



**FACULTAD DE CIENCIAS AGRARIAS
UNIVERSIDAD NACIONAL DE ROSARIO**

**ANÁLISIS IN SILICO DE LA VÍA DE LA ENZIMA LEUCOTRIENO A4
HIDROLASA EN LOS DISTINTOS ESTADIOS DEL DESARROLLO DEL
HEPATOCARCINOMA HUMANO**

LUCÍA OVIEDO BUSTOS

**TRABAJO FINAL PARA OPTAR AL TITULO DE ESPECIALISTA EN
BIOINFORMÁTICA**

DIRECTORA: AYLÉN AVILA

AÑO: 2025

**ANÁLISIS IN SILICO DE LA VÍA DE LA ENZIMA LEUCOTRIENO A4
HIDROLASA EN LOS DISTINTOS ESTADIOS DEL DESARROLLO DEL
HEPATOCARCINOMA HUMANO**

Lucía Oviedo Bustos

Médica – Universidad Nacional de Rosario

Este Trabajo Final es presentado como parte de los requisitos para optar al grado académico de Especialista en Bioinformática, de la Universidad Nacional de Rosario y no ha sido previamente presentada para la obtención de otro título en ésta u otra Universidad. El mismo contiene los resultados obtenidos en investigaciones llevadas a cabo en Rosario, durante el período comprendido entre marzo del 2023 y mayo del 2025, bajo la dirección de Aylén Avila.



Esp. Aylén Avila
33.335.958

Defendida:de 20__.

Abreviaturas y símbolos

5-LOX	5-lipooxigenasa
AA	Ácido araquidónico
AND	Ácido desoxirribonucleico
AJCC	Comité Conjunto Americano del Cáncer (American Joint Committee on Cancer)
ARN	Ácido ribonucleico
ARNm	Ácido ribonucleico mensajero
BCLC	Clasificación de Barcelona para el Carcinoma Hepatocelular
BLT1	Receptor 1 de leucotrieno B4
BLT2	Receptor 2 de leucotrieno B4
cDNA	ADN complementario
DEGs Genes)	Genes diferencialmente expresados (Differentially Expressed Genes)
FDR	Tasa de descubrimiento falso (<i>False Discovery Rate</i>)
GDC	Centro de Datos Genómicos (<i>Genomic Data Commons</i>)
GEO	Omnibus de Expresión Génica (Gene Expression Omnibus)
HCC	Carcinoma hepatocelular (Hepatocellular Carcinoma)
LTA4	Leucotrieno A4
LTA4H	Leucotrieno A4 hidrolasa
LTB4	Leucotrieno B4
NALFD	Enfermedad del hígado graso no alcohólico
NASH	Esteatohepatitis no alcohólica
ND	No Disponible
NES <i>Enrichment Score</i>)	Puntuación de Enriquecimiento Normalizada (<i>Normalized Enrichment Score</i>)
RNAseq	Secuenciación de ARN
TCGA	Atlas del Genoma del Cáncer (<i>The Cancer Genome Atlas</i>)
TNM	Clasificación de Tumor, Ganglio y Metástasis
TPM	Transcritos por millón (<i>Transcripts Per Million</i>)
VHB	Virus de la hepatitis B
VHC	Virus de la hepatitis C

Presentación a congresos

- “EFFECT OF THE INHIBITION OF LEUKOTRIENE A4 HYDROLASE (LTA4H) WITH SC 57461A IN LIVER CANCER PROLIFERATION”. Oviedo Bustos L, Comanzo CG, Vera MC, Livore VI, Palma NF, Ferretti AC, Frattini M, Ceballos MP, Quiroga AD, Álvarez ML. Reunión Científica Anual de la SAFIS. Desarrollada los días 25, 26 y 27 de octubre del 2023 en la ciudad de Buenos Aires (póster). Resumen publicado en la revista Physiological Mini-Reviews Vol. 16, 2023, pág. 59 (ISSN 1669-5410 Online).
- “EL ROL DE LA ENZIMA LEUCOTRIENO A4 HIDROLASA (LTA4H) EN LA PROGRESIÓN DEL CARCINOMA HEPATOCELULAR”. Oviedo Bustos L, Comanzo CG, Vera MC, Livore VI, Palma NF, Ferretti AC, Frattini M, Ceballos MP, Quiroga AD, Álvarez ML. Reunión Científica Anual de la SBR. Desarrollada los días 28 y 29 de noviembre del 2024 en la ciudad de Rosario(póster).

Resumen

Las proteínas LTA4H, LTB4R y LTB4R2 forman parte de una vía metabólica relacionada con procesos inflamatorios y proliferativos, y su papel en el cáncer, especialmente en el carcinoma hepatocelular (HCC), ha sido poco explorado. En este estudio, utilizamos herramientas bioinformáticas para analizar su expresión y relevancia biológica en el HCC.

Partimos de datos de expresión génica del conjunto TCGA-LIHC (HCC n=374 e hígado sano n=50), donde del total de genes (60.659) identificamos 8.066 genes sobreexpresados y 1.968 subexpresados (Fold Change ≥ 1 , FDR ≤ 0.01). Aunque LTA4H, LTB4R y LTB4R2 no mostraron cambios significativos en el análisis global de expresión diferencial, observamos que estos genes están significativamente sobreexpresados ($p < 2.2e-16$) en tumores de HCC en comparación con tejido sano. Esta sobreexpresión fue más evidente en las etapas I-III, según la clasificación AJCC, del HCC ($p = 0.072$), disminuyendo en la etapa IV ($p = 1.000$), lo que sugiere que podrían estar involucrados en la proliferación tumoral en fases tempranas de la enfermedad.

