividad y tendencia metastásica (Eroles et al., 2012; Visvader & Stingl, 2014).

Zhou et al. (2019) señalan que el CMTN presenta una mayor proporción de células madre de cáncer de mama, caracterizadas por su bajo grado de diferenciación y alto potencial de autorrenovación, lo que contribuye a la resistencia a fármacos. Asimismo, el estudio del Cancer Genome Atlas Network (2012) evidenció marcadas diferencias en la expresión génica entre el cáncer de mama basal y los demás subtipos, destacando además la similitud transcriptómica del fenotipo basal con otros tipos de cáncer, como el carcinoma seroso de ovario y el carcinoma escamoso de pulmón.

Por último, se observa que algunas muestras de CMnTN se agrupan junto con las de CMTN, lo cual puede explicarse por la presencia de un perfil de expresión de un tipo basal. En la **Figura 17** se aprecia claramente la separación del subtipo basal respecto a los demás subtipos tumorales. Aunque en diferentes estudios los términos "cáncer de mama triple negativo" y "subtipo basal" se utilizan como sinónimos, no son totalmente equivalentes (Zagami & Carey, 2022).

5.2. CMTN vs Control

La inflamación constituye una de las características distintivas del cáncer. Diversos mecanismos, como el daño físico a células vecinas sanas, la muerte celular en el centro necrótico, el reclutamiento de células inmunitarias que liberan citoquinas proinflamatorias y la formación de nuevos vasos sanguíneos inducida por hipoxia, contribuyen a mantener un estado inflamatorio crónico en el microambiente tumoral. Este proceso inflamatorio, a su vez, puede favorecer la progresión tumoral y la metástasis, estableciendo un circuito de retroalimentación entre inflamación y cáncer (Munn, 2017; Hiam-Galvez et al., 2021).

En este contexto, los interferones representan citoquinas clave con funciones reguladoras en los procesos moleculares, celulares y fisiológicos asociados a la respuesta inflamatoria (Lan et al., 2021). Durante el desarrollo tumoral, diversos eventos como la muerte celular en el núcleo necrótico del tumor o el estrés genómico provocan la liberación de moléculas endógenas de peligro (DAMPs, *Damage-Associated Molecular Patterns*). Estas son reconocidas por los receptores de reconocimiento de patrones (PRR, *Pattern Recognition Receptors*) del sistema inmunitario innato, lo que desencadena la producción de interferones tipo I y tipo III, típicamente unas de las primeras citoquinas inflamatorias en ser inducidas (Lazear et al., 2019; Murao et al., 2021).

Posteriormente, estos mismos interferones, junto con otras citoquinas del microambiente proinflamatorio, estimulan la activación de linfocitos y células citotóxicas, promoviendo la

producción de interferón gamma (IFNG), principalmente por células NK, NKT, T $CD4^+$ y T $CD8^+$, lo que a su vez representa una forma de interacción entre la respuesta inmunitaria innata y adaptativa (Jorgovanovic et al., 2020).

En el contexto tumoral, la correspondencia entre la abundancia de transcritos y los niveles proteicos pueden alterarse por diversos mecanismos, tales como las variaciones en el número de copias (CNA, del inglés *copy number alterations*) encontrados en el ADN, cuyo impacto suele ser más pronunciado a nivel transcriptómico que proteómico (Liu et al., 2016). No obstante, Johansson et al. (2019) describieron una correlación positiva entre ambos niveles en el cáncer de mama para proteínas metabólicas solubles y las involucradas en rutas de proliferación y señalización, incluyendo vías vinculadas a los interferones. Asimismo, se ha evidenciado que esta asociación es más estrecha en fenotipos con alta tasa proliferativa, como los subtipos basal y HER2-enriquecido, los cuales representan la casi totalidad de la cohorte CMTN evaluada en esta investigación (**Figura 13 y 17**). Por consiguiente, se asume en este trabajo que la sobreexpresión de un gen constituye un indicador fiable de la sobreexpresión de la proteína correspondiente en el microambiente tumoral.

Los resultados obtenidos en este estudio entre la comparación de muestras tumorales de CMTN frente a tejido Control (**Figura 18 y 19**) evidencian la sobreexpresión de genes pertenecientes a las tres vías de interferón, tanto tipo I, II, III; correlacionado con un evento inflamatorio asociado a cáncer relatado en líneas anteriores.

En particular, dentro de la vía de IFN-I, se observó la sobreexpresión concomitante de *IFNB1* e *IFNAR2*. Dado que IFNB1 es el subtipo con mayor afinidad por la subunidad IFNAR2, y que a mayor expresión del receptor se ha asociado con un incremento en la sensibilidad celular al interferón, este resultado sugiere una posible activación del eje IFNB1/IFNAR2 consistente con

una respuesta antiproliferativa, el cuál necesita mayores concentraciones y potencia de activación que una respuesta antiviral (Schreiber, 2017).

Así mismo, la sobreexpresión de genes como *LCK, PTPRC* y *VAV1*, integrantes de la vías de señalización de células inmunitarias, en especial de linfocitos T debido a la participación en la señalización mediada por TCR, puede estar relacionado con la infiltración de estas células en el microambiente tumoral del CMTN y una posible actividad de estas; sin embargo, el estado de activación o desactivación de la vía es una hipótesis limitada debido a que este proceso está relacionado principalmente con la fosforilación de estas proteínas (Bozso et al., 2020; Fernández-Aguilar et al., 2023; Neurath & Berg, 2024).

La actividad de la quinasa PI3K, perteneciente a la vía de señalización PI3K/Akt/mTOR, relacionada con la progresión y supervivencia celular, es regulada por subunidades como la PI3KR1 y PI3KR2. Debido a que PI3KR1 posee la capacidad de mantener inactiva a PI3K, considerado un supresor tumoral, su expresión es más abundante en tejidos no tumorales que en tejidos tumorales. Por otro lado, PI3KR2, quien puede mantener activa la vía ante estímulos bajos o falta de estímulos, es considerado un oncogén, cuya expresión es mayor en tejidos tumorales y está relacionada con tumores metastásicos como el CMTN, coincidiendo ambos resultados con los encontrados en este estudio (Vallejo-Díaz et al., 2019 y Fox et al., 2020). Una de las estrategias de activación de esta vía es mediada por los adaptadores como lo son IRS1 e IRS2; sin embargo, la expresión de ambas son dependientes de hormonas como estrógeno y progesterona y el CMTN no posee receptores hormonales (Gibson et al., 2007), esto podría ser la causa de su baja expresión en el análisis, y sugiere la activación de PI3K independiente de IRS1/IRS2.

En esta misma línea, el complejo de inicio de traducción eucariota eIF4F puede ser inducido tanto por PI3K/mTOR como por RAS/MAPK, otra vía de señalización relacionada con

la respuesta celular a estímulos externos (Chu et al., 2018). La sobreexpresión encontrada del gen *eIF4E*, codificante de la subunidad del complejo eIF4F, puede sugerir una modulación de la traducción mediada más por la vía de PI3K/mTOR que por RAS/MAPK, debido a que dos genes de esta última vía (*MAP2K6* y *RPS6KA5*) se encuentran subexpresados. Estos resultados se correlacionan con lo reportado por Ramón y Cajal et al. (2014), quienes mediante inmunohistoquímica observan mayor expresión de la proteína eIF4E y su forma fosforilada en cáncer de mama en comparación con tejido normal. Además, se ha reportado que la mayor expresión y fosforilación de eIF4E están relacionadas con una mayor progresión tumoral y peor pronóstico (Chu et al., 2018). Así también, Pu et al. (2018) encontraron que la baja expresión de *RPS6KA5* está asociada con el fenotipo de cáncer de mama triple negativo y con un mayor grado tumoral.

