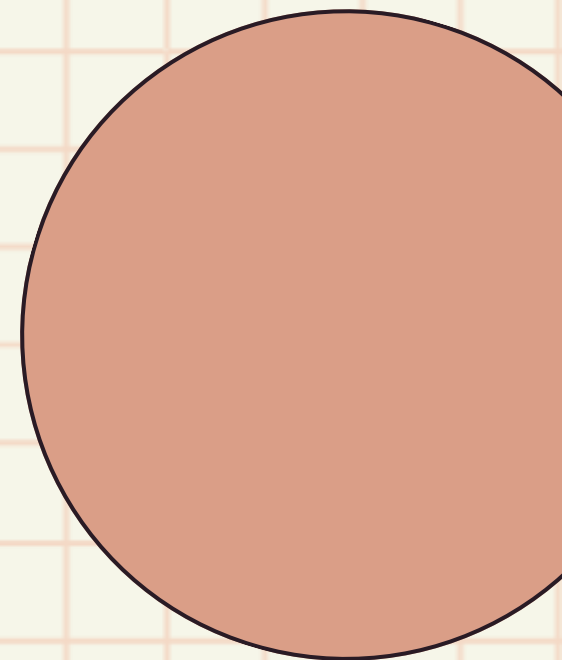


# PROYECTO FINAL

ANÁLISIS DE DATOS TCGA-BRCA

David Zaid Flores Jimenez  
Estrella Segobia Angel  
Daniela Villanueva Curiel



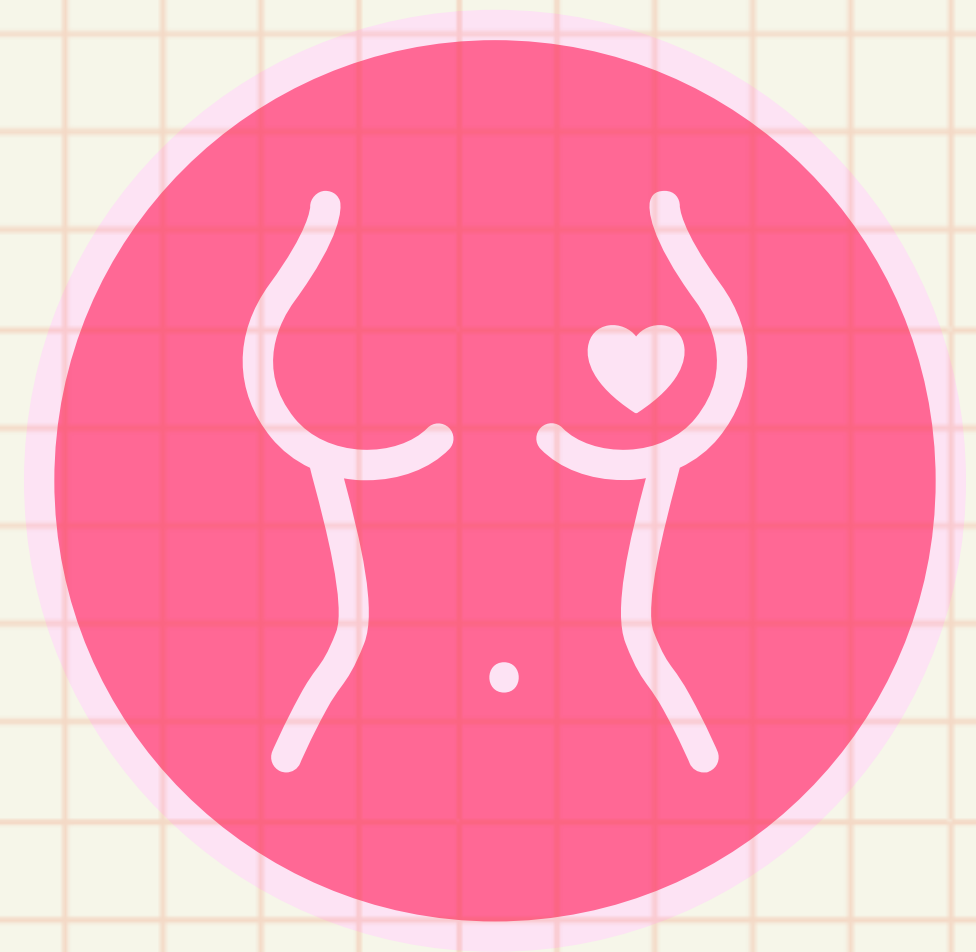
## INTRODUCCIÓN



El cáncer de mama es una de las principales causas de muerte por cáncer en mujeres a nivel mundial. Su estudio molecular ha permitido identificar distintos subtipos, cada uno con características clínicas y biológicas propias que afectan el pronóstico y la respuesta al tratamiento. Comprender los perfiles genéticos y de expresión génica es clave para avanzar hacia una medicina personalizada y mejorar el diagnóstico, pronóstico y tratamiento de esta enfermedad.

## OBJETIVO

Identificar genes diferencialmente expresados (DEGs), ontología de genes, redes de miRNA y proteínas con genes relevantes.



# SUBTIPOS MOLECULARES

## LUMINAL A

Tumores con receptores hormonales positivos (ER+), **baja proliferación y mejor pronóstico**. Suelen responder muy bien a **terapias hormonales** como el tamoxifeno.

## LUMINAL B

También ER+, pero con **+ proliferación celular** y a veces HER2+. Tienen un **pronóstico menos favorable** y requieren **hormonoterapia** combinada con **quimioterapia**.

