



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

DETERMINACIÓN DE FIRMAS TRANSCRIPCIONALES DE MACRÓFAGOS  
ASOCIADOS A TUMOR MEDIANTE ANÁLISIS BIOINFORMÁTICOS EN BASES DE  
DATOS EN LÍNEA

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
ERIC LÓPEZ HUERTA

TUTOR PRINCIPAL  
DR. EZEQUIEL MOISÉS FUENTES PANANÁ  
Unidad de Investigación en Virología y Cáncer.  
Hospital Infantil de México, Federico Gómez.

MIEMBROS DEL COMITÉ TUTOR  
DRA. LAURA CECILIA BONIFAZ ALFONZO  
Unidad de Investigación Médica en Inmunoquímica. Unidad Médica de Alta  
Especialidad (UMAE) Hospital de Especialidades, del Centro Médico Nacional (CMN)  
Siglo XXI.

DR. ENRIQUE HERNÁNDEZ LEMUS  
Departamento de Genómica Computacional. Instituto Nacional de Medicina Genómica  
(INMEGEN).

Ciudad de México. Marzo, 2022



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## **AGRADECIMIENTOS**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca de manutención recibida durante el periodo febrero 2019 a febrero 2021 (CVU: 967745).

Este proyecto fue financiado por el Consejo Nacional de Ciencia y Tecnología (CONACYT), Ciencia de Frontera (Proyecto: 40757) y por el Fondo de Apoyo a la investigación del Hospital infantil de México Federico Gómez (Proyecto: HIM-2014-053 y HIM 2018-076).

Para la realización de este trabajo se contó con el apoyo de Luis Aguilar, Alejandro De León, Carlos Flores y Jair García del Laboratorio Nacional de Visualización Científica Avanzada, UNAM.

A mi tutor principal

Dr. Ezequiel Moisés Fuentes Pananá

A los miembros del Comité Tutor

Dra. Laura Cecilia Bonifaz Alfonzo

Dr. Enrique Hernández Lemus

## **AGRADECIMIENTOS PERSONALES**

A mi tutor principal

Dr. Ezequiel Moisés Fuentes Pananá

Por su excelente guía y apoyo, así como por sus constantes cuestionamientos y siempre pertinentes preguntas. Por procurar el bienestar académico y personal de sus alumnos.

A los miembros del Comité Tutor

Dra. Laura Cecilia Bonifaz Alfonso

Dr. Enrique Hernández Lemus

Por su tiempo y disposición para asesorarme y sus valiosas contribuciones a este trabajo.

A los miembros del jurado

Dra. Marina Macías Silva

Dra. Elizabeth Ortiz Sánchez

Dra. Rosa María Gutiérrez Ríos

Dra. Silvia Jiménez Morales

Dr. Ernesto Soto Reyes Solís

Por sus observaciones que enriquecieron enormemente este trabajo.

## **AGRADECIMIENTOS PERSONALES**

A mis padres, María de Lourdes Huerta López y José López Rosas, por su incondicional apoyo e interés en mis decisiones. Por haber hecho todo dentro de sus posibilidades para proporcionarme las herramientas necesarias para alcanzar mis objetivos sin importar las adversidades. Por darme más oportunidades de las que alguna vez tuvieron.

A mis compañeros de laboratorio, Ceci, Yessi, Angélica, David, Moy, Axel, y Raúl con quienes compartí buenos momentos que siempre recordaré y por haber creado un ambiente de trabajo único, ameno y colaborativo.

A la UNAM por su excelente nivel académico, el cuál soy afortunado de haber recibido desde la licenciatura. “Por mi raza hablará el espíritu”

# ÍNDICE

ÍNDICE DE TABLAS.....	I
ÍNDICE DE FIGURAS .....	II
GLOSARIO DE TÉRMINOS Y ABREVIATURAS .....	IV
RESUMEN.....	1
ABSTRACT .....	3
1. INTRODUCCIÓN.....	5
1.1 Cáncer de mama .....	5
1.2 Microambiente tumoral.....	6
1.3 Funciones de macrófagos asociados a tumor (TAM) en cáncer .....	9
1.4 Relación de los TAM con un modelo de macrófagos M1/M2 .....	13
1.5 Problemas con la clasificación M1/M2 .....	15
1.6 Las nuevas perspectivas sobre la heterogeneidad de los TAM .....	16
2. PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACIÓN .....	17
3. PREGUNTA DE INVESTIGACIÓN.....	17
4. HIPÓTESIS .....	17
5. OBJETIVO GENERAL.....	18
5.1 Objetivos particulares.....	18
6. METODOLOGÍA.....	18
6.1 Adquisición y características de transcriptomas scRNA-seq provenientes de pacientes con cáncer de mama. ....	19
6.2 Procesamiento estándar de datos scRNA-seq .....	20
6.3 Anotación celular .....	22
6.4 <i>Weighted gene correlation network analysis</i> .....	22
6.5 Análisis de representación de vías celulares con GO y KEGG .....	23

6.6 Enriquecimiento de firmas transcripcionales.....	23
6.6.1 Enriquecimiento de la firma de las DC en células mieloides .....	24
6.6.2 Enriquecimiento de módulos en TAM.....	24
6.6.3 Enriquecimiento de las firmas en pacientes del TCGA.....	25
6.7 <i>Gene Set Enrichment Analysis</i> .....	26
6.8 Análisis de supervivencia.....	27
7. RESULTADOS .....	27
7.1 Identificación del infiltrado leucocitario en cáncer de mama .....	27
7.2 Módulos de genes presentes en TAM en cáncer de mama .....	33
7.3 Diferencias funcionales entre las poblaciones de TAM.....	38
7.4 Relación del modelo M1/M2 con las poblaciones de TAM encontradas .....	40
7.5 Expresión de los módulos en otras poblaciones celulares.....	43
7.6 Relevancia clínica de las firmas turquesa y anti-turquesa .....	47
8. DISCUSIÓN.....	48
8.1 Originalidad del trabajo realizado.....	48
8.2 Retos para identificar TAM en el TME de cáncer de mama.....	48
8.3 Funciones asociadas a cada firma de TAM. ....	49
8.4 Distribución de la expresión de los marcadores M1/M2 en las diferentes poblaciones de TAM. ....	53
8.5 Posible origen ontológico de las poblaciones de TAM encontradas .....	55
8.6 Limitaciones de este estudio.....	56
8.7 Implicaciones clínicas y terapéuticas. ....	57
9. CONCLUSIONES.....	59
10. PERSPECTIVAS .....	59
11. REFERENCIAS .....	60

## ÍNDICE DE TABLAS

Tabla 1. Recopilación de marcadores M1 inducidos por IFN- $\gamma$ , LPS y marcadores M2 inducidos por IL-4, IL-13 o IL-10.....	14
Tabla 2. Características de las muestras y de las publicaciones de las que se obtuvieron los datos scRNA-seq. ....	20
Tabla 3. Propuesta de firmas transcripcionales para identificar módulos y firmas funcionales de TAM presentes en cáncer de mama.....	44



## ÍNDICE DE FIGURAS

Figura 1.	Esquema de los diferentes componentes del microambiente tumoral. ....	7
Figura 2.	Esquema de las funciones de TAM en cáncer.....	12
Figura 3.	Resumen del flujo de trabajo y herramientas aplicadas durante el análisis bioinformático. ....	19
Figura 4.	Células presentes en el TME desplegadas en gráficos UMAP.....	29
Figura 5.	Mapa de calor de correlaciones obtenidas entre el transcriptoma de las células y el microarreglo de referencia mediante SingleR.....	30
Figura 6.	Expresión de marcadores canónicos de macrófagos, monocitos y DC en diversos tipos celulares. ....	31
Figura 7.	Mapa de calor del enriquecimiento de la firma DC en células mieloides. ..	32
Figura 8.	Módulos encontrados por WGCNA.....	34
Figura 9.	Análisis de procesos celulares representados por los genes pertenecientes a cada módulo. ....	35
Figura 10.	Mapas de calor del enriquecimiento de las 4 firmas encontradas con WGCNA en TAM. ....	36
Figura 11.	Patrón de expresión de las firmas turquesa, marrón y anti-turquesa.....	38
Figura 12.	Vías celulares enriquecidas según GSEA. ....	39
Figura 13.	GSEA de firma M1 y firma M2. ....	41
Figura 14.	Análisis de expresión diferencial entre las poblaciones TAM turquesa/marrón y TAM anti-turquesa. ....	42
Figura 15.	Expresión de marcadores M1/M2 en diferentes tipos celulares provenientes de tejido tumoral.....	43
Figura 16.	Distribución del ES de las firmas funcionales y de cada módulo en diferentes tipos celulares.....	45
Figura 17.	Distribución del ES de las firmas funcionales y de cada módulo de TAM en células tumorales y macrófagos. ....	46

Figura 18. Curvas de supervivencia de pacientes con cáncer de mama del TCGA enriquecidos con algún subtipo de TAM..... 47

Figura 19. Esquema de los programas transcripcionales encontrados en los TAM.... 53

Figura 20. Limitaciones de las plataformas scRNA-seq..... 57

## GLOSARIO DE TÉRMINOS Y ABREVIATURAS

- **CAF** (cancer-associated fibroblast): Fibroblastos asociados a cáncer.
- **CD**: *cluster* de designación
- **CCL**: quimiocinas con motivo c-c.
- **CMP** (*common myeloid progenitor*): progenitor común mieloide.
- **CPA**: célula presentadora de antígeno.
- **Dataset**: objeto que contiene una matriz de expresión y que puede estar acompañada por datos no transcriptómicos que sean de utilidad para identificar o catalogar la muestra o paciente.
- **DC**: (*dendritic cells*): Células dendríticas.
- **DESeq()**: función de la paquetería DESeq2. Realiza análisis de expresión diferencial haciendo una corrección del *fold change* de genes con pocas cuentas, esto evita que los genes con pocas cuentas tengan *fold change* mayores respecto a genes con más cuentas.
- **EGF** (*epidermal growth factor*): factor de crecimiento epidérmico.
- **Eigengene**: componente principal de la matriz de expresión de un módulo. Es un valor representativo de todo el perfil de expresión del módulo.
- **EMT** (*epithelial-mesenchymal transition*): transición epitelio mesénquima.
- **ES** (*enrichment score*): nivel de enriquecimiento.
- **fgsea()**: función de la paquetería *fastgsea* o *fgsea*. Calcula el enriquecimiento de una o más firmas de genes dentro de un grupo de muestras A respecto al grupo de muestras B. Para determinar el enriquecimiento utiliza la prueba Kolmogorov-Smirnov *random walk*. la cual asigna un puntaje de enriquecimiento a una determinada lista de genes de acuerdo con al valor de *fold change* de los genes pertenecientes a la firma. Si pocos genes de la firma tienen valores altos de *fold change*, el enriquecimiento disminuye conforme disminuyen los valores de *fold change*. Está enfocado en facilitar el análisis de enriquecimiento de firmas preestablecidas que describen vías celulares. Las firmas preestablecidas están

depositadas en bases de datos y pueden ser descargadas mediante funciones de la misma paquetería.

- **FindClusters():** función de la paquetería Seurat. Identifica grupos de células basándose en la similitud de su transcriptoma. Calcula *k-nearest neighbor* para determinar los grupos.
- **FindIntegrationAnchors():** función de la paquetería Seurat. Identifica genes en común entre todos los *datasets* que se desean. Estos genes en común se denominan *anchors*. Estos *anchors* serán usados posteriormente para integrar los *datasets* con `IntegrateData()`.
- **FindVariableFeatures():** función de la paquetería Seurat. Identifica genes que tienen más de 2 desviaciones estándar respecto a la dispersión media.
- **GDCdownload():** función de la paquetería *TCGAbiolinks*. Permite descargar los datos previamente solicitados con la función `GDCquery()` desde el *Genomic Data Commons*.
- **GDCquery():** función de la paquetería *TCGAbiolinks*. Permite interactuar con la interfaz del *Genomic Data Commons* para solicitar datos de expresión con ciertas especificaciones.
- **Gene membership:** representa el valor de correlación de la expresión del gene G con el “eigengene” del módulo M. Si el valor resultante para el *gene membership* del gene G respecto al módulo M es igual o cercano a 0, el gene G no es parte del módulo M. Por otra parte, si el valor es cercano a 1 o -1, el gene G correlaciona altamente con el módulo M, ya sea de manera positiva o negativa.
- **GEO:** *Gene Expression Omnibus*
- **GLOBOCAN** (*Global Cancer Observatory*): Observatorio global de cáncer
- **GO:** *Gene Ontology*
- **goana():** función de la paquetería *limma*. Realiza prueba de sobre representación de términos de *Gene Ontology* sobre uno o más grupos de genes.
- **GSEA:** *Gene Set Enrichment Analysis*
- **gsea():** función de la paquetería *phenoTest*. Calcula el enriquecimiento de una o más firmas de genes en el grupo de muestras A respecto al grupo de muestras B. Para determinar el enriquecimiento utiliza la prueba Kolmogorov-Smirnov *random*

*walk*. la cual asigna un puntaje de enriquecimiento a una determinada lista de genes de acuerdo con al valor de *fold change* de los genes pertenecientes a la firma. Si pocos genes de la firma tienen valores altos de *fold change*, el enriquecimiento disminuye conforme disminuyen los valores de *fold change*. La principal diferencia con *fgsea()* radica en que *gsea()* tiene integrado un análisis de expresión diferencial, por lo que no es necesario realizarlo previamente. Además *gsea()* no está enfocado en facilitar análisis de vías ni en proporcionar firmas de cada vía.

- **GSVA**: *Gene Set Variation Analysis*
- **gsva()**: función de la paquetería GSVA. Para cada muestra, calcula el enriquecimiento de los genes pertenecientes a la firma respecto a genes que no pertenecen a la firma dentro de una misma muestra. Para determinar el enriquecimiento utiliza la prueba Kolmogorov-Smirnov *random walk*, la cual asigna un puntaje de enriquecimiento a una determinada lista de genes de acuerdo con al valor de *fold change* de los genes pertenecientes a la firma. Si pocos genes de la firma tienen valores altos de *fold change*, el enriquecimiento disminuye conforme disminuyen los valores de *fold change*. La principal diferencia con *gsea()* o *fgsea()* radica en que *gsva()* no requiere de un análisis de expresión diferencial entre dos grupos de muestras, calcula el *fold change* de cada gene basándose en la desviación que tiene la expresión de dicho gene respecto a la media de expresión del resto de los genes dentro de una misma muestra o célula.
- **HER2**: receptor de factor de crecimiento epidérmico humano 2
- **HLA** (*human leukocyte antigens*): antígenos leucocitarios humanos.
- **HumanPrimaryCellAtlasData()**: función de la paquetería *celldex*. Descarga 713 microarreglos de expresión normalizados, provenientes de 713 muestras depositadas en el *Human Primary Cell Atlas*. Los microarreglos provienen en su mayoría de líneas celulares en reposo o estimuladas para su diferenciación. Cada muestra tiene asignada una de las 37 identidades celulares consideradas por los autores del *Human Primary Cell Atlas*.
- **IL**: interleucina
- **iNOS**: óxido nítrico sintasa.

- **IntegrateData():** función de la paquetería Seurat. Realiza la integración de diferentes *datasets* usando los resultados arrojados por la función `FindIntegrationAnchors()`.
- **INF:** interferón
- **isOutlier():** función de la paquetería Scater. Identifica a las células con una desviación media absoluta mayor a 3 dentro del grupo de datos especificados.
- **KEGG:** *Kyoto Encyclopedia of Genes and Genomes*
- **kegga():** función de la paquetería limma. Realiza prueba de sobre representación de vías celulares del *Kyoto Encyclopedia of Genes and Genomes* sobre uno o más grupos de genes.
- **LPS:** lipopolisacárido.
- **Matriz de expresión:** estructura de datos de 2 dimensiones en el que las filas corresponden a genes y las columnas corresponden a células, muestras, pacientes o condición.
- **MAPK** (*mitogen-activated protein kinase*): proteínas cinasas mitógeno-activadas
- **M-CSF** (*macrophage colony stimulator factor*): factor estimulante de colonias de macrófagos.
- **MEC:** matriz extracelular.
- **MHC** (*major histocompatibility complex*): complejo principal de histocompatibilidad.
- **MMP** (*matrix metalloproteinases*): metaloproteinasas de matriz.
- **Módulo:** Grupo de genes altamente correlacionados entre sí.
- **moduleEigengene():** función de la paquetería WGCNA. Calcula el componente principal de la matriz de expresión del módulo correspondiente, al cual se le denomina "eigengene". Este se considera como un valor representativo de todo el perfil de expresión del módulo.
- **MRC1:** receptor de manosa de macrófagos.
- **MsigDB:** *Molecular Signatures Database*
- **NES** (*normalized enrichment score*): nivel de enriquecimiento normalizado.
- **NF-κB:** factor nuclear-κB
- **NK:** *Natural Killer*.

- **NormalizeData():** función de la paquetería Seurat. Normaliza la matriz de cuentas. Las cuentas de cada célula son divididas entre el total de cuentas de esa misma célula y el resultado es multiplicado por un factor de escala = 10,000 para realizar una normalización a nivel de la célula y no a nivel poblacional. Por último, los datos son transformados con logaritmo natural.
- **PD-L** (*programmed death-ligand*): ligando de muerte programada
- **Paquetería:** conjunto de funciones programadas en lenguaje R, especializadas en el manejo y análisis de datos para satisfacer necesidades específicas.
- **RunUMAP():** Realiza análisis de reducción dimensional denominado *Uniform Manifold Aproximation and Projection*. Cada célula es representada con un punto; las similitudes y diferencias entre los transcriptomas de las células son plasmadas en un plano de 2 dimensiones.
- **RunPCA():** Realiza una reducción de la dimensionalidad de la matriz de células x genes, mediante un análisis de componentes principales
- **ScaleData():** Centra a cero la media de los valores de expresión de los genes. Para centrar los valores primero resta el promedio de expresión del gene del valor de expresión de dicho gene, después escala los valores de expresión centrada de cada gene dividiéndolos entre su desviación estándar.
- **scRNA-seq** (*Single-cell RNA sequencing*): secuenciación de RNA de célula única.
- **signedKME():** función de la paquetería WGCNA. Calcula el *gene membership* de un grupo de genes para uno o más módulos.
- **SingleR():** función de la paquetería SingleR. Permite identificar tipos celulares de forma no supervisada en datos de secuenciación de RNA de célula única, usando datasets de referencia que pertenecen a tipos celulares aislados; usualmente son líneas celulares.
- **TAM** (*Tumor-associated Macrophages*): macrófagos asociados a tumor.
- **TCGA:** *The Cancer Genome Atlas*.
- **TGF $\beta$**  (*transforming growth factor-beta*): factor de crecimiento transformante  $\beta$ .
- **TLR:** receptor *toll-like*
- **TME** (*tumor microenvironment*): microambiente tumoral.
- **TNF** (*tumor necrosis factor*): factor de necrosis tumoral

- **TNFR** (*tumor necrosis factor receptors*): receptores de factor de necrosis tumoral
- **Treg**: linfocitos T reguladores.
- **UMAP**: *Uniform Manifold Approximation and Projection* (UMAP)
- **VEGF** (Vascular Endothelial Growth Factor): factor de crecimiento endotelial vascular.
- **WGCNA**: *Weighted Gene Correlation Network Analysis*



## RESUMEN

En México el cáncer de mama es una de las principales causas de muerte por cáncer en mujeres. Durante los últimos años, el enfoque del estudio del cáncer ha dejado de estar centrado en las características intrínsecas del tumor, para enfocarse en la interacción que tienen las células tumorales con los diferentes tipos celulares presentes en el microambiente tumoral (TME, por sus siglas en inglés *tumor microenvironment*). Uno de los componentes más importantes del TME es el infiltrado leucocitario, dentro del cual destacan los macrófagos asociados a tumor (TAM, por sus siglas en inglés *tumor-associated macrophages*) porque se ha reportado que la densidad de TAM infiltrados en el tumor correlaciona con poca supervivencia en pacientes en diversos tipos de cáncer, incluido cáncer de mama. Los macrófagos tienen importantes funciones en el desarrollo y mantenimiento del tejido mamario, pero en un contexto tumoral, se ha reportado que estos pueden favorecer el desarrollo del cáncer participando en procesos de inflamación, angiogénesis, inmunoregulación y remodelación de la matriz extracelular (MEC), lo cual facilita la migración e invasión de las células tumorales.

En un intento de explicar las funciones pro-tumorales y anti-tumorales observadas en los TAM, se adoptó el modelo de activación de macrófagos M1 (anti-tumoral) y M2 (pro-tumoral). Sin embargo, evidencia proporcionada por nuevas tecnologías transcriptómicas como secuenciación de RNA de célula única, sugiere que el modelo de activación de macrófagos M1/M2 y sus respectivos marcadores, no son adecuados para describir las poblaciones de TAM presente en el TME.

Este proyecto se llevó a cabo totalmente *in silico* y se planteó el objetivo de analizar de forma no supervisada las diferentes poblaciones de TAM del TME y contrastar lo encontrado con el modelo M1/M2 y sus respectivos marcadores; para esto se recurrió a datos públicos de infiltrado leucocitario de cáncer de mama generados con secuenciación de RNA de célula única. Mediante herramientas bioinformáticas se lograron determinar grupos de genes altamente correlacionados entre sí, a los cuales se les denominó firmas transcripcionales; se calculó el nivel de enriquecimiento de dichas firmas en los TAM y se pudo determinar los procesos celulares llevados a cabo por cada población de TAM que expresaba dichas firmas. Los resultados sugieren que los TAM tienen un comportamiento

dicotómico anti-tumoral y pro-tumoral. La población de TAM que expresaba firmas transcripcionales asociadas a inflamación, sistema complemento, fagocitosis y presentación de antígenos extracelulares, estaba enriquecida en pacientes con cáncer mama con mayor supervivencia. Por otra parte, la población de TAM que expresaba una firma transcripcional asociada con exocitosis y remodelación de la MEC se encontró enriquecida en pacientes con menor supervivencia. Ninguna de las poblaciones de TAM antes mencionadas, pudo distinguirse una de la otra mediante su expresión de marcadores M1/M2, lo cual sugiere que el modelo de activación de macrófagos M1/M2 y sus respectivos marcadores, no son adecuados para describir las diferentes poblaciones de TAM que existen en el TME de cáncer de mama.

## ABSTRACT

Breast cancer is one of leading causes of cancer deaths among Mexican women. In the past few years, the study of cancer has shifted from a tumor centric focus to an interaction-based focus of the different cell types found in the tumor microenvironment (TME). Tumor infiltrated leukocytes are one of the foremost components of the TME because it influences the clinical outcome of a cancer patient, for instance, high densities of tumor-associated macrophages (TAM) correlate with poor survival in several types of cancers, such as breast cancer. TAM play important roles in the mammary gland development and maintenance, however, in cancer TAM could support tumor development by exerting functions such as inflammation, angiogenesis, immunoregulation and extracellular matrix (MEC) remodeling, the latter ends up facilitating tumor migration and invasion.

In order to explain the anti-tumoral and pro-tumoral functions observed in TAM, a model of macrophage M1 (anti-tumoral) and M2 (pro-tumoral) was adopted. Nonetheless, recent technologies like single-cell RNA sequencing suggest that the M1/M2 model of macrophage activation and its respective markers, are not adequate to describe the TAM populations observed in the TME.

This work was made totally *in silico* and it aimed to analyze, in an unsupervised manner, the different TAM populations present in the TME and thus compare the results with the M1/M2 model and its markers. To achieve this, we used publicly accessible data of tumor infiltrated leukocytes generated with single-cell RNA sequencing. Bioinformatic tools were implemented to determine highly correlated sets of genes, designated as transcriptional signatures. An enrichment score was calculated for each one of these signatures in TAM populations, as well as the cellular pathways activated in the same TAM enriched with these signatures. Results suggest that TAM are dichotomic, with anti-tumoral and pro-tumoral features. TAM expressing signatures associated with inflammation, complement system, phagocytosis and presentation of extracellular antigens, were enriched in breast cancer patients with good survival. On the other hand, TAM expressing a signature associated with exocytosis and MEC remodeling were enriched in patients with poor survival. None of the aforementioned TAM populations could be distinguished one from another by their expression of M1/M2 markers. This suggest that the M1/M2 macrophage

activation model and its markers, are not adequate to describe the different TAM populations found in the TME of breast cancer.

# 1. INTRODUCCIÓN

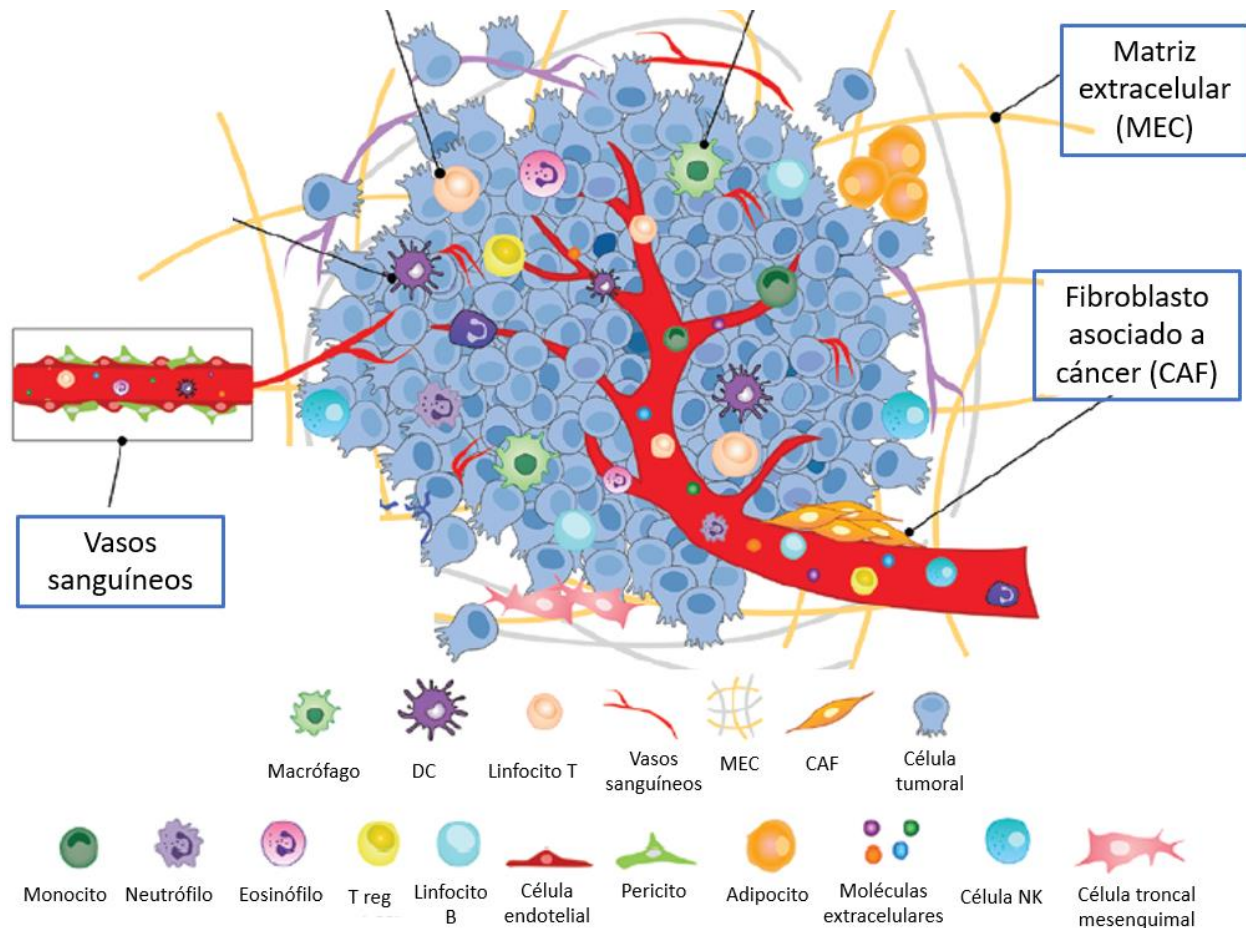
## 1.1 Cáncer de mama

El cáncer es un conjunto de enfermedades que se caracteriza por una proliferación desregulada de células anormales que pueden esparcirse a otros tejidos del cuerpo y causar estragos en la homeostasis del tejido [1]. Según datos estimados por el observatorio global de cáncer (GLOBOCAN, del inglés *Global Cancer Observatory*), el cáncer de mama es el primer lugar en incidencia y el quinto en mortalidad a nivel mundial (2020) [2]. En México, el cáncer de mama es el primer lugar en incidencia y también es el primer lugar en mortalidad por cancer en mujeres (2020) [2].

El cáncer de mama puede clasificarse de diversas formas. Por su tipo histológico se puede clasificar principalmente en invasivo ductal o invasivo lobulillar [3]. Puede clasificarse de acuerdo con su estado de diferenciación mediante el grado, el cual es un gradiente que va desde grado 1 (diferenciado) hasta grado 3 (pobremente diferenciado), siendo este último de mal pronóstico [4]. Basándose en la expresión de estrógeno, progesterona, del receptor de factor de crecimiento epidérmico humano 2 (HER2) y en el perfil transcripcional, el cáncer de mama se ha clasificado en los subtipos moleculares luminal A, luminal B, HER2-enriquecido y triple negativo, dentro del cual se encuentra el subtipo basal [5]. Se ha propuesto subclasificar a los cánceres triple negativo de acuerdo a su expresión de genes involucrados en ciclo celular, reparación de ADN, señalización del receptor de andrógeno, genes mesenquimales y por su nivel de infiltrado leucocitario [6]. Se ha observado que el nivel de infiltrado leucocitario es variable entre subtipos de cáncer de mama, por ejemplo, se ha reportado mayor infiltrado de macrófagos en cánceres de mama negativos a receptor de estrógeno y progesterona, así como en cánceres triple negativo basales en comparación con subtipos moleculares como luminal A [7,8,9] al mismo tiempo existen reportes contradictorios respecto a la correlación (positiva y negativa) que tienen los macrófagos con el subtipo HER2-enriquecido [8,9]. El mayor infiltrado de macrófagos en los cánceres triple negativo correlaciona con troncalidad, crecimiento, metástasis y con un microambiente inmunosupresor [10-14]; sin embargo, los macrófagos no son las únicas células del sistema inmune infiltradas en el estroma del cáncer de mama.