Por otro lado, podemos mencionar que los IFN-I inducen diversos estímulos según las proteínas de señalización activadas y la célula diana. Nuestros resultados sugieren una mayor señalización mediada por STAT1, el cuál tiende a ser más antiinflamatorio, proapoptótico y antiproliferativo, tanto en células cancerígenas como en células inmunitarias; al mismo tiempo se da una inhibición de la señalización mediada por STAT5, quienes son los principales inductores de la traducción de genes relacionados a interferón en linfocitos B y T CD4+, permitiendo su activación, diferenciación y supervivencia (Lazear et al., 2019), además, la vía de señalización puede ser obstaculizada por la baja expresión de *TYK2* impidiendo una respuesta completa a los IFN-I por las células diana (Schreiber, 2017).

Los interferones tipo III comparten múltiples características funcionales con los IFN-I, incluyendo la activación de vías de señalización similares y la inducción de genes con funciones antiproliferativas e inmunomoduladoras. Sin embargo, su acción está más restringida a superficies

epiteliales y presenta menor potencia biológica que los IFN-I (Lazear et al., 2019). Nuestros resultados muestran una marcada activación de la vía, evidenciada por la sobreexpresión simultánea de tres de los cuatro subtipos de IFN-III en tejido tumoral. Se ha descrito que IFNL1 e IFNL2 promueven la producción de citocinas proinflamatorias, favorecen la polarización de linfocitos T hacia un fenotipo Th1 y suprimen respuestas Th2 dependientes de células dendríticas. Asimismo, contribuyen al reclutamiento de células inmunitarias mediante la inducción de quimioquinas como CXCL9 y CXCL10, y a la activación de células efectoras a través de la producción de IFNG. Los IFN-III también modulan la actividad de neutrófilos, células NK, macrófagos y células B (Lazear et al., 2019; Manivasagam & Klein, 2021). En contraste, Lasfar et al. (2019) describen un rol dual, reportando que IFN-III puede promover la migración e invasión tumoral en cáncer de vejiga y en un modelo canino de cáncer de mama, sugiriendo una acción contexto-dependiente.

En la vía del interferón tipo II, la sobreexpresión de *IFNG* y del gen de la subunidad receptora *IFNGR2* concuerda con la elevada infiltración inmunológica característica del CMTN, dado que IFNG es producido por células NK, T $CD4^+$ y T $CD8^+$ (Jorgovanovic et al., 2020; Zagami & Carey, 2022). A diferencia de los IFN-I, cuya vía canónica depende del heterodímero STAT1/STAT2, la señalización del IFNG se basa en el homodímero STAT1, el cual también se encuentra sobreexpresado en las muestras, lo que sugiere una mayor activación de esta vía en el microambiente tumoral.

Esta activación se refleja en la inducción de citoquinas proinflamatorias (*IFNB1, IL1B*); quimiocinas que permiten el reclutamiento de células T y NK (*CXCL9, CXCL10*); moléculas de adhesión celular (*ICAM1*) y procesamiento/presentación de antígenos mediante MHC-I (*IRF1, HLA-B, TAP1*) (Jorgovanovic et al., 2020). La sobreexpresión de *PSMB9* podría indicar la

formación del inmunoproteasoma citoplasmático, favoreciendo la generación de péptidos óptimos para MHC-I en contextos inflamatorios (Geoffroy et al., 2023).

IFI6, principalmente inducido por IFN-I, se encuentra asociado a mayor capacidad metastásica en cáncer de mama a través del remodelamiento de actina, y su sobreexpresión se relaciona con peor pronóstico (Cheriyath et al., 2018).

En el caso de *IFIT2*, relacionado a apoptosis, Koh et al. (2019) reportaron una mayor expresión en tejido normal y tumor primario, con pérdida significativa en tejido metastásico y asociación entre alta expresión y mejor pronóstico en CMTN.

GBP1, fuertemente inducido por IFNG y con funciones antiangiogénicas, se expresa mayormente en células inmunitarias infiltrantes. Su alta expresión, junto con otros genes inducidos por IFNG, se ha asociado a un fenotipo inmunoactivado en CMTN y mejor pronóstico (Hunt et al., 2022), aunque también se ha observado su relación con metástasis cerebral en cáncer de mama RE− (Mustafa et al., 2018).

ISG15, involucrada en el proceso de modificación postraduccional de ISGylación, es inducida por IFN-I y II, así como otros mecanismos de inflamación como el daño a ADN (Villarroya-Beltri et al., 2017). Se ha encontrado sobreexpresado en cáncer de mama y asociado a progresión tumoral, en concordancia con los resultados en el presente estudio (Villarroya-Beltri et al., 2017; Tecalco-Cruz et al., 2020).

OAS1, involucrado en la degradación de ARN bicatenario, también aparece sobreexpresado, resultado que coincide con lo reportado por Zhang & Yu, (2020) quienes además han asociado la mayor expresión de este gen a un peor pronóstico.

Los resultados en el presente estudio coinciden también con Zhu et al. (2023), quienes reportan la sobreexpresión de *CYBB, CXCL10* e *ICAM1* en CMTN, además de correlaciones con T reguladoras y células supresoras mieloides.

En conjunto, esta primera comparación revela un fenotipo fuertemente inflamatorio en CMTN, caracterizado por una activación amplia de genes inducidos por interferones. Aunque muchos de estos genes participan en vigilancia inmunológica y antitumoral, varios están asociados a agresividad, progresión y metástasis, resaltando el rol dual de los interferones en cáncer.

5.3. CMTN vs TA

Como segundo análisis, se comparó CMTN frente a la totalidad del tejido adyacente (TA) disponible en la base de datos, proveniente de pacientes con distintos subtipos tumorales.

Se observó pequeñas diferencias entre los genes diferencialmente expresados entre CMTN frente a Control y TA total. Esto era esperable, considerando que el tejido adyacente, aunque fenotípicamente normal, no es equivalente a tejido sano. Diversos estudios han demostrado que el TA ya refleja señales del microambiente tumoral, mostrando alteraciones genómicas y transcriptómicas inducidas por el tumor (Gadaleta et al., 2022). Asimismo, Casbas-Hernandez et al. (2015) reportan que el subtipo molecular del tumor primario se refleja parcialmente en el transcriptoma del tejido adyacente, lo que explica que no todos los TA son comparables entre sí.

En este estudio, la alta similitud entre ambos contrastes (**Figura 20 y 21**) podría explicarse porque la mayoría de las muestras TA (42/51) provienen de pacientes con cáncer de mama tipo luminal, un subtipo caracterizado por un estado más diferenciado y una menor actividad inflamatoria en comparación con el CMTN. Debido a esta mayor “normalidad” relativa del subtipo luminal, su TA es transcriptómicamente más parecido al tejido sano.

En consecuencia, el CMTN contrasta marcadamente con ambos grupos de referencia, debido probablemente a que presenta un estado menos diferenciado, mayor infiltración inmunológica y elevada actividad de citocinas proinflamatorias, entre ellas los interferones (Visvader & Stingl, 2014; Zhou et al., 2019; Lan et al., 2021).