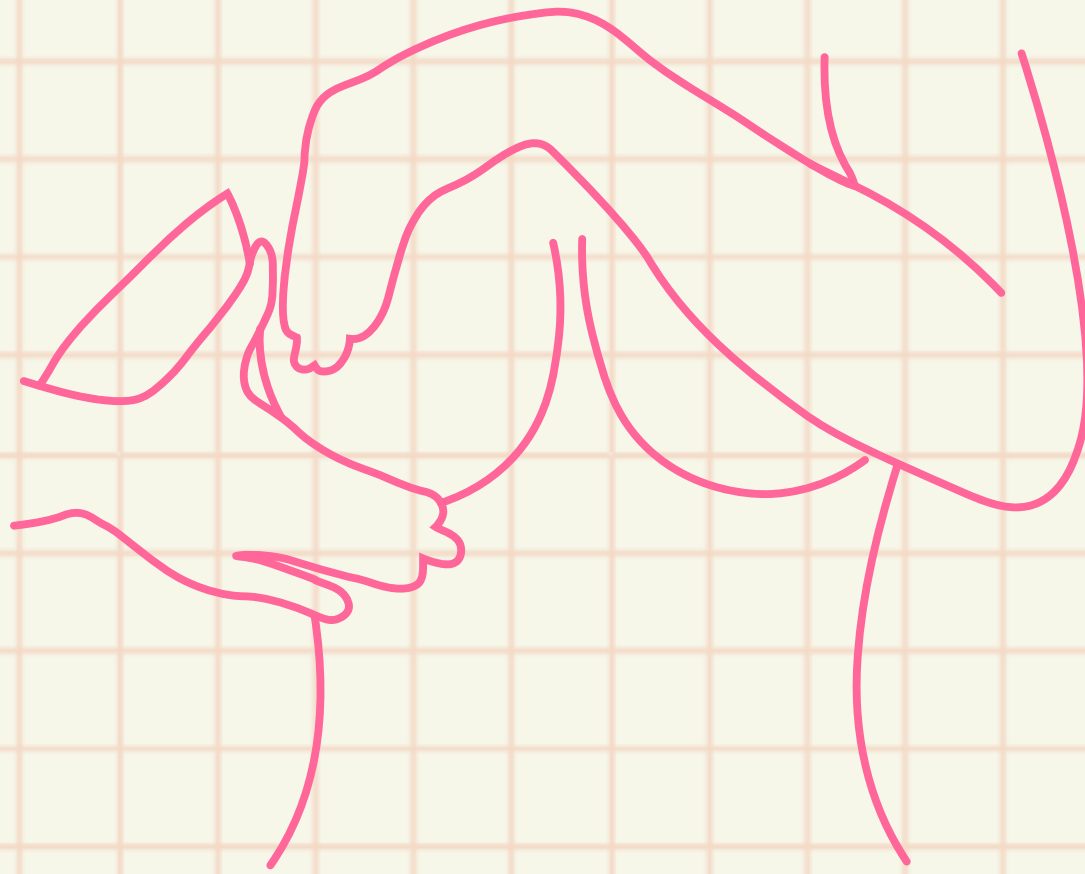
## HER2-ENRIQUECIDO

Alta expresión del **oncogén ERBB2 (HER2)**, generalmente **sin receptores hormonales**. Aunque **+ agresivo**, responde bien a **terapias dirigidas**.

# SUBTIPOS MOLECULARES

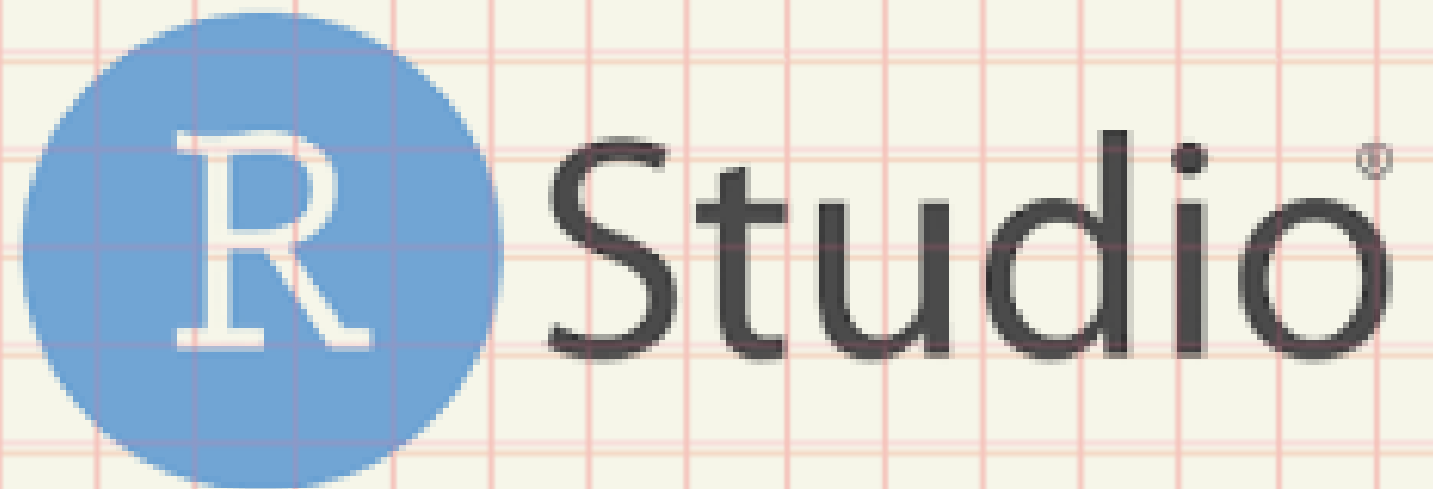
## BASAL

Subtipo **triple negativo** (ER-, PR-, HER2-), con **alta proliferación** y **mal pronóstico**. No hay terapias dirigidas disponibles, por lo que se trata con **quimioterapia**.



## NORMAL

Se parece a **tejido mamario sano**. Su relevancia clínica es incierta y a veces se considera un **artefacto técnico**.



