



**Universidad
Europea**

Máster en Bioinformática
ESTUDIO DE RUTAS DE
SEÑALIZACIÓN EN CÁNCER DE
OVARIO A NIVEL CELULAR

Autor: Marta Rodríguez Martínez

Tutor: Matías Marín Falco

Curso 2023-24

*“La ciencia siempre vale la pena, porque sus descubrimientos, tarde o temprano,
siempre se aplican”. Severo Ochoa.*

ÍNDICE GENERAL

Índice de ilustraciones	3
Índice de tablas	3
Resumen	4
Abstract	5
Introducción	6
1. CÁNCER DE OVARIO SEROSO DE ALTO GRADO (HGSOC)	6
1.1 Introducción al cáncer de ovario	6
1.1.1 Definición y tipos histológicos	6
1.1.2 Epidemiología, factores de riesgo e importancia clínica	8
1.2 Características moleculares del HGSOC	9
1.3 Rutas de señalización y relevancia en HGSOC	10
2. SECUENCIACIÓN DE ANÁLISIS DE scRNA- seq	11
2.1 Introducción a la secuenciación de ARN de células individuales (scRNA-seq)	11
2.1.1 Historia y evolución de las tecnologías de secuenciación de ARN	11
2.1.2 Procedimiento de scRNA-seq	12
2.1.3 Limitaciones desafíos y futuro en scRNA-seq	14
2.2 Impacto de scRNA- seq en el estudio del cáncer	15
2.3 Aplicación de scRNA-seq en HGSOC	16
3. BIOINFORMÁTICA/ BIOLOGÍA DE SISTEMAS	16
Hipótesis	19
Objetivos	19
Metodología	20
Resultados y Discusión	28
Conclusiones	40
Conclusions	41
Bibliografía	43
Declaración obligatoria del uso de herramientas de IA	47

Índice de ilustraciones

Ilustración 1	29
Ilustración 2	30
Ilustración 3	31
Ilustración 4	32
Ilustración 5	33
Ilustración 6	34
Ilustración 7	36
Ilustración 8	37

Índice de tablas

Tabla 1	28
---------	----

Resumen

El cáncer de ovario seroso de alto grado es el tipo más prevalente y mortal de cáncer de ovario. La quimiorresistencia ayuda a identificar las vías de señalización más significativas que contribuyan a la resistencia de las células tumorales a los tratamientos quimioterapéuticos. Mediante el análisis de datos de secuenciación de ARN de células individuales de pacientes antes y después del tratamiento quimioterapéutico se exploran cambios transcripcionales inducidos por la quimioterapia en las células tumorales y su interacción con el microambiente tumoral. En el análisis empleamos un enfoque de preprocesamiento de datos usando Seurat para filtrar y normalizar las cuentas de UMI, identificar tipos celulares, y aplicar reducción de dimensionalidad (PCA y UMAP) para agrupar células según sus perfiles de expresión génica. Además realizamos un análisis de expresión diferencial y de rutas de señalización mediante PROGENy para investigar las vías relevantes en la resistencia a la quimioterapia enfocándonos en las células epiteliales cancerosas.

Palabras clave: Cáncer de ovario seroso de alto grado; Quimiorresistencia; Secuenciación de ARN de células individuales;; Microambiente tumoral; Rutas de señalización

Abstract

High grade serous ovarian cancer is the most prevalent and deadly type of ovarian cancer.

Chemoresistance helps identify the most significant signaling pathways that contribute to the resistance of tumor cells to chemotherapeutic treatments. By analyzing single -cell RNA sequencing data from patients before and after chemotherapy treatment, chemotherapy - induced transcriptional changes in tumor cells and their interaction with the tumor microenvironment are explored. In the analysis we employed a data preprocessing approach using Seurat to filter and normalize UMI counts, identify cell types, and apply dimensionality reduction (PCA and UMAP) to cluster cells based on their gene expression profiles. We further performed differential expression and signaling pathway analysis using PROGENy to investigate relevant pathways in chemotherapy resistance focusing on cancer epithelial cells.

Keywords: High grade serous ovarian cancer; Chemoresistance; Single cell RNA sequencing; Tumor microenvironment; Signaling routes

Introducción

1. CÁNCER DE OVARIO SEROSO DE ALTO GRADO (HGSOC)

1.1 Introducción al cáncer de ovario

1.1.1 Definición y tipos histológicos

El cáncer de ovario continúa siendo una de las enfermedades más letales y desafiantes en el ámbito de oncología ginecológica (Momenimovahed *et al.*, 2019) que se caracteriza por la heterogeneidad tanto del microambiente tumoral como de las células tumorales. El diagnóstico tardío debido al crecimiento asintomático del tumor, las inefectivas pruebas de detección y la elevada resistencia a la quimioterapia junto con la frecuencia de recaídas complican su tratamiento (Yoneda *et al.*, 2012).

El cáncer de ovario se clasifica principalmente en tres tipos según su origen histológico: el cáncer epitelial, el cáncer de células germinales y el cáncer de cordones sexuales y estroma ovárico especializado. El primero de estos es el más común, mientras que los dos últimos tipos son menos frecuentes y solamente representan alrededor del 5% de todos los casos de cáncer de ovario (Stewart *et al.*, 2019).

Dentro de los tumores epiteliales, el carcinoma seroso ha sido tradicionalmente clasificado en tres grados, pero la clasificación de la OMS ha simplificado esta estructura, dividiéndolo en carcinoma seroso de alto grado (HGSOC, por sus siglas en inglés, High Grade Serous Ovarian Carcinoma) y carcinoma seroso de bajo grado (LGSOC, Low Grade Serous Ovarian Carcinoma) (Travis and Bramilla, 2014). Estas dos entidades son consideradas neoplasias distintas con diferencias moleculares significativas, lo que tiene importantes implicaciones en el desarrollo de tratamientos dirigidos específicos para cada tipo (LI *et al.*, 201; Herrera, 2018).

El carcinoma seroso de bajo grado tiende a ser estables, a nivel molecular, presentan una activación constante en la vía MAPK, lo que conlleva frecuentes mutaciones en genes como KRAS y BRAF entre otros (Castellanos, 2018).

El cáncer de ovario seroso de alto grado, conocido también como carcinoma seroso de alto grado es el tipo más prevalente y mortal de cáncer de ovario. Su origen puede ser el ovario, la trompa de Falopio o el peritoneo (Kim et al., 2018). La mayoría de los casos presentan metástasis peritoneales, complicando la identificación de su origen primario (Kim *et al.*, 2019).

Históricamente, se pensaba que el epitelio de la superficie ovárica era el principal sitio de origen del HGSC. Sin embargo, más recientemente se ha implantado la idea de que el epitelio de la trompa de Falopio es un precursor significativo, especialmente el carcinoma intraepitelial tubárico seroso (STIC). Este último se considera una lesión premaligna que podría progresar a HGSC. La clave para su diagnóstico es la identificación de un alto índice mitótico de las células que lo componen (Castellanos, 2018).

Entre el 60 y el 80% de los pacientes con cáncer de ovario seroso de alto grado responden positivamente al tratamiento inicial con cisplatino, sin embargo, a pesar de esta respuesta inicial positiva, con el tiempo, las células tumorales se adaptan haciendo que el platino deje de ser efectivo (Matthews *et al.*, 2021).

La complejidad de los mecanismos de resistencia al cisplatino en el tratamiento del cáncer de ovario seroso de alto grado (HGSOC) sugiere una falta de caracterización de la enfermedad y la necesidad de una mejor búsqueda de estrategias terapéuticas y tratamientos dirigidos (Freimund *et al.*, 20118).

Por ende, la quimiorresistencia representa un obstáculo importante en el tratamiento de HGSOC.

El uso de inhibidores de PARP es un avance para el tratamiento del HGSOC, especialmente en paciente con mutaciones BRCA o HRD, ofreciendo un mantenimiento efectivo después de la quimioterapia, aún así, los desafíos persisten incluyendo la necesidad de pruebas genéticas para optimizar el tratamiento y el

estudio de combinaciones de terapias para mejorar la respuesta en pacientes que no responden bien a los tratamientos actuales (Mahmood *et al.*, 2020).

1.1.2 Epidemiología, factores de riesgo e importancia clínica

La incidencia del cáncer de ovario varía significativamente en diferentes regiones del mundo, dependiendo de factores como la ubicación geográfica, la etnia y el desarrollo económico (Lisio *et al.*, 2019).

Según la Organización Mundial de la Salud (OMS), se estima que cada año se diagnostican alrededor de 225.500 nuevos casos de cáncer de ovario, y aproximadamente 140.200 mujeres fallecen debido a esta enfermedad. A nivel mundial, es la séptima causa más común de cáncer en mujeres y la octava causa principal de muerte por cáncer. En los países occidentales, el cáncer de ovario ocupa el quinto lugar entre las causas más frecuentes de muerte por cáncer en mujeres (Lisio *et al.*, 2019).

En cuanto al cáncer de ovario seroso de alto grado representa más del 60% de los cánceres de ovario epiteliales y es responsable de más del 70% de las muertes por cáncer de ovario (Lisio *et al.*, 2019).

El cáncer de ovario es más común en mujeres mayores de 65 años y es más frecuente después de la menopausia. La genética es un factor importante, tener antecedentes familiares de cáncer de ovario o de mama aumenta el riesgo y las mutaciones en los genes BRCA1 y BRCA2 (Kurman, 2013).

Además de los factores genéticos, el riesgo de desarrollar cáncer de ovario también se asocia con ciertos factores de estilo de vida, un inicio de la menstruación temprano y una menopausia tardía pueden aumentar el riesgo debido a la ovulación incesante. Factores ginecológicos como la endometriosis pueden aumentar el riesgo, especialmente para ciertos tipos de cáncer de ovario (Momenimovahed *et al.*, 2019). El embarazo y la maternidad ofrecen protección y reducen el riesgo de presentar cáncer de ovario (Gaona, 2020) Otros factores de riesgo incluyen la

obesidad, el tabaquismo y el uso de polvos de talco en el área perineal (Webb and Jordan 2017, Lisio *et al.*, 2019).

El tratamiento estándar del cáncer de ovario es la cirugía, para remover la mayor cantidad posible del tumor, aunque también se utiliza la quimioterapia basada en platino (Kelland, 2007) y más recientemente, la inmunoterapia.

1.2 Características moleculares del HGSOC

Las principales alteraciones genómicas se incluyen en las vías clave implicadas en la resistencia al cáncer como son las mutaciones en TP53, BRCA1 y BRCA2.

Más del 96% de los casos de HGSOC presenta mutaciones en TP53, lo que indica que esta alteración es fundamental para el desarrollo de la enfermedad (Lliberos *et al.*, 2024).

Además de TP53 los pacientes que albergan mutaciones de la línea germinal en BRCA1 y BRCA2 tienen un riesgo mucho mayor de desarrollar HGSOC, así como en otros genes menos frecuentes como CSMD3, NF1, CDK12 y RB1.

Otras vías moleculares alteradas en HGSOC incluyen mutaciones en genes involucrados en la señalización celular como KRAS, BRAF y PIK3CA, que afectan el crecimiento y la supervivencia celular. Aunque estas mutaciones no son tan comunes como las de TP53 y BRCA, contribuyen a la heterogeneidad de los tumores de ovario y pueden influir en el comportamiento clínico de la enfermedad (Lisio *et al.*, 2019).