En este punto, podemos resaltar que *SOCS1* se encuentra sobreexpresado en CMTN, sugiriendo una regulación negativa de la respuesta a los tres tipos de interferón en el microambiente tumoral (Jorgovanovic et al., 2020; Schreiber, 2017). Por su parte, *SOCS3*, inhibidor de IFN-II y relacionado a efectos antiproliferativos, también se encuentra sobreexpresado en CMTN, lo que sugiere dinámicas reguladoras adicionales (Xu et al., 2019).

5.4. CMTN vs TA (análisis pareado)

Para eliminar variabilidad derivada de diferencias interindividuales, se realizó un análisis pareado usando 7 pares de muestras CMTN–TA (**Figura 23 y 24**). Este enfoque permite comparar directamente tumor y tejido peritumoral del mismo paciente.

En este análisis, la sobreexpresión de *PI3KR2* junto con la subexpresión de *PI3KR1*, *IRS1* e *IRS2* sugiere una activación de la vía PI3K independiente de IRS1/IRS2. Las subunidades reguladoras p85α (*PI3KR1*) y p85β (*PI3KR2*) se asocian a las isoformas clase IA de PI3K (PI3Kα, PI3Kβ y PI3Kδ). PI3Kα y PI3Kβ median señalización de supervivencia celular, proliferación, metabolismo y angiogénesis en diversos tipos celulares, incluidas células cancerígenas; mientras que PI3Kδ participa principalmente en señalización inmunitaria, especialmente en linfocitos B y T (Bilanges et al., 2019). En nuestro análisis, PI3Kδ fue incluido pero no presentó diferencias significativas de expresión.

Cabe destacar que tanto los interferones tipo I como los tipo II constituyen estímulos capaces de activar la vía PI3K (Lazear et al., 2019). En estudios de líneas celulares, se ha observado

que la inhibición farmacológica de PI3K reduce la inducción de quimiocinas dependientes de IFNG, como CXCL9 y CXCL10, lo que resalta la interacción funcional entre las vías IFN–JAK–STAT y PI3K–AKT (Jorgovanovic et al., 2020).

Las quimiocinas CXCL9 y CXCL10, inducidas principalmente por IFNG, desempeñan funciones centrales en la quimiotaxis. Son secretadas por monocitos, fibroblastos y también células tumorales en respuesta a estímulos inflamatorios. Actúan sobre monocitos, linfocitos T, células NK, células dendríticas y, en algunos contextos, sobre las propias células cancerígenas, donde pueden promover proliferación y procesos metastásicos. Además, cumplen un papel clave en la polarización hacia la respuesta Th1 (Tokunaga et al., 2018).

Deng et al. (2019) realizaron un análisis similar entre tejido tumoral y tejido adyacente en CMTN para evaluar infiltración inmunológica. En sus resultados, el tumor presentó un mayor enriquecimiento de subtipos de células T, excepto las *naive*, y de células dendríticas, en comparación con el TA. Esto es coherente con los hallazgos en el presente estudio ya que una mayor producción de quimioquinas como CXCL9/CXCL10 suele asociarse con una mayor infiltración de células inmunitarias efectoras.

Un aspecto destacable de los resultados obtenidos es que ningún gen de moléculas de interferón se encontró diferencialmente expresado en el análisis pareado. Esto sugiere niveles de expresión similares entre el tejido tumoral y su TA correspondiente, lo cual refuerza la noción de que el TA está influenciado por el microambiente tumoral, reflejando parte de su programa transcriptómico (Gadaleta et al., 2022).

Una limitación importante de este análisis es el número de réplicas biológicas (n = 7 por grupo), un valor cercano al mínimo de seis réplicas recomendado por Schurch et al. (2016) para obtener estimaciones robustas de dispersión y minimizar falsos positivos en análisis de expresión

diferencial de genes. Incrementar el tamaño muestral aumentaría considerablemente la potencia para detectar verdaderas diferencias (Dawadi et al., 2025).

5.5. CMTN vs CMnTN

Mediante este último contraste, observamos que las muestras de CMTN presentan principalmente una mayor expresión de genes relacionados con IFN-II en comparación con otros subtipos de cáncer de mama. Esto se confirma por el mayor impacto de esta vía en los análisis de enriquecimiento mediante GSEA.

Esta sobreexpresión podría deberse a diversos factores, entre los cuales destacan: 1) la actividad de las hormonas progesterona y estrógeno, mediada por la expresión de sus receptores; Todorović-Raković et al. (2017), en un estudio retrospectivo, encontraron una relación inversa global entre la expresión de ARNm de IFNG y el receptor de progesterona, relación que también se observó para ambos receptores entre los 7 y 14 años de seguimiento. De manera concordante, Chavey et al. (2007) reportaron que los tumores que expresan RE o RP presentan niveles más bajos de diversas citoquinas, incluido IFNG. Asimismo, Walter et al. (2017) demostraron que el receptor de progesterona, incluso en ausencia de su ligando, puede mantener bajos los niveles de expresión de genes estimulados por interferón; en su análisis incluyen genes inducidos por IFN-I y/o IFN-II, como CXCL10; 2) Una mayor infiltración de células inmunitarias en el microambiente tumoral, actuando tanto como productoras como dianas de IFNG (Jorgovanovic et al., 2020); aunque el cáncer de mama suele mostrar menor infiltrado de células inmunitarias comparado con otros tumores sólidos (Zagami & Carey, 2022), el subtipo CMTN presenta una mayor frecuencia de tumores con infiltración inmunitaria elevada (Karn et al., 2020). A diferencia de otros subtipos, los tumores CMTN poseen una mayor proporción de células madre cancerígenas, asociadas con microambientes más inmunoactivos, con mayor infiltración y secreción de citoquinas (Zhou et al.,

2019). Este comportamiento es congruente con los resultados obtenidos, dado que la sobreexpresión de ZAP70 observada en CMTN se correlaciona con una mayor infiltración inmunitaria (Yu et al., 2022).

Comprender la influencia de la expresión de la vía asociada a IFNG en el subtipo CMTN es fundamental, ya que puede intervenir en características propias del tumor, como su fenotipo agresivo y su respuesta a terapias sistémicas. En esta línea, Cai et al. (2023) identificaron que la expresión de IFNG es un predictor de infiltración inmunitaria en muestras CMTN, mostrando correlación positiva con el enriquecimiento de células T CD4+, T CD8+, neutrófilos, células dendríticas y células B. Por otro lado, Heimes et al. (2020) reportaron que la expresión de un conjunto de genes asociados a la vía de IFNG se relaciona con un menor riesgo de metástasis en cánceres de mama altamente proliferativos, tales como CMTN, Luminal B y HER2+. Además, incluso la expresión individual del gen *IFNG* posee valor pronóstico solo en CMTN. Legrier et al. (2016) observaron que la activación temprana de genes inducibles por interferón tras el tratamiento con quimioterapia puede predecir la respuesta terapéutica, permitiendo diferenciar tempranamente tumores quimiosensibles de quimiorresistentes y evitando la prolongación innecesaria de ciclos en tumores CMTN y HER2+/RE−. Por su parte, Ayers et al. (2017) demostraron que una firma de genes relacionados con IFNG (*IDO1, CXCL9/10, HLA-DRA, STAT1, IFNG*) predice la respuesta a pembrolizumab en melanoma, carcinoma escamoso de cabeza y cuello y cáncer gástrico. Además, una firma de 18 genes inflamatorios asociados a células T, incluyendo *CXCL9* y *STAT1*, permite predecir la respuesta a este tratamiento en nueve tipos de cáncer diferentes.