## 1.2 Microambiente tumoral

El microambiente tumoral (TME, por sus siglas en inglés *tumor microenvironment*) es el entorno complejo y rico en células en el que se desarrolla y regula la progresión del cáncer. Además de las células tumorales, el TME tiene un componente celular formado por células del sistema inmunológico, este incluye linfocitos T y B, células *Natural Killer* (NK); células de linaje mieloide como lo son los macrófagos, monocitos, células dendríticas (DC, por sus siglas en inglés *dendritic cells*), neutrófilos y otras células polimorfonucleares; también está compuesto por células residentes de tejido como células epiteliales, células endoteliales que constituyen los vasos sanguíneos y por células estromales mesenquimales como los fibroblastos, pericitos, adipocitos, condrocitos y células troncales mesenquimales. El componente extracelular del TME incluye moléculas secretadas tanto por el tumor como por las células no tumorales, tales como factores de crecimiento, citocinas y quimiocinas. Por último, la matriz extracelular (MEC) es considerada parte del TME porque provee de soporte estructural a las células; está compuesta por diversas macromoléculas como fibronectina, colágeno, glucoproteínas y proteoglicanos, las cuales pueden ser producidas por diversos tipos celulares, pero principalmente por fibroblastos (Figura 1) [15-17].



**Figura 1. Esquema de los diferentes componentes del microambiente tumoral.**

Los vasos sanguíneos que irrigan al tumor con nutrientes y oxígeno pueden facilitar el infiltrado de diversas poblaciones de leucocitos, las cuales son atraídas por un microambiente inflamatorio. Las células tumorales, leucocitos infiltrantes de tumor, células residentes de tejido y la MEC forman interacciones entre sí que confieren de una gran complejidad al TME. CAF: fibroblastos asociados a cáncer. DC: células dendríticas. MEC: matriz extracelular. TME: microambiente tumoral. Imagen modificada de [16].

En el TME, las células pueden adquirir funciones pro-tumorales, esto es adquirir la capacidad de favorecer el desarrollo del cáncer. Por ejemplo, los fibroblastos asociados a cáncer (CAF, por sus siglas en inglés *cancer-associated fibroblasts*) están asociados a poca supervivencia en pacientes con cáncer pancreático, cabeza y cuello, esofágico y colorrectal [18-21]. Los CAF pueden remodelar la MEC y son fuente importante de factor de crecimiento endotelial vascular (VEGF) y del factor de crecimiento transformante  $\beta$  (TGF $\beta$ ), esto les permite promover procesos de invasión y angiogénesis, así como influir en el desarrollo del cáncer [15].

De todos los componentes del TME, el infiltrado inmune ha demostrado tener la mayor dualidad en cuanto a funciones pro-tumorales y anti-tumorales. Los leucocitos deben sus funciones anti-tumorales a su capacidad de efectuar o promover una respuesta citotóxica contra el cáncer. Dentro de los leucocitos con funciones anti-tumorales se encuentran principalmente las DC, los linfocitos T citotóxicos con el *cluster* de designación (CD)8<sup>+</sup>, los linfocitos T cooperadores CD4<sup>+</sup> Th1 y las células NK.

Las DC son capaces de endocitar, procesar y presentar antígenos tumorales para luego migrar a nódulos linfáticos y activar a los linfocitos Th1 y CD8<sup>+</sup> citotóxicos, que posteriormente migrarán al tumor y actuarán sinérgicamente para eliminar células tumorales. Esto es posible gracias a un eje de activación mediado por citocinas proinflamatorias como interleucina (IL)-12, interferón (IFN)- $\gamma$  y factor de necrosis tumoral (TNF)- $\alpha$ . Las DC son capaces de presentar antígenos a poblaciones de linfocitos T a través de la interacción entre el TCR de los linfocitos y el complejo principal de histocompatibilidad (MHC, por sus siglas en inglés *major histocompatibility complex*) en ratón o mediante la interacción con antígenos leucocitarios humanos (HLA, por sus siglas en inglés *human leukocyte antigens*) en humanos. Cuando las DC endocitan un antígeno extracelular y lo procesan, lo presentan en un contexto de HLA-II, con el cual pueden activar a los linfocitos T CD4<sup>+</sup> que eventualmente coordinarán y promoverán la respuesta adaptativa mediante la producción de citocinas. Las DC también son capaces de activar a los linfocitos T CD8<sup>+</sup> en un contexto de HLA-I, esto gracias a que, a diferencia de otras células presentadoras de antígeno (CPA), las DC pueden presentar antígenos extracelulares endocitados en un contexto de HLA-I en lugar de HLA-II; esto gracias al mecanismo de presentación cruzada de antígeno. Adicionalmente, la activación de los linfocitos T requiere de la participación de moléculas co-estimuladoras como CD80/86 (situadas en la superficie de las CPA) y de su respectivo receptor CD28 (situado en la superficie de los linfocitos T) [15-17].

Las células NK no necesitan previa activación de las DC, estas pueden inducir lisis citotóxica en células tumorales que carecen de HLA-I en su superficie o que expresan moléculas de superficie inducidas por estrés celular como *MICA/B* o *ULBP1*, las cuales



pueden ser reconocidas al interactuar con el receptor de las células NK, *NKG2D*. Las células NK también pueden producir IFN- $\gamma$ , TNF- $\alpha$ , perforinas y granzimas [22].

Otra población linfocítica importante en el TME son los linfocitos T reguladores (Treg), los cuales se han asociado con un papel pro-tumoral debido a que producen citocinas inmuno-reguladoras como IL-10 y TGF $\beta$  que suprimen la respuesta citotóxica anti-tumoral [15-17].

El estudio del infiltrado leucocitario se ha enfocado principalmente en la biología de las poblaciones linfocíticas y de cómo estas influyen en el desarrollo del cáncer; sin embargo, las poblaciones mieloides también juegan un papel importante en el desarrollo del cáncer. Una de las poblaciones de origen mieloides del TME que mayor interés ha despertado por su posible papel pro-tumoral, son los macrófagos asociados a tumor (TAM, por sus siglas en inglés *tumor-associated macrophages*). Evidencia acumulada por diversos estudios hechos en distintos tipos de cáncer, sugieren que la cantidad de TAM presente en el tumor correlaciona con mal pronóstico y poca supervivencia en la mayoría de los tipos de cáncer analizados [23]. Esto sugiere que los TAM tienen la capacidad de influir en el desarrollo del cáncer y, por lo tanto, en la supervivencia del paciente.

### **1.3 Funciones de macrófagos asociados a tumor (TAM) en cáncer**

Los macrófagos pueden tener diferentes orígenes ontológicos. Los macrófagos residentes de tejido se originan a partir de precursores embrionarios y mantienen la población celular mediante procesos de auto renovación [24-26]; los macrófagos también pueden derivarse de monocitos de médula ósea atraídos al tejido tanto en condiciones fisiológicas como por condiciones de inflamación [27]. Las células de los tejidos atraen monocitos mediante la secreción de quimiocinas con motivos c-c (CCL) como CCL-2, CCL-8 y CCL-7. Una vez en el tejido, los monocitos son diferenciados terminalmente a macrófagos por acción de citocinas como el M-CSF (del inglés *macrophage colony stimulating factor*). En condiciones fisiológicas los macrófagos tienen funciones importantes en el desarrollo y mantenimiento de los tejidos. Por ejemplo, los macrófagos tienen un papel central en el desarrollo de la glándula mamaria. Se ha propuesto que los macrófagos funcionan como intermediarios entre los niveles hormonales cambiantes durante el ciclo menstrual y la expansión de células troncales mamarias [28]. Así, los

macrófagos son parcialmente responsables de los cambios de tamaño de la glándula mamaria durante estos ciclos, pero también durante el embarazo y la lactancia.

Esta capacidad de los macrófagos de promover el crecimiento de los tejidos, sumado a la correlación positiva observada entre la cantidad de macrófagos y un mal pronóstico en pacientes con cáncer, apoya la teoría de que los TAM pueden favorecer funciones pro-tumorales principalmente. Estudios pioneros del grupo de investigación del Dr. Jeffrey W. Pollard comprobaron esta asociación. En estos estudios se observó que el silenciamiento de M-CSF resultó en un pobre infiltrado de TAM que retrasó los procesos de invasión y metástasis [29]. Adicionalmente, altos niveles de M-CSF y de su respectivo receptor *CSF1R*, caracterizan a los cánceres humanos con mal pronóstico, incluido el cáncer de mama [30]. Actualmente se conocen diversos mecanismos por los cuales los TAM podrían influir negativamente en el pronóstico del cáncer. A continuación, se describirán los siguientes fenómenos llevados a cabo por los TAM: contribución al mantenimiento de un microambiente inflamatorio crónico y generación de especies reactivas de oxígeno y nitrógeno; promoción de angiogénesis; inmunoregulación y facilitación de la migración e invasión a través de la remodelación de la MEC.

Se ha propuesto que los TAM participan en la iniciación y promoción del cáncer a través del mantenimiento de un microambiente inflamatorio crónico, mediante la secreción de mediadores proinflamatorios como TNF- $\alpha$ , IL-12 e IL-6; sin embargo, estas citocinas también participan en la activación de células linfoides con actividad anti-tumoral, por lo que resulta complicado catalogar a la producción de mediadores proinflamatorios como una función pro-tumoral o anti-tumoral. Adicionalmente los TAM producen especies reactivas de oxígeno y de nitrógeno que, en conjunto crean un ambiente mutagénico y promotor de crecimiento [31-32].

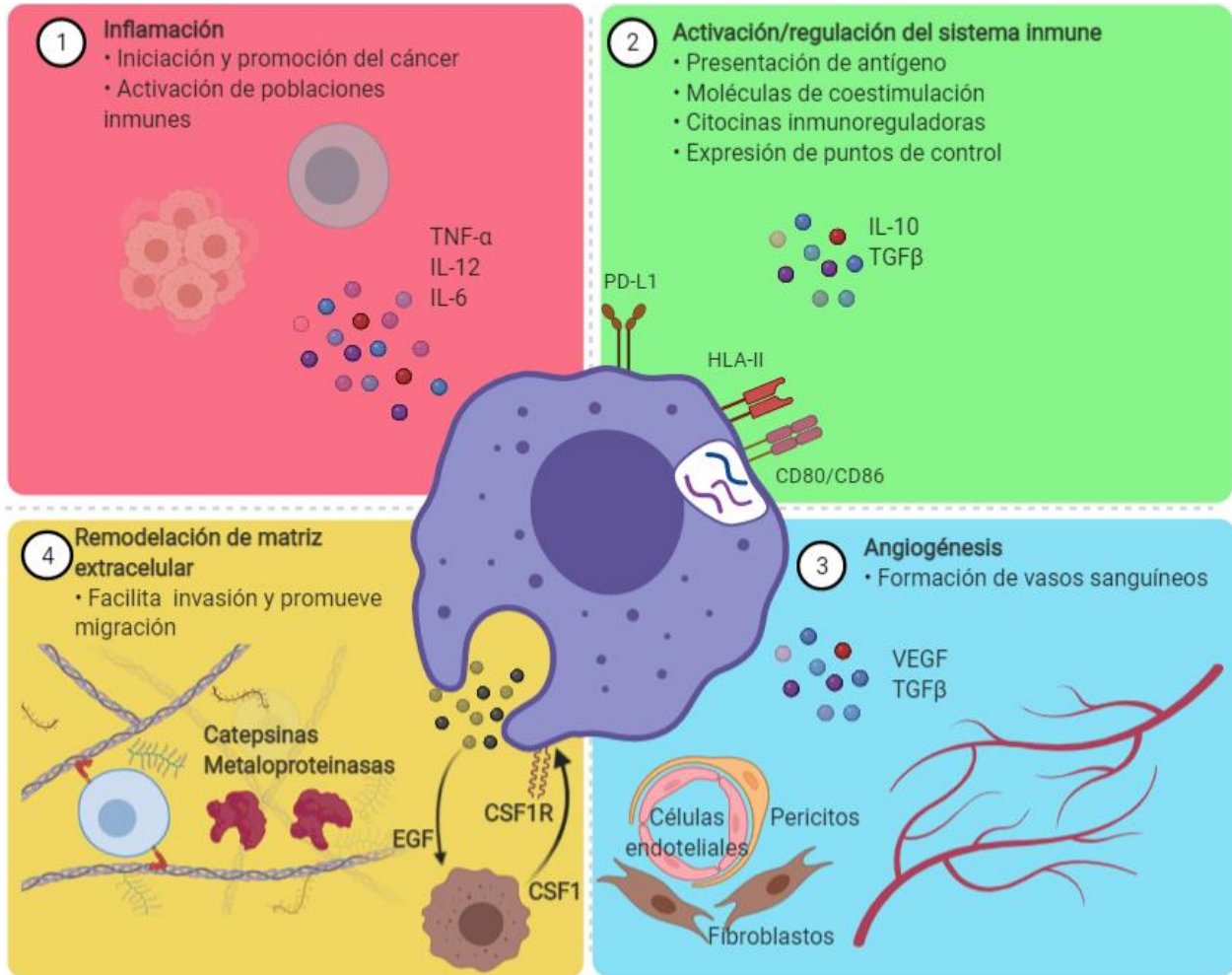
La angiogénesis es un proceso mediante el cual se forman nuevos vasos sanguíneos con el fin de suplir la demanda de oxígeno y nutrientes de las células tumorales. Se cree que los TAM contribuyen a este proceso, principalmente mediante la producción de factores pro angiogénicos como VEGF y TGF $\beta$  en áreas hipóxicas del tumor; de esta forma se reclutan células endoteliales, fibroblastos y pericitos que eventualmente permitirán la formación de nuevos vasos sanguíneos. Además, la angiogénesis requiere

de la degradación de la MEC, la cual es lograda mediante la secreción de enzimas proteolíticas capaces de remodelar la MEC como catepsinas y metaloproteinasas de matriz (MMP, por sus siglas en inglés *matrix metalloproteinases*), las cuales se ha reportado que pueden ser producidas por los TAM [32-34].

Relacionado a la remodelación de la MEC, se encuentra el proceso de invasión celular de las células tumorales. A parte de liberar catepsinas y MMP al TME, los TAM son fuente del factor de crecimiento epidérmico (EGF por sus siglas en inglés *epidermal growth factor*) en cáncer de mama. El EGF incrementa la migración de las células tumorales mediante un mecanismo en el que las células tumorales producen M-CSF y los TAM lo reconocen mediante el receptor *CSF1R* y responden con EGF formando un ciclo de retroalimentación positiva [35]. En conjunto estas funciones promueven la migración e invasión de las células y eventualmente metástasis [32].

Al igual que las DC, los macrófagos son CPA que pueden fagocitar antígenos extracelulares y presentarlos en moléculas HLA-II presentes en su superficie; también pueden expresar citocinas proinflamatorias y moléculas co-estimuladoras como CD80/86. En teoría esto permitiría a los TAM montar una respuesta anti-tumoral a través de estimular poblaciones linfoides con actividades citotóxicas, siempre y cuando los TAM hayan sido previamente activados por IFN- $\gamma$ . Sin embargo, se ha observado que los TAM tienen expresión reducida de IL-12, IL-18, óxido nítrico sintasa (iNOS) y HLA-II. Además, los TAM pueden contribuir a la supresión del sistema inmune mediante la secreción de citocinas inmunosupresoras como IL-10 y TGF $\beta$ , que suprimen las funciones citotóxicas de linfocitos T CD8<sup>+</sup>, células NK e inhiben la producción de IL-12 por parte de las DC, lo cual favorece la expansión y desarrollo de Treg. Quimiocinas derivadas de TAM como CCL-22, también favorecen el reclutamiento de más Treg al TME [31, 33, 36]. Otro mecanismo por el que los TAM pueden suprimir una respuesta citotóxica es a través de la expresión de ligandos de superficie que inhiben directamente la función de los linfocitos T, como el ligando de muerte programada 1 y 2 (PD-L1 y PD-L2, por sus siglas en inglés *programmed death-ligand*) [33].

En conjunto, la evidencia recabada sobre las diversas funciones de los TAM (Figura 2), sugieren que son un tipo celular altamente versátil que requiere de mayor investigación para entender su papel en el TME.



**Figura 2. Esquema de las funciones de TAM en cáncer.**

Las funciones de los TAM pueden clasificarse en 4 procesos. 1- Contribuyen a mantener un microambiente inflamatorio mediante la producción de citocinas proinflamatorias que pueden tanto promover el desarrollo del cáncer como promover la activación de células inmunes con funciones anti-tumorales. 2- Regulan al sistema inmunológico de manera positiva mediante presentación de antígeno (HLA-II) y moléculas coestimuladoras (CD80/CD86), pero también lo pueden regular de manera negativa mediante la producción de IL-10 y TGF $\beta$ , así como mediante la producción de puntos de control como PD-L1. 3- Favorecen el proceso de angiogénesis promoviendo el reclutamiento de fibroblastos, pericitos y células endoteliales mediante la secreción de VEGF y TGF $\beta$ . 4- Son capaces de degradar proteínas de la matriz extracelular secretando enzimas como catepsinas y metaloproteinasas. La remodelación de la matriz extracelular puede facilitar la invasión de las células tumorales a otros tejidos, además los TAM pueden promover la migración de estas células mediante un bucle de retroalimentación positiva conformado por EGF y CSF1/CSF1R. Imagen creada con la versión gratuita de Biorender.com.

## 1.4 Relación de los TAM con un modelo de macrófagos M1/M2

En 1992 Stein y colaboradores reportaron por primera vez que IL-4 inducía un fenotipo de activación alternativa en macrófagos distinto al inducido por IFN- $\gamma$ . El fenotipo inducido por IL-4 constaba de una mayor capacidad endocítica a través de una mayor actividad del receptor de manosa de macrófagos (MRC1), así como una mayor expresión de MHC-II en algunos macrófagos [37]. Mills y colaboradores extendieron el concepto de macrófagos alternativamente activados a un modelo dicotómico de macrófagos, proponiendo que estos pueden expresar 2 programas metabólicos opuestos [38]. En este estudio usaron dos cepas de ratón en donde la cepa C57BL/6 tenía mayor tendencia a producir respuestas Th1 junto con IFN- $\gamma$ , mientras que la cepa BALB/c tendía a producir respuestas Th2 e IL-4. Ante un estímulo con lipopolisacárido (LPS), los macrófagos provenientes de la cepa C57BL/6 produjeron más óxido nítrico y fueron denominados macrófagos M1; mientras que los macrófagos provenientes de la cepa BALB/c no produjeron óxido nítrico ante el mismo estímulo, pero incrementaban el metabolismo de arginina a ornitina y fueron denominados macrófagos M2. A partir de esta observación, se aceptó que IFN- $\gamma$  y LPS inducen en macrófagos la expresión de genes propios de un fenotipo M1, mientras que las interleucinas características de una respuesta Th2 como IL-4 e IL-13, inducen la expresión de genes propios de un fenotipo M2 en macrófagos.

Posteriormente, Mantovani A. y colaboradores [39] propusieron que los TAM debían considerarse como macrófagos M2 debido a que tienen características similares. Por ejemplo, los TAM al igual que los macrófagos M2, tienen baja producción de óxido nítrico y moléculas citotóxicas, poca capacidad de presentar antígeno y expresan poco IL-12, al mismo tiempo que tienen alta expresión de IL-10 y MCR1. Adicionalmente, los autores argumentaron que los linfocitos T infiltrantes de tumor tienen un perfil de expresión con alta producción de IL-4, IL-13 e IL-10, que podría promover un fenotipo M2 en los TAM. Además, hay evidencia que sugiere que los macrófagos M1 tienen funciones anti-tumorales, mientras que los macrófagos M2 tienen funciones pro-tumorales.

En la tabla 1 se resumen los genes que son frecuentemente aceptados como marcadores M1/M2. Los marcadores incluidos en la tabla 1 son genes inducidos en macrófagos o monocitos después de recibir uno o varios de los estímulos M1 (LPS, IFN- $\gamma$ ) o M2 (IL-4,

IL-13, IL-10), también contiene marcadores encontrados en TAM y que posteriormente fueron considerados como marcadores M2 debido a que se adoptó la idea de que los TAM tienen un fenotipo de macrófago M2.

**Tabla 1. Recopilación de marcadores M1 inducidos por IFN- $\gamma$ , LPS y marcadores M2 inducidos por IL-4, IL-13 o IL-10.**

Marcadores M1	Marcadores M2	Comentarios	Referencia
	<i>MRC1</i>		[37]
<i>NOS2</i> (iNOS)	<i>ARG1</i> (arginasa 1), <i>TGFB1</i>	La actividad de las enzimas aumentó, pero no se midió directamente la abundancia de las enzimas.	[38]
<i>IL12A</i> , <i>IL23A</i>	<i>IL10</i>		[40]
<i>CXCL9</i> , <i>CXCL10</i> , <i>CXCL11</i> , <i>CCL5</i> , <i>IRF1</i> , <i>TNF</i> , <i>IDO1</i> , <i>IL12B</i> , <i>IL6</i>	<i>MRC1</i> , <i>MSR1</i> ( <i>CD204</i> ), <i>CCL13</i> , <i>CTSC</i> , <i>TGFB1</i>		[41]
<i>HLA-II</i> , <i>CD80</i> , <i>CD86</i>		Estudio realizado con monocitos.	[42-43]
	<i>CD163</i>		[44-45]
	<i>CTSB</i> , <i>CTSS</i> , <i>CTSC</i> , <i>CTSH</i> , <i>CTSL</i> , <i>CTSZ</i> (catepsinas)		[46-47]
	<i>FN1</i>		[48]
	<i>CCL24</i>		[49]
	<i>CCL22</i>		[50]
	<i>CCL18</i>		[51]
	<i>CCL17</i>		[52]
	<i>IL1R2</i>	Reportado en monocitos	[53]
	<i>IL1RN</i> (receptor antagonista de IL-1)		[54]
	<i>CLEC7A</i>		[55]
<i>IL1B</i> (IL-1 $\beta$ ), <i>IL1A</i>		Marcadores aceptados como M1 por su actividad proinflamatoria y por inducir expresión de <i>NOS2</i>	[56]
	<i>MARCO</i>	Se encontró en TAM que expresaban marcadores M2	[57]
	<i>CD274</i> (PD-L1), <i>PDCD1LG2</i> (PD-L2)	Macrófagos aislados de ratones infectados con <i>Taenia crassiceps</i> . Los macrófagos también expresaban marcadores M2	[58]
	<i>EGF</i> , <i>CSF1R</i>	*	[59]
	<i>MMP12</i>	*	[60]

<i>MMP9</i>	*	[61-63]
<i>PLAUR</i> (receptor urocinasa de plasminógeno activador)	*	[64]
<i>VEGFC</i>	*	[65]

\*Marcador encontrado en TAM y posteriormente adoptado como marcador M2, aun cuando no hubiera reportes de ser inducido por IL-4, IL-13 o IL-10.

## 1.5 Problemas con la clasificación M1/M2

Existen varios problemas cuando se intenta simplificar el fenotipo y las funciones de los macrófagos en solo dos subtipos, los cuales tienen que encajar con el modelo de macrófagos M1/M2. Primero, los monocitos y los macrófagos son células que pueden responder a una gran variedad de estímulos con cambios morfológicos, fenotípicos y funcionales. Por ejemplo, los macrófagos además de fagocitar patógenos invasores, tienen como una de sus principales funciones, la eliminación de células envejecidas en condiciones no patológicas. Para esto los macrófagos expresan diversas moléculas que promueven la reparación o cicatrización de los tejidos, apoyan la proliferación celular y la angiogénesis y participan en procesos de inmunosupresión, como se mencionó anteriormente [31-34]. Segundo, existe una gran heterogeneidad de procedimientos experimentales y de marcadores utilizados para asignar un fenotipo que coincida con la función estudiada. Es por esto que, aunque es relativamente fácil inducir experimentalmente a los macrófagos para que realicen funciones de activación específicas, sigue siendo complicado asociar una función a un fenotipo específico. Finalmente, CD68 ha sido el principal marcador utilizado para reconocer TAMs, pero este marcador también es expresado por monocitos y por otras células del TME [66-67].

A diferencia de los linfocitos T CD8<sup>+</sup> y CD4<sup>+</sup> que se diferencian terminalmente a poblaciones efectoras específicas, los monocitos y los macrófagos son células muy plásticas que siguen respondiendo a diferentes estímulos, incluso después de haberse polarizado a un subtipo en específico [68]. Mosser y Edwards propusieron que los macrófagos existen en un espectro de polarización en el que los subtipos M1/M2 permanecen perfectamente separables solamente en los polos del espectro [69] mientras que en las partes intermedias del espectro de polarización pueden tener diferentes

fenotipos y funciones. De acuerdo con esta propuesta, estudios más recientes han observado que la polarización de monocitos con una amplia gama de señales genera macrófagos con programas transcripcionales igualmente diversos [70]. TAM aislados del modelo de cáncer de mama murino MMTV-PyMT no presentaron un programa transcripcional relacionado con el modelo M1/M2, pero se encontró que expresaban una firma transcripcional propia de la vía de señalización Notch [71]. En estudios *in vitro* [72], se ha encontrado que los macrófagos inducidos hacia una actividad pro-tumoral con líneas celulares de cáncer de mama triple negativo expresaban marcadores M1/M2 de manera indistinta.

Desde los primeros estudios que realizaron Mills y colaboradores [38] quedó claro que la clasificación M1 y M2 surgió con el fin de explicar las diferentes respuestas metabólicas que ellos observaron en los macrófagos, sin embargo, la adopción de este modelo para explicar diferencias funcionales entre TAM, no ha resultado igualmente exitosa.

## **1.6 Las nuevas perspectivas sobre la heterogeneidad de los TAM**

La llegada de nuevas tecnologías como secuenciación de RNA de célula única (scRNA-seq, por sus siglas en inglés *single-cell RNA sequencing*) ha permitido estudiar la heterogeneidad de las células con una resolución que no había sido posible con tecnologías previas, además tiene la ventaja de capturar el transcriptoma de cada célula de manera no supervisada. Evidencia arrojada recientemente por estudios hechos con tecnología scRNA-seq, indica que la heterogeneidad de los TAM es más compleja de lo que se creía con el modelo M1/M2 y que su fenotipo es difícil de clasificar [73-78]. Cada grupo de investigación reporta haber encontrado un número diferente de subgrupos de TAM con diferentes funciones en diferentes tipos de cáncer; sin embargo, todos coinciden en que ninguno de los grupos encontrados correlaciona con un fenotipo M1 ni M2 y que más bien, estos marcadores se expresan de forma indiscriminada en las distintas poblaciones de TAM [73-78]. En estudios proteómicos de célula única, tampoco se ha observado una división clara de los TAM basándose en el modelo de macrófagos M1/M2 [79-80].



## **2. PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACIÓN**

El cáncer de mama es un importante problema de salud pública y es el cáncer número uno en incidencia y mortalidad en mujeres mexicanas en edades laboralmente activas. Debido a su alta incidencia y mortalidad en nuestro país, es importante investigar y entender cómo los diferentes tipos celulares del TME influyen en el desarrollo del cáncer y la supervivencia del paciente. Se ha observado que los TAM correlacionan negativamente con la supervivencia del paciente. Aunque la evidencia sugiere que existen tanto macrófagos anti-tumorales como pro-tumorales, sus fenotipos asociados y las condiciones precisas que favorecen su formación, siguen siendo un área que necesita investigación exhaustiva. La dualidad observada en los TAM llevó a la adopción de un modelo dicotómico de activación de macrófagos M1/M2 para intentar explicar la biología de los TAM presentes en el TME; sin embargo, el modelo M1/M2 ha sido puesto a prueba con la reciente implementación de tecnologías scRNA-seq en el estudio del TME de diferentes tipos de cáncer. Gracias a la resolución a nivel celular y a los análisis no supervisados que ofrece scRNAseq, ha quedado evidenciado que los marcadores M1/M2 no logran describir con satisfacción las diferentes poblaciones de TAM observadas por estos estudios. Es importante mencionar que, la mayoría de los estudios que cuentan con datos scRNA-seq se han enfocado en describir superficialmente al TME o se ha enfocado en poblaciones de linfocitos T CD8<sup>+</sup>. Por lo tanto, es necesario realizar un análisis enfocado en las diferentes poblaciones de TAM presentes en el TME para determinar sus respectivas funciones celulares y posteriormente corroborar si estas poblaciones pueden explicarse mediante el modelo de macrófagos M1/M2. Este análisis también permitirá un mayor entendimiento del papel de los macrófagos en el desarrollo del cáncer y su relación con la supervivencia del paciente.

## **3. PREGUNTA DE INVESTIGACIÓN**

¿Es adecuado el modelo de macrófagos M1/M2 para describir las poblaciones y funciones celulares de los TAM presentes en el TME de cáncer de mama?