En los estudios recientes, se ha revelado que el HGSOC no es una única entidad, sino que presenta una diversidad molecular significativa. Un estudio clave, realizado por Tothill *et al.*, identificó cuatro subtipos moleculares principales de HGSOC, los cuales fueron denominados C1, C2, C4 y C5. El subtipo C1 está relacionado con una firma estromal reactiva y un mal pronóstico, mientras que el C2 mostró una mayor supervivencia al igual que C4, gracias a la activación de genes asociados con el sistema inmune. Sin embargo, C5, mostró características mesenquimales y una supervivencia inferior (Lisio *et al.*, 2019).

1.3 Rutas de señalización y relevancia en HGSOC

Un análisis integral de las mutaciones, las alteraciones del número de copias y los cambios en la expresión génica ha identificado diversas rutas de señalización clave en la patogénesis del HGSOC.

Las rutas de señalización en el cáncer de ovario seroso de alto grado (HGSOC) son clave para entender su progresión y desarrollar terapias dirigidas.

Entre las rutas más afectadas se encuentra la vía PI3K/AKT, causada por mutaciones o amplificaciones de genes como PIK3CA y AKT2, que están asociadas con una peor supervivencia, ya que está estrechamente vinculada a la resistencia a la quimioterapia. Esta ruta regula el crecimiento, la proliferación y la supervivencia celular. Los inhibidores de PI3K, AKT y mTOR han mostrado promesas en ensayos clínicos, especialmente en combinación con otros tratamientos como los PARPi (Llberos *et al.*, 2024).

La ruta de recombinación homóloga (HR) que repara rupturas de doble cadena de ADN se encuentra alterada en el cáncer de ovario principalmente debido a mutaciones en los genes BRCA1 y BRCA2. Estos defectos hacen que las células dependan de mecanismos alternativos, como la reparación por escisión de bases (BER), lo que ha permitido el desarrollo de tratamientos con inhibidores de PARP (PARPi) como olaparib y niraparib. Estos fármacos bloquean la BER, causando una acumulación de daño irreparable en células tumorales deficientes en BRCA, lo que lleva a su muerte (Llberos *et al.*, 2024).

El análisis de redes reguladoras de genes reveló que la vía JUN podría ser un factor clave en HGSOC. JUN forma parte del complejo AP-1, que regula la actividad de otros genes y, resultados en este estudio sugieren que JUN podría ser un motor importante de la carcinogénesis en el cáncer de ovario (Hao *et al.*, 2021).

2. SECUENCIACIÓN DE ANÁLISIS DE scRNA- seq

2.1 Introducción a la secuenciación de ARN de células individuales (scRNA-seq)

2.1.1 Historia y evolución de las tecnologías de secuenciación de ARN

La secuenciación de análisis de células individuales ha emergido como una herramienta poderosa para estudiar la heterogeneidad celular en cánceres sólidos, permitiendo la identificación de subpoblaciones celulares que podrían desempeñar un papel en la resistencia a los tratamientos. Este enfoque no solo proporciona información detallada sobre la expresión génica a nivel celular, sino que también permite el estudio de interacciones celulares complejas y redes de señalización.

A diferencia de las técnicas convencionales de secuenciación de ARN, que miden la expresión génica promedio en una población mixta de células, scRNA-seq permite analizar cambios transcripcionales y de señalización a nivel de una única célula, proporcionando una resolución sin precedentes para caracterizar tanto las células tumorales como las del microambiente (Jovic *et al.*, 2022).

Los primeros estudios de células individuales surgieron en la década de 1990, impulsados por el interés en la plasticidad celular y su funcionalidad variable según la localización. La secuenciación masiva (NGS) permitió un análisis cuantitativo del contenido celular, aunque las técnicas iniciales de secuenciación de ARN (bulk RNA-seq) no lograron distinguir los diferentes tipos de células presentes en una muestra. El desarrollo de tecnologías que pueden analizar grandes cantidades de células simultáneamente ha sido esencial para el avance de la transcriptómica de células individuales (Carangelo *et al.*, 2022).

2.1.2 Procedimiento de scRNA-seq

A día de hoy, el procedimiento experimental de scRNA-seq presenta un análisis estandarizado. La tecnología de scRNA-seq implica a grandes rasgos la separación de células individuales, seguida de la captura y secuenciación del ARN mensajero (ARNm) presente en cada célula.

Los protocolos de scRNA-seq involucran las siguientes etapas: aislamiento de células individuales de los tejidos, lisis celular para obtener el ARNm, retrotranscripción del ARN a ADN complementario (ADNc), amplificación del ADNc mediante reacción en cadena de la polimerasa (PCR) o por transcripción in vitro (IVT) y finalmente realizar la construcción y secuenciación de la biblioteca (Lu *et al.*, 2023).

Control de calidad

El primer paso en el análisis de datos scRNA-seq es el filtrado de calidad de las celdas. Este proceso implica eliminar celdas que tienen una baja calidad de secuenciación, que pueden ser causadas por la baja cantidad de material de ARN, contaminación o problemas técnicos. Se utilizan métricas como el número total de cuentas, la proporción de cuentas de genes mitocondriales y el número de genes detectados por célula para establecer criterios de inclusión y exclusión.

Normalización

Proceso fundamental para corregir las variaciones técnicas que pueden influir en los resultados.

Transformación

Después de la normalización, los datos generalmente se transforman mediante un enfoque logarítmico para estabilizar la varianza y facilitar comparaciones entre genes. Esta transformación permite que los datos sean más adecuadamente interpretados y analizados en etapas posteriores del proceso.

Imputación

Los conjuntos de datos scRNA-seq suelen tener una gran cantidad de valores faltantes, lo que puede distorsionar los resultados del análisis. Estos métodos utilizan la correlación entre genes para estimar estos valores faltantes, sin embargo, esta puede introducir ruidos y sesgos en los datos.

Selección de características y reducción de dimensionalidad

Dado que un conjunto de datos puede incluir miles de genes, es esencial realizar una reducción de dimensionalidad para facilitar su análisis. Esto se logra mediante

la identificación de genes altamente variables (HVG), que son aquellos que muestran la mayor variabilidad en su expresión. Herramientas como Seurat y Scanpy permiten seleccionar entre 1.000 y 1.500 HVG para su uso en el análisis posterior.

El análisis de Componentes Principales (PCA) es una técnica comúnmente utilizada para reducir la dimensionalidad y resumir la variabilidad correlacionadas en un número menor de variables no correlacionadas, llamadas componentes principales, preservando la mayor parte de la variación original.

Visualización y agrupamiento

La visualización de datos scRNA-seq se realiza a menudo utilizando t-SNE o UMAP, que permiten representar gráficamente la similitud entre células. Los algoritmos de agrupamiento se utilizan para identificar grupos de células con perfiles de expresión similares, lo que revela diferentes tipos celulares.

Anotación de clústeres

Una vez que se han definido los clústeres, es crucial identificar las poblaciones celulares mediante la detección de firmas genéticas. Esto se logra a través de pruebas de expresión diferencial que comparan genes entre los clústeres y el resto del conjunto de datos.

Análisis de expresión genética

Finalmente, el análisis de expresión genética permite profundizar en las diferencias moleculares entre las celdas. Se utilizan métodos de expresión diferencial y análisis de enriquecimiento de conjuntos de genes (GSEA) para identificar patrones de expresión relacionados con diferentes condiciones experimentales.

2.1.3 Limitaciones desafíos y futuro en scRNA-seq

A pesar de sus múltiples contribuciones, scRNA-seq también enfrenta limitaciones, como la dificultad para detectar mutaciones somáticas y la sensibilidad a la abundancia de transcripciones.

Además, en cuanto a la comprensión y clasificación de los tipos celulares tampoco están exentos de estos desafíos, ya que el proceso de disociar tejidos en células individuales no es sencillo. La disociación celular implica exponer los tejidos, y por ende las células, a una combinación de estrés isquémico, mecánico y enzimático. Las células reaccionan a estos tipos de estrés, lo cual puede influir en sus perfiles de expresión génica (Jovic *et al.*, 2022). Además, existe un sesgo de supervivencia, dado que las células que mantienen su integridad son las que se seleccionan preferentemente para el proceso de preparación de bibliotecas, lo que significa que las células más frágiles o comprometidas tienen menos probabilidad de aparecer en los datos finales.

El análisis de la comunicación celular en scRNA-Seq se basa mayormente en la detección de moléculas señalizadoras (como factores de crecimiento, citocinas, quimiocinas) y sus receptores en distintas células, para inferir redes de interacción celular. No obstante, este enfoque tiene limitaciones debido a la calidad y profundidad de los datos obtenidos, lo que lleva a análisis fundamentados en correlaciones, más que en evidencia directa de causalidad. Además, cuando múltiples señales y receptores están involucrados en diferentes poblaciones celulares, los métodos actuales de análisis como Cell-Talk no siempre logran capturar estas interacciones completamente complejas.

Sin embargo, la combinación de secuenciación de ARN de células individuales (scRNA-seq) con la secuenciación del transcriptoma espacial promete revolucionar

el estudio del microambiente tumoral (TME) al ofrecer una comprensión más detallada de las interacciones celulares (Ahmed *et al.*, 2022). Estos enfoques permiten no solo identificar la heterogeneidad celular dentro de los tumores, sino también cómo estas células se comunican y se reorganizan espacialmente, lo que es crucial para entender la progresión del tumor y la respuesta a los tratamientos.

La integración de estos datos multiescalares abre nuevas posibilidades en la identificación de subpoblaciones celulares específicas que contribuyen a la resistencia a la quimioterapia, una de las principales barreras en el tratamiento de HGSOE.

2.2 Impacto de scRNA- seq en el estudio del cáncer

El uso de la secuenciación de células individuales en la investigación del cáncer de ovario ha permitido una caracterización detallada de los patrones de expresión génica a nivel celular, lo que ha revelado las diferentes poblaciones celulares presentes en el microambiente tumoral, como las células cancerosas, estromales e inmunes. Este mapeo de alta precisión ha sido clave para entender cómo estas células contribuyen a la progresión del tumor, la invasión, la metástasis y la resistencia a los tratamientos.

Además, al ofrecer información sobre las rutas de señalización y la interacción entre células, esta técnica ayuda a identificar nuevos blancos terapéuticos y desarrollar enfoques de medicina.

La heterogeneidad en el cáncer de ovario se manifiesta en múltiples niveles siendo uno de los más relevantes la variabilidad molecular entre diferentes poblaciones celulares dentro del TME. Esta tecnología ha demostrado ser útil en la identificación de nuevos biomarcadores terapéuticos en el TME.

Las características del TME pueden variar significativamente entre los diferentes tipos de cáncer y son moldeadas por las interacciones entre las células tumorales, las células que rodean el tumor y las células del huésped que lo infiltran (Parry, 2023).

2.3 Aplicación de scRNA-seq en HGSOC

La secuenciación de ARN de células individuales (scRNA-seq) es una herramienta poderosa para desvelar detalles del microambiente tumoral (TME) en el cáncer seroso de ovario de alto grado (HGSOC) (Liu., *et al* 2022).

Hay un estudio que reveló empleando scRNA-seq que la inhibición de la vía JAK/STAT ejercía fuertes efectos antitumorales (Izar *et al.*, 2020). Recientemente se describió diversos fenotipos celulares estromales en la regulación del TME en HGSOC, como fibroblastos asociados al cáncer (CAF) mediados por TGF β , células endoteliales linfáticas y células mesoteliales (Olbrecht *et al.*, 2021, Xu *et al.*, 2022).

Un enfoque para modificar el TME que ha sido estudiado en inmunoterapias para cáncer de ovario, pero con resultados limitados consiste en influir en los comportamientos de las células inmunes, como la regulación de PD-L1, un importante regulador de la supresión de las células T (Mahoney *et al.*, 2015; Xu *et al.*, 2022).