Sin embargo, un trabajo fundamental realizado por Lehmann et al. (2011) evidenció la marcada heterogeneidad existente dentro del subtipo CMTN. Entre los seis subgrupos

transcriptómicamente definidos, el subgrupo “Inmunomodulatorio” se caracteriza por una mayor expresión de genes asociados a procesos inmunológicos, incluida la activación de la vía JAK/STAT. Posteriormente, un refinamiento de estas observaciones, sugiere que este perfil inmunomodulatorio responde principalmente a características extrínsecas del tumor, como una mayor infiltración de células inmunitarias (Lehmann et al., 2016), esto implica que la expresión de la firma de IFN-II, aunque elevada en el CMTN como grupo, puede variar significativamente entre sus subtipos internos. Dicha variabilidad representa un aspecto que requiere mayor investigación, dado que los tumores con una mayor actividad inmunológica, ya sea por infiltración celular o por mayor expresión de citocinas, suelen mostrar características intrínsecas particulares, como menor heterogeneidad clonal, baja carga de variaciones en el número de copias, menor tasa de mutación somática y menor carga de neoantígenos (Bianchini et al., 2022). En concordancia con lo señalado en líneas anteriores, esto también sugiere que la firma de IFN-II podría aportar valor pronóstico para identificar subgrupos específicos dentro del CMTN que podrían beneficiarse más de terapias inmunológicas, mostrar una mayor quimiosensibilidad, presentar mayor infiltración inmunitaria o poseer menor probabilidad de recurrencia metastásica, entre otros.

6. CONCLUSIONES

En conclusión, los resultados obtenidos en el presente estudio evidencian una mayor expresión de genes asociados a las vías de interferón tipo I, II y III en tumores CMTN en comparación con tejidos mamarios no tumorales, incluyendo tejido adyacente de pacientes con cáncer y tejido mamario sano, sugiriendo la presencia de un microambiente tumoral con una señalización inflamatoria mediada por interferones. Asimismo, la comparación entre tumores CMTN y CMnTN mostró diferencias específicas en la activación de estas vías, destacando la vía de interferón tipo II como la más representativa en los CMTN debido a la mayor expresión de

genes asociados y al enriquecimiento significativo de esta ruta biológica. Estos resultados sugieren que la señalización mediada por IFN-II tiene un papel crucial en las diferencias fenotípicas entre ambos subtipos tumorales.

7. RECOMENDACIONES

Los hallazgos de esta investigación pueden constituir un punto de partida para futuros estudios sobre el papel de los interferones en el cáncer de mama triple negativo. Se requieren análisis adicionales que profundicen en la heterogeneidad intrínseca de la expresión de la vía de IFN-II en dentro de este subtipo tumoral, así como en la identificación de las poblaciones celulares específicas responsables de esta mayor activación dentro del microambiente tumoral, con el fin de favorecer su adecuada aplicación y translación clínica.

8. REFERENCIAS BIBLIOGRÁFICAS

Agrawal, A., Balcı, H., Hanspers, K., Coort, S. L., Martens, M., Slenter, D. N., Ehrhart, F., Digles, D., Waagmeester, A., Wassink, I., Abbassi-Daloii, T., Lopes, E. N., Iyer, A., Acosta, J. M., Willighagen, L. G., Nishida, K., Riutta, A., Basaric, H., Evelo, C. T., … Pico, A. R. (2024). WikiPathways 2024: next generation pathway database. *Nucleic Acids Research*, *52*(D1), D679–D689. https://doi.org/10.1093/nar/gkad960

Andreou, M., Jąkalski, M., Duzowska, K., Filipowicz, N., Kostecka, A., Davies, H., Horbacz, M., Ławrynowicz, U., Chojnowska, K., Bruhn-Olszewska, B., Jankau, J., Śrutek, E., Las-Jankowska, M., Bała, D., Hoffman, J., Hartman, J., Pęksa, R., Skokowski, J., Jankowski, M., … Piotrowski, A. (2024). Prelude to malignancy: A gene expression signature in normal mammary gland from breast cancer patients suggests pre-tumorous alterations and is associated with adverse outcomes. *International Journal of Cancer. Journal International Du Cancer*, *155*(9), 1616–1628. https://doi.org/10.1002/ijc.35050

Aran, D., Camarda, R., Odegaard, J., Paik, H., Oskotsky, B., Krings, G., Goga, A., Sirota, M., & Butte, A. J. (2017). Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nature Communications*, *8*(1), 1077. https://doi.org/10.1038/s41467-017-01027-z

Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R., Albright, A., Cheng, J. D., Kang, S. P., Shankaran, V., Piha-Paul, S. A., Yearley, J., Seiwert, T. Y., Ribas, A., & McClanahan, T. K. (2017). IFN-γ–related mRNA profile predicts clinical

response to PD-1 blockade. *The Journal of Clinical Investigation*, *127*(8), 2930–2940. https://doi.org/10.1172/jci91190

Baillie, M., le Cessie, S., Schmidt, C. O., Lusa, L., Huebner, M., & Topic Group "Initial Data Analysis" of the STRATOS Initiative. (2022). Ten simple rules for initial data analysis. *PLoS Computational Biology*, *18*(2), e1009819. https://doi.org/10.1371/journal.pcbi.1009819

Bianchini, G., De Angelis, C., Licata, L., & Gianni, L. (2022). Treatment landscape of triple-negative breast cancer - expanded options, evolving needs. *Nature Reviews. Clinical Oncology*, *19*(2), 91–113. https://doi.org/10.1038/s41571-021-00565-2

Bilanges, B., Posor, Y., & Vanhaesebroeck, B. (2019). PI3K isoforms in cell signalling and vesicle trafficking. *Nature Reviews. Molecular Cell Biology*, *20*(9), 515–534. https://doi.org/10.1038/s41580-019-0129-z

Blighe, K. (2018). *EnhancedVolcano*. Bioconductor.

Bogdanovska-Todorovska, M., Kostadinova-Kunovska, S., Jovanovik, R., Krsteska, B., Kondov, G., Kondov, B., & Petrushevska, G. (2018). Correlation of immunohistochemistry and fluorescence in situ hybridization for HER-2 assessment in breast cancer patients: Single centre experience. *Open Access Macedonian Journal of Medical Sciences*, *6*(4), 593–599. https://doi.org/10.3889/oamjms.2018.124

Boukhaled, G. M., Harding, S., & Brooks, D. G. (2021). Opposing roles of type I interferons in cancer immunity. *Annual Review of Pathology*, *16*(1), 167–198. https://doi.org/10.1146/annurev-pathol-031920-093932

Bozso, S. J., Kang, J. J. H., & Nagendran, J. (2020). The role of competing mechanisms on Lck regulation. *Immunologic Research*, *68*(5), 289–295. https://doi.org/10.1007/s12026-020-09148-2

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *74*(3), 229–263. https://doi.org/10.3322/caac.21834