## **4. HIPÓTESIS**

La implementación de datos scRNA-seq y el aprovechamiento de la resolución a nivel celular que estos ofrecen, permitirán proponer una clasificación no supervisada de las

distintas poblaciones y respectivas funciones de los TAM presentes en el TME de cáncer de mama. Se espera que el modelo de macrófagos M1/M2 y sus respectivos marcadores, no serán adecuados para describir las diferentes poblaciones y funciones de los TAM.

## **5. OBJETIVO GENERAL**

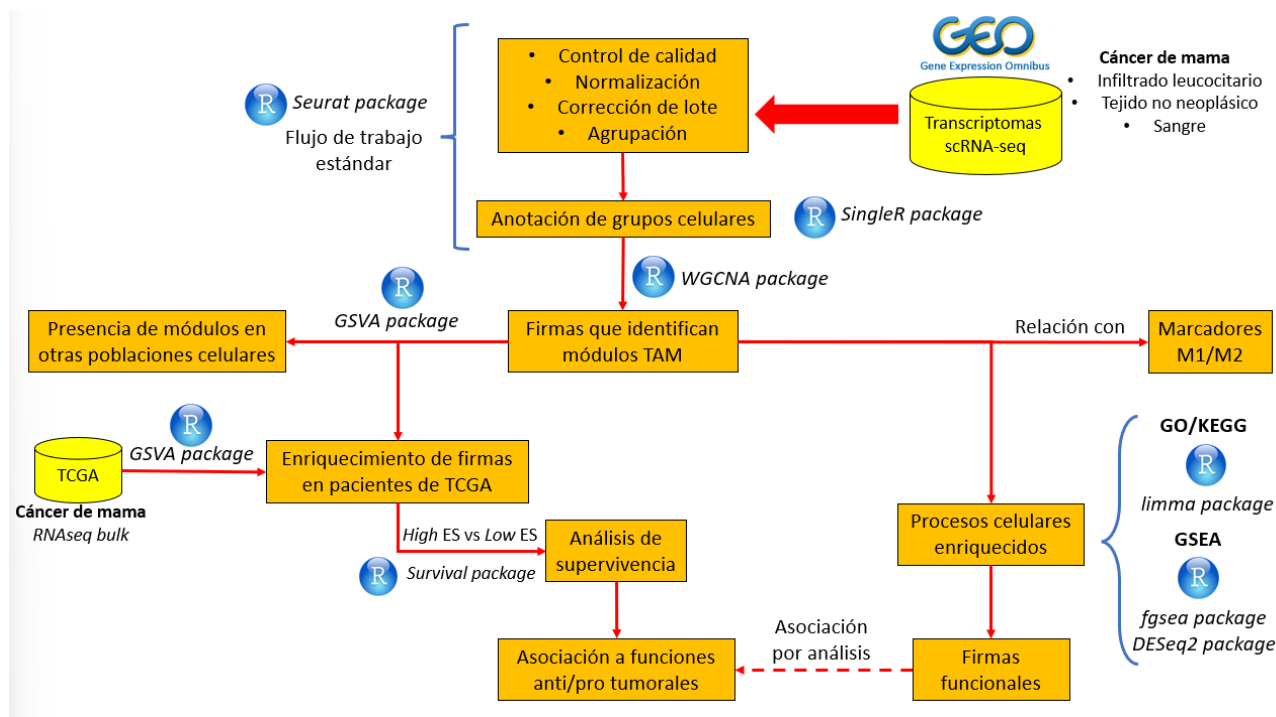
Utilizando datos públicos scRNA-seq provenientes de muestras de cáncer de mama, definir de manera no supervisada las poblaciones y funciones de los TAM presentes en el TME y mediante marcadores M1/M2, evaluar si estas poblaciones y sus respectivas funciones correlacionan con el modelo de macrófagos M1/M2.

### **5.1 Objetivos particulares**

1. Identificar y asignar identidad a las células que corresponden a macrófagos dentro de los datos scRNA-seq analizados.
2. Proponer una clasificación no supervisada de las diferentes poblaciones de TAM basada en sus funciones celulares.
3. Analizar la expresión de marcadores M1/M2 en la clasificación propuesta de TAM.
4. Evaluar la relevancia de las poblaciones de TAM en la supervivencia de pacientes del TCGA con cáncer de mama.

## **6. METODOLOGÍA**

La figura 3 resume el flujo del trabajo y las herramientas utilizadas para llevar a cabo la estrategia de análisis que a continuación se describe a detalle.



**Figura 3. Resumen del flujo de trabajo y herramientas aplicadas durante el análisis bioinformático.**

El análisis fue realizado con lenguaje de programación R. Se descargaron *datasets* scRNA-seq del *Gene Expression Omnibus* y se aplicó un flujo de trabajo estándar sobre estos datos con la paquetería *Seurat*. Se asignó identidad a cada célula con *SingleR* y se aplicó *WGCNA* sobre los TAM. Se exploraron los procesos celulares enriquecidos por los TAM que expresaban cada firma. Se determinó la relación que tenían los marcadores M1/M2 con los TAM que expresaban cada firma. Se determinó si las firmas encontradas eran expresadas por tipos celulares diferentes a los TAM. Por último, se determinaron los pacientes que tenían enriquecidas dichas firmas y se realizó un análisis de supervivencia comparando pacientes enriquecidos con diferentes firmas. GEO: *Gene Expression Omnibus*; scRNA-seq: *single-cell RNA sequencing*; WGCNA: *Weighted Gene Correlation Network Analysis*; TAM: macrófagos asociados a tumor; GSVAs: *Gene Set Variation Analysis*; ES: nivel de enriquecimiento.

## 6.1 Adquisición y características de transcriptomas scRNA-seq provenientes de pacientes con cáncer de mama.

Se realizó una búsqueda en PubMed (enero, 2019) de publicaciones con datos scRNA-seq en pacientes con cáncer de mama. Se encontraron dos publicaciones con datos disponibles de forma pública en GEO (del inglés, *Gene Expression Omnibus*) (Tabla 2).

El set de datos con el número de identificación GSE114725 [74] consta de 8 pacientes con cáncer de mama, de quienes en conjunto se recolectaron 30 muestras tumorales, 11 muestras de tejido no neoplásico y 9 muestras sanguíneas. El tejido no neoplásico fue obtenido de mastectomías profilácticas contralaterales de las mismas pacientes con

cáncer de mama. Sumando todas las muestras de los diferentes tejidos, resulta un total de 50 matrices de expresión (célula x genes) en formato .csv cuyo número de células secuenciadas por matriz iba desde 68 células hasta 2565 células, sumando un total de 52,759 células iniciales.

El segundo *set* de datos con el número de identificación GSE75688 [81], está conformado por 11 pacientes de cáncer de mama, cuyas muestras de tejido tumoral estaban todas contenidas en una misma matriz con formato .txt dando un total de 549 células iniciales; las cuales conforman tanto células tumorales como infiltrado leucocitario. El *set* de datos GSE75688 únicamente fue utilizado para llevar a cabo un análisis enfocado en las células tumorales que tenía disponibles, nunca fue integrado ni analizado junto con GSE114725.

Información adicional de los pacientes como edad, grado histológico, tamaño del tumor, expresión de receptores de progesterona y HER2, así como presencia o ausencia de metástasis, no fueron considerados para este análisis, pero se encuentran disponibles en el material suplementario de las respectivas publicaciones [74, 81]. No fueron reportados datos diferentes a los mencionados.

**Tabla 2. Características de las muestras y de las publicaciones de las que se obtuvieron los datos scRNA-seq.**

GEO ID	Células filtradas por CD45 <sup>+</sup>	Células totales	P	# muestras tumorales por P	# muestras no neoplásicas por P	# muestras sanguíneas por P	Año	Referencia
GSE114725	Sí	52,759	8	P1=4, P2=4, P3=3, P4=6, P5=4, P6=3, P7=3, P8=3	P1=4, P2=3, P3=3, P8=1	P1=2, P4=7	2018	[74]
GSE75688	No	549	11	NE	NA	NA	2017	[81]

**Nota:** No se especifica la zona del tumor de la que fue tomada cada muestra. No se especifica si cada muestra fue tomada en tiempos diferentes o si fue una misma muestra por paciente dividida en partes para secuenciación de RNA. P: pacientes; #: número de; NE: no especificado; NA: no aplica.

## 6.2 Procesamiento estándar de datos scRNA-seq

El manejo de los datos se llevó a cabo con lenguaje de programación R y con funciones de la paquetería Seurat versión 3.2 (<https://satijalab.org/seurat/>) con los parámetros fijados por *default*, a menos que se indique lo contrario. La gramática de todas las funciones del lenguaje de programación R requieren ser indicadas por paréntesis, dicha

gramática fue respetada en este trabajo para indicar cada función a continuación. Para ambos *datasets* se crearon objetos Seurat a partir de las matrices de cuentas, omitiendo células con menos de 200 genes detectados y genes que no fueron detectados en al menos 3 células.

El siguiente procedimiento fue realizado de forma paralela en ambos *datasets*, a excepción de la normalización y la corrección del error de lote, estos procedimientos no se realizaron en el segundo *dataset* GSE75688, debido a que las muestras ya estaban normalizadas e integradas. Para realizar el control de calidad, se calcularon los siguientes parámetros para cada célula: total de transcritos detectados, total de genes detectados y porcentaje de genes mitocondriales; estos parámetros son los recomendados y comúnmente usados en el control de calidad de los datos scRNA-seq [82]; ([https://satijalab.org/seurat/v3.2/pbmc3k\\_tutorial.html](https://satijalab.org/seurat/v3.2/pbmc3k_tutorial.html)). Mediante la función `isOutlier()` de paquetería `Scater` (<http://bioconductor.org/packages/release/bioc/html/scater.html>) se eliminaron células que tuvieron una desviación absoluta mayor a 3 respecto a la media, en al menos uno de los tres parámetros antes mencionados, según lo propuesto por Lun, A. T.L.; McCarthy, D. J.; Marion, J. C. [83]. Al excluir células, cuyos valores de genes detectados y transcritos detectados se desviaban 3 veces de la media, se están removiendo células que probablemente representaban dobletes celulares. De la misma forma, se removieron células en proceso de apoptosis o con algún tipo de daño celular basado en su alto porcentaje de genes mitocondriales detectados.

Previo al control de calidad, el primer *dataset* GSE114725 constaba de 52,759 células y después de realizar el control de calidad se removieron entre el 8-15% del total de células por cada tipo de tejido (sangre, tejido no neoplásico, tumor) por paciente (8 pacientes), con la excepción de la muestra tumoral 06 (BC06\_T), de la cual se removió el 27.3% de las células principalmente después de aplicar el filtro de transcritos detectados. El total de células restantes fue de 48,142 células. El segundo *dataset* GSE75688 constaba de 549 células, pero después de realizar el control de calidad, se conservaron el 89% dando un total de 489 células.

Cada matriz de genes (contenida en un objeto Seurat) fue normalizada de forma individual con `NormalizeData()`, posteriormente se determinaron los 2000 genes más

variables de cada matriz mediante `FindVariableFeatures()`. Los 2000 genes variables fueron usados como datos de entrada para realizar la corrección del error de lote y posteriormente, para integrar todas las matrices de datos mediante las funciones `FindIntegrationAnchors()` (fijando el parámetro `k.filter=50`) e `IntegrateData()`. Después de corregir el error de lote e integrar los datos, las células se agruparon usando solamente los 2000 genes variables mediante las funciones `ScaleData()`, `RunPCA()` y `FindClusters()`. Los gráficos UMAP fueron usados para la representación visual en dos dimensiones con `RunUMAP()`.

### 6.3 Anotación celular

Se recurrió a la paquetería `celldex` versión 1.0.0 (<http://bioconductor.org/packages/release/data/experiment/html/celldex.html>) para obtener el transcriptoma de referencia de 37 tipos celulares disponibles mediante la función `HumanPrimaryCellAtlasData()`; con esta función se obtienen transcriptomas de referencia que fueron previamente contruidos usando datos de 713 microarreglos de líneas celulares del Atlas de Células Primarias Humanas (del inglés, *Human Primary Cell Atlas*) [84].

Usando la paquetería `SingleR` versión 1.4.0 (<https://bioconductor.org/packages/release/bioc/html/SingleR.html>) [85] y los 37 transcriptomas referencia se procedió a etiquetar la identidad de las 48,142 células realizando análisis de correlación con los transcriptomas de referencia célula por célula mediante la función `SingleR()`, fijando el parámetro `method= "single"`. Todas las células fueron etiquetadas con alguna identidad celular, excepto un aproximado de 100 células que fueron catalogadas como desconocido ("*unknown*").

### 6.4 *Weighted gene correlation network analysis*

Se seleccionaron las células etiquetadas como macrófagos provenientes de tejido tumoral. Se eliminaron genes mitocondriales, ribosomales y sin identidad (secuencias numéricas), así como genes con un valor de expresión igual a cero en más del 99% de los macrófagos. Como resultado se obtuvo una matriz compuesta por 1354 TAM y 7247 genes, la cual fue usada como dato de entrada para realizar *Weighted gene correlation network analysis* (WGCNA) siguiendo las instrucciones de los propios desarrolladores de

la paquetería WGCNA versión 1.69 [86]

(<https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/>). Se construyó un *signed network* para encontrar correlaciones tanto positivas como negativas entre genes y se usó un *soft threshold* ( $power$ )=12 como lo recomiendan los autores

(<https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/faq.html>). Con esto se obtienen grupos de genes correlacionados entre sí, denominados módulos. Se obtuvo el *gene membership* (Revisar glosario de términos y abreviaciones) de cada gene con su respectivo módulo mediante las funciones `moduleEigengene()` y `signedKME()`. Los genes pertenecientes a cada módulo encontrado se catalogaron de acuerdo a su valor de *gene membership* (positivo o negativo) y posteriormente se seleccionaron sólo aquellos con un valor de *gene membership* mayor o igual a 0.4 para su respectivo módulo (procedimiento explicado a detalle en sección 7.2). El enriquecimiento medido en las secciones 6.6.2, 6.6.3 y 6.8 fue realizado con los genes representados en la tabla 3.

## 6.5 Análisis de representación de vías celulares con GO y KEGG

Con la paquetería `limma` versión 3.4.6.0 (<https://bioconductor.org/packages/release/bioc/html/limma.html>) se determinaron las vías celulares más representadas por todos los genes de cada módulo según Gene Ontology (GO) con la función `goana()` y *Kyoto Encyclopedia of Genes and Genomes* (KEGG) mediante la función `kegga()`.

## 6.6 Enriquecimiento de firmas transcripcionales

Para determinar el enriquecimiento conjunto de un grupo genes (firma) en una determinada matriz de expresión (célula x genes), se utilizó la función `gsva()` de la paquetería `gsva` versión 1.38.0 (<https://www.bioconductor.org/packages/release/bioc/html/GSVA.html>) [87] con sus parámetros fijados por *default*. Los valores de enriquecimiento arrojados por GSVA van de -1 a +1. A continuación se muestran las diferentes mediciones realizadas y los respectivos criterios escogidos para cada valor de enriquecimiento (ES del inglés, *enrichment score*) que se usaron para etiquetar a las células o pacientes.

### **6.6.1 Enriquecimiento de la firma de las DC en células mieloides**

Con el fin de corroborar que la identificación celular entre macrófagos, monocitos y DC por parte de SingleR se llevó a cabo satisfactoriamente, se construyó una firma de DC a partir de la literatura (*CD1C*, *THBD*, *CD209*, *CLEC10A*, *CLEC9A*, *FCER1A*, *ITGAX*, *ITGAM*) [66, 88-89]. Se calculó el ES de dicha firma y se determinó que había células originalmente etiquetadas como monocitos y macrófagos que, por su alto ES se trataba muy probablemente de DC en lugar de monocitos. Se cambió la identidad de 141 monocitos y 7 macrófagos a DC, debido a que presentaban un ES mayor a 0.1. También se renombraron 218 DC como monocitos debido a que tenían un ES menor a 0.1 de la firma DC y a que estas células formaban parte del grupo 5, el cual está conformado exclusivamente por monocitos. Solamente 12 células conservaron su identidad de DC. Estos cambios dieron como resultado 1354 macrófagos, 160 DC y 1741 monocitos en el tejido tumoral, mientras que la cantidad de células pertenecientes a tejido no neoplásico permanecieron sin cambio.

### **6.6.2 Enriquecimiento de módulos en TAM**

Una vez anotadas las células del TME con identidad de macrófagos, se usó el enriquecimiento de las firmas resultantes de WGCNA para etiquetar a cada TAM con su respectivo módulo. Se midió el enriquecimiento de tres grupos de genes denominados módulos turquesa (*Turquoise*), anti-turquesa (*Anti-turquoise*) y marrón (*Brown*). Estos módulos son grupos de genes altamente correlacionados entre sí y están conformados por 18 genes, 12 genes y 20 genes respectivamente. Las vías celulares representadas por cada módulo y el proceso mediante el cual se llegó a estos genes se describen en la sección 7.2. Los genes de cada módulo se encuentran especificados en la tabla 3 de la sección 7.5.

Se usaron dos límites diferentes de ES para etiquetar a los TAM con alguno de los módulos previamente mencionado.

- 1- Para realizar el mapa de calor de la figura 11, se decidió que si el ES de un determinado módulo para un TAM era mayor a 0.4, este se etiquetaba con dicho módulo, si un mismo TAM presentaba alto ES para más de un módulo, se etiquetaba con todos los módulos para los que obtuvo un alto ES. Si un TAM no



tenía un ES alto para ningún módulo, este no era etiquetado con ninguno de los módulos. Se eligió 0.4 como valor límite porque con este valor se lograba etiquetar a la mayoría de los TAM y se dejaba el menor número posible de TAM sin etiquetar. De esta forma se logró etiquetar el 92% de los TAM (n=1354), mientras que el 8% restante de las células no estaban enriquecidos ( $ES > 0.4$ ) con ninguna de las firmas turquesa, marrón o anti-turquesa.

- 2- Para poder llevar a cabo el análisis de expresión diferencial en conjunto con GSEA (*Gene Set Enrichment Analysis*) (explicado en sección 6.7), se necesitaba que los TAM estuvieran previamente etiquetados como anti-turquesa, turquesa, marrón o turquesa-marrón. Se decidió etiquetar a los TAM con el mismo procedimiento antes descrito, pero con un límite más estricto de ES (0.75). Con esta estrategia solamente se logró etiquetar al 26.4% de los TAM, resultando en 251 TAM anti-turquesa y 107 TAM etiquetados como turquesa, marrón o turquesa-marrón (107 es el resultado de sumar los TAM turquesa, marrón y turquesa-marrón). De esta forma, a pesar de que estos 358 TAM solamente representaban el 26.4%, se tenía la certeza de que eran los TAM más representativos de las firmas anti-turquesa, turquesa y marrón. Los 996 TAM restantes también estaban enriquecidos con alguna de las 3 firmas, pero en menor medida, por lo que no fueron incluidos en el análisis de expresión diferencial y GSEA.

### **6.6.3 Enriquecimiento de las firmas en pacientes del TCGA**

Se descargaron los datos de expresión de pacientes provenientes del TCGA (del inglés, *The Cancer Genome Atlas*) mediante la paquetería *TCGAbiolinks* (<https://bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html>) [90] y mediante las funciones `GDCquery()` y `GDCdownload()` (junio, 2020). Para la función `GDCquery()` se fijaron los siguientes parámetros: `project = "TCGA-BRCA"`, `data.category = "Transcriptome Profiling"`, `data.type = "Gene Expression Quantification"`, `workflow.type = "HTSeq - Counts"`, `legacy = FALSE` y las muestras a analizar fueron especificadas mediante el parámetro `barcode`. Se calculó el enriquecimiento de las firmas turquesa y anti-turquesa en 1106 pacientes de cáncer de mama. Si el valor de ES de una muestra, para una firma en específico formaba parte del cuartil superior de todos los valores de ES de las 1106 muestras, esta se etiquetaba como enriquecido con turquesa (*“high*

*turquoise*”) o enriquecido en anti-turquesa (“*high anti-turquoise*”), según fuera el caso. 79 pacientes fueron etiquetados como enriquecido con turquesa, 114 enriquecido con anti-turquesa y 66 como enriquecidos con turquesa/anti-turquesa (“*high turquoise/anti-turquoise*”) porque presentaban alto ES para ambas firmas. El resto de los 847 pacientes del TCGA no fueron considerados para el análisis bajo el criterio de selección elegido. No se aplicó el mismo criterio de selección para seleccionar las muestras con bajo o nulo enriquecimiento de las firmas turquesa y anti-turquesa (cuartil inferior de todos los valores de ES de las 1106 muestras).

## 6.7 Gene Set Enrichment Analysis

Previo al *Gene Set Enrichment Analysis* (GSEA), se realizó un análisis de expresión diferencial entre los 251 TAM etiquetados como anti-turquesa contra 107 TAM etiquetados ya sea como turquesa, marrón o turquesa-marrón. Se usó la función `DESeq()` de la paquetería `DESeq2` versión 1.30.0 (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>) [91]. GSEA requiere que se le proporcionen todos los valores de *fold change* (sin importar valores de p) [92] obtenidos como resultado de la función `DESeq()`, para posteriormente arrojar un valor de enriquecimiento con su propio valor de p (el cual no está relacionado a los valores individuales de *fold change* del análisis de expresión diferencial).

Un primer GSEA fue realizado para determinar el enriquecimiento y la significancia estadística perteneciente a las firmas de marcadores M1 y M2 en las poblaciones turquesa/marrón y anti-turquesa, con la paquetería `phenoTest` versión 1.38.0 (<https://bioconductor.org/packages/release/bioc/html/phenoTest.html>) mediante la función `gsea()`. Un segundo GSEA fue realizado para medir el enriquecimiento de genes pertenecientes a diversas vías celulares que pudieran estar representadas por los genes diferencialmente expresados entre estos dos grupos de TAM. Se utilizó la función `fgsea()` de la paquetería `fgsea` versión 1.16 (<https://bioconductor.org/packages/release/bioc/html/fgsea.html>). Para obtener las firmas correspondientes a cada vía celular, la paquetería recurre a la base de datos MSigDB (del inglés, Molecular Signatures Database), disponible de forma libre en (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>), la cual contiene diversas

colecciones de firmas génicas organizadas en 8 categorías y en donde cada firma está asociada a diferentes procesos, vías y poblaciones celulares. Se eligió la categoría *Hallmark gene set* (H), la cual consta de 50 firmas de genes, donde cada firma describe un proceso o vía celular diferente. El ES fue calculado para estas 50 firmas y sólo se consideraron las vías celulares con un NES (del inglés, *normalized enrichment score*) significativo ( $p\text{-value} < 0.05$ ).

## 6.8 Análisis de supervivencia

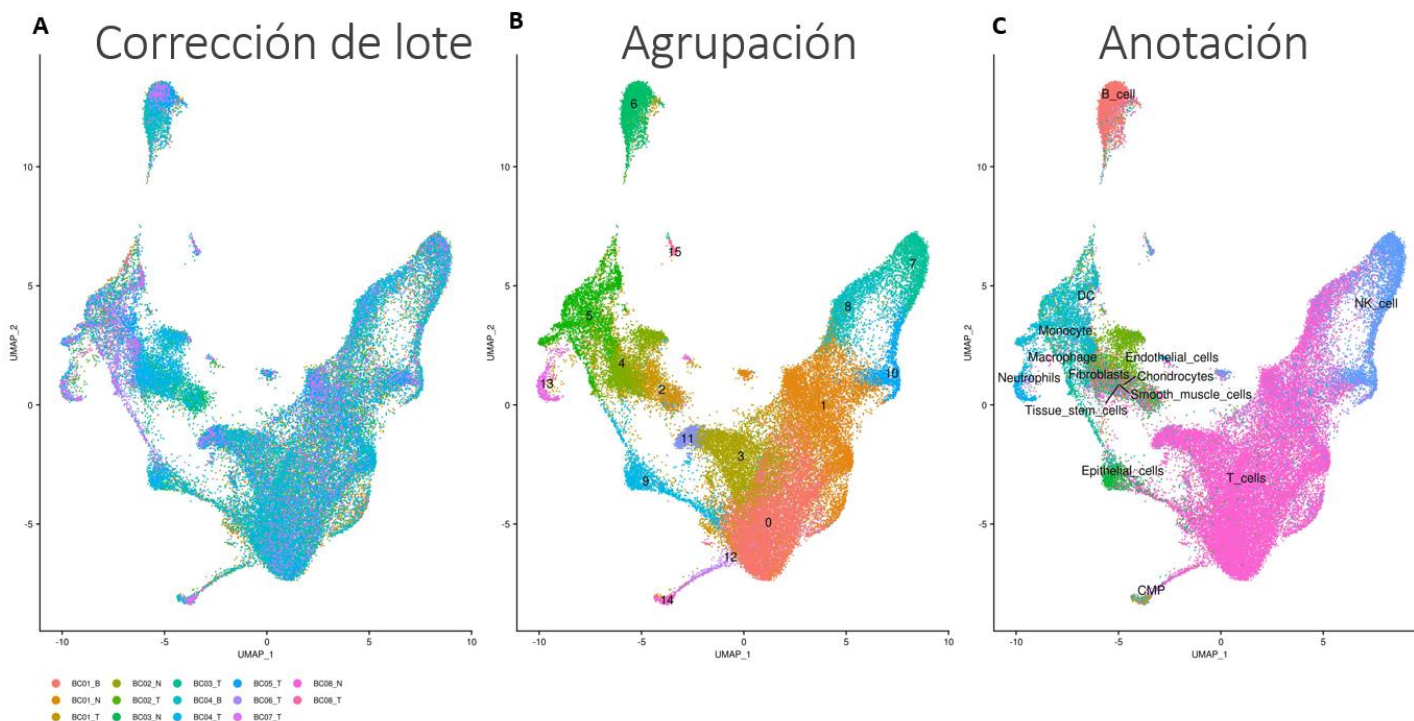
Se utilizaron las paqueterías *survival* versión 3.6 (<https://cran.r-project.org/web/packages/survival/index.html>) y *survminer* (<https://cran.r-project.org/web/packages/survminer/index.html>) [93] para comparar la supervivencia entre pacientes del TCGA con un alto enriquecimiento de la firma turquesa contra pacientes con alto enriquecimiento de la firma anti-turquesa (febrero, 2021). El cálculo del enriquecimiento fue determinado mediante GSVA.

## 7. RESULTADOS

### 7.1 Identificación del infiltrado leucocitario en cáncer de mama

Con el fin de poder explorar con mayor resolución la biología de los TAM, se decidió trabajar con datos scRNA-seq de pacientes con cáncer de mama, disponibles de forma pública. Se eligió realizar el análisis con el *set* de datos con más de 50,000 células (GSE114725) debido a que ofrece mayor robustez en análisis estadísticos. Previo a realizar el análisis sobre los datos scRNA-seq, se revisó la calidad de las células y se encontró que 48,142 células cumplían con los criterios de calidad (Revisar sección 6.2). Para explorar y poder visualizar las más de 48,000 células, se integraron los transcriptomas de los diferentes pacientes y sus respectivos tejidos de origen, además se corrigió el error de lote para evitar que las células resultaran agrupadas de acuerdo al paciente o al tipo de tejido del que provienen en lugar de ser agrupadas de acuerdo al tipo celular al que pertenecen. En la figura 4A-B se aprecia que las células no se agruparon de acuerdo al tejido o a los pacientes del que provienen, sino que más bien se agruparon de acuerdo al tipo celular al que pertenecen (Figura 4C), indicando que los

datos no cuentan con error de lote y que el transcriptoma de las células es de buena calidad. Se encontró que las células se agrupaban en 16 grupos etiquetados de 0 a 15 (Figura 4B); sin embargo, para conocer qué tipos celulares están representados por cada uno de los grupos encontrados, fue necesario realizar anotación celular usando la herramienta *SingleR* [85]. Esta estrategia de anotación permite asignar identidad a cada una de las células de forma *cluster*-independiente y sin necesidad de recurrir a grupos de marcadores canónicos para cada población. Esta herramienta asigna identidades celulares basándose en el nivel de correlación obtenida al comparar el transcriptoma de cada una de las células con los transcriptomas de referencia que tiene disponibles. Se identificaron un total de 14 tipos celulares diferentes (más 1 tipo celular desconocido) presentes en el infiltrado de cáncer de mama, tejido no neoplásico y en sangre, estos fueron células linfoides como linfocitos T (“*T\_cells*”), linfocitos B (“*B\_cell*”), células NK (“*NK\_cells*”), células de linaje mieloide como macrófagos (“*Macrophage*”), monocitos (“*Monocyte*”), DC, neutrófilos (“*Neutrophils*”), células de origen mesenquimal como fibroblastos (“*Fibroblast*”), condrocitos (“*Chondrocytes*”) y otras células residentes de tejido como células epiteliales (“*Epithelial\_cells*”), células endoteliales (“*Endothelial\_cell*”), células de músculo liso (“*Smooth\_muscle\_cells*”). Adicionalmente fueron encontradas células con características troncales del tejido (“*Tissue\_stem\_cells*”) y dos grupos de células anotadas putativamente como progenitor común mieloide (CMP, por sus siglas en inglés *common myeloid progenitor*) y desconocido (“*unknown*”) (Figura 4C). Los linfocitos T fueron la población más abundante y con mayor heterogeneidad al estar conformada por hasta 7 grupos de células. De manera esperada, los linfocitos T presentaron la mayor similitud con las células NK, las cuales están conformadas por 2 grupos. Los monocitos, DC, macrófagos y parte de los fibroblastos fueron los tipos celulares que mayor similitud presentaban entre sí debido a que se encontraron concentrados en los grupos 4 y 5. Células como condrocitos, fibroblastos, células endoteliales, epiteliales, de músculo liso y células troncales del tejido se encontraron en su gran mayoría en muestras de tejido no neoplásico contenidas en los grupos 2, parte del grupo 4 y parte del grupo 9. Los grupos 6, 13 y 14 están conformados cada uno exclusivamente por un único tipo celular, estos son linfocitos B, Neutrófilos y CMP, respectivamente. El grupo 15 contiene células anotadas como desconocidas.