Xu *et al* estudió el panorama transcriptómico unicelular de tumores HGSOC en etapas clínicas tempranas y avanzadas utilizando análisis de pseudotiempo y describiendo el desarrollo de células epiteliales en diferentes fases tumorales identificando además características relacionadas con la transición epitelio-mesénquima lo que podría contribuir al desarrollo de nuevas estrategias clínicas para el tratamiento del HGSOC.

Por lo tanto, es imprescindible identificar características celulares en distintas etapas del HGSOC mediante scRNA-seq, lo que ofrece información más precisa para desarrollar inmunoterapias dirigidas a fases específicas de la enfermedad.

3. BIOINFORMÁTICA/ BIOLOGÍA DE SISTEMAS

La biología de sistemas ha surgido como un campo interdisciplinario que mezcla la biología experimental y ciencias cuantitativas para desarrollar modelos integrales de procesos celulares, tanto en condiciones normales como patológicas. Entre las herramientas más empleadas en este ámbito, destacan las redes, que permiten representar las interacciones entre diferentes entidades biológicas de manera estructurada y cuantitativa, lo que ayuda en el desarrollo de nuevas estrategias terapéuticas.

Este enfoque es particularmente relevante en el cáncer, donde las redes de señalización celular se ven alteradas de manera compleja y diversa. La identificación de estas alteraciones específicas permite el diseño de terapias dirigidas, como los inhibidores de PARP en tumores con mutaciones en BRCA1/2, que han demostrado mejorar la respuesta terapéutica en pacientes seleccionados.

Actualmente el reto de esta ciencia radica en integrar grandes volúmenes de datos generados a partir de un solo tumor y traducirlos en estrategias terapéuticas personalizadas con potencial uso clínico. Las terapias dirigidas ofrecen una gran promesa para el tratamiento individualizado del cáncer, sin embargo, hay varios obstáculos que superar aún (Ram *et al.*, 2012).

Un desafío fundamental radica en cómo las redes de señalización interactúan de manera dinámica y cómo los tumores desarrollan mecanismos de resistencia a estos tratamientos dirigidos. La inhibición de una única vía de señalización como mTOR, puede activar otras rutas compensatorias o mecanismos de retroalimentación que favorecen la progresión tumoral, lo que ha llevado a resultados inesperados y adversos en ciertos pacientes. Por ello, la biología de sistemas propone un enfoque más holístico, integrando múltiples capas de información como datos genómicos, transcriptómicos y proteómicos para crear modelos predictivos más robustos que ayuden a anticipar las respuestas terapéuticas y los posibles mecanismos de resistencia.

En conclusión, es trascendental comprender cómo funcionan las redes de señalización en el cáncer para identificar los mejores objetivos terapéuticos basados en las alteraciones específicas de cada tumor.

La capacidad de interpretar de manera precisa estos datos “ómicos” y traducirlos en intervenciones clínicas personalizadas depende en gran medida del desarrollo de tecnologías avanzadas de bioinformática. El análisis computacional permite la integración de datos de diversas fuentes y la identificación de patrones moleculares clave, facilitando la selección de combinaciones terapéuticas que apunten a múltiples vías al mismo tiempo, reduciendo así las probabilidades de resistencia.

A medida que estas estrategias se perfeccionan, la biología de sistemas promete ser un pilar fundamental en la creación de tratamientos cada vez más personalizados y eficaces para el cáncer incluyendo el HGSOC.

Masood et al estudió las subpoblaciones de células T en pacientes con cáncer de ovario seroso de alto grado (HGSOC) y observó que la variabilidad en cómo responden a la inmunoterapia está vinculada a ciertos patrones de expresión genética y funciones celulares. Esto implica que identificar estas subpoblaciones podría ayudar a predecir la efectividad de los tratamientos, lo que permitiría desarrollar estrategias terapéuticas más adaptadas a las necesidades individuales de los pacientes (Masood *et al.*, 2020).

La selección del tratamiento más adecuado sigue siendo compleja debido a la falta de marcadores predictivos efectivos, para paliar este desafío recientemente se ha desarrollado una herramienta computacional para pacientes con cáncer de ovario seroso de alto grado denominada ALISON (digitAI twIn Simulator Ovarian cancer), un innovador modelo computacional en sílico diseñado para estudiar el comportamiento celular simulando la heterogeneidad celular donde interactúan células normales y cancerosas. Al modelar tanto poblaciones celulares homogéneas como heterogéneas, el estudio demostró que la heterogeneidad celular es más representativa de las fases avanzadas de la enfermedad, mientras que las heterogéneas reflejan mejor la complejidad de las etapas más avanzadas, donde la resistencia a la quimioterapia suele ser mayor.

El estudio también simuló tratamientos combinados con platino y paclitaxel, dos de los principales fármacos utilizados en el tratamiento del HGSOC. Las simulaciones revelan diferencias en la respuesta al tratamiento entre los distintos pacientes,

particularmente en aquellos con cáncer en etapas intermedias, lo que podría reflejar una transición entre estados de sensibilidad y resistencia a los medicamentos. Dado que el estudio es muy reciente se necesita una validación más exhaustiva para confirmar la utilidad y precisión de ALISON en escenarios clínicos y experimentales (Cortesi *et al.*, 2024).

Hipótesis

La resistencia a la quimioterapia en el cáncer de ovario seroso de alto grado (HGSOC) está mediada por una combinación de cambios transcripcionales inducidos por el tratamiento en las células tumorales y alteraciones en las interacciones con el microambiente tumoral. La identificación de subpoblaciones celulares específicas y rutas moleculares activadas post-tratamiento permitirá el desarrollo de biomarcadores predictivos y nuevas estrategias terapéuticas para superar la quimiorresistencia.

Objetivos

Objetivo Primario:

Identificar patrones transcripcionales específicos y subpoblaciones celulares que contribuyan a la resistencia a la quimioterapia en HGSOC mediante análisis de datos de secuenciación de ARN de células individuales antes y después del tratamiento.

Objetivos Secundarios:

-Analizar la heterogeneidad celular pre y post-quimioterapia para identificar rutas de señalización celulares asociadas con la resistencia.

-Caracterizar los cambios transcripcionales específicos en células tumorales, estromales e inmunes inducidos por la quimioterapia

Metodología

Información de muestras de pacientes

Los pacientes que participaron en este estudio clínico dieron su consentimiento informado por escrito y fueron tratados en el Hospital Universitario de Turku, Finlandia, como parte del proyecto DECIDER (NCT04846933).

Las muestras clínicas utilizadas en este estudio provienen de perfiles de expresión de células individuales obtenidos mediante secuenciación de alto rendimiento (scRNA-seq). Estas muestras prospectivas corresponden a tejidos derivados de 11 pacientes con cáncer de ovario seroso de alto grado (HGSOC) antes y después de la quimioterapia.

Preprocesamiento y filtrado de datos scRNA-seq reales con Seurat

Se utilizó el software R studio, varios paquetes R descargados de las versiones disponibles para la versión de R (4.4.1) incluido el paquete Seurat (versión 5.1.0) https://satijalab.org/seurat/articles/install_v5.

El procedimiento tiene varias etapas claves que antes de aplicar filtros más específicos para realizar la asignación del tipo celular, primero debemos realizar un análisis exploratorio.

Paso 1: Carga de datos

El primer paso en el análisis de datos scRNA-seq es la carga de los archivos.

Utilizamos dos archivos principales: uno con las cuentas de UMI (unidad de molécula individual) (GSE165897_UMIcounts_HGS0C.tsv.gz) y otro con información

asociada a cada célula (GSE165897_cellinfo_HGS0C.tsv.gz). Estos archivos fueron descomprimidos y cargados en el entorno de R utilizando la función *fread()* del paquete *data-table*, que permite manejar archivos grandes de manera eficiente.

Se tomó un subconjunto de los datos UMI para trabajar con las primeras 1000 filas.

A continuación, se llevó a cabo un procesamiento básico de los nombres de los genes, donde se reemplazaron guiones bajos por guiones comunes, un paso importante para garantizar la consistencia de los nombres en futuras manipulaciones de datos. Además, se revisó la presencia de valores faltantes (NA) en la matriz de expresión génica, que es común en este tipo de datos, y se reemplazaron con ceros para evitar errores en el análisis posterior.

Creamos un objeto *Seurat v5*, un paquete de R disponible en CRAN para todas las plataformas, diseñado para manejar datos de scRNA-seq en el que contenga las cuentas UMI para cada célula, así como los metadatos asociados. Para ello, los datos de UMI fueron convertidos en una matriz (para que los genes correspondan a filas y las células a columnas). La matriz resultante fue utilizada como entrada para la creación del objeto *Seurat*, lo que permitió manejar los datos de manera eficiente para el análisis posterior. Tras cargar la información adicional de cada célula se añadieron los metadatos al objeto *Seurat* para poder ser empleados posteriormente en el análisis.

Paso 2: Filtrado de células de baja calidad y dobletes

El primer filtrado elimina células de baja calidad, que pueden ser identificadas por la alta proporción de genes mitocondriales y dobletes (dos o más células atrapadas en una única gota), basándonos en combinaciones de marcadores celulares específicos.

Este filtrado elimina las células con combinaciones de marcadores contradictorios (indicativo de dobletes) y las que tienen más del 15% de sus UMI provenientes de genes mitocondriales.

Un paso importante en el preprocesamiento de datos de scRNA-seq es el cálculo del porcentaje de genes mitocondriales. Las células que muestran un porcentaje elevado de transcripción mitocondrial suelen estar en proceso de muerte celular, por lo que se calculó el porcentaje de genes mitocondriales utilizando la función de *PercentageFeatureSet()*. Este cálculo fue agregado como un nuevo metadato en el objeto Seurat, proporcionando un criterio adicional para el filtrado de células.

Paso 3: Identificación preliminar de tipos celulares

Normalización de los datos

Después de realizar el filtrado, el siguiente paso fue la normalización de los datos. Los datos de expresión se normalizaron utilizando la función *NormalizeData()* del paquete Seurat, que aplica una transformación logarítmica para ajustar las cuentas UMI entre células y hacer los datos comparables. Este proceso es clave para corregir las diferencias en el número total de cuentas entre células y permitir la identificación de patrones transcripcionales de manera robusta.

Identificación de características variables

Posteriormente, se llevó a cabo la identificación de características variables entre las células mediante la función *FindVariableFeatures()*. Este paso identificó los genes cuya expresión varía más entre las células, lo que es indicativo de su relevancia en el proceso biológico que se está estudiando. Se seleccionan los 2.000 genes más variables para su análisis en profundidad ya que estos genes son los más probables de estar implicados en las diferencias fenotípicas entre las células, especialmente en lo que respecta a la resistencia a la quimioterapia en el cáncer de HGSOC.

Escalado de datos

En el análisis de secuenciación de ARN de células individuales (scRNA-seq), el escalado es una etapa clave para preparar los datos para técnicas de reducción de dimensionalidad y agrupamiento. Utilizamos la función *ScaleData()* del paquete Seurat para centrar y escalar los valores de expresión génica en cada célula. Esto significa que, para cada gen, se restablecerá la media de las expresiones de ese gen en todas las células y luego se dividirá por la desviación estándar. Este proceso asegura que los genes con diferentes rangos de expresión sean comparables, lo que es especialmente útil en el análisis de componentes principales (PCA) que se realiza a continuación. El escalado es necesario porque algunos algoritmos de reducción de dimensionalidad, como el PCA, son sensibles a las magnitudes de los datos y podrían sesgar los resultados hacia genes con mayores niveles de expresión.