## LIBRERIAS EN RSTUDIO Y BASES DE DATOS QUE SI USAMOS

-

**TCGAbiolinks**

-

**TCGA (The Cancer  
Genome Atlas)**

-

**igraph**

-

**pheatmap**

-

**STRINGdb**

## TEMAS VISTU DURANTE CLASE QUE SI USAMOS

-

**Gene ontology**

-

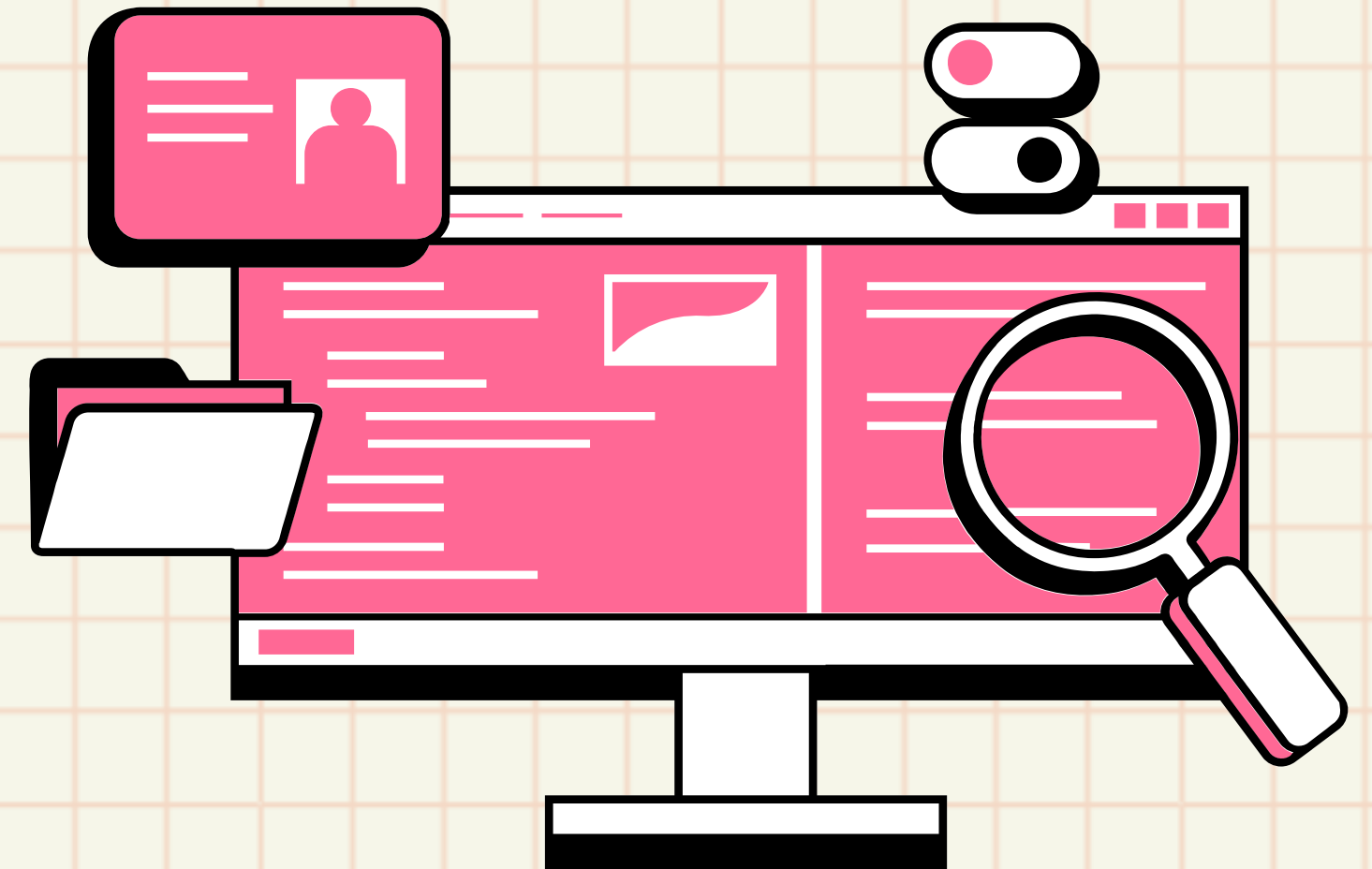
**Red de proteínas/miRNA**

-

**Volcano plot**

-

**HeatMap**



**MARKDOWN**



# RESULTADOS PRINCIPALES

**Identificación de genes diferencialmente expresados entre tumor y tejido normal.**

**Destacan genes como FN1 y COL1A1, asociados a invasión y remodelación de la matriz extracelular.**

**Análisis GO revela funciones alteradas en adhesión, migración y señalización celular.**

**Construcción de redes de interacción miRNA-gen y proteína-proteína para visualizar mecanismos clave en la progresión tumoral**



# Exploring Breast Cancer-Associated Genes: A Comprehensive Analysis and Competitive Endogenous RNA Network Construction

CharuMeena H<sup>1</sup>, Sagaya Jansi R<sup>1\*</sup>, Aishwarya S<sup>1</sup>, Balamurugan Shanmugaraj<sup>2</sup>,  
Praveen Kumar Panthagani<sup>3</sup>

*1. Department of Bioinformatics, Stella Maris College, Chennai, Tamil Nadu, India.*

*2. Department of Biotechnology, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.*



*3. Molecular Biology Department, Ampath Lab, Hyderabad, Telangana.*

OXFORD

<https://doi.org/10.1093/bio/bbaeb>

Case Study

## Comprehensive bioinformatics and machine learning analyses for breast cancer staging using TCGA dataset

Saurav Chandra Das<sup>1,2</sup>, Wahia Tasnim<sup>3</sup>, Humayan Kabir Rana<sup>3</sup>, Uzzal Kumar Acharjee <sup>1,\*</sup>, Md. Manowarul Islam <sup>1</sup>,  
Rabea Khatun<sup>3</sup>

<sup>1</sup>Department of Computer Science and Engineering, Jagannath University, Dhaka-1100, Bangladesh