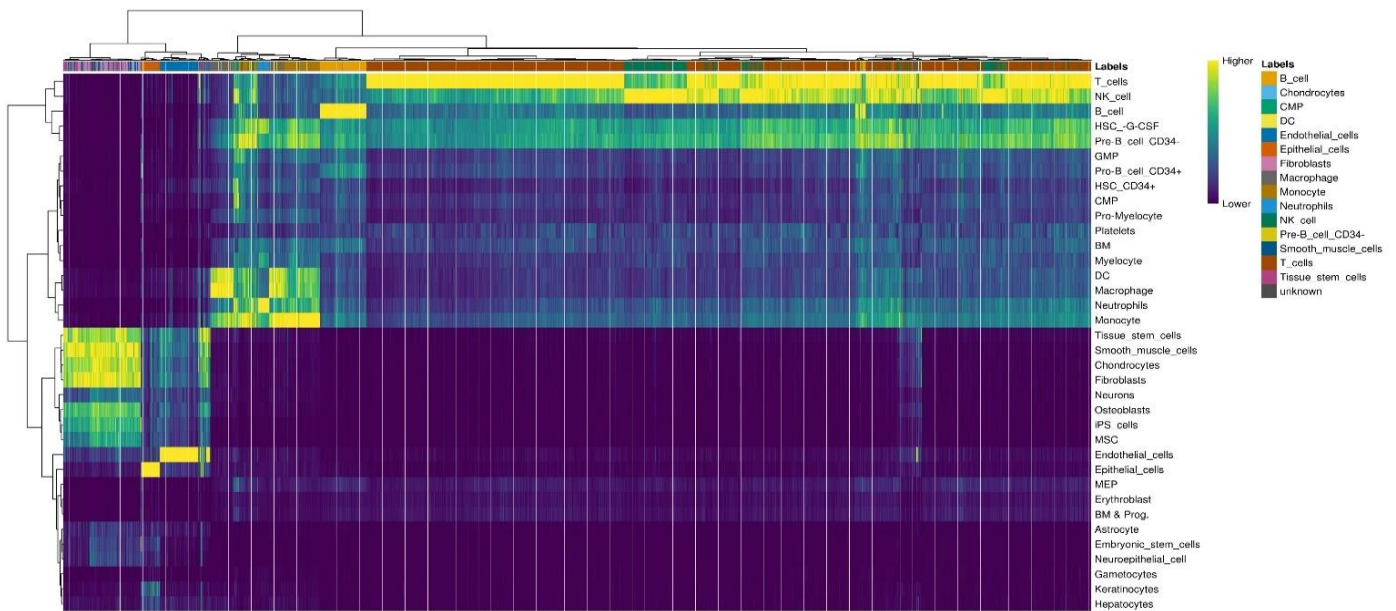


**Figura 4. Células presentes en el TME desplegadas en gráficos UMAP.**

Células provenientes del estudio con la clave GSE114725. **A.** Células etiquetadas por tejido y por paciente; **B.** Células etiquetadas por clusters; **C.** Células etiquetadas por tipo celular. La nomenclatura debajo del panel izquierdo indica si la muestra proviene de sangre (B), tejido no neoplásico (N) o tejido tumoral (T). Número total de células: 48142. Número de células por cada tipo celular: Linfocitos B=2493, Condrocitos=443, DC=334, Células endoteliales=1979, Células epiteliales=871, Fibroblastos=1474, Macrófagos=1541, Monocitos=2126, Neutrófilos=595, Células NK=5686, Células de músculo liso= 862, Linfocitos T=28201, Células tróncas de tejido=1120, CMP=166, desconocido=251. TME: microambiente tumoral. UMAP: *Uniform Manifold Approximation and Projection*.

Aunque la estrategia adoptada por *SingleR* permite obtener una asignación de identidades más robusta en comparación a otras estrategias como, por ejemplo, etiquetar manualmente cada grupo de células, basándose únicamente en la expresión de un puñado de marcadores canónicos para cada población celular, se decidió revisar la calidad de la correlación obtenida para cada identidad celular asignada en la figura 4C. El mapa de calor mostrado en la figura 5, indica que la correlación obtenida para células endoteliales, células epiteliales, neutrófilos, linfocitos B, linfocitos T y células NK es muy alta y mutuamente excluyente de los demás tipos celulares; por lo tanto, la asignación de identidad de estos tipos celulares fue robusta. Sin embargo, es importante mencionar que la mitad de los linfocitos T también tienen alta correlación con el transcriptoma de células NK y viceversa. Esto probablemente se debe a que, los linfocitos T que mejor se

distinguen de las células NK son linfocitos de muestras sanguíneas, mientras que los linfocitos que más similitud presentan con las células NK, pertenecen a tejido tumoral. Esto puede deberse a que ambos tipos celulares comparten programas transcripcionales citotóxicos en un microambiente tumoral. Células pertenecientes a muestras de tejido no neoplásico como fibroblastos, células de músculo liso, células tróncas de tejido y condrocitos también tuvieron una correlación alta y excluyente de otros tipos celulares, pero con valores de correlación muy similares entre sí, indicando que es difícil diferenciar una población de las otras debido a su mismo origen mesenquimal.



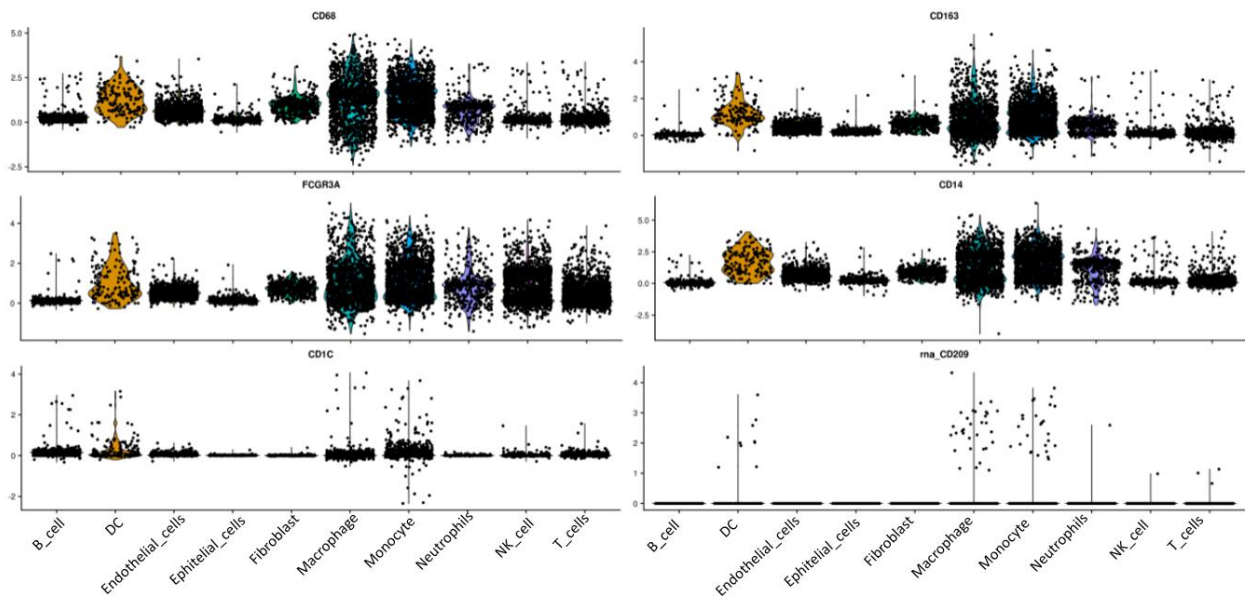
**Figura 5. Mapa de calor de correlaciones obtenidas entre el transcriptoma de las células y el microarreglo de referencia mediante SingleR.**

SingleR fue aplicado sobre el transcriptoma de células provenientes de tejido tumoral, sangre y tejido no neoplásico, pertenecientes a pacientes con cáncer de mama (GSE114725). La intensidad del color indica el nivel de correlación entre el transcriptoma de cada célula (columnas) con el microarreglo de referencia (filas). En las filas están desplegados los 36 transcriptomas de referencia que le fueron proporcionados a SingleR a través de celldex. Solamente 14 de 36 transcriptomas de referencia fueron identificados dentro del *dataset*.

El transcriptoma de células identificadas como monocitos provenientes de muestras sanguíneas, fueron las que mayores diferencias presentaron respecto a células identificadas como macrófagos y DC provenientes tanto de sangre como de tejido tumoral. No se identificó ni un solo macrófago en muestras sanguíneas, indicando que



*SingleR* asignó adecuadamente identidad a monocitos provenientes de sangre y no fueron confundidos con macrófagos provenientes de tejido no neoplásico o tumoral. Sin embargo, en el tejido tumoral el transcriptoma de macrófagos, monocitos y DC muestran alta similitud entre sí. Además, estas 3 poblaciones expresan indistintamente marcadores canónicos, comúnmente usados en la literatura para identificar: macrófagos (*CD68*, *CD163*), monocitos (*CD14*, *FCGR3A/CD16*) y DC (*CD1C*, *CD209/DC-SIGN*) (Figura 6), estos 2 últimos marcadores fueron detectados en menor medida, posiblemente porque las DC no son abundantes en el TME. Esto indica que macrófagos, monocitos y DC son tipos celulares que poseen gran plasticidad y que no es posible diferenciarlas usando únicamente unos pocos marcadores de forma individual.

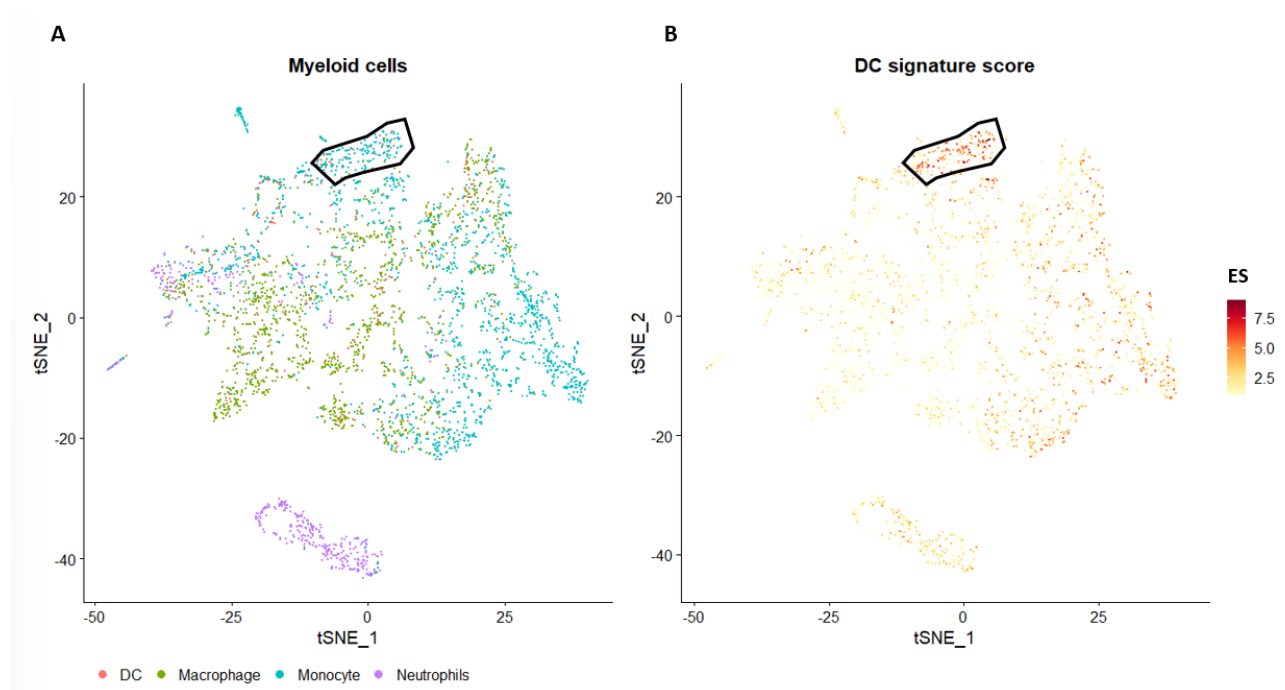


**Figura 6. Expresión de marcadores canónicos de macrófagos, monocitos y DC en diversos tipos celulares.**

Células provenientes de tejido tumoral de pacientes con cáncer de mama (GSE114725). Número total de cada tipo celular: Linfocitos B=982, DC=160, Células endoteliales=1645, Células epiteliales=503, Fibroblastos=482, Macrófagos=1354, Monocitos=1741, Neutrófilos=469, Células NKs=2818, Linfocitos T=11501.

Adicionalmente, células identificadas como DC también presentan alta correlación con los transcriptomas de referencia pertenecientes a macrófagos y monocitos, por lo tanto, existe la posibilidad de que el desempeño de *SingleR* para diferenciar DC de macrófagos o monocitos, no sea suficientemente bueno. Con el fin de corroborar la exactitud con la

que fueron identificados estos tipos celulares altamente similares en un contexto tumoral, se procedió a medir el enriquecimiento de una firma propia de DC, conformada por marcadores curados de la literatura (Revisar sección 6.6.1). El ES de la firma DC dio un valor bajo para la mayoría de las células identificadas originalmente como DC; sin embargo, un pequeño grupo de células identificadas originalmente como monocitos, obtuvieron el mayor valor de ES para esta firma y fueron reclasificadas como DC y viceversa (Figura 7A-B).



**Figura 7. Mapa de calor del enriquecimiento de la firma DC en células mieloides.**

**A.** Gráfico tSNE de macrófagos, DC, monocitos y neutrófilos encontrados en el TME de pacientes con cáncer de mama provenientes del estudio con la clave GSE114725. **B.** Mapa de calor del nivel de enriquecimiento de la firma DC usando el mismo gráfico tSNE de A. En la parte superior en B, se encuentra un área delineada de negro la cual señala a las células con el mayor enriquecimiento de la firma DC, el área delineada en A es la misma que en B y se puede apreciar que las células con mayor enriquecimiento de la firma DC son en su mayoría monocitos.

Inicialmente, SingleR identificó 180 macrófagos, 104 DC y 244 monocitos provenientes de tejido no neoplásico, así como 1361 macrófagos, 230 DC y 1664 monocitos en muestras de tejido tumoral. Basándose en los resultados del enriquecimiento de la firma DC, se decidió cambiar la identidad de 141 monocitos y 7 macrófagos a DC, 218 DC a monocitos, quedando como resultado final 1354 macrófagos, 160 DC y 1741 monocitos

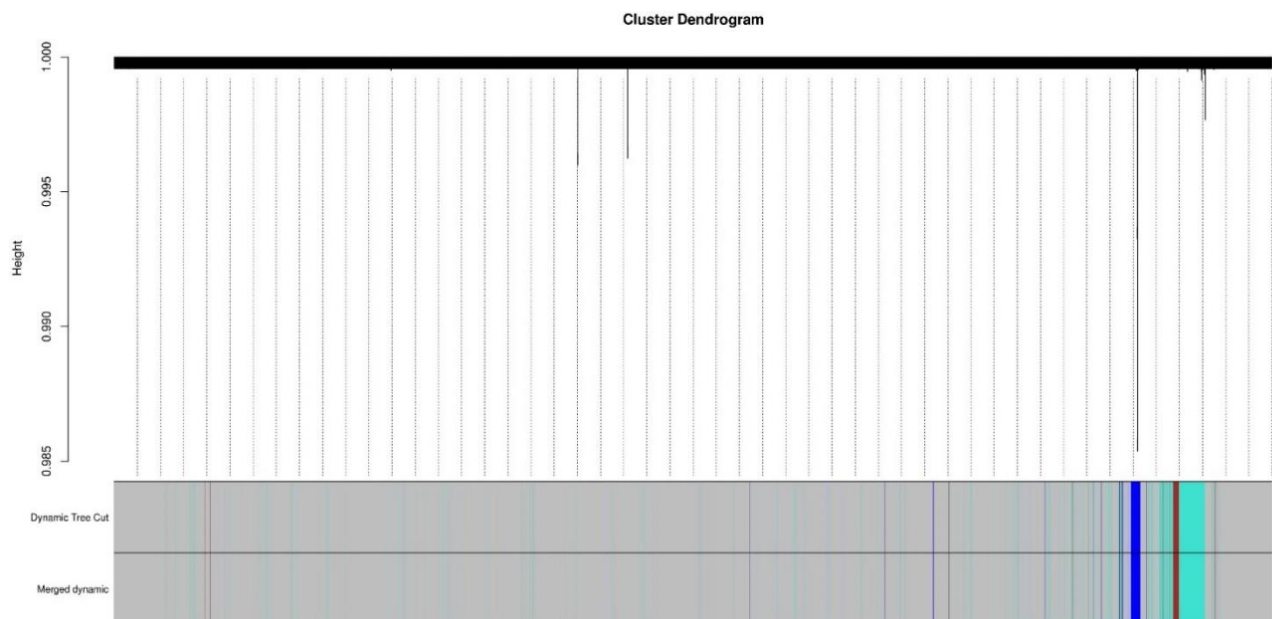


en tejido tumoral (Revisar sección 6.6.1). Las células provenientes de tejido no neoplásico permanecieron sin cambio.

## 7.2 Módulos de genes presentes en TAM en cáncer de mama

Una vez definidos los tipos celulares presentes en el TME, el análisis se enfocó en los macrófagos provenientes de tejido tumoral (TAM). Se contaba con un total de 1354 TAM, cada uno con 7247 genes que cumplían la condición de estar expresados en más del 1% de los TAM (Revisar sección 6.4).

Con el fin de encontrar patrones transcripcionales que agruparan a los genes de acuerdo a sus altos niveles de correlación y a las vías de señalización compartidas entre sí, se recurrió a WGCNA, con el cual se esperaba identificar grupos de genes que aporten información acerca de las funciones que están llevando a cabo los TAM. WGCNA se aplicó sobre 1354 TAM, cada uno con 7247 genes y se encontraron 3 grupos de genes denominados módulos turquesa (*turquoise*), módulo azul (*blue*) y módulo marrón (*brown*) (Figura 8), los cuales se conforman por 352, 85 y 49 genes respectivamente. Estos módulos transcripcionales potencialmente agrupan macrófagos de acuerdo a distintas funciones. El análisis también señaló que hay similitud entre los módulos turquesa y marrón pero no la suficiente (al menos 80% similares) como para agruparlos dentro de un solo módulo.

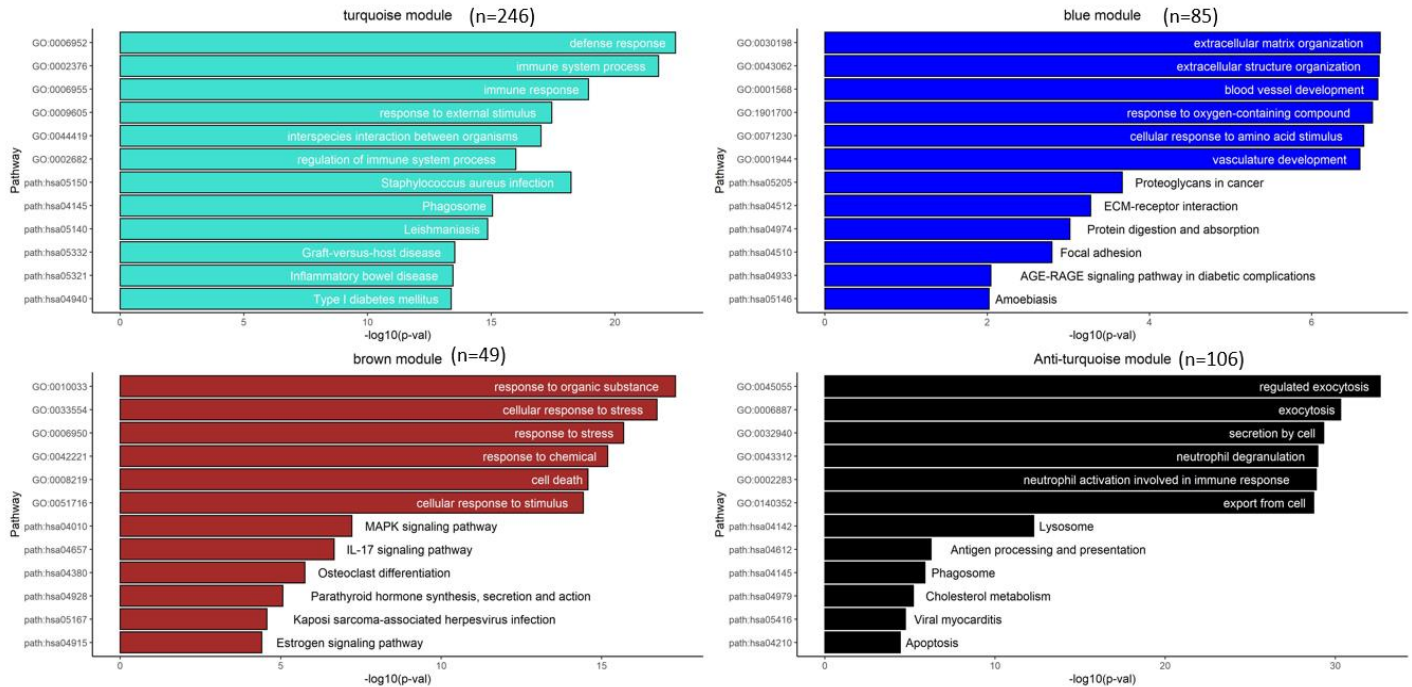


### Figura 8. Módulos encontrados por WGCNA.

WGCNA fue aplicado sobre la matriz de expresión de 1354 TAM. La sección *Dynamic Tree Cut* muestra el módulo (turquesa, azul o marrón) en el que fue agrupado cada gene. Si el gene es compartido indistintamente entre módulos o si no es suficientemente abundante entre todas las células, este no se agrupa dentro de ningún módulo y se representa de color gris. La sección *Merged dynamic* indica cuáles módulos fueron combinados por tener al menos 80% de similitud. En este caso, ningún módulo fue combinado dentro de otro módulo. WGCNA: *Weighted gene correlation network analysis*; TAM: macrófagos asociados a tumor.

Dentro del módulo turquesa había 106 genes que tenían un *gene membership* negativo (Revisar glosario de términos y abreviaturas), lo cual indica que en general, la expresión de estos genes tiene una correlación negativa con el resto de los 246 genes del módulo turquesa. Por esta razón se optó por separar los 106 genes con *gene membership* negativo y se agruparon en un cuarto módulo denominado anti-turquesa (*anti-turquoise*) dejando al módulo turquesa con 246 genes.

Para saber cuáles son las vías celulares más representadas por todos los genes de cada módulo encontrado en los TAM, se realizó un análisis de sobrerrepresentación de vías celulares mediante GO y KEGG (Figura 9). Se encontró que los genes del módulo turquesa tienen participación en vías relacionadas a la respuesta inmunológica, específicamente aquellos en los que participan moléculas HLA clase II y procesos citotóxicos, como el rechazo a trasplante. El módulo azul participa en vías relacionadas a la organización de la MEC, así como en la generación de vasos sanguíneos. Los genes marrón participan en respuesta al estrés y en procesos de muerte celular, así como vías de señalización de proteínas cinasas mitógeno-activadas (MAPK por sus siglas en inglés *mitogen-activated protein kinase*). Por último, el módulo anti-turquesa representa procesos de secreción y exportación extracelular.



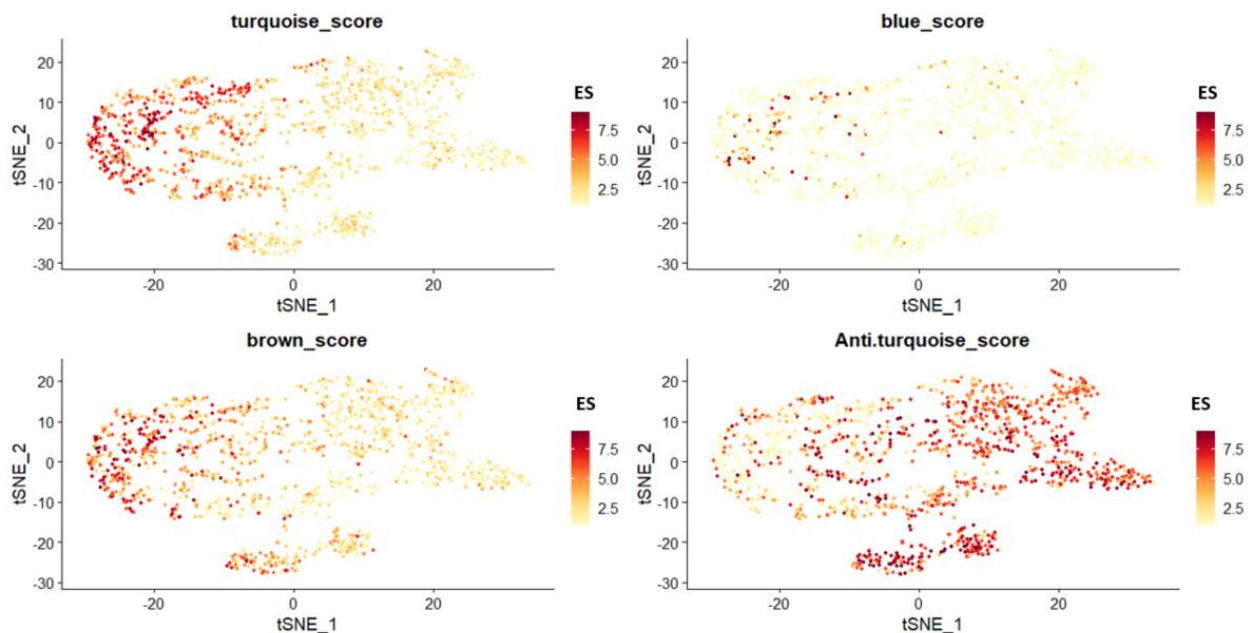
**Figura 9. Análisis de procesos celulares representados por los genes pertenecientes a cada módulo.**

Se muestran los 6 procesos con mayor significancia estadística de GO, seguidos de los 6 procesos con mayor significancia estadística de KEGG. En la parte superior se indica la totalidad del número de genes que conforma cada módulo, los cuales se usaron para realizar el análisis con GO y KEGG. GO (*Gene Ontology*); KEGG (*Kyoto Encyclopedia of Genes and Genomes*).

Con el objetivo de obtener una lista de genes más robusta y reducida para cada módulo, se optó por conservar solamente aquellos genes con un valor de *gene membership* mayor o igual a 0.4 (menor o igual a -0.4 para genes con *gene membership* negativo) para su respectivo módulo. Después de llevar a cabo este filtro se conservaron 18 genes turquesa, 15 azul, 20 marrón y 12 anti-turquesa (especificados en sección 7.5), los cuales se trataron como firmas para identificar a cada módulo debido a que están representado a los genes con mayor relevancia para cada módulo.

Cada uno de los 4 módulos representa vías celulares distintas que están presentes en los TAM; sin embargo, aún hay que determinar si estos 4 módulos pueden coexistir dentro de un mismo TAM o si son excluyentes entre sí. Previo a la bipartición del módulo turquesa original (352 genes) en turquesa (246 genes) y anti-turquesa (106 genes), WGCNA indicaba que el módulo marrón guardaba similitud con el módulo turquesa original (Figura 8), dicha similitud no fue suficiente para que ambos módulos fueran

considerados como uno mismo, esto quiere decir que el módulo marrón solamente guarda similitud con turquesa o anti-turquesa, adicionalmente, el módulo azul parecía no tener similitud con el resto de los módulos. Para determinar cómo se comparten o excluyen estos módulos en los TAM, se procedió a medir el enriquecimiento de las 4 firmas de cada módulo en los 1354 TAM mediante GSVA (Figura 10). Se observó que hay TAM que comparten un ES alto de las firmas turquesa y marrón al mismo tiempo que tienen un ES bajo de la firma anti-turquesa, de igual manera, los TAM con un ES alto de la firma anti-turquesa tienen bajo ES tanto de la firma turquesa como de la firma marrón. Esto indica que las firmas turquesa y marrón coexisten dentro de un mismo TAM al mismo tiempo que ambas son excluyentes de la firma anti-turquesa. El ES para la firma azul fue alto solamente en un número muy reducido de células. Esto sugiere que el programa transcripcional de los TAM tiene un comportamiento principalmente dicotómico.