Paso 4: Reducción de Dimensionalidad mediante PCA

Análisis de PCA

Una vez escalados los datos, se procede a realizar el análisis de componentes principales (PCA). El PCA es una técnica de reducción de dimensionalidad que permite resumir las variaciones más importantes en los datos en un conjunto reducido de “componentes principales”. Se aplicó la función *RunPCA()* sobre las características variables previamente identificadas, utilizando solo los genes que presentaban mayor variabilidad entre las células para asegurar que el PCA captura las diferencias biológicamente relevantes. El resultado del PCA nos proporciona un conjunto de componentes principales que explican la mayor parte de la variación en los datos. Se examinan los primeros cinco componentes para evaluar su contribución a la variabilidad general y visualizar los genes que más influyen en ellos. Los primeros componentes generalmente capturan las diferencias más significativas entre las células, lo que nos permite identificar patrones de expresión subyacentes relacionados con la heterogeneidad celular, en este caso, asociado a la resistencia a la quimioterapia.

Búsqueda de vecinos más cercanos

Una vez que tenemos los resultados del PCA, el siguiente paso es realizar el agrupamiento de las células. Primero, se calculan los vecinos más cercanos para cada célula utilizando la función *FindNeighbors()*. Este paso crea una estructura de gráfico, donde cada célula está conectada con sus vecinos más similares en función de las primeras dimensiones del PCA (en este caso, se utilizan las primeras 10 dimensiones). La identificación de los vecinos más cercanos (SNN) es fundamental para realizar el posterior agrupamiento o “clustering” de las células.

Agrupamiento de células

Después de definir los vecinos, se procede a agrupar las células en clústeres con la función *FindClusters()*, que utiliza un algoritmo de agrupamiento de gráficos basado en la optimización de la modularidad (algoritmo de Lovaina). Se especifica una resolución de 0.5, que controla el tamaño y número de clústeres; una resolución más alta produciría más clústeres, mientras que una resolución más baja agruparía las células en menos clústeres.

El resultado es un conjunto de comunidades (clústeres) que representan grupos de células con perfiles de expresión génica similares. Estos clústeres pueden corresponder a diferentes subpoblaciones celulares o estados celulares. Finalmente, se visualizan las primeras filas de los identificadores *Idents()* de los clústeres para confirmar que se han formado correctamente.

En este paso, el algoritmo Louvain identificó tres comunidades en las células analizadas. El número de nodos (células) y las conexiones entre ellos son generados a partir del grafo de vecinos más cercanos, permitiendo dividir las células en clústeres coherentes.

Visualización de resultados

Una vez que se han identificado los clusters, se visualizan los resultados de la reducción de dimensionalidad mediante PCA utilizando la función *DimPlot()*. Esta visualización permite observar cómo se agrupan las células en el espacio de reducción y proporciona información sobre la heterogeneidad celular en el conjunto de datos.

Se genera un gráfico de codo (ElbowPlot) para ayudar a determinar el número óptimo de clusters en el análisis. Este gráfico permite observar la variación explicada por cada componente principal y es útil para identificar un punto donde agregar más clusters no mejora significativamente el modelo.

Paso 5: Análisis UMAP para Visualización

Para interpretar visualmente los clústeres generadores en un espacio de menor dimensión, se utilizó la técnica UMAP (Uniform Manifold Approximation and Projection). UMAP es una herramienta de reducción de dimensionalidad que proyecta datos de alta dimensión en un espacio de dos dimensiones, preservando tanto la estructura global como local de los datos. Es útil especialmente en scRNA-seq porque facilita la visualización de clústeres celulares en gráficos bidimensionales.

Hemos utilizado este código : `DimPlot(seurat_obj, reduction = "umap", group.by = "cell_type", label = TRUE) + NoLegend()` para visualizar el gráfico diferenciado por tipo de célula (EOC, Immune y Stromal).

Paso 6: Análisis de Clústers

Conteo de célula por clúster y tipo celular

Para obtener el número de células en cada clúster y su tipo celular se generó una tabla de frecuencia con la función: `table(seurat_obj$seurat_clusters, seurat_obj$cell_type)`

Paso 7: Análisis de expresión diferencial

Identificación de marcadores específicos para cada tipo celular

Para caracterizar los clústeres y vincularlos con tipos celulares específicos, se realizó una identificación de marcadores genéticos específicos en cada clúster utilizando la función *FindAllMarkers*, lo que permitió seleccionar los genes diferencialmente expresados en cada clúster capturando todos los marcadores característicos.

Paso 8: Aplicación de filtros de calidad específicos

Asignación de etiquetas a los clústeres basada en marcadores conocidos

Agregamos filtros para las células epiteliales cancerosas basados en los umbrales definidos para `nCount_RNA`, `nFeature_RNA`, `percent.mt` (porcentaje de UMI mitocondriales) (Zhang *et al.*, 2022).

Este es el código que utilizamos:

```
eoc_cells <- subset(seurat_obj, subset = cell_type == "EOC" & nCount_RNA > 8192  
& nFeature_RNA > 1552 & percent.mt < 12).
```

Los 10 genes con mayor expresión se seleccionaron mediante la función: `top10 <- cluster_markers %>% group_by(cluster) %>% top_n(10, avg_log2FC)`. Estos genes se visualizaron en un Heatmap mediante la función *DoHeatmap*.

Paso 9: Filtración células epiteliales cancerosas

En nuestro caso usamos la información de la columna `cell_type` y `cell_subtype` de nuestro data frame `cell_info` para seleccionar solamente las células epiteliales cancerosas (EOC).

Realizamos el preprocesamiento de los datos en las células filtradas que son las que nos interesan para proceder a estudiar las rutas de señalización y visualizamos

las células epiteliales en el UMAP del objeto completo con los filtros anteriormente explicados

Al realizar el análisis de expresión diferencial filtramos y seleccionamos solo las células epiteliales cancerosas (EOC) capturando así las dinámicas relevantes en respuesta a la quimioterapia

Paso 10: Análisis de rutas de señalización

Mediante PROGENy (Pathways RespOnsive GENes) estimamos la actividad de las vías de señalización relevantes basándonos en firmas genéticas de consenso obtenidas a partir de experimentos de perturbación. Esta herramienta bioinformática es diseñada para inferir la actividad de distintas vías de señalización en datos transcriptómicos, tanto de muestras masivas como de células individuales.

La implementación de PROGENy se realizó en R, utilizando el paquete disponible en Bioconductor (Schubert *et al.*, 2018) <https://saezlab.github.io/progeny/index.html>.

En nuestro caso, queremos analizar y visualizar los puntajes de actividad de las vías de señalización (pathways) en células epiteliales (EOC) antes y después del tratamiento.

El procedimiento sería: aislar las células epiteliales, calcular los puntajes de progeny, escalar los puntajes de progeny, transformar a un Dataframe, unión con la información de fase de tratamiento para analizar diferencias entre las condiciones, resumir los puntajes de Progeny por vía y fase de tratamiento calculando el promedio y la desviación estándar y crear un mapa de calor o heatmap en el que se muestra la actividad de las diferentes vías de señalización en células epiteliales diferenciando entre las fases de tratamiento.

Paso 11: Ajustar un modelo lineal

Finalmente para conocer que pathways están diferencialmente activados de manera significativa probamos a realizar una regresión lineal. Un p-valor menor a 0.05

indicaría que la actividad de un pathway específico cambia significativamente entre las fases de tratamiento.

El modelo lineal se ajusta usando la variable `treatment_phase` como variable independiente y `avg` como variable dependiente de puntuación promedio para cada vía de señalización (pathway) comparando las puntuaciones promedio entre las dos fases de tratamiento (`treatment-naive` y `post-NACT`).

Si no hay suficientes observaciones o el modelo no es significativo los resultados se retornan como NA. Los modelos de regresión requieren una cantidad mínima de datos para calcular los parámetros de manera confiable.

Al tener únicamente dos observaciones (una para cada fase de tratamiento) nos viene a decir que los modelos no tienen suficientes grados de libertad para calcular los coeficientes correctamente y procedemos a realizar método no paramétrico como la prueba de Wilcoxon.

Resultados y Discusión

ANÁLISIS DE COMUNIDADES MEDIANTE EL ALGORITMO DE LOUVAIN

Utilizando el optimizador de modularidad (versión 1.3.0) desarrollado por Ludo Waltman y Nees Jan van Eck, hemos identificado 9 clusters o grupos celulares, los cuales están asociados con los tipos celulares EOC (células epiteliales cancerosas), Immune (células inmunes) y Stromal (células estromales).

A continuación, para obtener el número de células por clústers se utilizó la función: `table(seurat_obj$seurat_clusters, seurat_obj$cell_type)`.

	EOC	Immune	Stromal
0	4	13812	21
1	5	11556	19
2	4	7604	13
3	41	34	5299
4	4238	18	37
5	7	8	2630
6	2504	0	1
7	1984	3	2
8	19	1900	23

Tabla 1. Número de células por clusters

ANÁLISIS DE REDUCCIÓN DE DIMENSIONALIDAD USANDO UMAP

DimPlot ha generado un gráfico de dispersión (scatter plot) donde cada punto representa una célula y los colores representan diferentes clústers o grupos celulares.

Nos muestra nueve grupos distintos en los datos que están agrupados compartiendo características funcionales o de tipo celular (Figura 1).

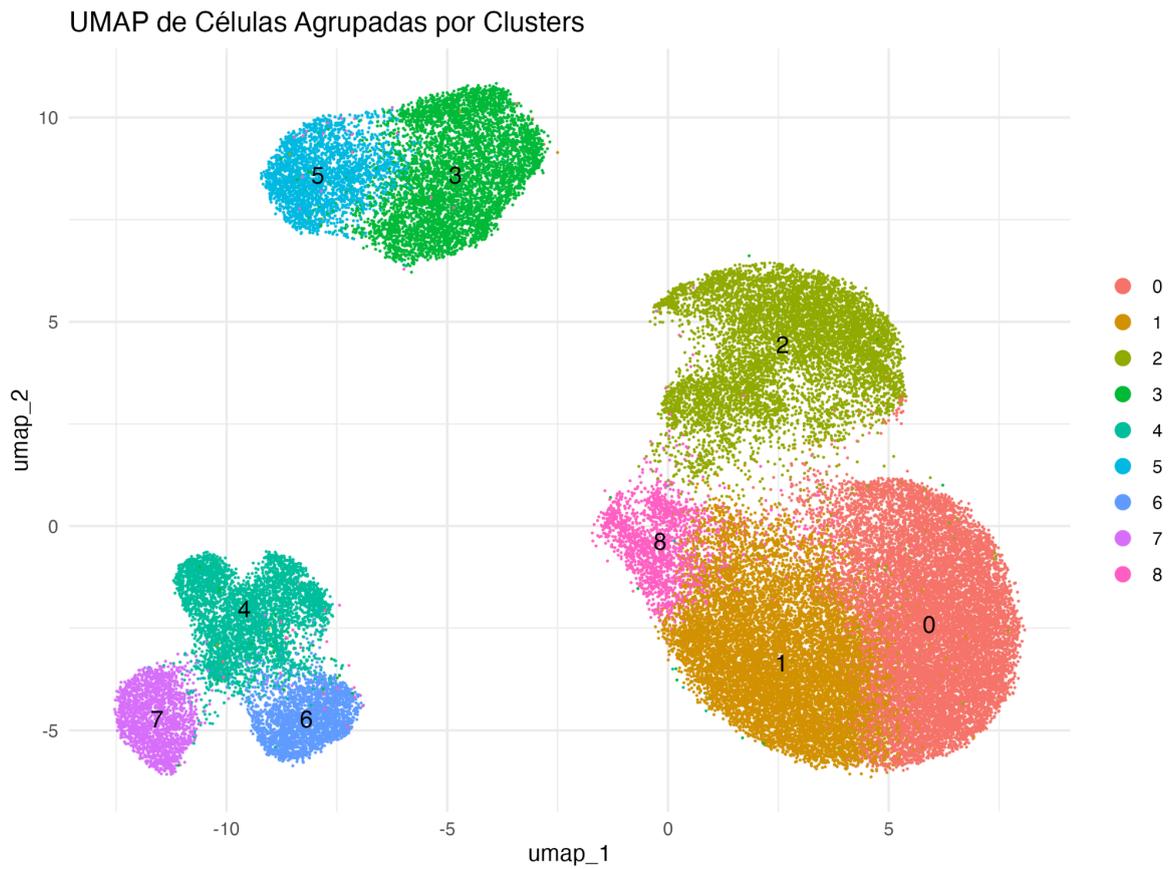


Figura 1. UMAP de células tumorales, estromales e inmunes en HGSOc agrupadas en 9 comunidades.