Cai, J., Zhong, M., Xu, J., Cheng, H., & Xu, S. (2023). Codelivery of triptolide and IFN-γ to boost antitumor immunity for triple-negative breast cancer. *International Immunopharmacology*, *120*(110346), 110346. https://doi.org/10.1016/j.intimp.2023.110346

Cancer Genome Atlas Network. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*(7418), 61–70. https://doi.org/10.1038/nature11412

Casbas-Hernandez, P., Sun, X., Roman-Perez, E., D'Arcy, M., Sandhu, R., Hishida, A., McNaughton, K. K., Yang, X. R., Makowski, L., Sherman, M. E., Figueroa, J. D., & Troester, M. A. (2015). Tumor intrinsic subtype is reflected in cancer-adjacent tissue. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *24*(2), 406–414. https://doi.org/10.1158/1055-9965.epi-14-0934

Chavey, C., Bibeau, F., Gourgou-Bourgade, S., Burlinchon, S., Boissière, F., Laune, D., Roques, S., & Lazennec, G. (2007). Oestrogen receptor negative breast cancers exhibit high cytokine content. *Breast Cancer Research: BCR*, *9*(1), R15. https://doi.org/10.1186/bcr1648

Chen, X., Zhao, W., Huang, Y., Luo, S., Tang, X., & Yi, Q. (2024). Association of GATA3 expression in triple-positive breast cancer with overall survival and immune cell infiltration. *Scientific Reports*, *14*(1), 17795. https://doi.org/10.1038/s41598-024-68788-8

Cheriyath, V., Kaur, J., Davenport, A., Khalel, A., Chowdhury, N., & Gaddipati, L. (2018). G1P3 (IFI6), a mitochondrial localised antiapoptotic protein, promotes metastatic potential of breast cancer cells through mtROS. *British Journal of Cancer*, *119*(1), 52–64. https://doi.org/10.1038/s41416-018-0137-3

Chu, J., Ramon Y Cajal, S., Sonenberg, N., & Pelletier, J. (2018). Eukaryotic initiation factor 4F-sidestepping resistance mechanisms arising from expression heterogeneity. *Current Opinion in Genetics & Development*, *48*, 89–96. https://doi.org/10.1016/j.gde.2017.11.002

Crawley, M. J. (2013). *The R book* (Second edition). Wiley.

Dawadi, P., Pokharel, B., Shrestha, A., Niraula, D., Naeem, A., Miura, S., Roy, M., & Nepal, S. (2025). From bench to bytes: a practical guide to RNA sequencing data analysis. *Frontiers in Genetics*, *16*(1697922), 1697922. https://doi.org/10.3389/fgene.2025.1697922

Deng, L., Lu, D., Bai, Y., Wang, Y., Bu, H., & Zheng, H. (2019). Immune profiles of tumor microenvironment and clinical prognosis among women with triple-negative breast cancer. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *28*(12), 1977–1985. https://doi.org/10.1158/1055-9965.EPI-19-0469

Ensenyat-Mendez, M., Llinàs-Arias, P., Orozco, J. I. J., Íñiguez-Muñoz, S., Salomon, M. P., Sesé, B., DiNome, M. L., & Marzese, D. M. (2021). Current triple-negative breast cancer subtypes: Dissecting the most aggressive form of breast cancer. *Frontiers in Oncology*, *11*, 681476. https://doi.org/10.3389/fonc.2021.681476

Eroles, P., Bosch, A., Pérez-Fidalgo, J. A., & Lluch, A. (2012). Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer Treatment Reviews*, *38*(6), 698–707. https://doi.org/10.1016/j.ctrv.2011.11.005

Fan, F., Mo, H., Zhang, H., Dai, Z., Wang, Z., Qu, C., Liu, F., Zhang, L., Luo, P., Zhang, J., Liu, Z., Cheng, Q., & Ding, F. (2022). HOXA5: A crucial transcriptional factor in cancer and a potential therapeutic target. *Biomedecine & Pharmacotherapie [Biomedicine & Pharmacotherapy]*, *155*(113800), 113800. https://doi.org/10.1016/j.biopha.2022.113800

Fernández-Aguilar, L. M., Vico-Barranco, I., Arbulo-Echevarria, M. M., & Aguado, E. (2023). A story of kinases and adaptors: The role of Lck, ZAP-70 and LAT in switch panel governing T-cell development and activation. *Biology*, *12*(9), 1163. https://doi.org/10.3390/biology12091163

Fox, M., Mott, H. R., & Owen, D. (2020). Class IA PI3K regulatory subunits: p110-independent roles and structures. *Biochemical Society Transactions*, *48*(4), 1397–1417. https://doi.org/10.1042/BST20190845

Gadaleta, E., Thorn, G. J., Ross-Adams, H., Jones, L. J., & Chelala, C. (2022). Field cancerization in breast cancer. *The Journal of Pathology*, *257*(4), 561–574. https://doi.org/10.1002/path.5902

Gagnon-Bartsch, J. A., & Speed, T. P. (2012). Using control genes to correct for unwanted variation in microarray data. *Biostatistics (Oxford, England)*, *13*(3), 539–552. https://doi.org/10.1093/biostatistics/kxr034

Gao, J. J., & Swain, S. M. (2018). Luminal A breast cancer and molecular assays: A review. *The Oncologist*, *23*(5), 556–565. https://doi.org/10.1634/theoncologist.2017-0535

Geoffroy, K., Araripe Saraiva, B., Viens, M., Béland, D., & Bourgeois-Daigneault, M.-C. (2023). Increased expression of the immunoproteasome subunits PSMB8 and PSMB9 by cancer cells correlate with better outcomes for triple-negative breast cancers. *Scientific Reports*, *13*(1), 2129. https://doi.org/10.1038/s41598-023-28940-2

Geoffroy, K., & Bourgeois-Daigneault, M.-C. (2020). The pros and cons of interferons for oncolytic virotherapy. *Cytokine & Growth Factor Reviews*, *56*, 49–58. https://doi.org/10.1016/j.cytogfr.2020.07.002

Gibson, S. L., Ma, Z., & Shaw, L. M. (2007). Divergent roles for IRS-1 and IRS-2 in breast cancer metastasis. *Cell Cycle (Georgetown, Tex.)*, *6*(6), 631–637. https://doi.org/10.4161/cc.6.6.3987

Goh, W. W. B., Wang, W., & Wong, L. (2017). Why batch effects matter in omics data, and how to avoid them. *Trends in Biotechnology*, *35*(6), 498–507. https://doi.org/10.1016/j.tibtech.2017.02.012

Gómez-Rázuri, K., Abad-Licham, M., Astigueta, J., & Moreno, J. (2024). Cáncer de mama en el norte del Perú: subtipos moleculares y HER2 low. *Revista peruana de medicina experimental y salud publica*, 62–68. https://doi.org/10.17843/rpmesp.2024.411.13424

GTEx Consortium. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, *45*(6), 580–585. https://doi.org/10.1038/ng.2653

Gur-Dedeoglu, B., Konu, O., Bozkurt, B., Ergul, G., Seckin, S., & Yulug, I. G. (2009). Identification of endogenous reference genes for qRT-PCR analysis in normal matched breast tumor tissues. *Oncology Research*, *17*(8), 353–365. https://doi.org/10.3727/096504009788428460

Hang, Y., Burns, J., Shealy, B. T., Pauly, R., Ficklin, S. P., & Feltus, F. A. (2022). Identification of condition-specific regulatory mechanisms in normal and cancerous human lung tissue. *BMC Genomics*, *23*(1), 350. https://doi.org/10.1186/s12864-022-08591-9