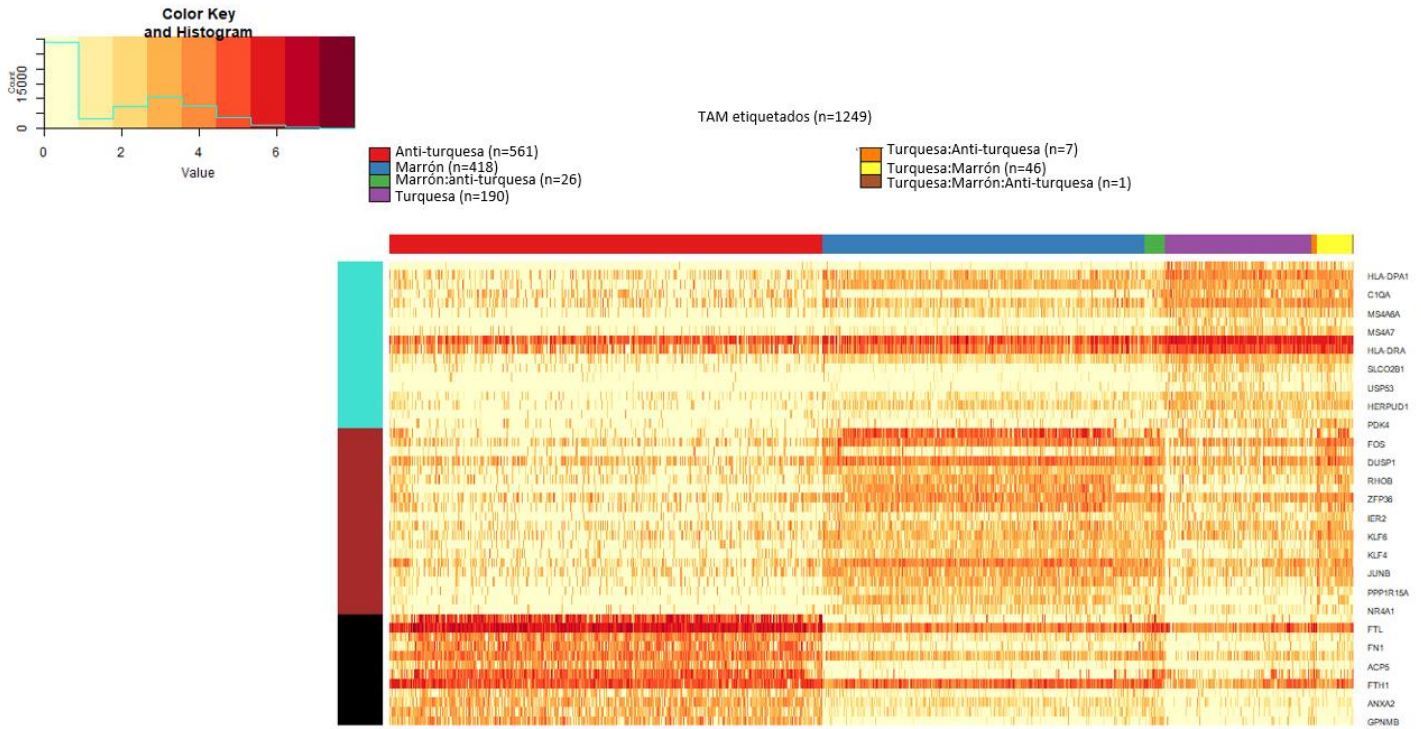
**Figura 10. Mapas de calor del enriquecimiento de las 4 firmas encontradas con WGCNA en TAM.**

Se despliega en 2 dimensiones de la estructura de los datos transcriptómicos de 1354 células anotadas como TAM. t-SNE\_1 y t-SNE\_2 representan gradientes de similitud entre cada célula, sin embargo, no son interpretables como medida cuantitativa. TAM: macrófagos asociados a tumor.

Debido al escaso número de células enriquecidas con la firma azul y a que posteriormente en este análisis (Revisar sección 7.5) se notó que la firma azul está más enriquecida en

fibroblasto que en TAM, se decidió continuar el análisis solamente con las firmas turquesa, marrón y anti-turquesa.

Los TAM fueron etiquetados con alguno de los 3 módulos analizados según sus respectivos valores de ES obtenidos para las firmas de cada módulo. Con la estrategia utilizada (Revisar sección 6.6.2), se logró etiquetar a 1249 TAM (92%) con al menos uno de los 3 módulos, la mayoría de los TAM etiquetados tuvieron alto ES de la firma anti-turquesa (n=561), marrón (n=418), turquesa (n=190) y 46 con alto ES para las firmas turquesa y marrón. De forma inesperada, hubo 26 TAM etiquetados como “marrón:anti-turquesa”, adicionalmente hubo un número muy reducido de TAM etiquetados como “turquesa:anti-turquesa” (n=7) y sólo 1 macrófago etiquetado con los 3 módulos. En conjunto, los TAM que no se explican mediante el comportamiento dicotómico previamente observado, representan el 2.5% del total de los TAM (Figura 11). El patrón de expresión de la firma de cada módulo varía de tal forma que permite distinguir principalmente 2 subgrupos de TAM. Los TAM etiquetados como anti-turquesa, expresan de manera exclusiva la firma anti-turquesa, sin embargo, los TAM etiquetados como turquesa, marrón y “turquesa:marrón” comparten alta expresión tanto de las firmas turquesa como marrón. En general se repite el comportamiento dicotómico observado en figura 10. Las observaciones previas sugieren que las firmas turquesa y marrón convergen en los mismos TAM a pesar ser firmas diferentes, al mismo tiempo, la firma anti-turquesa permanece excluyente de ambas, por lo tanto, parece más adecuado englobar a los TAM turquesa, marrón y “turquesa:marrón” en una misma población de TAM denominada turquesa/marrón. Con lo anterior se concluye que podrían existir principalmente 2 poblaciones de TAM en el TME de cáncer de mama.



**Figura 11. Patrón de expresión de las firmas turquesa, marrón y anti-turquesa.**

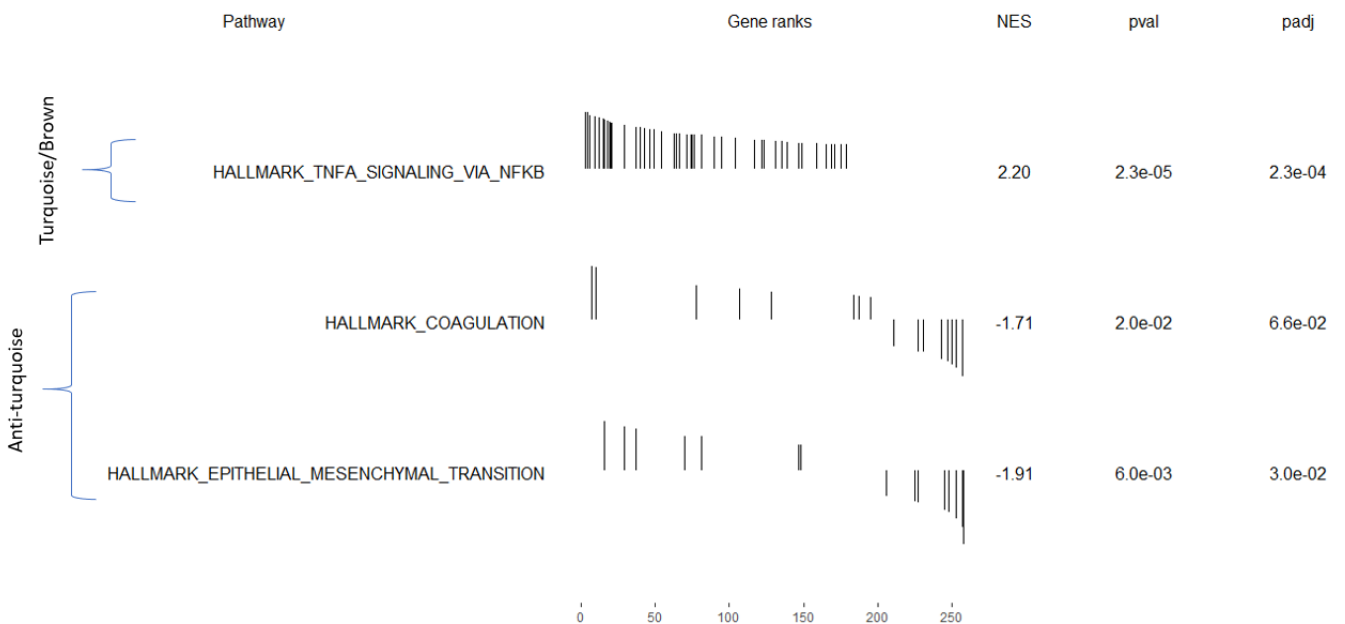
Las firmas turquesa, marrón y anti-turquesa constan de 18, 20 y 12 respectivamente. Cada columna representa uno de los 1249 TAM a los que se les asignó su respectivo módulo, las barras horizontales de colores y la leyenda encima del mapa de calor, indican el módulo con el que fue etiquetado cada TAM, los TAM etiquetados con más de un módulo tienen “:” como separador. Las barras verticales de colores del lado izquierdo indican el módulo al que pertenece cada gene, siendo el color negro el color asignado para señalar los genes anti-turquesa.

### 7.3 Diferencias funcionales entre las poblaciones de TAM

Para determinar si los TAM turquesa/marrón y los TAM anti-turquesa, además de presentar diferencias en la expresión de sus respectivas firmas, también presentan diferencias en las funciones celulares que llevan a cabo en el TME, se realizó un análisis de expresión diferencial junto con GSEA (del inglés, *Gene Set Enrichment Analysis*) comparando TAM turquesa/marrón contra TAM anti-turquesa. Para realizar la comparación pertinente entre TAM turquesa/marrón y TAM anti-turquesa, se decidió considerar la propuesta de que los macrófagos tienen un espectro de polarización en donde los fenotipos sólo son perfectamente distinguibles en los polos del espectro [69]. De forma análoga se decidió comparar los TAM situados en los polos de valores de ES de las firmas turquesa/marrón y anti-turquesa; para esto se escogieron parámetros de ES más estrictos para trabajar solamente con los TAM con el mayor enriquecimiento de sus



respectivas firmas (Revisar sección 6.6.2). Con la estrategia adoptada se etiquetaron 251 TAM anti-turquesa, 51 turquesa, 44 marrón y 12 TAM con alto ES (0.75) de las firmas turquesa y marrón. Esta estrategia redujo a 26.4% la cantidad de TAM analizados, pero solo así se asegura que se están comparando TAM con transcriptomas diametralmente opuestos. Había 264 genes diferencialmente expresados entre las dos poblaciones de TAM; dichos genes indican que cada población tiene activadas diferentes vías celulares (Figura 12). Los TAM turquesa/marrón tienen activadas funciones celulares asociadas a respuestas anti-tumorales como la vía de TNF $\alpha$ , mientras que los TAM anti-turquesa tienen enriquecidas vías de coagulación y transición epitelio mesénquima (EMT por sus siglas en inglés *epithelial to mesenchymal transition*) en las cuales participan genes con capacidad para remodelar la MEC. Con estos resultados se puede concluir que las diferencias entre los TAM turquesa/marrón y anti-turquesa, no están limitadas a la expresión de los genes pertenecientes a sus respectivas firmas, si no que en general sus transcriptomas y sus funciones celulares también reflejan un comportamiento dicotómico.



**Figura 12. Vías celulares enriquecidas según GSEA.**

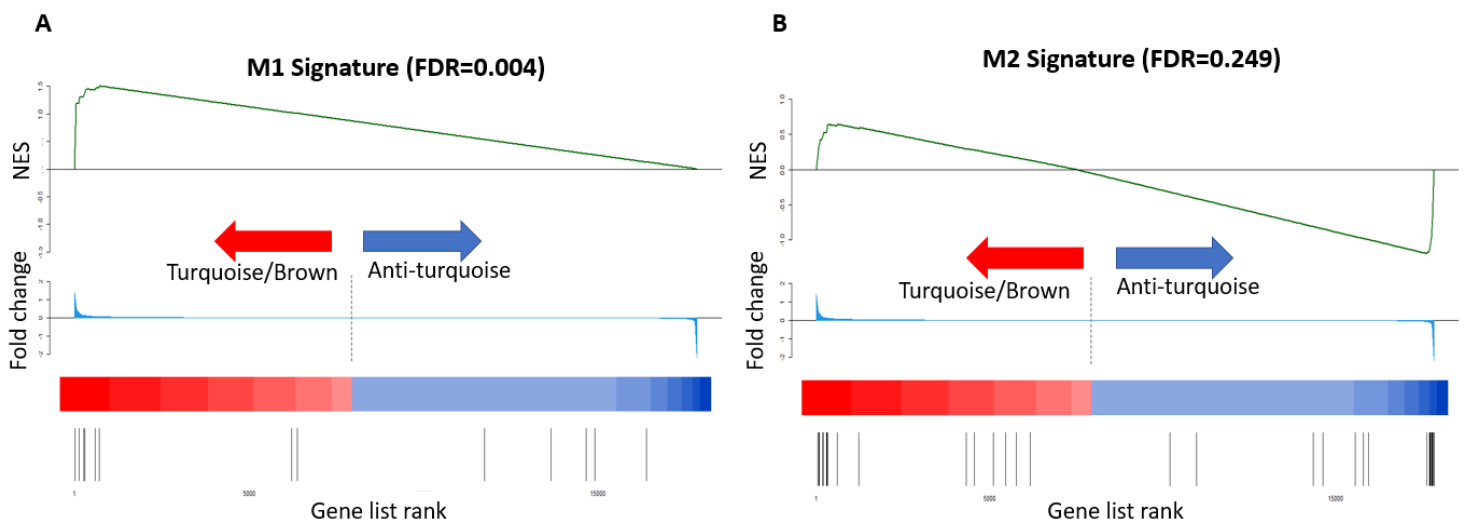
Se compararon TAM “turquoise/marrón” (NES positivo) contra TAM anti-turquesa (NES negativo) usando los 264 genes diferencialmente expresados ( $p\text{-value} < 0.01$ ) entre estos grupos. NES: *Normalized enrichment score*.

## 7.4 Relación del modelo M1/M2 con las poblaciones de TAM encontradas

Con el fin de explorar la posibilidad de que el comportamiento dicotómico observado en los TAM pudiera correlacionar con el modelo M1/M2, se usó GSEA para determinar si los marcadores M1/M2 se encontraban estadísticamente representados dentro de los módulos TAM turquesa/marrón o TAM anti-turquesa.

Para generar la firma de marcadores M1/M2, se tomaron los genes comúnmente aceptados como marcadores M1/M2 (Tabla 1), además se añadieron las variantes de algunos marcadores como *ARG2* (variante de arginasa-1), *TGFB2* y *TGFB3* (variantes de *TGFβ*), *VEGFA* y *VEGFB* (variantes de *VEGF*), *CTSA* y *CTSD* (ejemplos adicionales de catepsinas). Adicionalmente, se escogieron *HLA-DQA1*, *HLA-DPB1* y *HLA-DRA* como genes representativos de HLA-II y fueron añadidos a la lista de marcadores M1. Posterior a la adición de dichos genes se obtuvo una lista de 19 marcadores M1, sin embargo, *IL12B* y *NOS2* no se encontraban disponibles en este set de datos, por lo que la firma M1 se redujo a 17 genes, mientras que la firma M2 resultó con 37 genes.

Según GSEA, la firma transcripcional M1 alcanzó un enriquecimiento estadísticamente significativo (FDR=0.004) en los TAM turquesa/marrón, mientras que los marcadores pertenecientes a la firma transcripcional M2 se encontró enriquecida indistintamente tanto en TAM turquesa/marrón como en anti-turquesa (FDR=0.249) (Figura 13A-B).





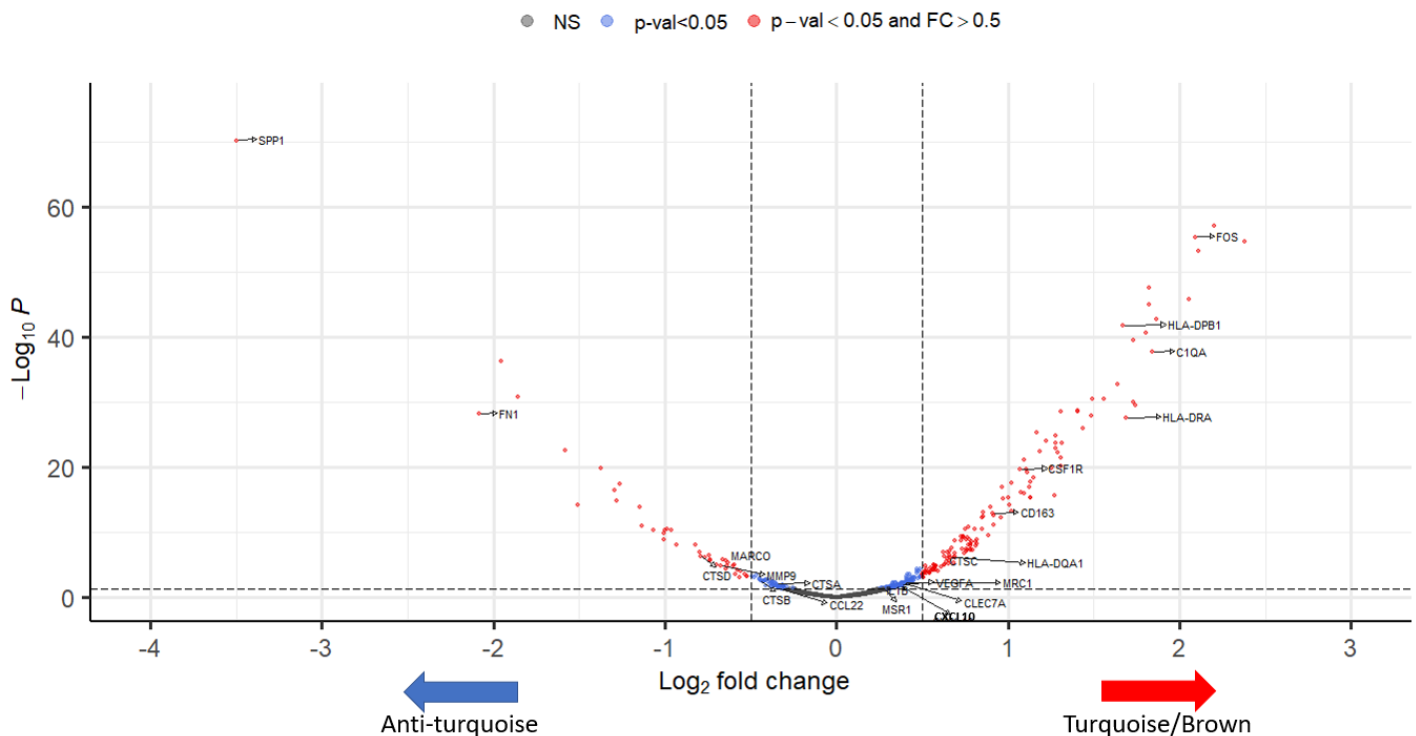
### Figura 13. GSEA de firma M1 y firma M2.

A. enriquecimiento de firma M1; B. enriquecimiento de firma M2. NES: *Normalized enrichment score*; GSEA: *Gene Set Enrichment Analysis*; FDR: *False Discovery Rate*

El análisis de expresión diferencial realizado en conjunto con GSEA reveló que el enriquecimiento de la firma M1 observado en los TAM turquesa/marrón se debió principalmente a genes HLA-II como *HLA-DRA*, *HLA-DPB1*, *HLA-DQA1* y en menor medida por *IL-1 $\beta$*  y *CXCL10* (*fold change*=1.29; *p-value*=0.017 para ambos) (Figura 14). Ningún otro marcador M1 se expresaba diferencialmente en ninguno de las poblaciones. Los TAM anti-turquesa expresan diferencialmente marcadores M2 como *CCL22*, *CTSA*, *CTSB*, *MMP9*, *MARCO*, *CTSD* y *FN1* (*fold change* < -1.25; *p-value*<0.05), pero el enriquecimiento de la firma M2 no fue significativo porque los TAM turquesa/marrón también estaban enriquecidos significativamente con algunos marcadores M2 como *CLEC7A*, *MRC1*, *CTSC*, *CD163*, *CSF1R* y *VEGFA* (*fold change* > 1.3; *p-value* < 0.05). En general los genes pertenecientes a la firma M2 se encontraron enriquecidos indistintamente tanto en TAM turquesa/marrón como en anti-turquesa, confirmando lo observado en la figura 13B.

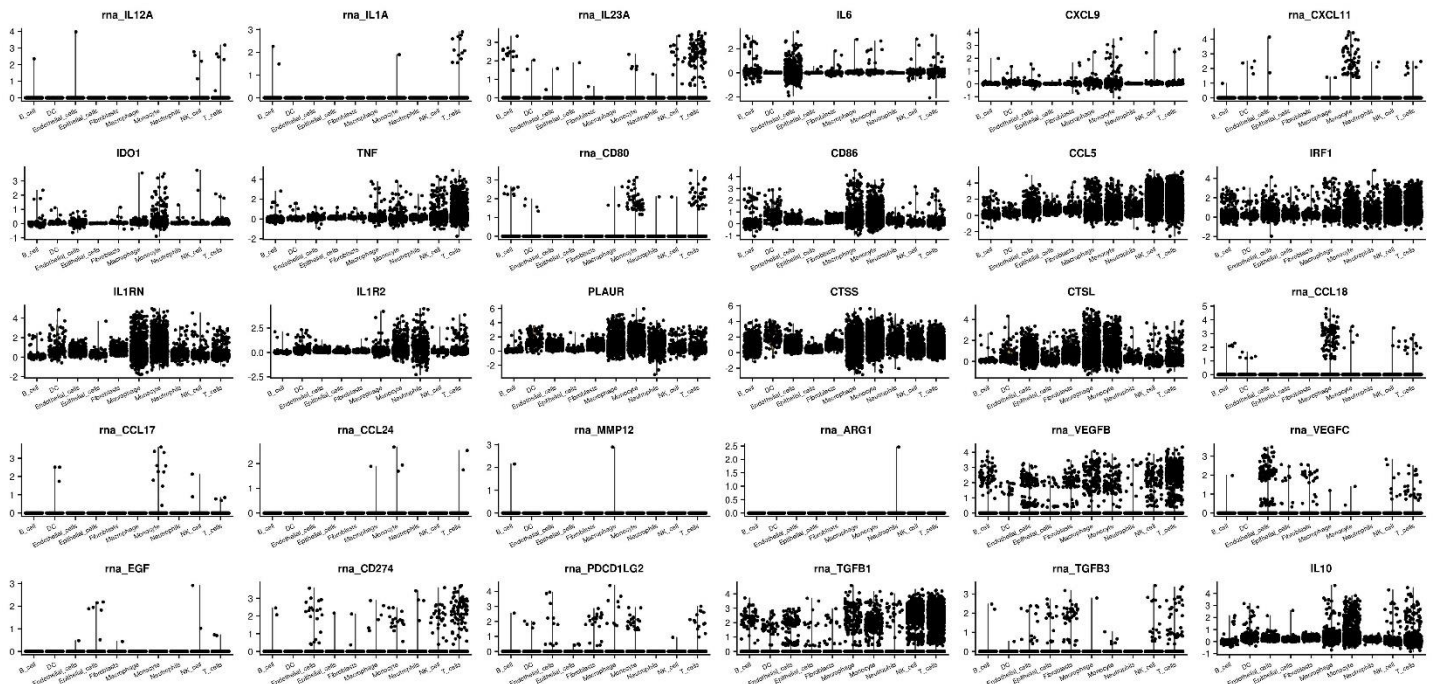
De los 17 marcadores M1 más los 37 marcadores M2 analizados, se observó que 36 (66.7%) marcadores M1/M2 no alcanzaron valores estadísticamente significativos según el análisis de expresión diferencial realizado entre los TAM turquesa/marrón y anti-turquesa. Se eligieron 30 marcadores M1/M2 para averiguar la razón por la que no se encontraron diferencialmente expresados en ninguno de las poblaciones TAM. Se decidió explorar la expresión de estos 30 marcadores en otros 9 tipos celulares presentes en el TME para determinar si la expresión de estos marcadores se comportaba igual en otros tipos celulares o si se trataba de una cuestión exclusiva de los TAM. Se encontró que 17 de los 30 marcadores M1/M2 elegidos, no se expresaban diferencialmente debido a que su expresión era nula en macrófagos o en su defecto, se expresan en muy pocas de estas células (Figura 15), entre estos se encuentran los marcadores M1: *IL12A*, *IL1A*, *IL23*, *IL6*, *CXCL9*, *CXCL11*, *IDO1* y *CD80*; marcadores M2: *CCL17*, *CCL24*, *MMP12*, *ARG1*, *VEGFC*, *EGF*, *CD274*, *PDCD1LG2* y *TGFB3*. De los 17 marcadores anteriores, no se detectó expresión de 7 marcadores (*IL12A*, *IL1A*, *CCL17*, *CCL24*, *MMP12*, *ARG1* y *EGF*) en ningún tipo celular, indicando que existen dificultades técnicas para detectar

transcritos de algunos genes que posiblemente, no son los suficientemente abundantes para ser detectados con tecnologías de alto desempeño, pero de baja profundidad como scRNA-seq. Los 13 marcadores restantes (*TNF*, *CD86*, *CCL5*, *IRF1*, *IL1RN*, *IL1R2*, *PLAUR*, *CTSS*, *CTSL*, *CCL18*, *VEGFB*, *TGFB1* e *IL10*) sí presentan expresión tanto en macrófagos como en otros tipos celulares. Esto indica que estos 13 marcadores propuestos como M1 o M2 no se asocian a los TAM turquesa/marrón o anti-turquesa que tienen el mayor enriquecimiento de estas firmas y no pueden ser usados para identificarlos.



**Figura 14. Análisis de expresión diferencial entre las poblaciones TAM turquesa/marrón y TAM anti-turquesa.**

Se encuentran indicados los marcadores M1 (*HLA-DRA*, *HLA-DPB1*, *HLA-DQA1*, *CXCL10* e *IL1B*) y M2 (*MSR1*, *CLEC7A*, *VEGFA*, *MRC1*, *CTSC*, *CD163*, *CSF1R*, *CTSA*, *CTSB*, *MMP9*, *MARCO*, *CTSD* y *FN1*) que obtuvieron un  $FC > 0.5$  o un  $p\text{-val} < 0.05$ . También se resalta el gene más representativo de cada módulo (*SPP1*, *C1QA* y *FOS*), seleccionados de acuerdo con sus altos valores de *gene membership* para los módulos anti-turquesa, turquesa y marrón respectivamente. Nótese que marcadores M2 como *VEGFA* y *CTSC* están enriquecidos significativamente en los TAM turquesa/marrón.  $FC$   $\log_2 \text{ fold change}$ ; NS: No significativo.



**Figura 15. Expresión de marcadores M1/M2 en diferentes tipos celulares provenientes de tejido tumoral.**

Firma M1: *IL12A*, *IL1A*, *IL23A*, *IL6*, *CXCL9*, *CXCL11*, *IDO1*, *TNF*, *CD80*, *CD86*, *CCL5* e *IRF1*. Firma M2: *IL1RN*, *IL1R2*, *PLAUR*, *CTSS*, *CTSL*, *CCL18*, *CCL17*, *CCL24*, *MMP12*, *ARG1*, *VEGFB*, *VEGFC*, *EGF*, *CD274*, *PDCD1LG2*, *TGFB1*, *TGFB3* e *IL10*. Número total de cada tipo celular: Linfocitos B=982, DC=160, Células endoteliales=1645, Células epiteliales=503, Fibroblastos=482, Macrófagos=1354, Monocitos=1741, Neutrófilos=469, Células NK=2818 y Linfocitos T=11501.

## 7.5 Expresión de los módulos en otras poblaciones celulares

Como se mencionó antes, se filtraron los genes de cada módulo de acuerdo a su valor de *gene membership*. De esta forma las firmas turquesa, marrón y anti-turquesa quedaron con 18, 20 y 12 genes respectivamente. Adicional a estas firmas, se aprovechó el parámetro “*Leading Edge*” arrojado como parte de los resultados de GSEA para definir los genes más importantes de las vías enriquecidas TNF $\alpha$ , coagulación y EMT, de esta forma, se definió una firma para cada vía, las cuales se presentan como firmas funcionales asociadas a las firmas que identifican cada módulo de WGCNA (Tabla 3).