También podemos visualizar el UMAP separado por tipo celular diferenciándose en células epiteliales cancerosas, estromales o inmunes facilitando la comparación de la distribución de los tipos celulares en el mismo espacio tridimensional.

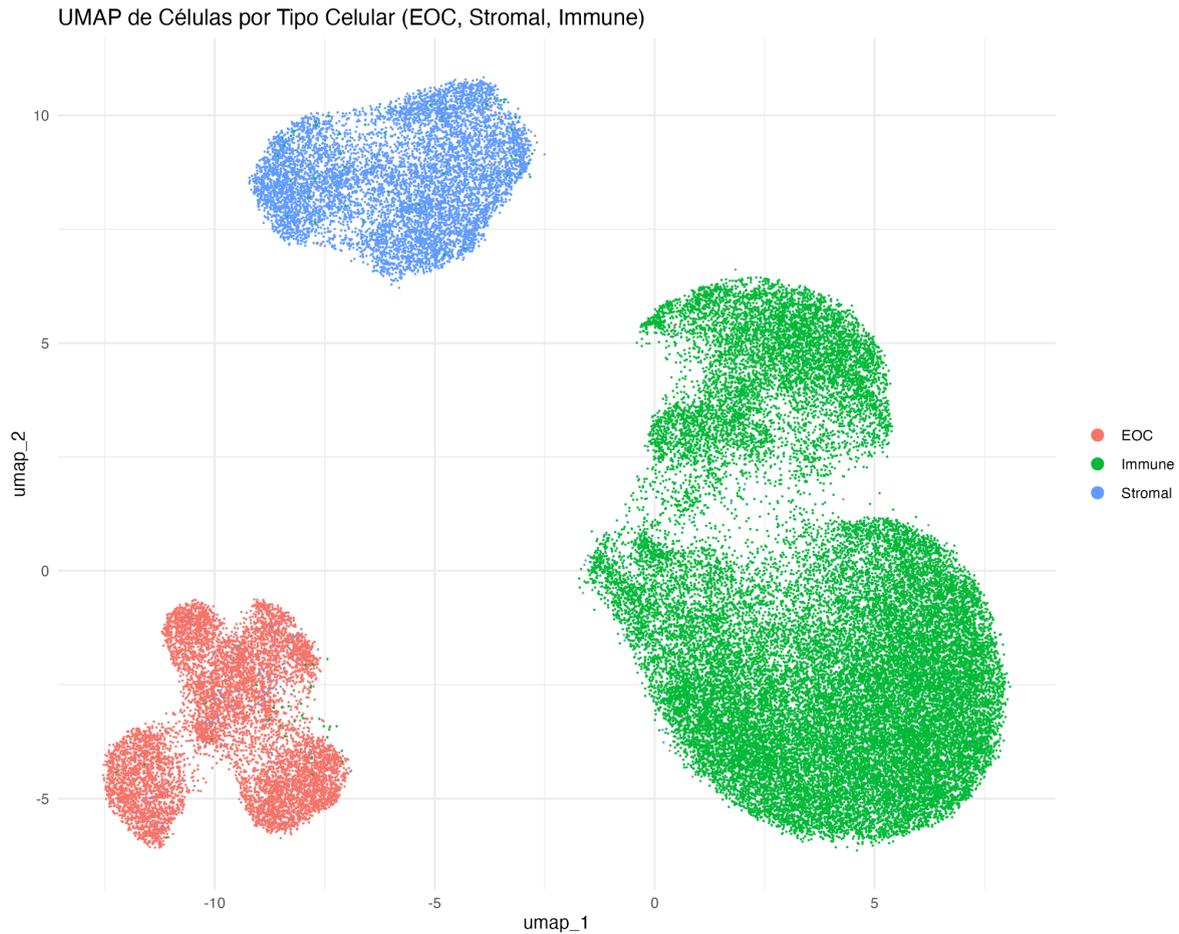


Figura 2. UMAP de células por tipo celular: EOC, Estromal e Inmune.

Ambos gráficos nos proporcionan información valiosa sobre la heterogeneidad celular y cómo los diferentes tipos de células se distribuyen y agrupan entre sí.

Como nuestro trabajo se enfoca en conocer las células epiteliales cancerosas y conocer las rutas de señalización más significativas para observar cambios antes y después del tratamiento, también obtuvimos el UMAP filtrando solo las células epiteliales cancerosas en el objeto completo.

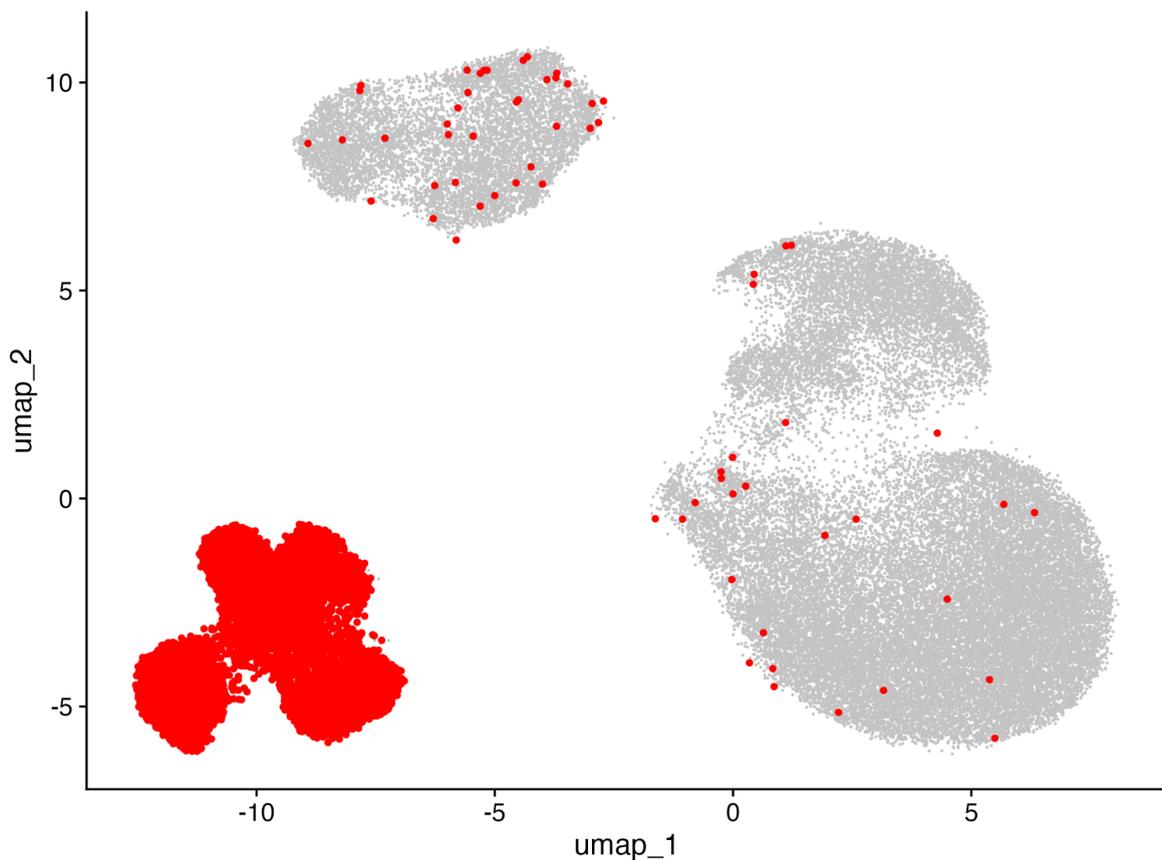


Figura 3. Visualización en UMAP filtrando las células epiteliales cancerosas

ANÁLISIS DE EXPRESIÓN DIFERENCIAL

Para conocer cómo responden las células cancerosas al tratamiento, queremos identificar aquellos genes que presentan una expresión más significativa tanto en las muestras que no presentan ningún tratamiento previo (treatment-naive) como en aquellas en las que el paciente ha adquirido quimioterapia neoadyuvante (post-NACT).

Para esto, realizamos un análisis de expresión diferencial en el subconjunto específico de células epiteliales y buscamos genes identificados con sobreexpresión pudiendo estar implicados con la resistencia a la quimioterapia en el HGSOC (Figura 4).

Seleccionamos los 10 genes con mayor cambio en la expresión entre las fases de tratamiento (treatment-naive y post-NACT) lo cual nos permite obtener un

subconjunto de genes que reflejan las diferencias más destacadas entre ambos grupos filtrando los genes con $av_log2FC > 0.25$ y $p_val_adj < 0.05$.

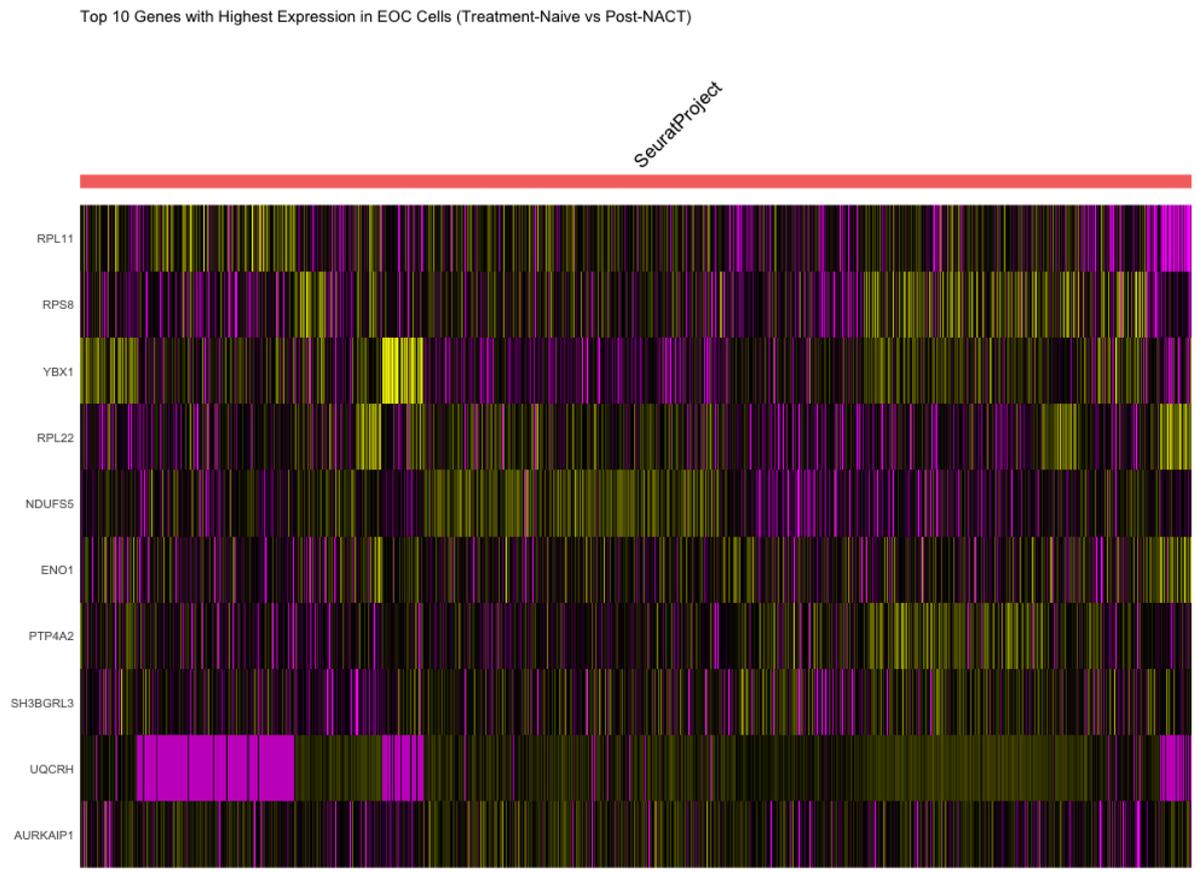


Figura 4. Heatmap de los genes más expresados de células cancerosas epiteliales (EOC) en todo el objeto Seurat.