Heath, A. P., Ferretti, V., Agrawal, S., An, M., Angelakos, J. C., Arya, R., Bajari, R., Baqar, B., Barnowski, J. H. B., Burt, J., Catton, A., Chan, B. F., Chu, F., Cullion, K., Davidsen, T., Do, P.-M., Dompierre, C., Ferguson, M. L., Fitzsimons, M. S., … Grossman, R. L. (2021). The NCI genomic data commons. *Nature Genetics*, *53*(3), 257–262. https://doi.org/10.1038/s41588-021-00791-5

Heimes, A.-S., Härtner, F., Almstedt, K., Krajnak, S., Lebrecht, A., Battista, M. J., Edlund, K., Brenner, W., Hasenburg, A., Sahin, U., Gehrmann, M., Hengstler, J. G., & Schmidt, M. (2020). Prognostic significance of interferon-γ and its signaling pathway in early breast cancer depends on the molecular subtypes. *International Journal of Molecular Sciences*, *21*(19), 7178. https://doi.org/10.3390/ijms21197178

Hiam-Galvez, K. J., Allen, B. M., & Spitzer, M. H. (2021). Systemic immunity in cancer. *Nature Reviews. Cancer*, *21*(6), 345–359. https://doi.org/10.1038/s41568-021-00347-z

Hunt, E. N., Kopacz, J. P., & Vestal, D. J. (2022). Unraveling the role of Guanylate-binding proteins (GBPs) in breast cancer: A comprehensive literature review and new data on prognosis in breast cancer subtypes. *Cancers*, *14*(11), 2794. https://doi.org/10.3390/cancers14112794

Jo, J., Choi, S., Oh, J., Lee, S.-G., Choi, S. Y., Kim, K. K., & Park, C. (2019). Conventionally used reference genes are not outstanding for normalization of gene expression in human cancer

research. *BMC Bioinformatics*, *20*(Suppl 10), 245. https://doi.org/10.1186/s12859-019-2809-2

Johansson, H. J., Socciarelli, F., Vacanti, N. M., Haugen, M. H., Zhu, Y., Siavelis, I., Fernandez-Woodbridge, A., Aure, M. R., Sennblad, B., Vesterlund, M., Branca, R. M., Orre, L. M., Huss, M., Fredlund, E., Beraki, E., Garred, Ø., Boekel, J., Sauer, T., Zhao, W., … Lehtiö, J. (2019). Breast cancer quantitative proteome and proteogenomic landscape. *Nature Communications*, *10*(1), 1600. https://doi.org/10.1038/s41467-019-09018-y

Jorgovanovic, D., Song, M., Wang, L., & Zhang, Y. (2020). Roles of IFN-γ in tumor progression and regression: a review. *Biomarker Research*, *8*(1), 49. https://doi.org/10.1186/s40364-020-00228-x

Karn, T., Denkert, C., Weber, K. E., Holtrich, U., Hanusch, C., Sinn, B. V., Higgs, B. W., Jank, P., Sinn, H. P., Huober, J., Becker, C., Blohmer, J.-U., Marmé, F., Schmitt, W. D., Wu, S., van Mackelenbergh, M., Müller, V., Schem, C., Stickeler, E., … Loibl, S. (2020). Tumor mutational burden and immune infiltration as independent predictors of response to neoadjuvant immune checkpoint inhibition in early TNBC in GeparNuevo. *Annals of Oncology*, *31*(9), 1216–1222. https://doi.org/10.1016/j.annonc.2020.05.015

Kılıç, Y., Çelebiler, A. Ç., & Sakızlı, M. (2014). Selecting housekeeping genes as references for the normalization of quantitative PCR data in breast cancer. *Clinical & Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*, *16*(2), 184–190. https://doi.org/10.1007/s12094-013-1058-5

Koh, S. Y., Moon, J. Y., Unno, T., & Cho, S. K. (2019). Baicalein suppresses stem cell-like characteristics in radio- and chemoresistant MDA-MB-231 human breast cancer cells through up-regulation of IFIT2. *Nutrients*, *11*(3), 624. https://doi.org/10.3390/nu11030624

Lan, T., Chen, L., & Wei, X. (2021). Inflammatory cytokines in cancer: Comprehensive understanding and clinical progress in gene therapy. *Cells (Basel, Switzerland)*, *10*(1), 100. https://doi.org/10.3390/cells10010100

Larsen, T. V., Daugaard, T. F., Gad, H. H., Hartmann, R., & Nielsen, A. L. (2023). PD-L1 and PD-L2 immune checkpoint protein induction by type III interferon in non-small cell lung cancer cells. *Immunobiology*, *228*(3), 152389. https://doi.org/10.1016/j.imbio.2023.152389

Lasfar, A., Zloza, A., Silk, A. W., Lee, L. Y., & Cohen-Solal, K. A. (2019). Interferon lambda: Toward a dual role in cancer. *Journal of Interferon & Cytokine Research: The Official*

Journal of the International Society for Interferon and Cytokine Research, *39*(1), 22–29. https://doi.org/10.1089/jir.2018.0046

Lazear, H. M., Schoggins, J. W., & Diamond, M. S. (2019). Shared and distinct functions of type I and type III interferons. *Immunity*, *50*(4), 907–923. https://doi.org/10.1016/j.immuni.2019.03.025

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K., & Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews. Genetics*, *11*(10), 733–739. https://doi.org/10.1038/nrg2825

Legrier, M.-E., Bièche, I., Gaston, J., Beurdeley, A., Yvonnet, V., Déas, O., Thuleau, A., Château-Joubert, S., Servely, J.-L., Vacher, S., Lassalle, M., Depil, S., Tucker, G. C., Fontaine, J.-J., Poupon, M.-F., Roman-Roman, S., Judde, J.-G., Decaudin, D., Cairo, S., & Marangoni, E. (2016). Activation of IFN/STAT1 signalling predicts response to chemotherapy in oestrogen receptor-negative breast cancer. *British Journal of Cancer*, *114*(2), 177–187. https://doi.org/10.1038/bjc.2015.398

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of Clinical Investigation*, *121*(7), 2750–2767. https://doi.org/10.1172/jci45014

Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., Moses, H. L., Sanders, M. E., & Pietenpol, J. A. (2016). Refinement of triple-negative breast cancer molecular subtypes: Implications for neoadjuvant chemotherapy selection. *PloS One*, *11*(6), e0157368. https://doi.org/10.1371/journal.pone.0157368

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell*, *165*(3), 535–550. https://doi.org/10.1016/j.cell.2016.03.014

Liu, Y., Cai, J., Fahira, A., Zhuang, K., Wang, J., Zhang, Z., Yan, L., Liu, Y., Ouyang, D., & Huang, Z. (2025). Unveiling the impact of interferon genes on the immune microenvironment of triple-negative breast cancer: identification of therapeutic targets. *Frontiers in Bioinformatics*, *5*(1629526), 1629526. https://doi.org/10.3389/fbinf.2025.1629526

Lopes-Ramos, C. M., Chen, C.-Y., Kuijjer, M. L., Paulson, J. N., Sonawane, A. R., Fagny, M., Platig, J., Glass, K., Quackenbush, J., & DeMeo, D. L. (2020). Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell Reports*, *31*(12), 107795. https://doi.org/10.1016/j.celrep.2020.107795