**Tabla 3. Propuesta de firmas transcripcionales para identificar módulos y firmas funcionales de TAM presentes en cáncer de mama.**

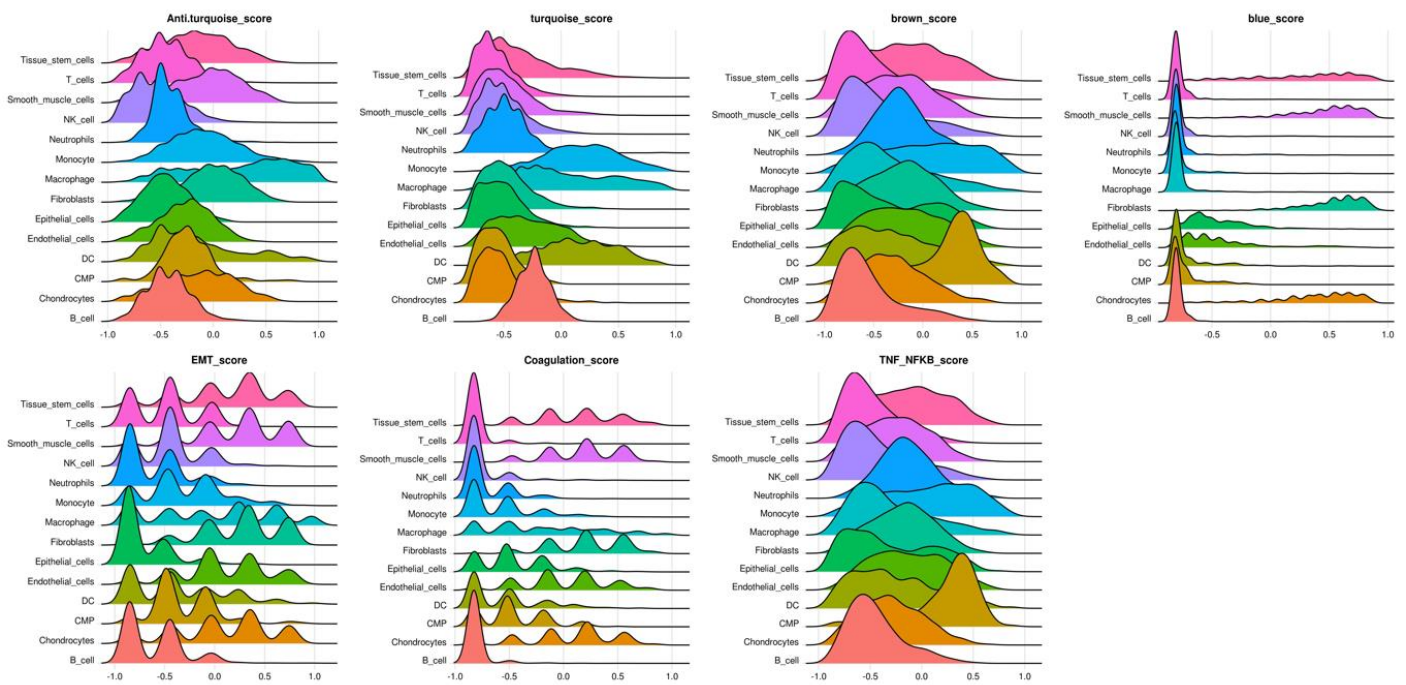
	Anti-turquesa	Turquesa	Marrón
<b>Firmas para identificar módulos</b>			
		C3 HLA-DPA1 CEBPD	HSPA1A FOS EGR1
		C1QA HLA-DPB1 MS4A6A	DUSP1 MCL1 RHOB
SPP1 FTL CSTB FN1 VIM		CX3CR1 MS4A7 CD74	DNAJB1 ZFP36 JUN IER2
ACP5 APOE FTH1 CD52		HLA-DRA ITM2B SLC02B1	MAFB KLF6 FOSB KLF4
ANXA2 LGALS3 GPNMB		GPR34 USP53 CSF1R	SRGN JUNB KLF2
		HERPUD1 SORL1 PDK4	PPP1R15A HSPA1B
			NR4A1
<b>Firmas funcionales asociadas</b>			
Transición epitelio	SPP1 VIM FLNA TIMP3		/
mesénquima	FN1		/
Coagulación	GSN CD9 TIMP3 SPARC		/
	FN1 APOC1 MMP9		
Vía de TNF $\alpha$	/	KLF6 PPP1R15A GADD45B NFKBIA SGK1 DUSP1 EGR1	
		NR4A1 FOSB KLF2 ZFP36 SAT1 BTG1 KLF4 IER3 MCL1	
		RHOB BTG2 IER2 IER5 FOS JUNB EIF1 JUN CEBPD	
		MARCKS	

**Nota:** Los genes están ordenados de mayor a menor valor de *gene membership* para el caso de las firmas que identifican módulos. Las firmas funcionales están conformadas por el “Leading Edge” de cada vía celular enriquecida según GSEA.

La firma azul (conformada por *C1R DCN APOD MGP COL1A1 CTGF FBLN1 COL1A2 COL3A1 CCDC80 HBB COL6A3 STC2 PDGFRA HSPG2*) no fue considerada para los análisis de expresión diferencial y GSEA porque estaba enriquecida en un número muy reducido de células (n=28) (Figura 10), pero fue rescatada para el siguiente análisis.

Mediante GSVA se calculó el ES de las 4 firmas que identifican cada módulo, así como sus vías celulares asociadas en las más de 48000 células de nuestro set de datos, esto con el fin de saber si estas firmas son exclusivas de TAM o se comparten con otras poblaciones celulares. En la figura 16 se observa que, tomando como referencia un ES mayor a 0.5, la firma anti-turquesa es la única que está enriquecida casi exclusivamente en los TAM. La firma turquesa está enriquecida en TAM, monocitos y en DC, pero no es una firma común de poblaciones mieloides ya que no está compartida con neutrófilos. Se aprecia que la firma marrón es una firma general de activación y por lo tanto es la menos específica para distinguir un TAM, de hecho, el enriquecimiento de esta firma es mayor

en otros tipos celulares como monocitos y CMP. Por último, la firma azul que estaba presente solamente en 28 células etiquetadas como TAM, está exclusivamente enriquecida en 4 tipos celulares con transcriptomas y funciones muy similares, como fibroblastos, células de músculo liso, células troncales de tejido y condrocitos, tal como se había observado en la figura 5. El ES de las firmas funcionales coagulación y EMT, ambas asociadas a la firma específica anti-turquesa, es disperso y no parece ser específico de un tipo celular en particular, indicando que, aunque estos procesos son característicos de un TAM anti-turquesa, no son exclusivos de estos. A pesar de que la firma funcional TNF $\alpha$  es característica tanto de los TAM turquesa como de los TAM marrón, la distribución de su enriquecimiento, es exactamente igual al de la firma marrón, por lo que se concluye que, aunque las firmas marrón y TNF $\alpha$  son características de una población particular de TAM, estas firmas también están activas en varios tipos celulares.

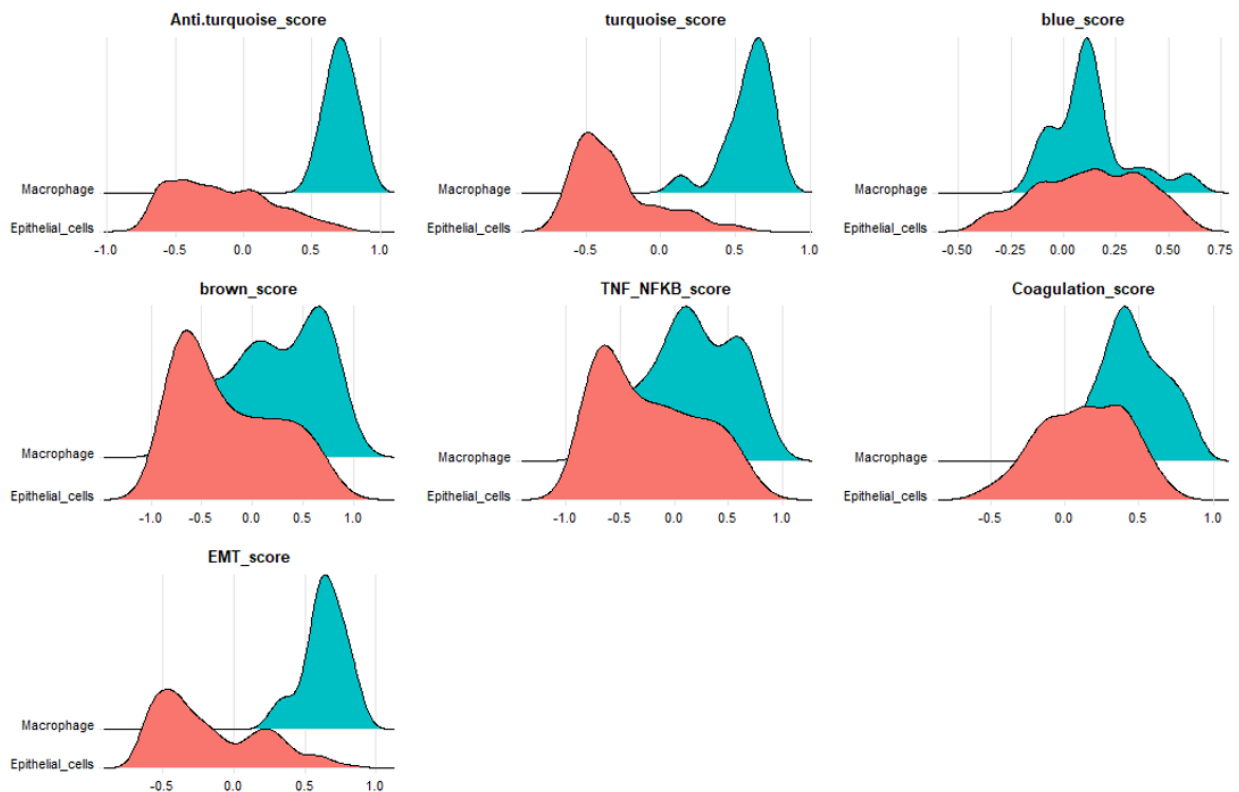


**Figura 16. Distribución del ES de las firmas funcionales y de cada módulo en diferentes tipos celulares.**

Se representa el ES de 48142 células pertenecientes a distintos tipos celulares de infiltrado tumoral, tejido no neoplásico y sangre. En el eje X están representados los valores de enriquecimiento obtenidos con GSVA, los cuales están dentro de un rango de -1 a 1. Células provenientes del estudio con la clave GSE114725.

Para explorar la posibilidad de que estas firmas pudieran también expresarse en células tumorales, se recurrió al *set* de datos con el ID GSE75688 (Tabla 2), el cual cuenta con

un total de 549 células que incluye tanto células tumorales como infiltrado leucocitario. Se calculó el enriquecimiento de estas firmas en 270 células tumorales y en 17 TAM (Figura 17) y se encontró que las firmas turquesa y anti-turquesa mantuvieron su especificidad en los TAM, al mismo tiempo que se observa que hay células tumorales que tienen valores de ES tan altos como la mayoría de las células TAM para las firmas marrón y  $TNF\alpha$ . Se observó el mismo comportamiento para las firmas de coagulación y EMT, aunque este fue menor. Estos resultados sugieren que la firma anti-turquesa es la única que puede considerarse exclusiva de TAM, mientras que la firma turquesa es una firma compartida por TAM, monocitos y DC, las cuales son poblaciones ontológicamente similares.



**Figura 17. Distribución del ES de las firmas funcionales y de cada módulo de TAM en células tumorales y macrófagos.**

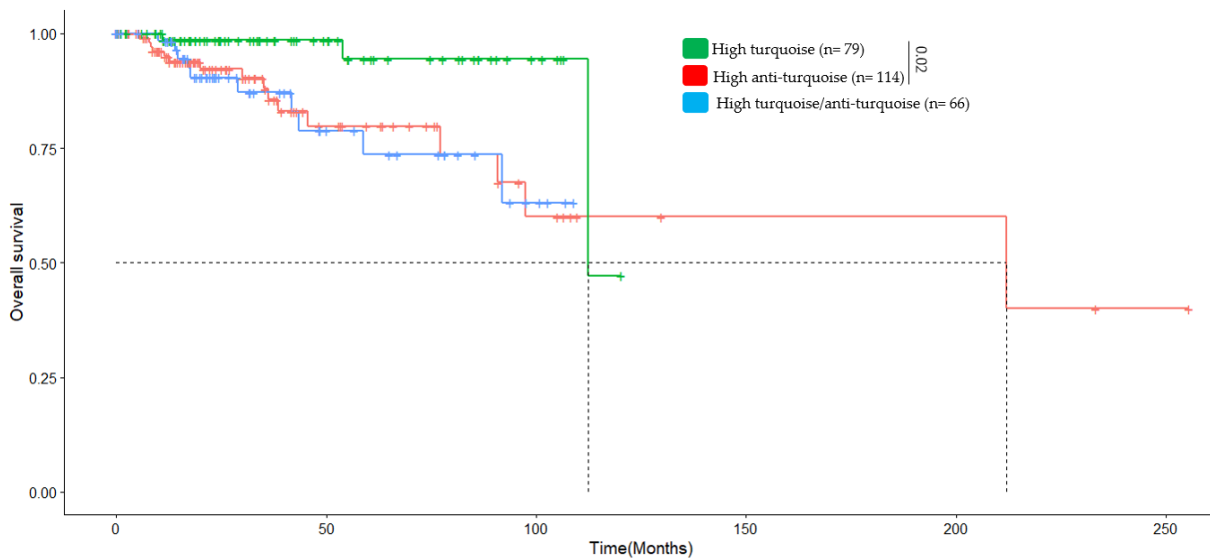
Se representa el ES de 17 TAM y 270 células tumorales (“Epithelial\_cells”). En el eje X están representados los valores de enriquecimiento obtenidos con GSVA, los cuales están dentro de un rango de -1 a 1. Células provenientes del estudio con la clave GSE75688. TAM: macrófagos asociados a tumor.



## 7.6 Relevancia clínica de las firmas turquesa y anti-turquesa

La firma marrón es compartida por diferentes poblaciones de leucocitos e incluso por células tumorales, sin embargo, los TAM que expresan la firma marrón también expresan la firma turquesa, por lo tanto, esta última puede utilizarse como medidor indirecto de la firma marrón y su respectiva vía asociada de TNF $\alpha$ .

Aprovechando la especificidad de la firma anti-turquesa en los TAM (asociada a coagulación y EMT) y la especificidad de la firma turquesa (asociada a la vía TNF $\alpha$ ) por TAM/monocito/DC, se calculó el enriquecimiento de estas 2 firmas en pacientes del TCGA con cáncer de mama con el fin de saber si estas tienen algún impacto en la supervivencia del paciente. Los pacientes con alto enriquecimiento de la firma turquesa (*"High turquoise"*) presentaron mayor supervivencia respecto a aquellos pacientes con alto enriquecimiento de la firma anti-turquesa (*"High anti-turquoise"*) (Figura 18); los pacientes que presentaron alto ES para ambas firmas (*"High turquoise/anti-turquoise"*) tuvieron una supervivencia similar a los pacientes *High anti-turquoise*, posiblemente porque la presencia de TAM anti-turquesa influye en la supervivencia del paciente de manera dominante, independientemente de si coexisten con otras poblaciones de TAM.



**Figura 18. Curvas de supervivencia de pacientes con cáncer de mama del TCGA enriquecidos con algún subtipo de TAM.**

Los pacientes están agrupados de acuerdo a su ES de las firmas correspondientes a TAM turquesa (*"high turquoise"*), anti-turquesa (*"high anti-turquoise"*) o de ambos (*"high turquoise/anti-turquoise"*). *High turquoise* vs *high anti-turquoise*: *p-value*=0.02.

## 8. DISCUSIÓN

### 8.1 Originalidad del trabajo realizado.

Se tiene registro de 4 estudios que han implementado tecnología scRNA-seq para comprender el infiltrado leucocitario en cáncer de mama (febrero, 2021) [94]. En 2 de estos estudios solamente se capturó el transcriptoma de linfocitos T y B [95-96] y en los 2 estudios restantes, los cuales se emplearon en este análisis, se capturó el transcriptoma de los diferentes tipos celulares en el TME [74, 81]. Estos últimos 2 estudios realizaron un análisis general del TME y un análisis enfocado en los linfocitos T; sin embargo, ninguno intentó describir a las poblaciones de macrófagos que encontraron y se limitaron a mencionar que ninguna de sus poblaciones de TAM correlacionaba con marcadores M1 ni M2 [74]. Bao, X. y colaboradores [97] se plantearon usar los datos públicos de uno de los estudios scRNA-seq [81] para analizar la biología de los TAM en cáncer de mama. Analizaron células tumorales, linfocitos T, linfocitos B y TAM, de estos últimos concluyeron que correlacionaban con una firma M2; sin embargo, el número reducido de macrófagos disponibles en ese estudio en específico (menos de 100), les impidió llevar a cabo análisis más sofisticados. Por lo tanto, el análisis realizado en este trabajo es el primero en presentar un análisis enfocado exclusivamente en los TAM de cáncer de mama y en implementar un número robusto de células para describir sus funciones celulares.

### 8.2 Retos para identificar TAM en el TME de cáncer de mama.

La identificación y etiquetado de los TAM resultó especialmente difícil debido a las similitudes que comparten con otras células del mismo linaje celular, como lo son los monocitos y las DC. La similitud entre el transcriptoma de monocitos y de DC fue tal, que *SingleR* identificó erróneamente alrededor de 200 células como monocitos, cuando en realidad eran DC y viceversa; esto puede deberse a que tanto macrófagos como DC, pueden derivarse a partir de monocitos en condiciones inflamatorias, además comparten expresión de marcadores que previamente se consideraban específicos de cada población (CD68, CD163, CD14, CD11c, CD11b) [66-67, 98-99]. Esto tiene relevancia ya que los estudios que reportaron a los TAM como una de las poblaciones más abundantes del TME en diferentes tipos de cáncer [18], lo hicieron mediante la identificación de



macrófagos, usando técnicas de inmunohistoquímica y CD68 como marcador, lo cual pudo llevar a identificar conjuntamente varias poblaciones y a sobre estimar la abundancia de los TAM. Relacionado a esto, se encontró que los monocitos infiltrantes de tumor (n=1741) eran ligeramente más abundantes que los TAM (n=1354), esto llama la atención porque los monocitos al ingresar al tejido y exponerse a los estímulos del TME, pueden diferenciarse a macrófagos o a DC [100]. Se ha reportado que los monocitos circulantes de sangre pueden permanecer con mínimas características de diferenciación hacia macrófagos o DC aún después de haber infiltrado tejido, tanto en condiciones fisiológicas [101-102] como en tejido inflamado [103]; estos últimos se denominan monocitos inflamatorios y se ha observado que conservan marcadores de monocito al mismo tiempo que adquieren funciones comúnmente asociadas a macrófagos y DC, como por ejemplo, mayor expresión de moléculas MHC-II, secreción de interleucinas pro inflamatorias, producción de TNF y especies reactivas de oxígeno y nitrógeno, así como la capacidad de migrar a órganos linfoides secundarios [103-107]. Se desconoce el tiempo que requieren los monocitos para diferenciarse a macrófagos en un sistema *in vivo* [107], por lo tanto, también existe la posibilidad de que los monocitos infiltrantes de tumor encontrados en el TME de cáncer de mama se encuentren en proceso de diferenciación hacia TAM.

### **8.3 Funciones asociadas a cada firma de TAM.**

Con este análisis se logró determinar que existen principalmente 3 programas transcripcionales o módulos en los TAM; de acuerdo con su valor de *gene membership* se pudieron determinar los genes con mayor relevancia para cada módulo. A continuación, se ejemplifican 2 genes con alto *membership* para cada módulo: *SPP1* y *FN1* para el módulo anti-turquesa; *C1QA* y *HLA-DPA1* para el módulo turquesa y *FOS* y *JUN* para el módulo marrón. Debido a que los módulos turquesa y marrón convergen en una misma función celular, se puede decir que los TAM siguen un comportamiento dicotómico. No se ha reportado algo similar para los TAM en cáncer de mama, ni siquiera por los mismos grupos de investigación que generaron los datos scRNA-seq usados en este análisis [74, 81]. Sin embargo, el comportamiento dicotómico de los TAM, así como algunos de los genes con mayor *gene membership* de cada módulo, ya habían sido observados por otros grupos de investigación que recurrieron a tecnología scRNA-seq.

Zhang, L y colaboradores [73] analizaron macrófagos de cáncer colorrectal y encontraron dos poblaciones de TAM, una de ellas expresaba *SPP1* como marcador y tenía activadas vías relacionadas al proceso de angiogénesis, a la regulación negativa de la respuesta inmunológica y a la metástasis (población *SPP1*<sup>+</sup>). La segunda población de TAM (población *C1QC*<sup>+</sup>) expresaba genes de complemento como *C1QA*, *C1QB* y *C1QC* y moléculas HLA-II como marcadores, además esta población presentaba alta expresión de factores de transcripción como *FOS* y *JUN* respecto a la población *SPP1*<sup>+</sup>, lo cual apoya que los módulos turquesa y marrón son dos firmas diferentes que convergen en una misma población de TAM. Por último, su análisis mostraba que la población *C1QC*<sup>+</sup> interactuaba con linfocitos Th1 y estaba asociada a una mayor supervivencia respecto a la población *SPP1*<sup>+</sup> en pacientes del TCGA con cáncer colorrectal. Lavin, Y. y colaboradores [108] realizaron un análisis scRNA-seq en muestras de adenocarcinoma de pulmón. En ese estudio se encontró que había un grupo de TAM especialmente enriquecido en pacientes con menor supervivencia, además esta misma población de TAM presentaba alta expresión de genes pertenecientes a la firma anti-turquesa (*SPP1*, *MARCO* y *APOE*) respecto a otra población de TAM que no caracterizaron, pero estaba enriquecida con genes presentes en las firmas turquesa y marrón como *C1QA*, *CSF1R*, *JUN*, *MAFB* y *DUSP1*. En conjunto los estudios mencionados previamente, apoyan los marcadores y el comportamiento dicotómico encontrado en los TAM en este trabajo.

Adicionalmente, en este análisis se encontró un cuarto módulo denominado azul, que contenía genes implicados en procesos angiogénicos y estaba enriquecido en células de origen mesenquimal como fibroblastos, condrocitos, células mesenquimales troncales y en menor medida en células endoteliales. Esto concuerda con el papel que tienen los fibroblastos y otras células de origen mesenquimal como los pericitos, en la formación de vasos sanguíneos [109-110]. Este módulo fue encontrado en fibroblastos que fueron inicialmente identificados como macrófagos, esto probablemente se debe a que los fibroblastos expresan marcadores de macrófagos (CD68) y de monocitos (CD14, CD16), lo cual ya se observó tanto en la literatura [111] como en este trabajo.

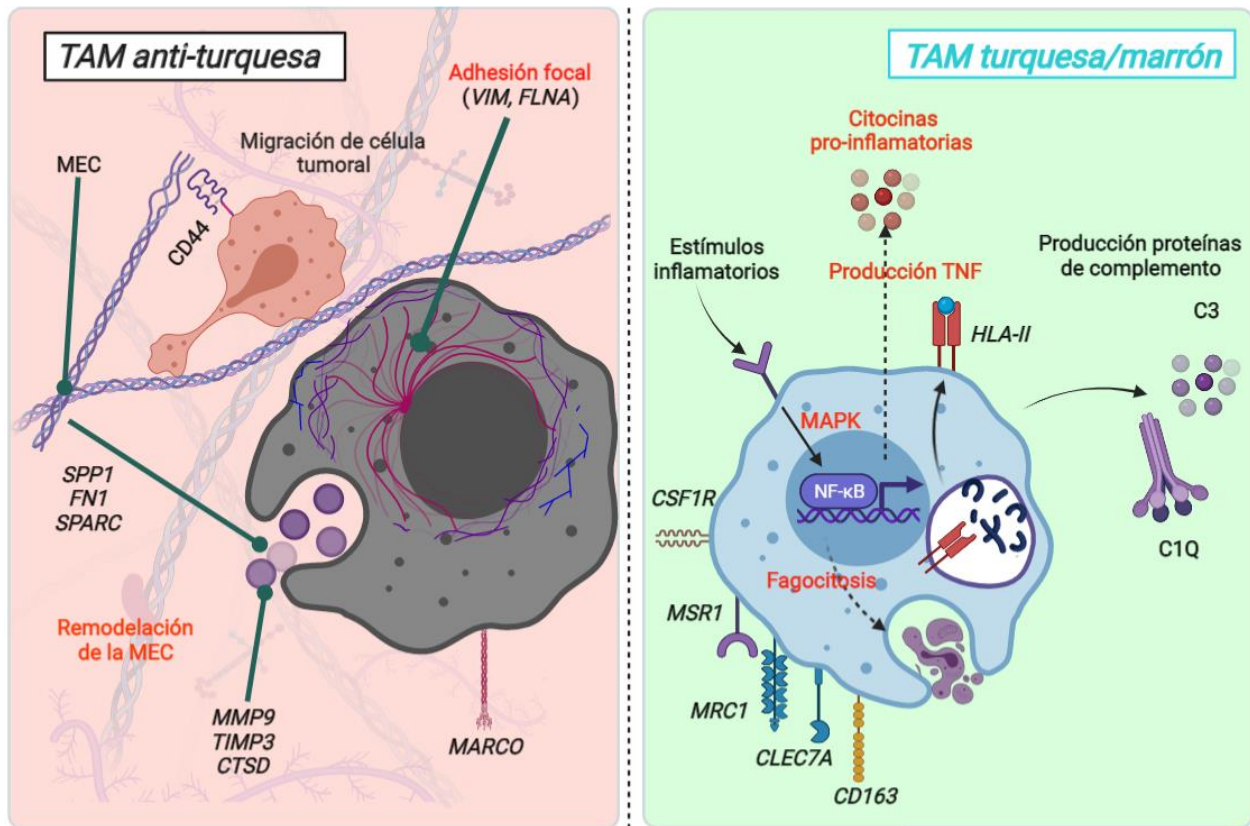
La firma marrón contiene genes que participan en la vía de señalización MAPK como *HSPA1A*, *FOS*, *JUN*, *DUSP1*, etc. La vía de señalización MAPK es ubicua en células de

mamíferos y controla procesos celulares como inflamación, proliferación, diferenciación, apoptosis y supervivencia. Esta vía puede ser activada por factores de crecimiento y por estrés celular; esto último incluye LPS, choque de calor y citocinas proinflamatorias [112]. En células del sistema inmune innato los receptores de factor de necrosis tumoral (TNFR), receptores de interleucina y receptores *toll-like* (TLR) desencadenan la señalización de la vía MAPK y pueden derivar en la producción de TNF a través del factor nuclear- $\kappa$ B (NF- $\kappa$ B), así como en la producción de iNOS y en la regulación de la fagocitosis [113-114]. Aunado a esto, los TAM que expresan la firma marrón también expresan la firma turquesa, la cual contiene genes que codifican para moléculas HLA-II. Se sabe que los macrófagos pueden presentar antígenos exógenos en un contexto HLA-II después de fagocitarlos, además se observó que los TAM turquesa/marrón expresaban diferencialmente genes que participan en la fagocitosis como *CD163*, *MRC1*, *MSR1* y *CLEC7A* [115]. Adicionalmente, los TAM turquesa/marrón expresan algunos genes del sistema complemento (*C3*, *C1QA*). Hay reportes contradictorios sobre las funciones asociadas al sistema complemento en el TME; se le han asignado tanto funciones pro-tumorales como anti-tumorales en términos de crecimiento tumoral y metástasis, por lo que se cree que las funciones del sistema complemento son contexto dependiente [116]. No se sabe con exactitud el mecanismo por el cual el sistema complemento es activado en el TME, pero se cree que podría ser activado por la vía clásica mediante el reconocimiento de inmunoglobulinas G (adheridas a células tumorales) por la proteína C1q, pero esta cascada de activación no deriva en la formación del complejo de ataque a la membrana [116]. En el contexto particular de este análisis se propone los TAM turquesa/marrón podrían promover el microambiente inflamación mediante la liberación de anafilotoxinas C3a y C5a mientras que C3b promovería la opsonización de células apoptóticas. Todo lo anterior sugiere que las funciones de los TAM turquesa/marrón consisten en promoción del microambiente inflamatorio, así como en la fagocitosis de cuerpos apoptóticos y presentación de antígenos tumorales en un contexto HLA-II. Las funciones anteriores son asociadas como funciones antitumorales y de buen pronóstico; además los TAM turquesa/marrón se encontraron enriquecidos en muestras de cáncer de mama con mayor supervivencia en pacientes del TCGA.

Los TAM anti-turquesa expresan firmas de EMT y coagulación; ambas vías contienen genes implicados en la remodelación de la MEC como *SPP1*, *FN1*, *MMP9*, *TIMP3* y *SPARC*. La remodelación de la MEC es un proceso que ocurre de manera fisiológica durante la reparación del tejido [117] y este incluye la síntesis de proteínas extracelulares que dan sostén a la estructura de la MEC como osteopontina, fibronectina y la proteína secretada ácida y rica en cisteína, codificadas por los genes *SPP1*, *FN1* y *SPARC* respectivamente, los cuales participan en la interacción con otras proteínas de la MEC y regulan la adhesión celular [118-120]. La remodelación de la MEC también implica la degradación de la matriz provisional que se forma al inicio del proceso de reparación, dicha degradación es llevada cabo por la actividad de catepsinas (*CTSA*, *CTSB*, *CTSD*) [121] y por MMP como *MMP9*, aunque también es acompañada por reguladores de MMP como *TIMP3* [117, 119, 122]. En cáncer este proceso se encuentra desregulado, así como los componentes de la MEC [123]. *SPP1* y *FN1* regulan la adhesión celular a la MEC y la migración a través de la MEC mediante la interacción con integrinas, además promueven la expresión de diversas MMP que degradan componentes de la MEC [118-119, 122]. Interesantemente, *SPP1* puede promover la migración e incrementar la viabilidad de las células tumorales, al interactuar con el receptor de superficie CD44, el cual es un marcador de troncalidad en células tumorales [124-125]. Por lo tanto, la desregulación de genes involucrados en la remodelación de la MEC, podría facilitar la migración e invasión de células tumorales.

Adicionalmente, los TAM anti-turquesa expresan *VIM* y *FLNA*, los cuales codifican para proteínas que, junto con integrinas, participan en la estabilización de la estructura que permite la adhesión focal de las células a la MEC, además *VIM* es un regulador de la migración celular [126].

Lo anterior sugiere que los TAM anti-turquesa interactúan íntimamente con la MEC y se movilizan por dicha estructura al mismo tiempo que facilitan la migración e invasión de las células tumorales mediante la exportación de componentes extracelulares capaces de remodelar la MEC. La figura 19 resume el modelo propuesto de las funciones que llevan a cabo las poblaciones de TAM encontradas.



**Figura 19. Esquema de los programas transcripcionales encontrados en los TAM.**

El panel izquierdo resume los genes expresados por los TAM anti-turquesa, involucrados en la remodelación de la MEC, así como en la adhesión focal a esta misma. *SPP1* además de ser un componente de la MEC, podría interaccionar con CD44 y promover así la migración de células tumorales. El panel derecho esquematiza los procesos celulares que puede estar llevando a cabo la población de TAM turquesa. Ante estímulos inflamatorios estos macrófagos señalizan a través de la vía MAPK y el factor de transcripción NF-κB para promover la síntesis de TNF y otras citocinas pro inflamatorias. La vía MAPK también regula el proceso de fagocitosis, lo cual podría verse reflejado con una mayor expresión de *MSR1*, *MRC1*, *CD163* y *CLEC7A*. Antígenos extracelulares capturados dentro de fagosomas son procesados y cargados en moléculas HLA-II para posteriormente ser presentados a linfocitos T CD4<sup>+</sup>. La síntesis de proteínas del sistema complemento podría ser en respuesta los estímulos inflamatorios del TME o podría promover más inflamación mediante anafilotoxinas como C3a. MEC: matriz extracelular. Imagen creada con la versión gratuita de Biorender.com.