Comparamos la expresión génica entre células en las fases “treatment-naive” y “post-NACT” para identificar genes con cambios significativos en su expresión y filtramos aquellos genes diferencialmente expresados para encontrar los más significativos, seleccionando los 10 genes principales con mayor cambio de expresión (Figura 5).

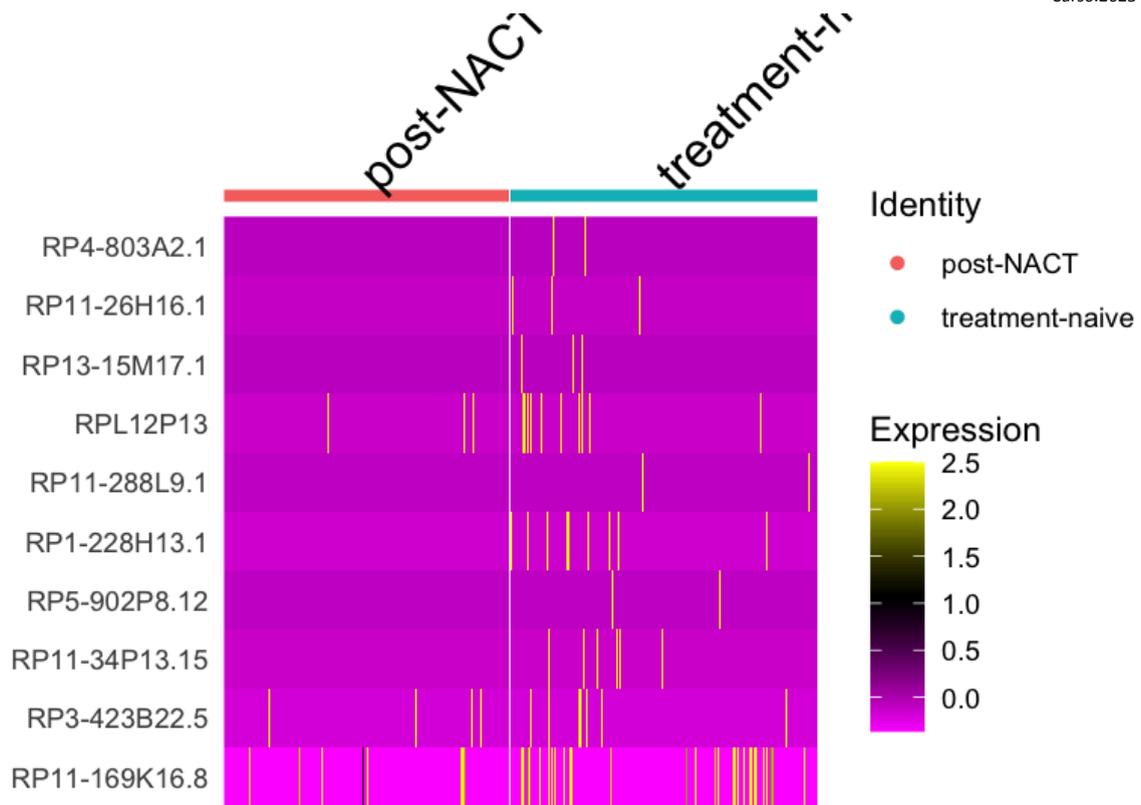


Figura 5. Heatmap de los genes diferencialmente expresados en células epiteliales cancerosas según fase de tratamiento

ANÁLISIS DE RUTAS DE SEÑALIZACIÓN

Para conocer cómo cambia la activación de diferentes vías (pathways) de señalización en las células epiteliales de HGSOC antes y después de tratamiento con quimioterapia neoadyuvantes procedemos a analizar las vías de señalización.

Proceso de Análisis de Datos de Progeny en células epiteliales de cáncer de ovario

El mapa de calor es una herramienta visual que nos permite resumir a modo de gráfico la información de activación de vías de señalización (Figura 5).

Los colores en el mapa de calor representan los puntajes de actividad de las vías de señalización, que va de azul oscuro (bajo nivel de actividad) a rojo (alto nivel de actividad).

Conocer las vías de señalización alteradas en el tumor nos ayuda a identificar terapias dirigidas y a optimizar la efectividad de los tratamientos al elegir dianas precisas.

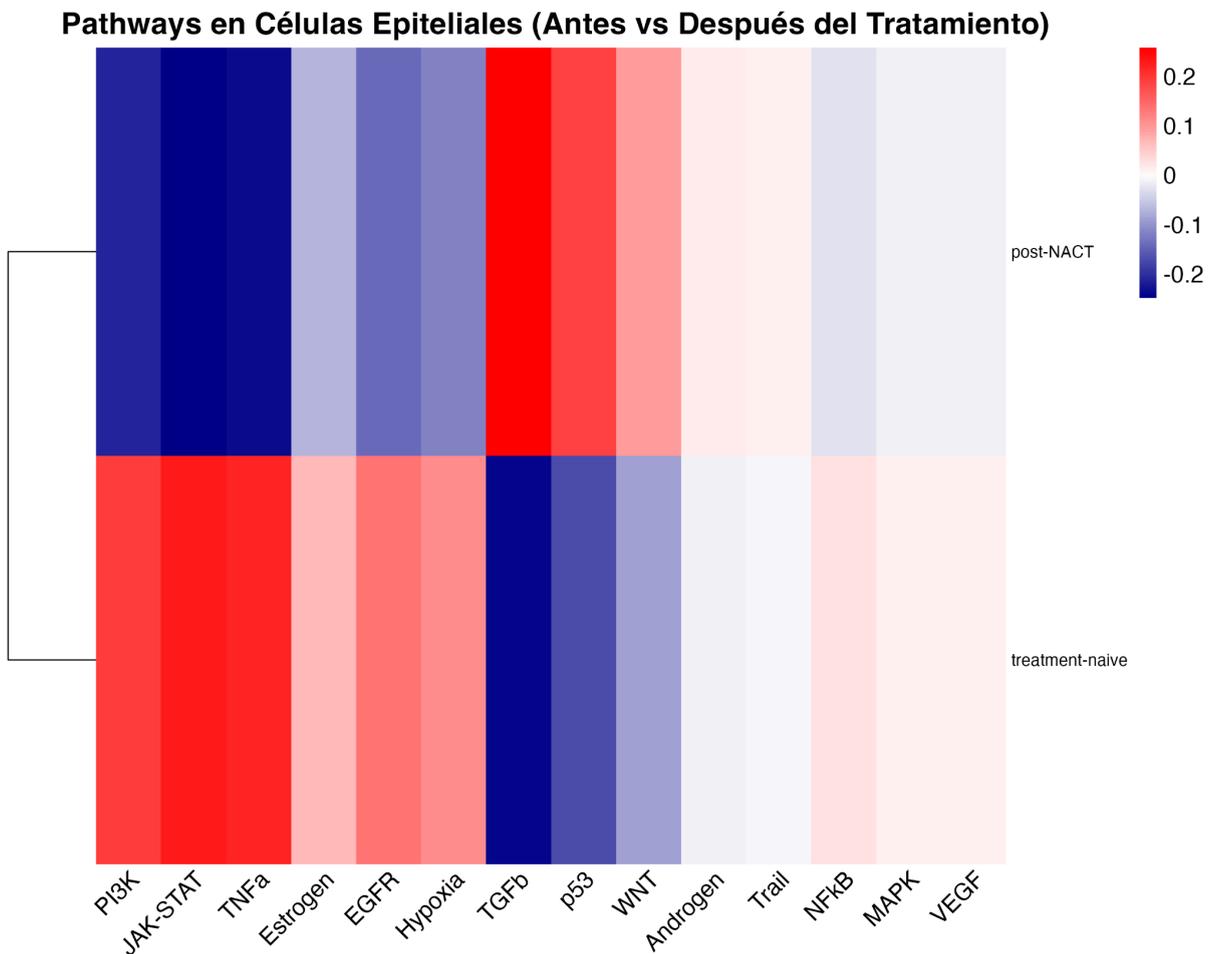


Figura 6. Comparación entre fases de tratamiento en la activación de vías de señalización en células epiteliales cancerosas.

Las vías JAK-STAT, PI3K, TNFa, EGFR e Hypoxia se inactivan o tienen menos actividad después del tratamiento a diferencia de las vías WNT, TGFb, p53 las cuales se intensifican con el tratamiento (Figura 4).

Es lógico observar que las vías que están comúnmente asociadas con la proliferación y supervivencia celular se inactivan después del tratamiento ya que las células tumorales buscan reducir la proliferación como mecanismo de resistencia.

En concordancia con los resultados de Rinne *et al* (Rinne *et al.*, 2021) la desregulación de la vía PI3K/AKT/mTOR es un mecanismo clave en el desarrollo y progresión del cáncer de ovario y puede ser un mecanismo de resistencia a la quimioterapia.

La hipoxia en el TME incrementa la resistencia a la quimioterapia en el HGSOC a través del factor HIF-1 α . Su sobreexpresión está asociada a la resistencia al platino (Wilczyński *et al.*, 2023).

Una disminución en la hipoxia podría ser un indicador de que el tratamiento está mejorando el suministro de oxígeno al tumor o una respuesta favorable al tratamiento, un hallazgo relevante que podría ser útil para planificar estrategias terapéuticas futuras.

El factor del crecimiento endotelial vascular (VEGF), siendo VEGF-A la principal vía de señalización en el desarrollo, homeostasis y cáncer y VEGFR3 que promueve la linfangiogénesis relacionada con la metástasis del cáncer de ovario a través del sistema linfático (Cairns, 2023) se han permanecido en niveles bajos cercanos a 0 tanto antes como después de NACT haciendo ver que el tratamiento no responde de forma significativa a esta ruta. Indica un bajo rol de la angiogénesis ya que como bien hemos dicho anteriormente se está adaptando mediante rutas inmunosupresoras.

Sin embargo, la intensificación de las vías anteriormente citadas pueden indicar una respuesta adaptativa de las células cancerosas ante el estrés del tratamiento.

Un aumento de intensificación en la vía de TFGB puede contribuir a la inmunosupresión (Kandalajt *et al.*, 2022).