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lyng, M. B., Laenkholm, A.-V., Pallisgaard, N., & Ditzel, H. J. (2008). Identification of genes for normalization of real-time RT-PCR data in breast carcinomas. *BMC Cancer*, *8*(1), 20. https://doi.org/10.1186/1471-2407-8-20

Maltseva, D. V., Khaustova, N. A., Fedotov, N. N., Matveeva, E. O., Lebedev, A. E., Shkurnikov, M. U., Galatenko, V. V., Schumacher, U., & Tonevitsky, A. G. (2013). High-throughput identification of reference genes for research and clinical RT-qPCR analysis of breast cancer samples. *Journal of Clinical Bioinformatics*, *3*(1), 13. https://doi.org/10.1186/2043-9113-3-13

Manivasagam, S., & Klein, R. S. (2021). Type III interferons: Emerging roles in autoimmunity. *Frontiers in Immunology*, *12*, 764062. https://doi.org/10.3389/fimmu.2021.764062

MAQC Consortium, Shi, L., Reid, L. H., Jones, W. D., Shippy, R., Warrington, J. A., Baker, S. C., Collins, P. J., de Longueville, F., Kawasaki, E. S., Lee, K. Y., Luo, Y., Sun, Y. A., Willey, J. C., Setterquist, R. A., Fischer, G. M., Tong, W., Dragan, Y. P., Dix, D. J., … Slikker, W., Jr. (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnology*, *24*(9), 1151–1161. https://doi.org/10.1038/nbt1239

Marín Sánchez, O., Vivas-Ruiz, D., Neira, M., Sandoval, G. A., Marín- Machuca, O., Rodriguez-Landauro, A. J., & Chacón, R. D. (2019). Rol de los interferones tipo I y tipo III: Una revisión de conceptos. *Ágora Revista Científica*, *6*(2), e6. https://doi.org/10.21679/arc.v6i2.133

Medrano, R. F. V., Hunger, A., Mendonça, S. A., Barbuto, J. A. M., & Strauss, B. E. (2017). Immunomodulatory and antitumor effects of type I interferons and their application in cancer therapy. *Oncotarget*, *8*(41), 71249–71284. https://doi.org/10.18632/oncotarget.19531

Munn, L. L. (2017). Cancer and inflammation. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, *9*(2). https://doi.org/10.1002/wsbm.1370

Murao, A., Aziz, M., Wang, H., Brenner, M., & Wang, P. (2021). Release mechanisms of major DAMPs. *Apoptosis: An International Journal on Programmed Cell Death*, *26*(3–4), 152–162. https://doi.org/10.1007/s10495-021-01663-3

Mustafa, D. A. M., Pedrosa, R. M. S. M., Smid, M., van der Weiden, M., de Weerd, V., Nigg, A. L., Berrevoets, C., Zeneyedpour, L., Priego, N., Valiente, M., Luider, T. M., Debets, R.,

Martens, J. W. M., Foekens, J. A., Sieuwerts, A. M., & Kros, J. M. (2018). T lymphocytes facilitate brain metastasis of breast cancer by inducing Guanylate-Binding Protein 1 expression. *Acta Neuropathologica*, *135*(4), 581–599. https://doi.org/10.1007/s00401-018-1806-2

Neurath, M. F., & Berg, L. J. (2024). VAV1 as a putative therapeutic target in autoimmune and chronic inflammatory diseases. *Trends in Immunology*, *45*(8), 580–596. https://doi.org/10.1016/j.it.2024.06.004

Nguyen, T.-M., Shafi, A., Nguyen, T., & Draghici, S. (2019). Identifying significantly impacted pathways: a comprehensive review and assessment. *Genome Biology*, *20*(1), 203. https://doi.org/10.1186/s13059-019-1790-4

Ogony, J., Choi, H. J., Lui, A., Cristofanilli, M., & Lewis-Wambi, J. (2016). Interferon-induced transmembrane protein 1 (IFITM1) overexpression enhances the aggressive phenotype of SUM149 inflammatory breast cancer cells in a signal transducer and activator of transcription 2 (STAT2)-dependent manner. *Breast Cancer Research: BCR*, *18*(1), 25. https://doi.org/10.1186/s13058-016-0683-7

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.

Oza, V. H., Whitlock, J. H., Wilk, E. J., Uno-Antonison, A., Wilk, B., Gajapathy, M., Howton, T. C., Trull, A., Ianov, L., Worthey, E. A., & Lasseigne, B. N. (2023). Ten simple rules for using public biological data for your research. *PLoS Computational Biology*, *19*(1), e1010749. https://doi.org/10.1371/journal.pcbi.1010749

Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O., & Botstein, D. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(16), 9212–9217. https://doi.org/10.1073/pnas.96.16.9212

Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A. L., Brown, P. O., & Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, *406*(6797), 747–752. https://doi.org/10.1038/35021093

Pingwara, R., Witt-Jurkowska, K., Ulewicz, K., Mucha, J., Tonecka, K., Pilch, Z., Taciak, B., Zabielska-Koczywas, K., Mori, M., Berardozzi, S., Botta, B., Rygiel, T. P., & Krol, M. (2017). Interferon lambda 2 promotes mammary tumor metastasis via angiogenesis extension and stimulation of cancer cell migration. *Journal of Physiology and Pharmacology: An Official Journal of the Polish Physiological Society*, *68*(4), 573–583.

Platanias, L. C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nature Reviews. Immunology*, *5*(5), 375–386. https://doi.org/10.1038/nri1604

Provance, O. K., & Lewis-Wambi, J. (2019). Deciphering the role of interferon alpha signaling and microenvironment crosstalk in inflammatory breast cancer. *Breast Cancer Research: BCR*, *21*(1), 59. https://doi.org/10.1186/s13058-019-1140-1

Pu, X., Storr, S. J., Ahmad, N. S., Rakha, E. A., Green, A. R., Ellis, I. O., & Martin, S. G. (2018). High nuclear MSK1 is associated with longer survival in breast cancer patients. *Journal of Cancer Research and Clinical Oncology*, *144*(3), 509–517. https://doi.org/10.1007/s00432-018-2579-7

Ramon Y Cajal, S., De Mattos-Arruda, L., Sonenberg, N., Cortes, J., & Peg, V. (2014). The intra-tumor heterogeneity of cell signaling factors in breast cancer: p4E-BP1 and peIF4E are diffusely expressed and are real potential targets. *Clinical & Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*, *16*(11), 937–941. https://doi.org/10.1007/s12094-014-1203-9

Risso, D., Ngai, J., Speed, T. P., & Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, *32*(9), 896–902. https://doi.org/10.1038/nbt.2931

Sceneay, J., Goreczny, G. J., Wilson, K., Morrow, S., DeCristo, M. J., Ubellacker, J. M., Qin, Y., Laszewski, T., Stover, D. G., Barrera, V., Hutchinson, J. N., Freedman, R. A., Mittendorf, E. A., & McAllister, S. S. (2019). Interferon signaling is diminished with age and is associated with immune checkpoint blockade efficacy in triple-negative breast cancer. *Cancer Discovery*, *9*(9), 1208–1227. https://doi.org/10.1158/2159-8290.CD-18-1454

Schreiber, G. (2017). The molecular basis for differential type I interferon signaling. *The Journal of Biological Chemistry*, *292*(18), 7285–7294. https://doi.org/10.1074/jbc.R116.774562

Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., Blaxter, M., & Barton, G. J. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA (New York, N.Y.)*, *22*(6), 839–851. https://doi.org/10.1261/rna.053959.115

Shetty, S. R., Kar, T., & Das, A. (2025). Epidermal growth factor receptor mutations in breast Cancer: Therapeutic challenges and way forward. *Bioorganic Chemistry*, *154*(108037), 108037. https://doi.org/10.1016/j.bioorg.2024.108037

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Tan, T. W., Tong, J. C., Khan, A. M., de Silva, M., Lim, K. S., & Ranganathan, S. (2010). Advancing standards for bioinformatics activities: persistence, reproducibility, disambiguation and Minimum Information About a Bioinformatics investigation (MIABi). *BMC Genomics*, *11 Suppl 4*, S27. https://doi.org/10.1186/1471-2164-11-S4-S27

Tecalco-Cruz, A. C., Ramírez-Jarquín, J. O., & Cruz-Ramos, E. (2020). Regulation and action of interferon-stimulated gene 15 in breast cancer cells. *Human Cell*, *33*(4), 954–962. https://doi.org/10.1007/s13577-020-00414-x

Tilli, T. M., Castro, C. da S., Tuszynski, J. A., & Carels, N. (2016). A strategy to identify housekeeping genes suitable for analysis in breast cancer diseases. *BMC Genomics*, *17*(1), 639. https://doi.org/10.1186/s12864-016-2946-1

Todorović-Raković, N., Radulovic, M., Vujasinović, T., Milovanović, J., & Nikolić-Vukosavljević, D. (2017). The prognostic time dependence of intra-tumoural IFNγ mRNA and protein in patients with breast cancer followed for 14 years after surgery and radiotherapy, without subsequent systemic therapy. *European Cytokine Network*, *28*(4), 151–156. https://doi.org/10.1684/ecn.2018.0402

Tokunaga, R., Zhang, W., Naseem, M., Puccini, A., Berger, M. D., Soni, S., McSkane, M., Baba, H., & Lenz, H.-J. (2018). CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. *Cancer Treatment Reviews*, *63*, 40–47. https://doi.org/10.1016/j.ctrv.2017.11.007

Vallejo-Díaz, J., Chagoyen, M., Olazabal-Morán, M., González-García, A., & Carrera, A. C. (2019). The opposing roles of PIK3R1/p85α and PIK3R2/p85β in cancer. *Trends in Cancer*, *5*(4), 233–244. https://doi.org/10.1016/j.trecan.2019.02.009

Villarroya-Beltri, C., Guerra, S., & Sánchez-Madrid, F. (2017). ISGylation - a key to lock the cell gates for preventing the spread of threats. *Journal of Cell Science*, *130*(18), 2961–2969. https://doi.org/10.1242/jcs.205468

Visvader, J. E., & Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes & Development*, *28*(11), 1143–1158. https://doi.org/10.1101/gad.242511.114

Walter, K. R., Goodman, M. L., Singhal, H., Hall, J. A., Li, T., Holloran, S. M., Trinca, G. M., Gibson, K. A., Jin, V. X., Greene, G. L., & Hagan, C. R. (2017). Interferon-stimulated genes are transcriptionally repressed by PR in breast cancer. *Molecular Cancer Research: MCR*, *15*(10), 1331–1340. https://doi.org/10.1158/1541-7786.MCR-17-0180

Weng, L., Xie, Q., & Ma, R. (2025). IFN-γ: a beneficial immune marker in triple-negative breast cancer. *MEDS Clinical Medicine*, *6*(1), 127–138. https://doi.org/10.23977/medsc.2025.060118

Woo, S.-H., & Jung, B. C. (2023). Big data analytics in RNA-sequencing. *Korean Journal of Clinical Laboratory Science*, *55*(4), 235–243. https://doi.org/10.15324/kjcls.2023.55.4.235

Xu, J.-Z., Shao, C.-C., Wang, X.-J., Zhao, X., Chen, J.-Q., Ouyang, Y.-X., Feng, J., Zhang, F., Huang, W.-H., Ying, Q., Chen, C.-F., Wei, X.-L., Dong, H.-Y., Zhang, G.-J., & Chen, M. (2019). circTADA2As suppress breast cancer progression and metastasis via targeting miR-203a-3p/SOCS3 axis. *Cell Death & Disease*, *10*(3), 175. https://doi.org/10.1038/s41419-019-1382-y

Yaghoobi, H., Azizi, H., Oskooei, V. K., Taheri, M., & Ghafouri-Fard, S. (2018). Assessment of expression of interferon γ (IFN-G) gene and its antisense (IFNG-AS1) in breast cancer. *World Journal of Surgical Oncology*, *16*(1), 211. https://doi.org/10.1186/s12957-018-1508-1

Yu, J., Gao, W., Gao, S., Wen, S., Zhao, Y., Shang, Z., Wang, Y., & Niu, Y. (2022). Identification of LAT/ZAP70 characterized immune subtypes of prostate cancer. *World Journal of Urology*, *40*(11), 2817–2824. https://doi.org/10.1007/s00345-022-04170-5

Zagami, P., & Carey, L. A. (2022). Triple negative breast cancer: Pitfalls and progress. *NPJ Breast Cancer*, *8*(1), 95. https://doi.org/10.1038/s41523-022-00468-0

Zavala, V. A., Casavilca-Zambrano, S., Navarro-Vásquez, J., Tamayo, L. I., Castañeda, C. A., Valencia, G., Morante, Z., Calderón, M., Abugattas, J. E., Gómez, H. L., Fuentes, H. A., Liendo-Picoaga, R., Cotrina, J. M., Neciosup, S. P., Roque, K., Vásquez, J., Mas, L., Gálvez-Nino, M., Fejerman, L., & Vidaurre, T. (2023). Breast cancer subtype and clinical characteristics in women from Peru. *Frontiers in Oncology*, *13*, 938042. https://doi.org/10.3389/fonc.2023.938042

Zhang, Y., & Yu, C. (2020). Prognostic characterization of OAS1/OAS2/OAS3/OASL in breast cancer. *BMC Cancer*, *20*(1), 575. https://doi.org/10.1186/s12885-020-07034-6

Zhou, J., Chen, Q., Zou, Y., Chen, H., Qi, L., & Chen, Y. (2019). Stem cells and cellular origins of breast cancer: Updates in the rationale, controversies, and therapeutic implications. *Frontiers in Oncology*, *9*, 820. https://doi.org/10.3389/fonc.2019.00820

Zhou, L., Zhang, Y., Wang, Y., Zhang, M., Sun, W., Dai, T., Wang, A., Wu, X., Zhang, S., Wang, S., & Zhou, F. (2020). A dual role of type I interferons in antitumor immunity. *Advanced Biosystems*, *4*(11), e1900237. https://doi.org/10.1002/adbi.201900237

Zhu, A., Ibrahim, J. G., & Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics (Oxford, England)*, *35*(12), 2084–2092. https://doi.org/10.1093/bioinformatics/bty895

Zhu, J., Min, N., Gong, W., Chen, Y., & Li, X. (2023). Identification of hub genes and biological mechanisms associated with non-alcoholic fatty liver disease and triple-negative breast cancer. *Life (Basel, Switzerland)*, *13*(4), 998. https://doi.org/10.3390/life13040998

Bibliografía