## 8.4 Distribución de la expresión de los marcadores M1/M2 en las diferentes poblaciones de TAM.

Previamente se ha analizado la expresión de marcadores M1/M2 en poblaciones de TAM de diferentes tipos de cánceres a nivel scRNA-seq [73-78]. Debido a que el análisis de los marcadores M1/M2 no era el objetivo principal de ninguno de esos estudios, no se hizo una revisión exhaustiva de estos, por lo que hay variaciones entre las listas de marcadores usados en cada estudio, además en todos estos estudios el análisis se limitó

a calcular el enriquecimiento conjunto de los marcadores M1 o M2 y se concluyó que ninguna población de TAM estaba significativamente más enriquecida para ninguna de las firmas M1 o M2 respecto al resto de las poblaciones de TAM. En este trabajo se desglosó la expresión individual de cada marcador entre las poblaciones de TAM y entre las diferentes poblaciones del TME de cáncer de mama, lo cual permitió observaciones relevantes. Por ejemplo, los TAM turquesa/marrón expresaban IL-1 $\beta$ , CXCL10 y moléculas HLA-II (marcadores M1) mientras que los únicos marcadores M2 que expresaron fueron aquellos involucrados en procesos de fagocitosis (*CD163*, *MRC1*, *MSR1* y *CLEC7A*). Esto tiene sentido porque los antígenos que son presentados en un contexto HLA-II, tuvieron que haber sido previamente fagocitados, además estos mismos TAM tienen activada la vía MAPK, la cual regula inflamación y fagocitosis [114]. Zhang, L y colaboradores también encontraron una mayor expresión de *MRC1* y *CD163* por parte de su población macrófagos C1QC<sup>+</sup> [73]. Esto sugiere que en un contexto tumoral las moléculas HLA-II, los marcadores inflamatorios y los asociados a fagocitosis, pueden coexistir en un mismo macrófago, independientemente de si se trata de marcadores M1 o M2, ya que todo parece formar parte de un mismo proceso celular orquestado por la vía MAPK y el factor de transcripción NF $\kappa$ B en una misma población de TAM (Figura 19).

Otra observación derivada de este análisis fue que los TAM anti-turquesa no expresaban marcadores M2 asociados a otra función que no fuera reparación de tejido (*SPP1*, *FN1*, *MMP9*, *TIMP3*, *SPARC*, *CTSA*, *CTSB*, *CTSD*) [117-122] como por ejemplo, angiogénesis (VEGF, TGF $\beta$ ) o inmunosupresión (IL-10, TGF $\beta$ , PD-L1, PD-L2), de esta última función solamente expresaban *CCL22*, el cual participa en la atracción de Treg [36]. Por lo tanto, las poblaciones de TAM del cáncer de mama, parecen ser inflamatorios con expresión de HLA-II o ser TAM con funciones de reparación de tejido, pero no inmunosupresores o angiogénicos.

Cuando se desglosó la expresión de los marcadores M1/M2 en los diferentes tipos celulares, se pudo notar que los TAM no son la población con mayor expresión de estos marcadores, de hecho, marcadores angiogénicos como VEGF y TGF $\beta$  se expresan en fibroblastos y células endoteliales, lo cual concuerda con la importancia que tienen ambos tipos celulares en la formación de vasos sanguíneos [127-128]. Interesantemente, VEGF

fue identificado como marcador de TAM debido a que co-localizaba con CD68 [65], el cual es un marcador compartido por fibroblastos [111].

Las observaciones anteriores sugieren que, aunque existen dos poblaciones de TAM en cáncer de mama que presentan algunas características del fenotipo M1 (inflamación y presentación de antígeno) o del fenotipo M2 (remodelación de la MEC), estos no pueden considerarse como macrófagos M1/M2 porque la gran mayoría de los marcadores no conservan la división propuesta por el modelo M1/M2. Es importante considerar que los marcadores de este modelo son en su mayoría, genes inducidos en macrófagos que fueron estimulados con INF- $\gamma$  y LPS o IL-4 e IL-13 de manera *in vitro* o son genes inducidos en respuesta a parásitos que promueven respuestas tipo Th2 en modelos murinos [37-58]. Es claro que en el TME coexisten múltiples estímulos que en conjunto son mucho más complejos que los estímulos a los que puede ser sometido un macrófago de manera *in vitro* y considerablemente diferentes a los generados en la respuesta contra parásitos. De acuerdo con lo observado en este trabajo, se puede concluir que la división de macrófagos de acuerdo los marcadores del modelo M1/M2, no se extiende a los TAM presentes el TME de cáncer de mama.

### **8.5 Posible origen ontológico de las poblaciones de TAM encontradas**

Se encontraron similitudes entre los monocitos infiltrantes de tumor y los TAM turquesa. Estos monocitos presentaban enriquecimiento tanto de la firma turquesa como de la firma marrón, además expresaban marcadores M1 asociados a inflamación y a activación de linfocitos T como *CXCL9*, *CXCL11*, *IDO1*, *TNF*, *CD80* y *CD86*, los cuales se esperaban encontrar expresados en macrófagos. Esto, aunado a la dificultad para distinguir a los monocitos de poblaciones como TAM y DC debido a la similitud de sus transcriptomas dentro del TME, nos lleva a hipotetizar que las similitudes encontradas, probablemente se deben a que los monocitos infiltrantes de tumor, provenientes de sangre periférica, podrían dar lugar a TAM con alta expresión de genes pertenecientes a las firmas turquesa y marrón. Pero si los TAM turquesa/marrón provienen de monocitos de sangre periférica, ¿cuál es el origen ontológico de los TAM anti-turquesa? Zhang, L. y colaboradores [73] se apoyaron de análisis de inferencia de trayectoria y propusieron un atractivo modelo de diferenciación entre las poblaciones TAM SPP1<sup>+</sup> y C1QA<sup>+</sup> que encontraron. Según su

modelo, los TAM C1QA<sup>+</sup> pueden provenir de monocitos de sangre periférica que se infiltraron en el tumor, mientras que los TAM SPP1<sup>+</sup> probablemente provienen principalmente de macrófagos residentes de tejido. Se sabe que los macrófagos residentes de tejido tienen un precursor embrionario y que no provienen de monocitos postnatales [26]. Curiosamente, se ha reportado que la expresión de MARCO se mantiene sin cambios, independientemente de los estímulos del microambiente en macrófagos de origen embrionario [129] y MARCO fue el único marcador M2 asociado a fagocitosis que no fue expresado por los TAM turquesa/marrón, sino que más bien se encontró expresado por TAM anti-turquesa.

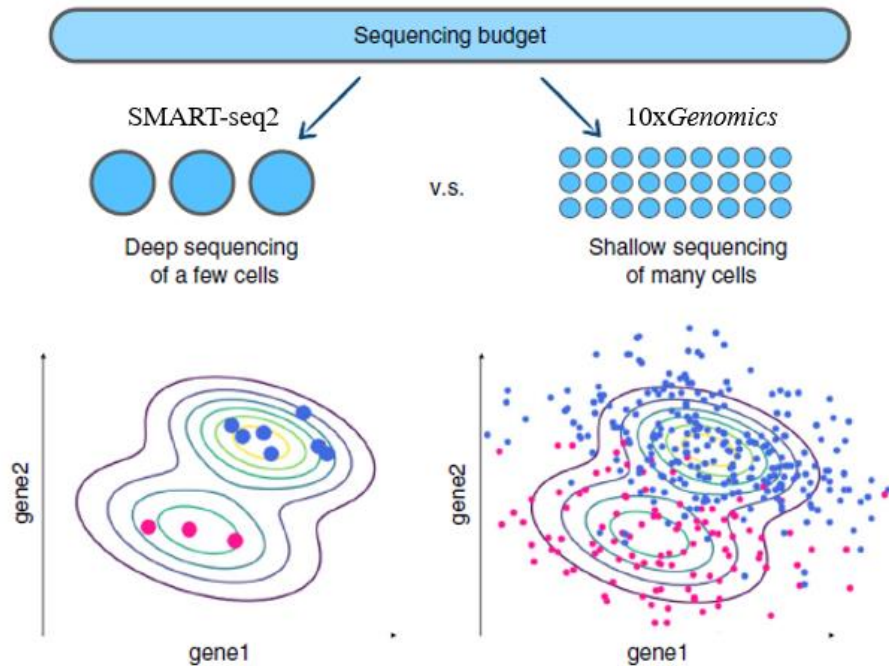
Sin embargo, se requiere de validación experimental para corroborar este modelo de diferenciación de TAM que propone orígenes ontológicos diferentes para cada población.

## **8.6 Limitaciones de este estudio.**

A pesar de haber explorado la expresión de los marcadores M1/M2 en un universo de más de 20,000 células infiltrantes de tumor, se observó que la expresión de *IL12A*, *IL1A*, *CCL17*, *CCL24*, *MMP12*, *ARG1* y *EGF* era nula o escasa en todos los tipos celulares. Esto parece indicar que se trata de una cuestión técnica más que de una cuestión biológica, evidenciando que este análisis heredó limitaciones intrínsecas de las tecnologías scRNA-seq. Actualmente, cuando se realiza scRNA-seq se tiene que equilibrar el número de células que se quiere secuenciar y la profundidad de secuenciación de cada célula. Una mayor profundidad de secuenciación (mayor número de genes detectados) se obtiene a costa del número de células que se puede secuenciar (mayor número de poblaciones recuperadas) [130] (Figura 20). Plataformas como 10xGenomics son capaces de capturar decenas de miles de células de manera eficiente, a costa de perder información de los genes menos abundantes, mientras que plataformas como Smart-seq2 se encuentran en la situación opuesta (secuenciación profunda de pocas células) [130-132]. Los datos analizados de este trabajo fueron generados con la plataforma 10xGenomics [74]; por lo tanto, los marcadores M1/M2 de los que no se encontró transcrito alguno en ninguna de las poblaciones, se deben muy probablemente a las limitaciones técnicas de esta plataforma. Esto también quiere decir que se pudo haber perdido información de genes menos abundantes que no fueron analizados, pero



que podrían tener relevancia en la biología de los TAM. Se requiere extender este análisis a otros estudios que cuenten con datos generados con otras plataformas o incluso extenderlo a estudios que cuenten con técnicas proteómicas, con el fin de corroborar los resultados obtenidos. Además, se requiere que las observaciones obtenidas mediante herramientas bioinformáticas como las de este trabajo, sean validadas experimentalmente.



**Figura 20. Limitaciones de las plataformas scRNA-seq.**

Con un presupuesto fijo se escoge entre secuenciar la mayor parte del transcriptoma de pocas células o secuenciar solamente los genes más abundantes de un gran número de células. La elección de la plataforma depende de las necesidades y objetivos de cada grupo de investigación. Secuenciar con 10xGenomics permitiría capturar poblaciones celulares poco abundantes, mientras que SMART-seq2 permitiría capturar el transcriptoma de cada célula con poca pérdida de información. Imagen modificada de [130].

## 8.7 Implicaciones clínicas y terapéuticas.

Los estudios que correlacionaron una mayor abundancia de TAM con menor supervivencia en varios tipos de cáncer [18], lo hicieron mediante la identificación inespecífica de TAM, monocitos y DC usando el marcador CD68; además tampoco se consideró la posibilidad de que existieran diferentes poblaciones de TAM. En este análisis se observó que la firma anti-turquesa perteneciente a los TAM anti-turquesa estaba

asociada a una menor supervivencia y que la firma turquesa perteneciente a TAM/monocito/DC estaba asociada a una mayor supervivencia en pacientes con cáncer de mama del TCGA, sin embargo, cuando ambas firmas estaban enriquecidas en un mismo paciente, se asociaron con una menor supervivencia. Esto parece indicar que, aunque existe una población de TAM asociado con mayor supervivencia, la población de TAM asociada con menor supervivencia podría influir de manera predominantemente negativa sobre la supervivencia del paciente, incluso aunque ambas poblaciones de TAM se encuentren presentes. Esto podría explicar por qué los primeros estudios que evaluaron la relación entre TAM y supervivencia, los encontraron asociados con menor supervivencia aun cuando estos fueron identificados con CD68 y sin ninguna separación en subpoblaciones. Importantemente, los TAM enriquecidos con la firma turquesa también expresan CSF1R y MSR1, los cuales suelen ser blanco de estrategias terapéuticas que buscan impedir el reclutamiento o eliminar poblaciones de macrófagos M2 [133-135]. Sin embargo, el efecto antitumoral que se observa al implementar moléculas dirigidas contra macrófagos que expresan estos marcadores M2, muy probablemente se debe a una eliminación indiscriminada de las diferentes poblaciones de TAM [135]. Cabe resaltar que la firma turquesa también se encontró enriquecida en poblaciones de buen pronóstico como las DC, las cuales también podrían expresar CSF1R y MSR1. De esta forma, eliminar células que expresen CSF1R o MSR1 podría tener resultados clínicos desfavorables para el paciente.

Además de impedir el reclutamiento de macrófagos al TME mediante inhibidores de CSF1R [133] y de la eliminación global de TAM mediante biomoléculas que inducen apoptosis en células con capacidades fagocíticas [135-136], existe otro tipo de estrategia que consiste en repolarizar a los TAM M2 hacia un fenotipo M1 inflamatorio. Por ejemplo, se planea lograr la repolarización mediante agonistas de CD40 [137-138], silenciamiento de VEGF [139] o mediante el bloqueo de la polarización hacia M2 causada por IL-13 [140], entre otros acercamientos [136]. Todas estas estrategias se basan en los marcadores establecidos por el modelo de macrófagos M1/M2, por esta razón es importante establecer si estos marcadores son realmente puntos de referencia adecuados para diseñar terapia anti-tumorales, ya que como se observó en este trabajo,

los TAM con funciones pro-tumorales y los TAM con funciones antitumorales expresan tanto marcadores M1 y M2.

## **9. CONCLUSIONES**

Con las observaciones obtenidas a partir de este análisis se propone que, las poblaciones de TAM del TME de cáncer de mama siguen un modelo dicotómico en el que existe una población que expresa vías inflamatorias y de presentación de antígeno, se asocia a funciones antitumorales y a mayor supervivencia. Al mismo tiempo existe otra población de TAM con mayor expresión de genes que remodelan la MEC, se asocia a funciones pro-tumorales y a menor supervivencia. Sin embargo, ninguna de las 2 poblaciones de TAM encontradas pueden ser identificadas con marcadores M1/M2 porque en su mayoría, se encuentran expresados de forma indistinta en una misma población de TAM o en su defecto tienen escasa expresión en macrófagos. Por lo tanto, el modelo de activación de macrófagos M1/M2 y sus respectivos marcadores, no son adecuados para clasificar las diferentes poblaciones de TAM que existen en el TME de cáncer de mama. Además, estas observaciones brindan consideraciones importantes para el diseño de terapias actualmente dirigidas contra macrófagos M2 en cáncer.

## **10. PERSPECTIVAS**

Las observaciones de este análisis estuvieron limitadas a los TAM de un solo tipo de cáncer y al rendimiento y profundidad de una sola plataforma scRNA-seq. Quedan preguntas acerca de si las poblaciones de TAM observadas en cáncer de mama, son conservadas a través de diferentes tipos de cáncer o si la presencia o abundancia de estas poblaciones está en función del tejido. Reportes en cáncer colorrectal y de pulmón parecen indicar que las poblaciones de TAM encontradas en este análisis no son exclusivas de cáncer de mama. Se planea extender este análisis a otros 7 tipos de cáncer que ya cuentan con datos scRNA-seq de manera pública, así como complementar el análisis de expresión de los marcadores M1/M2 con otras plataformas scRNA-seq como SMART-seq2, que son capaces de recuperar el transcriptoma de las células de manera más profunda.

## 11. REFERENCIAS

1. NIH, National Institutes of Health. *National Cancer Institute* [Internet]. <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>, (2020).
2. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2020). Global Cancer Observatory: Cancer Today. Lyon, France: *International Agency for Research on Cancer*. Disponible en: <https://gco.iarc.fr/today>. (Consultado en abril, 2021).
3. Tsang, J., & Tse, G. M. (2020). Molecular Classification of Breast Cancer. *Advances in anatomic pathology*, 27(1), 27–35. <https://doi.org/10.1097/PAP.000000000000232>
4. Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, 19(5), 403–410. <https://doi.org/10.1111/j.1365-2559.1991.tb00229.x>
5. Prat, A., Pineda, E., Adamo, B., Galván, P., Fernández, A., Gaba, L., Díez, M., Viladot, M., Arance, A., & Muñoz, M. (2015). Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast (Edinburgh, Scotland)*, 24 Suppl 2, S26–S35. <https://doi.org/10.1016/j.breast.2015.07.008>
6. Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., Moses, H. L., Sanders, M. E., & Pietersen, 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>
7. Medrek, C., Pontén, F., Jirstrom, K., & Leandersson, K. (2012). The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. *BMC cancer*, 12, 306. <https://doi.org/10.1186/1471-2407-12-306>
8. Mahmoud SM, Lee AH, Paish EC, Macmillan RD, Ellis IO, Green AR. Tumour-infiltrating macrophages and clinical outcome in breast cancer. *J Clin Pathol*. 2012;65(2):159-163. doi:10.1136/jclinpath-2011-200355
9. Lindsten, T., Hedbrant, A., Ramberg, A., Wijkander, J., Solterbeck, A., Eriksson, M., Delbro, D., & Erlandsson, A. (2017). Effect of macrophages on breast cancer cell proliferation, and on expression of hormone receptors, uPAR and HER-2. *International journal of oncology*, 51(1), 104–114. <https://doi.org/10.3892/ijo.2017.3996>
10. Zhang, W. J., Wang, X. H., Gao, S. T., Chen, C., Xu, X. Y., Sun, Q., Zhou, Z. H., Wu, G. Z., Yu, Q., Xu, G., Yao, Y. Z., & Guan, W. X. (2018). Tumor-associated macrophages correlate with phenomenon of epithelial-mesenchymal transition and contribute to poor prognosis in triple-negative breast cancer patients. *The Journal of surgical research*, 222, 93–101. <https://doi.org/10.1016/j.jss.2017.09.035>
11. Zhang, Q., Le, K., Xu, M., Zhou, J., Xiao, Y., Yang, W., Jiang, Y., Xi, Z., & Huang, T. (2019). Combined MEK inhibition and tumor-associated macrophages depletion suppresses tumor growth in a triple-negative breast cancer mouse model. *International immunopharmacology*, 76, 105864. <https://doi.org/10.1016/j.intimp.2019.105864>

12. Zhou, J., Wang, X. H., Zhao, Y. X., Chen, C., Xu, X. Y., Sun, Q., Wu, H. Y., Chen, M., Sang, J. F., Su, L., Tang, X. Q., Shi, X. B., Zhang, Y., Yu, Q., Yao, Y. Z., & Zhang, W. J. (2018). Cancer-Associated Fibroblasts Correlate with Tumor-Associated Macrophages Infiltration and Lymphatic Metastasis in Triple Negative Breast Cancer Patients. *Journal of Cancer*, *9*(24), 4635–4641. <https://doi.org/10.7150/jca.28583>
13. Yuan, Z. Y., Luo, R. Z., Peng, R. J., Wang, S. S., & Xue, C. (2014). High infiltration of tumor-associated macrophages in triple-negative breast cancer is associated with a higher risk of distant metastasis. *OncoTargets and therapy*, *7*, 1475–1480. <https://doi.org/10.2147/OTT.S61838>
14. Sami, E., Paul, B. T., Koziol, J. A., & ElShamy, W. M. (2020). The Immunosuppressive Microenvironment in BRCA1-IRIS-Overexpressing TNBC Tumors Is Induced by Bidirectional Interaction with Tumor-Associated Macrophages. *Cancer research*, *80*(5), 1102–1117. <https://doi.org/10.1158/0008-5472.CAN-19-2374>
15. Bejarano, L., Jordão, M. J. C., & Joyce, J. A. (2021). Therapeutic targeting of the tumor microenvironment. *Cancer Discovery*, *11*(4), 933–959. <https://doi.org/10.1158/2159-8290.CD-20-1808>
16. Wu, T., & Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Letters*, *387*, 61–68. <https://doi.org/10.1016/j.canlet.2016.01.043>
17. Quail, D. F., & Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nature Medicine*, *19*(11), 1423–1437. <https://doi.org/10.1038/nm.3394>
18. Fujita, H., Ohuchida, K., Mizumoto, K., Nakata, K., Yu, J., Kayashima, T., Cui, L., Manabe, T., Ohtsuka, T., & Tanaka, M. (2010). alpha-Smooth Muscle Actin Expressing Stroma Promotes an Aggressive Tumor Biology in Pancreatic Ductal Adenocarcinoma. *Pancreas*, *39*(8), 1254–1262. <https://doi.org/10.1097/MPA.0b013e3181dbf647>
19. Underwood, T. J., Hayden, A. L., Derouet, M., Garcia, E., Noble, F., White, M. J., Thirdborough, S., Mead, A., Clemons, N., Mellone, M., Uzoho, C., Primrose, J. N., Blaydes, J. P., & Thomas, G. J. (2015). Cancer-associated fibroblasts predict poor outcome and promote periostin-dependent invasion in oesophageal adenocarcinoma. *The Journal of pathology*, *235*(3), 466–477. <https://doi.org/10.1002/path.4467>
20. Qin, X., Yan, M., Wang, X., Xu, Q., Wang, X., Zhu, X., Shi, J., Li, Z., Zhang, J., & Chen, W. (2018). Cancer-associated Fibroblast-derived IL-6 Promotes Head and Neck Cancer Progression via the Osteopontin-NF-kappa B Signaling Pathway. *Theranostics*, *8*(4), 921–940. <https://doi.org/10.7150/thno.22182>
21. Becht, E., de Reyniès, A., Giraldo, N. A., Pilati, C., Buttard, B., Lacroix, L., Selves, J., Sautès-Fridman, C., Laurent-Puig, P., & Fridman, W. H. (2016). Immune and Stromal Classification of Colorectal Cancer Is Associated with Molecular Subtypes and Relevant for Precision Immunotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research*, *22*(16), 4057–4066. <https://doi.org/10.1158/1078-0432.CCR-15-2879>
22. Wang, J., & Matosevic, S. (2020). Functional and metabolic targeting of natural killer cells to solid tumors. *Cellular Oncology*, *43*(4), 577–600. <https://doi.org/10.1007/s13402-020-00523-7>

23. Bingle, L., Brown, N. J., & Lewis, C. E. (2002). The role of tumour-associated macrophages in tumour progression: Implications for new anticancer therapies. *Journal of Pathology*, 196(3), 254–265. <https://doi.org/10.1002/path.1027>
24. Ginhoux, F., & Jung, S. (2014). Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nature reviews. Immunology*, 14(6), 392–404. <https://doi.org/10.1038/nri3671>
25. Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M. B., Leboeuf, M., Becker, C. D., See, P., Price, J., Lucas, D., Greter, M., Mortha, A., Boyer, S. W., Forsberg, E. C., Tanaka, M., van Rooijen, N., García-Sastre, A., Stanley, E. R., Ginhoux, F., Frenette, P. S., ... Merad, M. (2013). Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity*, 38(4), 792–804. <https://doi.org/10.1016/j.immuni.2013.04.004>
26. Blériot, C., Chakarov, S., & Ginhoux, F. (2020). Determinants of Resident Tissue Macrophage Identity and Function. *Immunity*, 52(6), 957–970. <https://doi.org/10.1016/j.immuni.2020.05.014>
27. Coillard, A., & Segura, E. (2019). *In vivo* Differentiation of Human Monocytes. *Frontiers in immunology*, 10, 1907. <https://doi.org/10.3389/fimmu.2019.01907>
28. Kannan, N., & Eaves, C. J. (2018). Macrophages stimulate mammary stem cells. *Science*, 360(6396), 1401–1402. <https://doi.org/10.1126/science.aau1394>
29. Lin, E. Y., Nguyen, A. V., Russell, R. G., & Pollard, J. W. (2001). Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. *Journal of Experimental Medicine*, 193(6), 727–739. <https://doi.org/10.1084/jem.193.6.727>
30. Kacinski B. M. (1995). CSF-1 and its receptor in ovarian, endometrial and breast cancer. *Annals of medicine*, 27(1), 79–85. <https://doi.org/10.3109/07853899509031941>
31. Qian, B. Z., & Pollard, J. W. (2010). Macrophage Diversity Enhances Tumor Progression and Metastasis. *Cell*, 141(1), 39–51. <https://doi.org/10.1016/j.cell.2010.03.014>
32. Cassetta, L., & Pollard, J. W. (2018). Targeting macrophages: Therapeutic approaches in cancer. *Nature Reviews Drug Discovery*, 17(12), 887–904. <https://doi.org/10.1038/nrd.2018.169>
33. Williams, C. B., Yeh, E. S., & Soloff, A. C. (2016). Tumor-associated macrophages: Unwitting accomplices in breast cancer malignancy. *Npj Breast Cancer*, 2(1). <https://doi.org/10.1038/npjbcancer.2015.25>
34. van Dalen, F. J., van Stevendaal, M., Fennemann, F. L., Verdoes, M., & Ilina, O. (2018). Molecular Repolarisation of Tumour-Associated Macrophages. *Molecules (Basel, Switzerland)*, 24(1), 9. <https://doi.org/10.3390/molecules24010009>
35. Goswami, S., Sahai, E., Wyckoff, J. B., Cammer, M., Cox, D., Pixley, F. J., Stanley, E. R., Segall, J. E., & Condeelis, J. S. (2005). Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop. *Cancer research*, 65(12), 5278–5283. <https://doi.org/10.1158/0008-5472.CAN-04-1853>

36. Curiel, T. J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., ... Zou, W. (2004). Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nature Medicine*, 10(9), 942–949. <https://doi.org/10.1038/nm1093>
37. Stein, M., Keshav, S., Harris, N., & Gordon, S. (1992). Interleukin 4 potentially enhances murine macrophage mannose receptor activity: A marker of alternative immunologic macrophage activation. *Journal of Experimental Medicine*, 176(1), 287–292. <https://doi.org/10.1084/jem.176.1.287>
38. Mills, C. D., Kincaid, K., Alt, J. M., Heilman, M. J., & Hill, A. M. (2000). M-1/M-2 Macrophages and the Th1/Th2 Paradigm. *The Journal of Immunology*, 164(12), 6166–6173. <https://doi.org/10.4049/jimmunol.164.12.6166>
39. Mantovani, A., Sozzani, S., Locati, M., Allavena, P., & Sica, A. (2002). Macrophage polarization: Tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends in Immunology*, 23(11), 549–555. [https://doi.org/10.1016/S1471-4906\(02\)02302-5](https://doi.org/10.1016/S1471-4906(02)02302-5)
40. Verreck, F. A. W., De Boer, T., Langenberg, D. M. L., Hoeve, M. A., Kramer, M., Vaisberg, E., ... Ottenhoff, T. H. M. (2004). Human IL-23-producing type 1 macrophages promote but IL-10-producing type 2 macrophages subvert immunity to (myco)bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13), 4560–4565. <https://doi.org/10.1073/pnas.0400983101>
41. Martinez, F. O., Gordon, S., Locati, M., & Mantovani, A. (2006). Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. *The Journal of Immunology*, 177(10), 7303–7311. <https://doi.org/10.4049/jimmunol.177.10.7303>
42. de Waal Malefyt, R., Figdor, C. G., Huijbens, R., Mohan-Peterson, S., Bennett, B., Culpepper, J., Dang, W., Zurawski, G., & de Vries, J. E. (1993). Effects of IL-13 on phenotype, cytokine production, and cytotoxic function of human monocytes. Comparison with IL-4 and modulation by IFN-gamma or IL-10. *Journal of immunology (Baltimore, Md. : 1950)*, 151(11), 6370–6381.
43. Freedman, A. S., Freeman, G. J., Rhyhart, K., & Nadler, L. M. (1991). Selective induction of B7/BB-1 on interferon-gamma stimulated monocytes: a potential mechanism for amplification of T cell activation through the CD28 pathway. *Cellular immunology*, 137(2), 429–437. [https://doi.org/10.1016/0008-8749\(91\)90091-o](https://doi.org/10.1016/0008-8749(91)90091-o)
44. Buechler, C., Ritter, M., Orsó, E., Langmann, T., Klucken, J., & Schmitz, G. (2000). Regulation of scavenger receptor CD163 expression in human monocytes and macrophages by pro- and antiinflammatory stimuli. *Journal of Leukocyte Biology*, 67(1), 97–103. <https://doi.org/10.1002/jlb.67.1.97>
45. Genin, M., Clement, F., Fattaccioli, A., Raes, M., & Michiels, C. (2015). M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide. *BMC cancer*, 15, 577. <https://doi.org/10.1186/s12885-015-1546-9>
46. Gocheva, V., Wang, H. W., Gadea, B. B., Shree, T., Hunter, K. E., Garfall, A. L., ... Joyce, J. A. (2010). IL-4 induces cathepsin protease activity in tumor-