Un estudio reciente de Sun *et al.*, describe la regulación positiva en las vías de señalización WNT, TGF β y MAPK reforzando la idea que sugiere que las células cancerosas epiteliales de HGSOC presentan un comportamiento adaptativo que les permite sobrevivir y proliferar en condiciones desfavorables como en la quimioterapia (Sun *et al.*, 2022).

Podemos ver en un gráfico de cajas como varía la actividad individual de cada pathway en función de las diferentes fases de tratamiento. Vemos puntos individuales para cada muestra en cada fase de tratamiento (en el eje x) y su correspondiente puntuación de activación en el eje y.

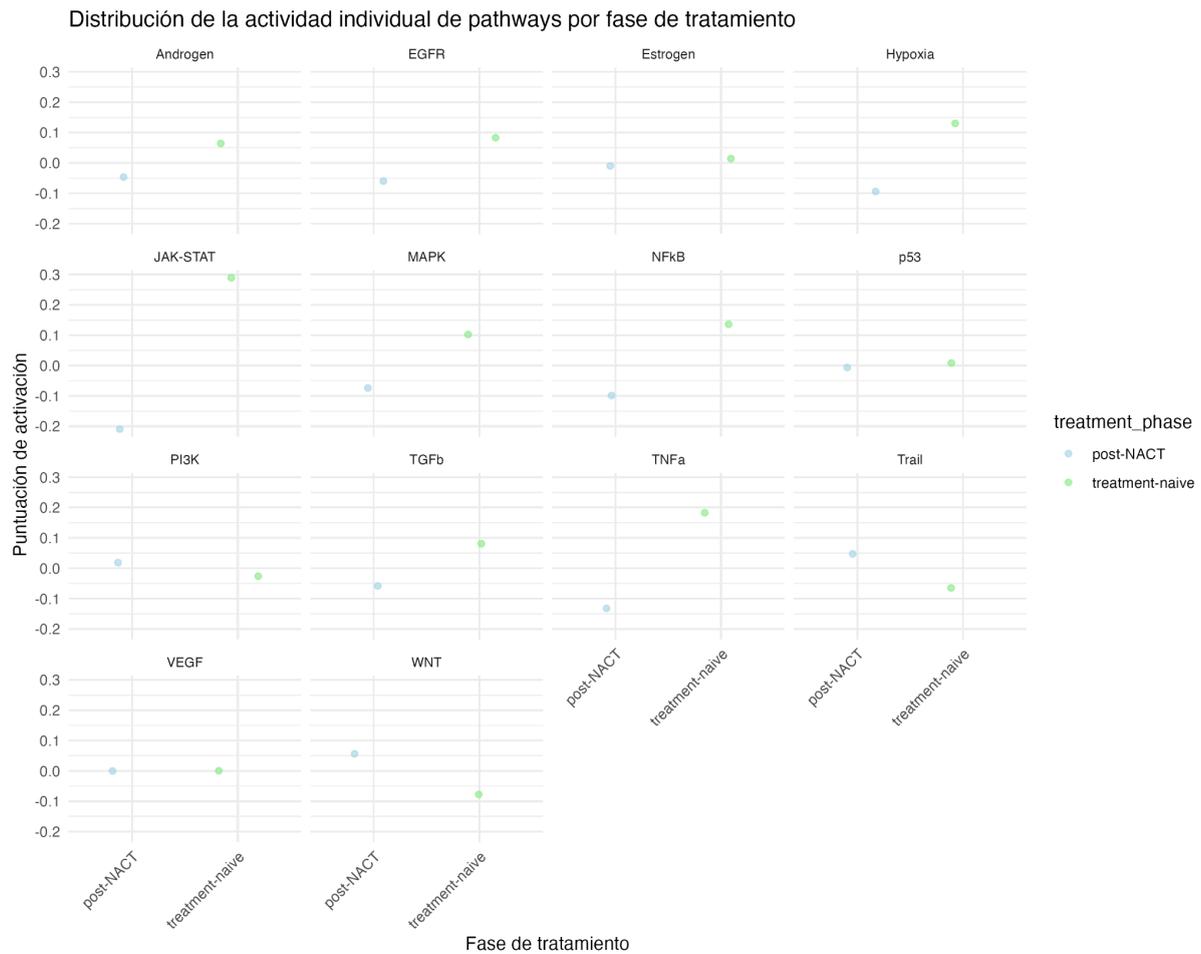


Figura 7. Distribución de la actividad de vías de señalización en función de la fase de tratamiento (naive y post-NACT).

La dispersión de los puntos en cada faceta muestra la variabilidad de la puntuación de activación de cada pathway.

A continuación, calculamos los puntajes de progeny a partir de células epiteliales (EOC) con genes diferencialmente expresados, lo cual variará el análisis de las vías de señalización.

Activación de Pathways en Células Epiteliales (Antes vs Después del Tratamiento)

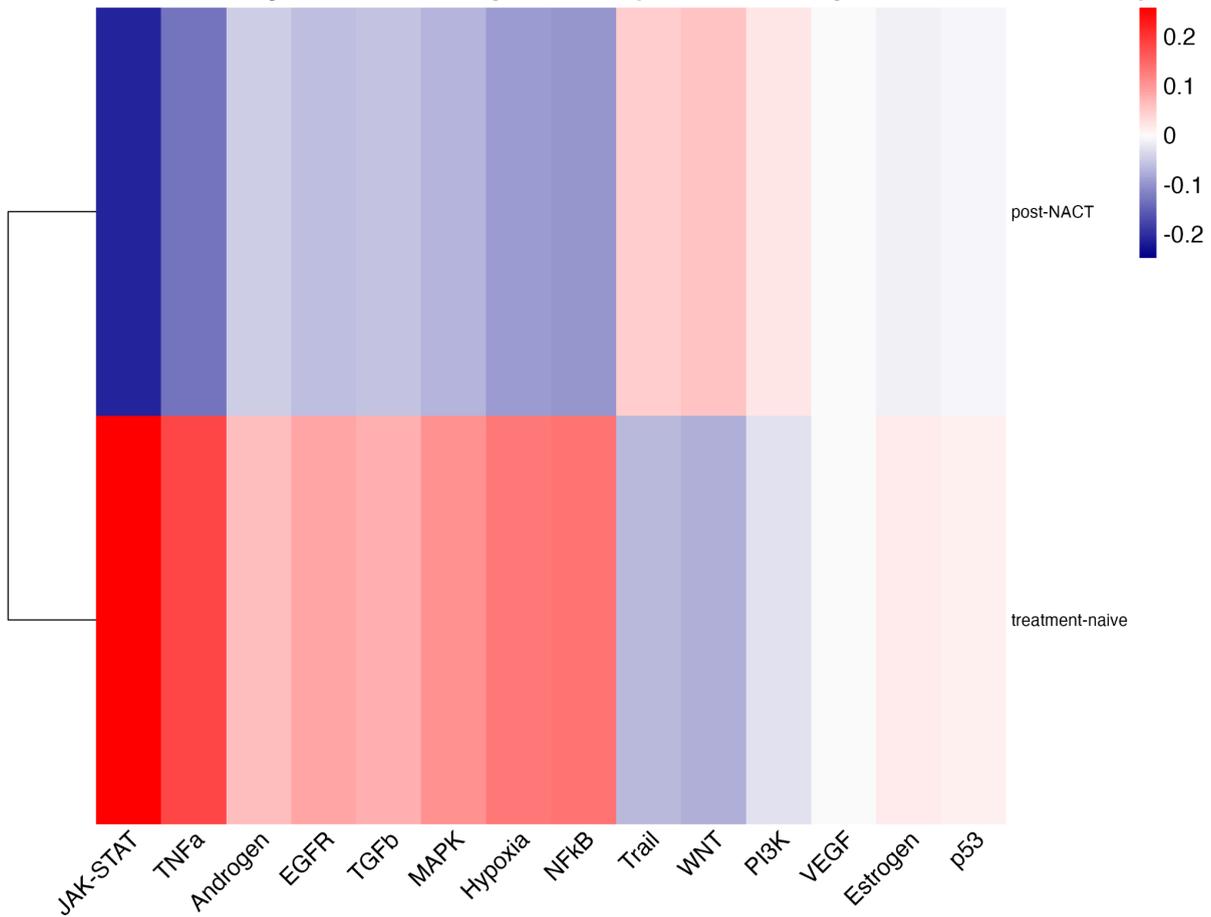


Figura 8. Comparación entre fases de tratamiento en la activación de vías de señalización en EOC en genes diferencialmente expresados.

El color más intenso y marcado en las vías JAK-STAT y TNFa tanto antes como después del tratamiento indica que estas vías están experimentando un cambio significativo en su actividad siendo fuertemente moduladas por la quimioterapia ayudando a mejorar estrategias terapéuticas en el futuro (Figura 8). Lo más seguro es que estas rutas contienen genes cuya expresión está significativamente regulada por el tratamiento.

Sin embargo, las otras vías presentan tonos más suaves que indican que no están tan influenciadas por la quimioterapia o su regulación es más sutil indicando que no son fundamentales en la respuesta o resistencia al tratamiento (Figura 8).

ANÁLISIS ESTADÍSTICO

Prueba de Wilcoxon

El resultado que obtuvimos indica que no se encontró ninguna diferencia significativa entre los dos grupos de la variable treatment-phase en relación con la variable avg.

Obtuvimos un p-value = 1, por lo que, la prueba de Wilcoxon no mostró diferencias significativas entre las fases de tratamiento para la variable de activación de las vías de señalización.

El resumen de los datos muestra que la variable avg tiene los siguientes valores de resumen:

-Mínimo (-0.0067312)

-1er Cuartil (-0.0029923)

-Mediana (0.0007466)

-3er Cuartil (0.0044855)

-Máximo (0.0082244)

Podemos observar que los datos son relativamente homogéneos o centrados en torno a 0 y la diferencia entre el valor máximo y el mínimo es pequeña, lo que sugiere que hay poca variabilidad en los datos, que explicaría el por qué no se observa una diferencia significativa entre los grupos treatment_phase.

Este resultado podría ser inesperado, ya que generalmente se esperaría que el tratamiento altere la activación de estas vías.

Es posible que la falta de diferencia significativa sea un reflejo de la complejidad del proceso de adaptación tumoral, en el que la actividad de las vías de señalización puede verse modulada de formas sutiles pero importantes.

Conclusiones

El estudio de las rutas de señalización en el cáncer de ovario seroso de alto grado (HGSOC) a nivel celular mediante secuenciación de ARN de células individuales (scRNA-seq) nos ha permitido observar las alteraciones moleculares que ocurren antes y después de la quimioterapia en las células tumorales, estromales e inmunes. Hemos trabajado con muestras de datos de 11 pacientes con HGSOC obtenidos antes y después del tratamiento, que a través del análisis de datos scRNA-seq y el uso de herramientas bioinformáticas avanzadas, hemos identificado rutas de señalización y subpoblaciones celulares clave en la resistencia a la quimioterapia permitiendo una mejor comprensión en el estudio de futuras terapias dirigidas.

Primera: Nuestros resultados revelan la activación de rutas de señalización correlacionadas con la quimiorresistencia en el cáncer de ovario seroso de alto grado (HGSOC), indicando una asociación de estas vías en la modulación de la respuesta a la quimioterapia. En particular, observamos una regulación diferencial de varias rutas clave, como las involucradas en la proliferación celular, la supervivencia, la invasión y la regulación del microambiente tumoral.

Segunda: Nuestros resultados son coincidentes con los obtenidos por otros autores para las vías PI3K, WNT y TGF β en la quimiorresistencia en HGSOC. Sin embargo, nuestros resultados difieren para la vía MAPK, ya que no se observó una activación significativa de esta ruta en nuestras muestras, a pesar de su implicación previamente reportada en la resistencia a la quimioterapia. Esta distinción podría reflejar diferencias en los perfiles moleculares de los subtipos tumorales o en los métodos de análisis utilizados en estudios anteriores, lo que resalta la complejidad y heterogeneidad de los mecanismos de quimiorresistencia en HGSOC.