<sup>2</sup>Department of Internet of Things and Robotics Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University, Bangladesh, Kaliakair, Gazipur-1750, Bangladesh

<sup>3</sup>Department of Computer Science and Engineering, Green University of Bangladesh, Narayanganj-1461, Dhaka, Bangladesh

\*Corresponding author. E-mail: [uzzal@cse.jnu.ac.bd](mailto:uzzal@cse.jnu.ac.bd)

# COMPARACIÓN CON OTROS ARTÍCULOS

## A GRANDES RASGOS...

- Diversas investigaciones recientes han demostrado el valor del análisis multiómico aplicado a datos de cáncer de mama del consorcio TCGA, permitiendo descubrir genes centrales (hub genes), rutas funcionales enriquecidas y subtipos biológicamente distintos que no pueden ser detectados mediante análisis uniómico (**Hosseinpour et al., 2024; Nguyen et al., 2020**). Herramientas desarrolladas en R, como TCGAbiolinks, limma, DESeq2, clusterProfiler y WGCNA, ofrecen un marco robusto y reproducible para este tipo de análisis integrativo.

# Bibliografía

**Hosseinpour, Z., Tavirani, M. R., & Akbari, M. E. (2024). Integrated Analysis of Multi-Omic Data Reveals Regulatory Mechanisms and Network Characteristics in Breast Cancer. Medical Journal of the Islamic Republic of Iran, 38, 63.**

**Nguyen, Q. H., Nguyen, H., Nguyen, T., & Le, D. H. (2020). Multi-omics analysis detects novel prognostic subgroups of breast cancer. Frontiers in genetics, 11, 574661.**

**Das, S. C., Tasnim, W., Rana, H. K., Acharjee, U. K., Islam, M. M., & Khatun, R. (2025). Comprehensive bioinformatics and machine learning analyses for breast cancer staging using TCGA dataset. Briefings in Bioinformatics, 26(1), bbae628.**

**Shanmugaraj, B., & Panthagani, P. K. (2025). Exploring Breast Cancer-Associated Genes: A Comprehensive Analysis and Competitive Endogenous RNA Network Construction. Archives of Razi Institute.**