- associated macrophages to promote cancer growth and invasion. *Genes and Development*, 24(3), 241–255. <https://doi.org/10.1101/gad.1874010>
47. Yan, D., Wang, H. W., Bowman, R. L., & Joyce, J. A. (2016). STAT3 and STAT6 Signaling Pathways Synergize to Promote Cathepsin Secretion from Macrophages via IRE1 $\alpha$  Activation. *Cell Reports*, 16(11), 2914–2927. <https://doi.org/10.1016/j.celrep.2016.08.035>
  48. Gratchev, A., Guillot, P., Hakiy, N., Politz, O., Orfanos, C. E., Schledzewski, K., & Goerdts, S. (2001). Alternatively activated macrophages differentially express fibronectin and its splice variants and the extracellular matrix protein  $\beta$ IG-H3. *Scandinavian Journal of Immunology*, 53(4), 386–392. <https://doi.org/10.1046/j.1365-3083.2001.00885.x>
  49. Watanabe, K., Jose, P. J., & Rankin, S. M. (2002). Eotaxin-2 Generation Is Differentially Regulated by Lipopolysaccharide and IL-4 in Monocytes and Macrophages. *The Journal of Immunology*, 168(4), 1911–1918. <https://doi.org/10.4049/jimmunol.168.4.1911>
  50. Bonecchi, R., Sozzani, S., Stine, J. T., Luini, W., D'Amico, G., Allavena, P., ... Mantovani, A. (1998). Divergent Effects of Interleukin-4 and Interferon- $\gamma$  on Macrophage-Derived Chemokine Production: An Amplification Circuit of Polarized T Helper 2 Responses. *Blood*, 92(8), 2668–2671. <https://doi.org/10.1182/blood.v92.8.2668>
  51. Kodelja, V., Müller, C., Politz, O., Hakij, N., Orfanos, C. E., & Goerdts, S. (1998). Alternative macrophage activation-associated CC-chemokine-1, a novel structural homologue of macrophage inflammatory protein-1 alpha with a Th2-associated expression pattern. *Journal of immunology (Baltimore, Md. : 1950)*, 160(3), 1411–1418.
  52. Imai, T., Nagira, M., Takagi, S., Kakizaki, M., Nishimura, M., Wang, J., ... Yoshie, O. (1999). Selective recruitment of CCR4-bearing T(h)2 cells toward antigen-presenting cells by the CC chemokines thymus and activation-regulated chemokine and macrophage-derived chemokine. *International Immunology*, 11(1), 81–88. <https://doi.org/10.1093/intimm/11.1.81>
  53. Colotta, F., Saccani, S., Giri, J. G., Dower, S. K., Sims, J. E., Introna, M., & Mantovani, A. (1996). Regulated expression and release of the IL-1 decoy receptor in human mononuclear phagocytes. *Journal of immunology (Baltimore, Md. : 1950)*, 156(7), 2534–2541.
  54. Scotton, C. J., Martinez, F. O., Smelt, M. J., Sironi, M., Locati, M., Mantovani, A., & Sozzani, S. (2005). Transcriptional Profiling Reveals Complex Regulation of the Monocyte IL-1 $\beta$  System by IL-13. *The Journal of Immunology*, 174(2), 834–845. <https://doi.org/10.4049/jimmunol.174.2.834>
  55. Willment, J. A., Lin, H.-H., Reid, D. M., Taylor, P. R., Williams, D. L., Wong, S. Y. C., ... Brown, G. D. (2003). Dectin-1 Expression and Function Are Enhanced on Alternatively Activated and GM-CSF-Treated Macrophages and Are Negatively Regulated by IL-10, Dexamethasone, and Lipopolysaccharide. *The Journal of Immunology*, 171(9), 4569–4573. <https://doi.org/10.4049/jimmunol.171.9.4569>
  56. Corbett, J. A., Sweetland, M. A., Jr., J. R. L., & McDaniel, M. L. (1993). A 1-hour pulse with IL-1 $\beta$  induces formation of nitric oxide and inhibits insulin secretion by rat islets of Langerhans: evidence for a tyrosine kinase signaling



- mechanism. *The FASEB Journal*, 7(2), 369–374.  
<https://doi.org/10.1096/fasebj.7.2.8440413>
57. Georgoudaki, A. M., Prokopec, K. E., Boura, V. F., Hellqvist, E., Sohn, S., Östling, J., ... Karlsson, M. C. I. (2016). Reprogramming Tumor-Associated Macrophages by Antibody Targeting Inhibits Cancer Progression and Metastasis. *Cell Reports*, 15(9), 2000–2011.  
<https://doi.org/10.1016/j.celrep.2016.04.084>
  58. Terrazas, L. I., Montero, D., Terrazas, C. A., Reyes, J. L., & Rodríguez-Sosa, M. (2005). Role of the programmed Death-1 pathway in the suppressive activity of alternatively activated macrophages in experimental cysticercosis. *International Journal for Parasitology*, 35(13), 1349–1358.  
<https://doi.org/10.1016/j.ijpara.2005.06.003>
  59. Goswami, S., Sahai, E., Wyckoff, J. B., Cammer, M., Cox, D., Pixley, F. J., Stanley, E. R., Segall, J. E., & Condeelis, J. S. (2005). Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop. *Cancer research*, 65(12), 5278–5283.  
<https://doi.org/10.1158/0008-5472.CAN-04-1853>
  60. Kerkelä, E., Ala-aho, R., Klemi, P., Grénman, S., Shapiro, S. D., Kähäri, V. M., & Saarialho-Kere, U. (2002). Metalloelastase (MMP-12) expression by tumour cells in squamous cell carcinoma of the vulva correlates with invasiveness, while that by macrophages predicts better outcome. *Journal of Pathology*, 198(2), 258–269. <https://doi.org/10.1002/path.1198>
  61. Nielsen, B. S., Sehested, M., Kjeldsen, L., Borregaard, N., Rygaard, J., & Danø, K. (1997). Expression of matrix metalloprotease-9 in vascular pericytes in human breast cancer. *Laboratory investigation; a journal of technical methods and pathology*, 77(4), 345–355.
  62. Nielsen, B. S., Timshel, S., Kjeldsen, L., Sehested, M., Pyke, C., Borregaard, N., & Danø, K. (1996). 92 kDa type IV collagenase (MMP-9) is expressed in neutrophils and macrophages but not in malignant epithelial cells in human colon cancer. *International Journal of Cancer*, 65(1), 57–62.  
[https://doi.org/10.1002/\(SICI\)1097-0215\(19960103\)65:1<57::AID-IJC10>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0215(19960103)65:1<57::AID-IJC10>3.0.CO;2-F)
  63. Coussens, L. M., Tinkle, C. L., Hanahan, D., & Werb, Z. (2000). MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. *Cell*, 103(3), 481–490. [https://doi.org/10.1016/s0092-8674\(00\)00139-2](https://doi.org/10.1016/s0092-8674(00)00139-2)
  64. Pyke, C., Graem, N., Ralfkiaer, E., Rønne, E., Høyer-Hansen, G., Brønner, N., & Danø, K. (1993). Receptor for urokinase is present in tumor-associated macrophages in ductal breast carcinoma. *Cancer research*, 53(8), 1911–1915.
  65. Ding, M., Fu, X., Tan, H., Wang, R., Chen, Z., & Ding, S. (2012). The effect of vascular endothelial growth factor C expression in tumor-associated macrophages on lymphangiogenesis and lymphatic metastasis in breast cancer. *Molecular Medicine Reports*, 6(5), 1023–1029.  
<https://doi.org/10.3892/mmr.2012.1043>
  66. Naeim, F. (2008). Principles of Immunophenotyping. *Hematopathology*, 27–55.  
<https://doi.org/10.1016/B978-0-12-370607-2.00002-8>
  67. Schupp, J., Krebs, F. K., Zimmer, N., Trzeciak, E., Schuppan, D., & Tuettgenberg, A. (2019). Targeting myeloid cells in the tumor sustaining

- microenvironment. *Cellular immunology*, 343, 103713.  
<https://doi.org/10.1016/j.cellimm.2017.10.013>
68. Edwards, J. P., Zhang, X., Frauwirth, K. A., & Mosser, D. M. (2006). Biochemical and functional characterization of three activated macrophage populations. *Journal of leukocyte biology*, 80(6), 1298–1307.  
<https://doi.org/10.1189/jlb.0406249>
69. Mosser, D. M., & Edwards, J. P. (2008). Exploring the full spectrum of macrophage activation. *Nature reviews. Immunology*, 8(12), 958–969.  
<https://doi.org/10.1038/nri2448>
70. Xue, J., Schmidt, S. V., Sander, J., Draffehn, A., Krebs, W., Quester, I., De Nardo, D., Gohel, T. D., Emde, M., Schmidleithner, L., Ganesan, H., Nino-Castro, A., Mallmann, M. R., Labzin, L., Theis, H., Kraut, M., Beyer, M., Latz, E., Freeman, T. C., Ulas, T., ... Schultze, J. L. (2014). Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity*, 40(2), 274–288. <https://doi.org/10.1016/j.immuni.2014.01.006>
71. Franklin, R. A., Liao, W., Sarkar, A., Kim, M. V., Bivona, M. R., Liu, K., Pamer, E. G., & Li, M. O. (2014). The cellular and molecular origin of tumor-associated macrophages. *Science (New York, N.Y.)*, 344(6186), 921–925.  
<https://doi.org/10.1126/science.1252510>
72. Chimal-Ramírez, G. K., Espinoza-Sánchez, N. A., Chávez-Sánchez, L., Arriaga-Pizano, L., & Fuentes-Pananá, E. M. (2016). Monocyte Differentiation towards Protumor Activity Does Not Correlate with M1 or M2 Phenotypes. *Journal of immunology research*, 2016, 6031486. <https://doi.org/10.1155/2016/6031486>
73. Zhang, L., Li, Z., Skrzypczynska, K. M., Fang, Q., Zhang, W., O'Brien, S. A., ... Yu, X. (2020). Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. *Cell*, 181(2), 442-459.e29.  
<https://doi.org/10.1016/j.cell.2020.03.048>
74. Azizi, E., Carr, A. J., Plitas, G., Cornish, A. E., Konopacki, C., Prabhakaran, S., ... Pe'er, D. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell*, 174(5), 1293-1308.e36.  
<https://doi.org/10.1016/j.cell.2018.05.060>
75. Zhang, Q., He, Y., Luo, N., Patel, S. J., Han, Y., Gao, R., ... Zhang, Z. (2019). Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. *Cell*, 179(4), 829-845.e20. <https://doi.org/10.1016/j.cell.2019.10.003>
76. Sathe, A., Grimes, S. M., Lau, B. T., Chen, J., Suarez, C., Huang, R. J., ... Ji, H. P. (2020). Single-Cell Genomic Characterization Reveals the Cellular Reprogramming of the Gastric Tumor Microenvironment. *Clinical Cancer Research*, 26(11), 2640–2653. <https://doi.org/10.1158/1078-0432.CCR-19-3231>
77. Müller, S., Kohanbash, G., Liu, S. J., Alvarado, B., Carrera, D., Bhaduri, A., ... Diaz, A. (2017). Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biology*, 18(1), 1–14.  
<https://doi.org/10.1186/s13059-017-1362-4>
78. Peng, J., Sun, B. F., Chen, C. Y., Zhou, J. Y., Chen, Y. S., Chen, H., ... Wu, W. (2019). Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Research*, 29(9), 725–738. <https://doi.org/10.1038/s41422-019-0195-y>

79. Friebel, E., Kapolou, K., Unger, S., Núñez, N. G., Utz, S., Rushing, E. J., ... Becher, B. (2020). Single-Cell Mapping of Human Brain Cancer Reveals Tumor-Specific Instruction of Tissue-Invading Leukocytes. *Cell*, *181*(7), 1626-1642.e20. <https://doi.org/10.1016/j.cell.2020.04.055>
80. Wagner, J., Rapsomaniki, M. A., Chevrier, S., Anzeneder, T., Langwieder, C., Dykgers, A., ... Bodenmiller, B. (2019). A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell*, *177*(5), 1330-1345.e18. <https://doi.org/10.1016/j.cell.2019.03.005>
81. Chung, W., Eum, H. H., Lee, H. O., Lee, K. M., Lee, H. B., Kim, K. T., Ryu, H. S., Kim, S., Lee, J. E., Park, Y. H., Kan, Z., Han, W., & Park, W. Y. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nature communications*, *8*, 15081. <https://doi.org/10.1038/ncomms15081>
82. Ilicic, T., Kim, J. K., Kolodziejczyk, A. A., Bagger, F. O., McCarthy, D. J., Marioni, J. C., & Teichmann, S. A. (2016). Classification of low quality cells from single-cell RNA-seq data. *Genome biology*, *17*, 29. <https://doi.org/10.1186/s13059-016-0888-1>
83. Lun, A. T., McCarthy, D. J., & Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research*, *5*, 2122. <https://doi.org/10.12688/f1000research.9501.2>
84. Mabbott, N. A., Baillie, J. K., Brown, H., Freeman, T. C., & Hume, D. A. (2013). An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC genomics*, *14*, 632. <https://doi.org/10.1186/1471-2164-14-632>
85. Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R. P., Wolters, P. J., Abate, A. R., Butte, A. J., & Bhattacharya, M. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature immunology*, *20*(2), 163–172. <https://doi.org/10.1038/s41590-018-0276-y>
86. Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*, *9*, 559. <https://doi.org/10.1186/1471-2105-9-559>
87. Hänzelmann, S., Castelo, R., & Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC bioinformatics*, *14*, 7. <https://doi.org/10.1186/1471-2105-14-7>
88. Collin, M., & Bigley, V. (2018). Human dendritic cell subsets: an update. *Immunology*, *154*(1), 3–20. <https://doi.org/10.1111/imm.12888>
89. Segura, E. (2016). Review of mouse and human dendritic cell subsets. *Methods in Molecular Biology*, *1423*, 3–15. [https://doi.org/10.1007/978-1-4939-3606-9\\_1](https://doi.org/10.1007/978-1-4939-3606-9_1)
90. Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., Ceccarelli, M., Bontempi, G., & Noushmehr, H. (2016). TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic acids research*, *44*(8), e71. <https://doi.org/10.1093/nar/gkv1507>

91. 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>
92. 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>
93. Therneau, T.M. survival: A Package for Survival Analysis in R. R package versión 3.2-7. CRAN (2020). Disponible en: <https://CRAN.R-project.org/package=survival>. (Consultado 19 febrero 2021).
94. Yuan, X., Wang, J., Huang, Y., Shangguan, D., & Zhang, P. (2021). Single-Cell Profiling to Explore Immunological Heterogeneity of Tumor Microenvironment in Breast Cancer. *Frontiers in immunology*, 12, 643692. <https://doi.org/10.3389/fimmu.2021.643692>
95. Savas, P., Virassamy, B., Ye, C., Salim, A., Mintoff, C. P., Caramia, F., Salgado, R., Byrne, D. J., Teo, Z. L., Dushyanthen, S., Byrne, A., Wein, L., Luen, S. J., Poliness, C., Nightingale, S. S., Skandarajah, A. S., Gyorki, D. E., Thornton, C. M., Beavis, P. A., Fox, S. B., ... Loi, S. (2018). Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nature medicine*, 24(7), 986–993. <https://doi.org/10.1038/s41591-018-0078-7>
96. Lu, Y., Zhao, Q., Liao, J. Y., Song, E., Xia, Q., Pan, J., Li, Y., Li, J., Zhou, B., Ye, Y., Di, C., Yu, S., Zeng, Y., & Su, S. (2020). Complement Signals Determine Opposite Effects of B Cells in Chemotherapy-Induced Immunity. *Cell*, 180(6), 1081–1097.e24. <https://doi.org/10.1016/j.cell.2020.02.015>
97. Bao, X., Shi, R., Zhao, T., Wang, Y., Anastasov, N., Rosemann, M., & Fang, W. (2021). Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels tumour heterogeneity plus M2-like tumour-associated macrophage infiltration and aggressiveness in TNBC. *Cancer immunology, immunotherapy : CII*, 70(1), 189–202. <https://doi.org/10.1007/s00262-020-02669-7>
98. Geissmann, F., Manz, M. G., Jung, S., Sieweke, M. H., Merad, M., & Ley, K. (2010). Development of monocytes, macrophages, and dendritic cells. *Science (New York, N.Y.)*, 327(5966), 656–661. <https://doi.org/10.1126/science.1178331>
99. Liu, K., & Nussenzweig, M. C. (2010). Origin and development of dendritic cells. *Immunological reviews*, 234(1), 45–54. <https://doi.org/10.1111/j.0105-2896.2009.00879.x>
100. Coillard, A., & Segura, E. (2019). *In vivo* Differentiation of Human Monocytes. *Frontiers in immunology*, 10, 1907. <https://doi.org/10.3389/fimmu.2019.01907>
101. Jakubzick, C., Gautier, E. L., Gibbings, S. L., Sojka, D. K., Schlitzer, A., Johnson, T. E., Ivanov, S., Duan, Q., Bala, S., Condon, T., van Rooijen, N., Grainger, J. R., Belkaid, Y., Ma'ayan, A., Riches, D. W., Yokoyama, W. M., Ginhoux, F., Henson, P. M., & Randolph, G. J. (2013). Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to

lymph nodes. *Immunity*, 39(3), 599–610.  
<https://doi.org/10.1016/j.immuni.2013.08.007>

- 102.** Desch, A. N., Gibbings, S. L., Goyal, R., Kolde, R., Bednarek, J., Bruno, T., Slansky, J. E., Jacobelli, J., Mason, R., Ito, Y., Messier, E., Randolph, G. J., Prabagar, M., Atif, S. M., Segura, E., Xavier, R. J., Bratton, D. L., Janssen, W. J., Henson, P. M., & Jakubzick, C. V. (2016). Flow Cytometric Analysis of Mononuclear Phagocytes in Nondiseased Human Lung and Lung-Draining Lymph Nodes. *American journal of respiratory and critical care medicine*, 193(6), 614–626. <https://doi.org/10.1164/rccm.201507-1376OC>
- 103.** Zigmund, E., Varol, C., Farache, J., Elmaliah, E., Satpathy, A. T., Friedlander, G., Mack, M., Shpigel, N., Boneca, I. G., Murphy, K. M., Shakhar, G., Halpern, Z., & Jung, S. (2012). Ly6C hi monocytes in the inflamed colon give rise to proinflammatory effector cells and migratory antigen-presenting cells. *Immunity*, 37(6), 1076–1090. <https://doi.org/10.1016/j.immuni.2012.08.026>
- 104.** Sponaas, A. M., Freitas do Rosario, A. P., Voisine, C., Mastelic, B., Thompson, J., Koernig, S., Jarra, W., Renia, L., Mauduit, M., Potocnik, A. J., & Langhorne, J. (2009). Migrating monocytes recruited to the spleen play an important role in control of blood stage malaria. *Blood*, 114(27), 5522–5531. <https://doi.org/10.1182/blood-2009-04-217489>
- 105.** Langlet, C., Tamoutounour, S., Henri, S., Luche, H., Ardouin, L., Grégoire, C., Malissen, B., & Williams, M. (2012). CD64 expression distinguishes monocyte-derived and conventional dendritic cells and reveals their distinct role during intramuscular immunization. *Journal of immunology (Baltimore, Md. : 1950)*, 188(4), 1751–1760. <https://doi.org/10.4049/jimmunol.1102744>
- 106.** Flores-Langarica, A., Marshall, J. L., Bobat, S., Mohr, E., Hitchcock, J., Ross, E. A., Coughlan, R. E., Khan, M., Van Rooijen, N., Henderson, I. R., Maclennan, I. C., & Cunningham, A. F. (2011). T-zone localized monocyte-derived dendritic cells promote Th1 priming to Salmonella. *European journal of immunology*, 41(9), 2654–2665. <https://doi.org/10.1002/eji.201141440>
- 107.** Jakubzick, C. V., Randolph, G. J., & Henson, P. M. (2017). Monocyte differentiation and antigen-presenting functions. *Nature reviews. Immunology*, 17(6), 349–362. <https://doi.org/10.1038/nri.2017.28>
- 108.** Lavin, Y., Kobayashi, S., Leader, A., Amir, E. D., Elefant, N., Bigenwald, C., Remark, R., Sweeney, R., Becker, C. D., Levine, J. H., Meinhof, K., Chow, A., Kim-Shulze, S., Wolf, A., Medaglia, C., Li, H., Rytlewski, J. A., Emerson, R. O., Solovyov, A., Greenbaum, B. D., ... Merad, M. (2017). Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell*, 169(4), 750–765.e17. <https://doi.org/10.1016/j.cell.2017.04.014>
- 109.** Newman, A. C., Nakatsu, M. N., Chou, W., Gershon, P. D., & Hughes, C. C. (2011). The requirement for fibroblasts in angiogenesis: fibroblast-derived matrix proteins are essential for endothelial cell lumen formation. *Molecular biology of the cell*, 22(20), 3791–3800. <https://doi.org/10.1091/mbc.E11-05-0393>
- 110.** Bergers G, Song S. The role of pericytes in blood-vessel formation and maintenance. *Neuro Oncol.* 2005 Oct;7(4):452-64. doi: 10.1215/S1152851705000232. PMID: 16212810; PMCID: PMC1871727.

111. Kunz-Schughart LA, Weber A, Rehli M, Gottfried E, Brockhoff G, Krause SW, et al. [The “classical” macrophage marker CD68 is strongly expressed in primary human fibroblasts]. *Verh Dtsch Ges Pathol* (2003) 87:215–23.
112. Zhang, W., & Liu, H. T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell research*, 12(1), 9–18. <https://doi.org/10.1038/sj.cr.7290105>
113. Schröfelbauer, B., & Hoffmann, A. (2011). How do pleiotropic kinase hubs mediate specific signaling by TNFR superfamily members?. *Immunological reviews*, 244(1), 29–43. <https://doi.org/10.1111/j.1600-065X.2011.01060.x>
114. Arthur, J. S., & Ley, S. C. (2013). Mitogen-activated protein kinases in innate immunity. *Nature reviews. Immunology*, 13(9), 679–692. <https://doi.org/10.1038/nri3495>
115. Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2016). GenBank. *Nucleic acids research*, 44(D1), D67–D72. <https://doi.org/10.1093/nar/gkv1276>
116. Roumenina, L. T., Daugan, M. V., Petitprez, F., Sautès-Fridman, C., & Fridman, W. H. (2019). Context-dependent roles of complement in cancer. *Nature reviews. Cancer*, 19(12), 698–715. <https://doi.org/10.1038/s41568-019-0210-0>
117. Reinke, J. M., & Sorg, H. (2012). Wound repair and regeneration. *European surgical research. Europäische chirurgische Forschung. Recherches chirurgicales europeennes*, 49(1), 35–43. <https://doi.org/10.1159/000339613>
118. Lamort, A. S., Giopanou, I., Psallidas, I., & Stathopoulos, G. T. (2019). Osteopontin as a Link between Inflammation and Cancer: The Thorax in the Spotlight. *Cells*, 8(8), 815. <https://doi.org/10.3390/cells8080815>
119. Patten, J., & Wang, K. (2021). Fibronectin in development and wound healing. *Advanced drug delivery reviews*, 170, 353–368. <https://doi.org/10.1016/j.addr.2020.09.005>
120. Ramu, I., Buchholz, S. M., Patzak, M. S., Goetze, R. G., Singh, S. K., Richards, F. M., Jodrell, D. I., Sipos, B., Ströbel, P., Ellenrieder, V., Hessmann, E., & Neesse, A. (2019). SPARC dependent collagen deposition and gemcitabine delivery in a genetically engineered mouse model of pancreas cancer. *EBioMedicine*, 48, 161–168. <https://doi.org/10.1016/j.ebiom.2019.09.024>
121. Fonović, M., & Turk, B. (2014). Cysteine cathepsins and extracellular matrix degradation. *Biochimica et biophysica acta*, 1840(8), 2560–2570. <https://doi.org/10.1016/j.bbagen.2014.03.017>
122. Lund, S. A., Giachelli, C. M., & Scatena, M. (2009). The role of osteopontin in inflammatory processes. *Journal of cell communication and signaling*, 3(3-4), 311–322. <https://doi.org/10.1007/s12079-009-0068-0>
123. Elgundi, Z., Papanicolaou, M., Major, G., Cox, T. R., Melrose, J., Whitelock, J. M., & Farrugia, B. L. (2020). Cancer Metastasis: The Role of the Extracellular Matrix and the Heparan Sulfate Proteoglycan Perlecan. *Frontiers in oncology*, 9, 1482. <https://doi.org/10.3389/fonc.2019.01482>
124. Icer, M. A., & Gezmen-Karadag, M. (2018). The multiple functions and mechanisms of osteopontin. *Clinical biochemistry*, 59, 17–24. <https://doi.org/10.1016/j.clinbiochem.2018.07.003>

125. Morath, I., Hartmann, T. N., & Orian-Rousseau, V. (2016). CD44: More than a mere stem cell marker. *The international journal of biochemistry & cell biology*, 81(Pt A), 166–173. <https://doi.org/10.1016/j.biocel.2016.09.009>
126. Danielsson, F., Peterson, M. K., Caldeira Araújo, H., Lautenschläger, F., & Gad, A. (2018). Vimentin Diversity in Health and Disease. *Cells*, 7(10), 147. <https://doi.org/10.3390/cells7100147>
127. Kellouche, S., Mourah, S., Bonnefoy, A., Schoëvaert, D., Podgorniak, M. P., Calvo, F., ... Dosquet, C. (2007). Platelets, thrombospondin-1 and human dermal fibroblasts cooperate for stimulation of endothelial cell tubulogenesis through VEGF and PAI-1 regulation. *Experimental Cell Research*, 313(3), 486–499. <https://doi.org/10.1016/j.yexcr.2006.10.023>
128. Paunescu, V., Bojin, F. M., Tatu, C. A., Gavriiuc, O. I., Rosca, A., Gruia, A. T., ... Vermesan, S. (2011). Tumour-associated fibroblasts and mesenchymal stem cells: More similarities than differences. *Journal of Cellular and Molecular Medicine*, 15(3), 635–646. <https://doi.org/10.1111/j.1582-4934.2010.01044.x>
129. Gibbings, S. L., Goyal, R., Desch, A. N., Leach, S. M., Prabagar, M., Atif, S. M., Bratton, D. L., Janssen, W., & Jakubzick, C. V. (2015). Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. *Blood*, 126(11), 1357–1366. <https://doi.org/10.1182/blood-2015-01-624809>
130. Zhang, M. J., Ntranos, V., & Tse, D. (2020). Determining sequencing depth in a single-cell RNA-seq experiment. *Nature communications*, 11(1), 774. <https://doi.org/10.1038/s41467-020-14482-y>
131. Xu, G., Liu, Y., Li, H., Liu, L., Zhang, S., & Zhang, Z. (2020). Dissecting the human immune system with single cell RNA sequencing technology. *Journal of leukocyte biology*, 107(4), 613–623. <https://doi.org/10.1002/JLB.5MR1019-179R>
132. Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., Bharadwaj, R., ... Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature communications*, 8, 14049. <https://doi.org/10.1038/ncomms14049>
133. Pathria, P., Louis, T. L., & Varner, J. A. (2019). Targeting Tumor-Associated Macrophages in Cancer. *Trends in immunology*, 40(4), 310–327. <https://doi.org/10.1016/j.it.2019.02.003>
134. Mantovani, A., Marchesi, F., Malesci, A., Laghi, L., & Allavena, P. (2017). Tumour-associated macrophages as treatment targets in oncology. *Nature reviews. Clinical oncology*, 14(7), 399–416. <https://doi.org/10.1038/nrclinonc.2016.217>
135. Zhan, X., Jia, L., Niu, Y., Qi, H., Chen, X., Zhang, Q., Zhang, J., Wang, Y., Dong, L., & Wang, C. (2014). Targeted depletion of tumour-associated macrophages by an alendronate-glucomannan conjugate for cancer immunotherapy. *Biomaterials*, 35(38), 10046–10057. <https://doi.org/10.1016/j.biomaterials.2014.09.007>
136. Sawa-Wejksza, K., & Kandefers-Szerszeń, M. (2018). Tumor-Associated Macrophages as Target for Antitumor Therapy. *Archivum immunologiae et*

*therapiae experimentalis*, 66(2), 97–111. <https://doi.org/10.1007/s00005-017-0480-8>

137. Lum, H. D., Buhtoiarov, I. N., Schmidt, B. E., Berke, G., Paulnock, D. M., Sondel, P. M., & Rakhmievich, A. L. (2006). Tumoristatic effects of anti-CD40 mAb-activated macrophages involve nitric oxide and tumour necrosis factor-alpha. *Immunology*, 118(2), 261–270. <https://doi.org/10.1111/j.1365-2567.2006.02366.x>
138. Wiehagen, K. R., Girgis, N. M., Yamada, D. H., Smith, A. A., Chan, S. R., Grewal, I. S., Quigley, M., & Verona, R. I. (2017). Combination of CD40 Agonism and CSF-1R Blockade Reconditions Tumor-Associated Macrophages and Drives Potent Antitumor Immunity. *Cancer immunology research*, 5(12), 1109–1121. <https://doi.org/10.1158/2326-6066.CIR-17-0258>
139. Conde, J., Bao, C., Tan, Y., Cui, D., Edelman, E. R., Azevedo, H. S., Byrne, H. J., Artzi, N., & Tian, F. (2015). Dual targeted immunotherapy via *in vivo* delivery of biohybrid RNAi-peptide nanoparticles to tumour-associated macrophages and cancer cells. *Advanced functional materials*, 25(27), 4183–4194. <https://doi.org/10.1002/adfm.201501283>
140. Ding, L., Liang, G., Yao, Z., Zhang, J., Liu, R., Chen, H., Zhou, Y., Wu, H., Yang, B., & He, Q. (2015). Metformin prevents cancer metastasis by inhibiting M2-like polarization of tumor associated macrophages. *Oncotarget*, 6(34), 36441–36455. <https://doi.org/10.18632/oncotarget.5541>