Tercera: El análisis de la diferencia en la actividad de los pathways entre las fases de tratamiento no mostró resultados estadísticamente significativos.

A pesar de la activación de ciertas rutas, la prueba de Wilcoxon arrojó un valor p de 1, lo que indica que no existen diferencias claras en la actividad de los caminos entre las fases de tratamiento. Esto podría sugerir que, en el contexto de las muestras analizadas, la quimioterapia no induce modificaciones significativas en la activación de las rutas de señalización, o que los efectos en la señalización son más complejos y dependen de factores adicionales, como la heterogeneidad celular o las interacciones microambientales.

Por lo tanto, la búsqueda de metodologías más robustas y de herramientas bioinformáticas avanzadas es necesario para mejorar la precisión de los análisis y la identificación de patrones biológicos relevantes en la mejora para la resistencia de la quimioterapia en el HGSO.

Conclusions

The study of signaling pathways in high-grade serous ovarian cancer (HGSO) at the cellular level using single-cell RNA sequencing (scRNA-seq) has shown us the molecular alterations that allow them to occur before and after chemotherapy in tumor, stromal and immune cells. We have worked with data samples from 11 patients with HGSO obtained before and after treatment, which through scRNA-seq data analysis and the use of advanced bioinformatics tools, we have identified key signaling pathways and cellular subpopulations in resistance to chemotherapy allowing a better understanding in the study of future targeted therapies.

First: Our results reveal the activation of signaling pathways correlated with chemoresistance in HGSO, indicating an association of these pathways in modulating the response to chemotherapy. In particular, we observed differential regulation of several key pathways, such as those involved in cell proliferation, survival, invasion and regulation of the tumor microenvironment.

Second: Our results are consistent with those obtained by other authors for the PI3K, WNT and TGF β pathways in chemoresistance in HGSOC. However, our results differ for the MAPK pathway, as no significant activation of this pathway was observed in our samples, despite its previously reported involvement in chemotherapy resistance. This distinction could reflect differences in the molecular profiles of tumor subtypes or in the analysis methods used in previous studies, highlighting the complexity and heterogeneity of chemoresistance mechanisms in HGSOC.

Third: The analysis of the difference in pathway activity between the treatment phases did not show statistically significant results. Despite the activation of certain pathways, the Wilcoxon test returned a p value of 1, indicating that there are no clear differences in pathway activity between treatment phases. This could suggest that, in the context of the samples analyzed, chemotherapy does not induce significant modifications in the activation of signaling pathways, or that the effects on signaling are more complex and depend on additional factors, such as cellular heterogeneity or microenvironmental interactions.

Therefore, the search for more robust methodologies and more advanced bioinformatics tools is necessary to improve the precision of the analyzes and the identification of relevant biological patterns in the improvement for chemotherapy resistance in HGSOC.

Para finalizar suscribo las siguientes palabras:

“Lo que más amo de la ciencia es que a medida que aprendes, realmente no obtienes respuestas, solo tienes mejores preguntas” (John Green)

Bibliografía

Ahmed R, Zaman T, Chowdhury F, Mraiche F, Tariq M, Ahmad IS, Hasan A. Secuenciación de ARN de células individuales con transcriptómica espacial de tejidos cancerosos. *Int J Mol Sci.* 2022; 23 :3042.

Cairns, E. S. (2023). *HGSOC subtypes and treatment interaction: implications for clinical implementation of the PrOTYPE classifier* (Doctoral dissertation, University of British Columbia).

Carangelo, G., Magi, A., & Semeraro, R. (2022). From multitude to singularity: An up-to-date overview of scRNA-seq data generation and analysis. *Frontiers in Genetics, 13*, 994069.

Cortesi, M., Liu, D., Powell, E., Barlow, E., Warton, K., Giordano, E., & Ford, C. E. (2024). Development and validation of a computational tool to predict treatment outcomes in cells from High-Grade Serous Ovarian Cancer patients. *bioRxiv*, 2024-10

Freimund, A. E., Beach, J. A., Christie, E. L., & Bowtell, D. D. (2018). Mechanisms of drug resistance in high-grade serous ovarian cancer. *Hematology/Oncology Clinics, 32*(6), 983-996

Hao, Q., Li, J., Zhang, Q., Xu, F., Xie, B., Lu, H., ... & Zhou, X. (2021). Single-cell transcriptomes reveal heterogeneity of high-grade serous ovarian carcinoma. *Clinical and Translational Medicine, 11*(8), e500.

Herrera Cañizares, F. (2018). Microambiente tumoral en el cáncer de ovario seroso de alto grado: identificación de biomarcadores inmunes y análisis de su valor pronóstico.

Izar, B., Tirosh, I., Stover, E. H., Wakiro, I., Cuoco, M. S., Alter, I., ... & Regev, A. (2020). A single-cell landscape of high-grade serous ovarian cancer. *Nature medicine, 26*(8), 1271-1279.

Jovic, D., Liang, X., Zeng, H., Lin, L., Xu, F., & Luo, Y. (2022). Single-cell RNA sequencing technologies and applications: A brief overview. *Clinical and translational medicine, 12*(3), e694.

Kim, J., Park, E. Y., Kim, O., Schilder, J. M., Coffey, D. M., Cho, C. H., & Bast Jr, R. C. (2018). Cell origins of high-grade serous ovarian cancer. *Cancers, 10*(11), 433.

Kandalaft, L. E., Dangaj Laniti, D., & Coukos, G. (2022). Immunobiology of high-grade serous ovarian cancer: lessons for clinical translation. *Nature Reviews Cancer*, 22(11), 640-656.

Kurman, R. J. (2013). Origin and molecular pathogenesis of ovarian high-grade serous carcinoma. *Annals of Oncology*, 24, x16-x21.

L. Kelland, El resurgimiento de la quimioterapia contra el cáncer basada en platino. *Nat. Rev. Cancer* 7 , 573–584 (2007).

Lisio, M.-A.; Fu, L.; Goyeneche, A.; Gao, Z.-h.; Telleria, C. High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints. *Int. J. Mol. Sci.* 2019, 20, 952.

Liu, C., Zhang, Y., Li, X., & Wang, D. (2022). Ovarian cancer-specific dysregulated genes with prognostic significance: scRNA-Seq with bulk RNA-Seq data and experimental validation. *Annals of the New York Academy of Sciences*, 1512(1), 154-173.

Lliberos, C., Richardson, G., & Papa, A. (2024). Oncogenic Pathways and Targeted Therapies in Ovarian Cancer. *Biomolecules*, 14(5), 585.

Lu, J., Sheng, Y., Qian, W., Pan, M., Zhao, X., & Ge, Q. (2023). scRNA-seq data analysis method to improve analysis performance. *IET nanobiotechnology*, 17(3), 246-256

Mahoney, K. M., Rennert, P. D., & Freeman, G. J. (2015). Combination cancer immunotherapy and new immunomodulatory targets. *Nature reviews Drug discovery*, 14(8), 561-584.

Mahmood, R. D., Morgan, R. D., Edmondson, R. J., Clamp, A. R., & Jayson, G. C. (2020). First-line management of advanced high-grade serous ovarian cancer. *Current Oncology Reports*, 22, 1-14.

Masood, F., Khan, W., & Uddin, R. (2020). Computational-based identification and analysis of globally expressed differential genes in high-grade serous ovarian carcinoma cell lines. *Computational Biology and Chemistry*, 88, 107333.

Matthews, B. G., Bowden, N. A., & Wong-Brown, M. W. (2021). Epigenetic mechanisms and therapeutic targets in chemoresistant high-grade serous ovarian cancer. *Cancers*, 13(23), 5993.

Momenimovahed, Z., Tiznobaik, A., Taheri, S., & Salehiniya, H. (2019). Ovarian cancer in the world: epidemiology and risk factors. *International journal of women's health*, 287-299.

Olbrecht, S., Busschaert, P., Qian, J., Vanderstichele, A., Loverix, L., Van Gorp, T., ... & Vergote, I. High-grade serous tuboovarian cancer refined with single-cell RNA sequencing: specific cell subtypes influence survival and determine molecular subtype classification. *Genome Med.* 202.

Parry, T. W. (2023). Use of scRNA-seq to characterise the tumour microenvironment of high grade serous ovarian carcinoma (HGSOC).

Ram, P. T., Mendelsohn, J., & Mills, G. B. (2012). Bioinformatics and systems biology. *Molecular oncology*, 6(2), 147-154.

Rinne N, Christie EL, Ardasheva A, Kwok CH, Demchenko N, Low C, Tralau-Stewart C, Fotopoulou C, Cunnea P. Targeting the PI3K/AKT/mTOR pathway in epithelial ovarian cancer, therapeutic treatment options for platinum-resistant ovarian cancer. *Cancer Drug Resist.* 2021 Apr 14;4(3):573-595.

Schubert, M., Klinger, B., Klünemann, M., Sieber, A., Uhlitz, F., Sauer, S., ... & Saez-Rodriguez, J. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nature communications*, 9(1), 20.

Stewart, C., Ralyea, C., & Lockwood, S. (2019, April). Ovarian cancer: an integrated review. In *Seminars in oncology nursing* (Vol. 35, No. 2, pp. 151-156). WB Saunders.

Sun, Y., Yao, L., Wang, C., Xiong, B., Guo, J., Wang, L., ... & Liu, S. (2022). Involvement of Cancer stem cells in Chemo-resistant relapse of epithelial ovarian.

Webb, P. M., & Jordan, S. J. (2017). Epidemiology of epithelial ovarian cancer. *Best practice & research Clinical obstetrics & gynaecology*, 41, 3-14.

Wilczyński, J., Paradowska, E., Wilczyńska, J., & Wilczyński, M. (2023). Prediction of Chemoresistance—How Preclinical Data Could Help to Modify Therapeutic Strategy in High-Grade Serous Ovarian Cancer. *Current Oncology*, 31(1), 229-249.

Xu, J., Fang, Y., Chen, K., Li, S., Tang, S., Ren, Y., ... & Lu, W. (2022). Single-cell RNA sequencing reveals the tissue architecture in human high-grade serous ovarian cancer. *Clinical Cancer Research*, 28(16), 3590-3602

Yoneda A, Lendorf ME, Couchman JR, Mulhaupt HA. Breast and ovarian cancers: a study and possible functions of cell surface heparan sulfate proteoglycans. *J Histochem Cytochem*. 2012.

Zhang, Y., Wang, D., Peng, M., Tang, L., Ouyang, J., Xiong, F., ... & Xiong, W. (2021). Single-cell RNA sequencing in cancer research. *Journal of Experimental & Clinical Cancer Research*, 40, 1-17.

Declaración obligatoria del uso de herramientas de IA

El autor declara el uso de la herramienta de inteligencia artificial (IA) ChatGPT de OpenAI (modelo GPT-4) para el apoyo en tareas de investigación y redacción.

-Asistencia en redacción científica y desarrollo de capítulos: La IA ha contribuido en el apoyo a la redacción del resumen y de las conclusiones.

Uso dado: Apoyo en la estructura y redacción de secciones

-Análisis de datos y justificación de técnicas bioinformáticas. Con el prompt y recomendaciones metodológicas de estudios previos se ha utilizado ChatGPT para obtener orientación en la selección de herramientas de filtrado y análisis de datos de scRN-seq.

Uso dado: Orientación en la herramienta de filtrado de datos y análisis de datos.