Además, exploramos las interacciones funcionales e inmunológicas de estos genes. Los genes correlacionados con LTA4H, LTB4R y LTB4R2 (con coeficientes de Spearman moderados, Rho entre ± 0.3 y 0.5) se asociaron con procesos como la traducción, la remodelación de la cromatina y el transporte celular. En cuanto a la infiltración inmune, LTA4H mostró una correlación significativa con la presencia de macrófagos y células dendríticas (análisis con TIMER), así como una correlación negativa con el microambiente tumoral (xCell: $\rho = -0.165$, $p = 0.0014$). Por su parte, LTB4R2 presentó la correlación negativa más fuerte con el compartimento inmune ($\rho = -0.1729$, $p = 0.0008$). Estos resultados sugieren que la vía de LTA4H podría estar modulando el microambiente tumoral, aunque de manera compleja y dependiente del contexto celular.

Finalmente, mediante análisis GSEA, descubrimos que altos niveles de LTA4H están asociados con vías relacionadas con la proliferación tumoral ("CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP"; NES=1.821, $p=0.0$), la apoptosis ("KEGG_APOPTOSIS"; NES=1.451, $p=0.0$) y un peor pronóstico ("HOSHIDA_LIVER_CANCER_SURVIVAL_UP"; NES=1.299, $p=0.003$). Esto indica que la sobreexpresión de LTA4H podría estar impulsando la progresión del HCC a través de mecanismos que favorecen la proliferación y la apoptosis, lo que se traduce en una menor supervivencia de los pacientes.

En conclusión, nuestros hallazgos sugieren que LTA4H, LTB4R y LTB4R2 desempeñan un papel importante en las etapas iniciales del HCC, influyendo tanto en la proliferación tumoral como en la modulación del microambiente

inmunológico. Estos genes representan candidatos prometedores para futuras investigaciones enfocadas en el desarrollo de terapias y biomarcadores pronósticos para el HCC.

Palabras clave: Bioinformática, Carcinoma Hepatocelular, LTA4H, Microambiente Tumoral, Proliferación.

Abstract

In silico analysis of the leukotriene a4 hydrolase pathway in different stages of human hepatocellular carcinoma development.

The proteins LTA4H, LTB4R, and LTB4R2 are key components of a metabolic pathway involved in inflammatory and proliferative processes, with potential roles in cancer. In this study, we investigated their expression and biological relevance in hepatocellular carcinoma (HCC) using bioinformatic approaches.

We analyzed gene expression data from the TCGA-LIHC dataset, identifying 8,066 upregulated and 1,968 downregulated genes (Fold Change ≥ 1 , FDR ≤ 0.01). Although LTA4H, LTB4R, and LTB4R2 did not meet the criteria for differential expression in the global analysis, they were significantly overexpressed ($p < 2.2e-16$) in HCC compared to healthy tissue. This overexpression was more pronounced in stages I-III (AJCC classification), decreasing in stage IV ($p = 0.072-1.000$), suggesting a role in early tumor proliferation.

Functional and immune correlations were also explored. Genes associated with LTA4H, LTB4R, and LTB4R2 (moderate Spearman coefficients: $Rho \pm 0.3-0.5$) were enriched in processes such as translation, chromatin remodeling, and cellular transport. Additionally, LTA4H expression correlated with macrophage and dendritic cell infiltration (TIMER) and showed a negative correlation with the tumor microenvironment (xCell: $\rho = -0.165$, $p = 0.0014$). LTB4R2 exhibited the strongest negative correlation with the immune compartment ($\rho = -0.1729$, $p = 0.0008$), indicating a potential role in modulating the tumor immune microenvironment.

GSEA analysis revealed that high LTA4H expression was associated with proliferation ("CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP"; NES=1.821, $p=0.0$), apoptosis ("KEGG_APOPTOSIS"; NES=1.451, $p=0.0$), and poor prognosis ("HOSHIDA_LIVER_CANCER_SURVIVAL_UP"; NES=1.299, $p=0.003$), suggesting that LTA4H may drive HCC progression through proliferative and apoptotic mechanisms.

In conclusion, LTA4H, LTB4R, and LTB4R2 play significant roles in early HCC stages, influencing tumor proliferation and immune microenvironment modulation. These proteins emerge as promising candidates for further research into therapeutic targets and prognostic biomarkers in HCC.

Índice

1. Introducción
 - 1.1. Cáncer de hígado
 - 1.2. Microambiente tumoral
 - 1.3. Vía LTA4H
 - 1.4. Desregulación del ciclo celular y muerte celular en el cáncer
 - 1.5. Bioinformática
 - 1.5.1. Obtención y cuantificación de datos de expresión génica
 - 1.5.2. Análisis de expresión génica diferencial
 - 1.5.3. Análisis de enriquecimiento funcional
 - 1.5.4. Bases de datos públicas
 - 1.5.5. Impacto de la bioinformática en la investigación biomédica
 2. Objetivos
 - 2.1. Objetivo general
 - 2.2. Objetivos específicos
 3. Metodología
 - 3.1. Obtención y preprocesamiento de datos
 - 3.2. Comparación de la expresión génica según estadio tumoral
 - 3.3. Análisis de la expresión diferencial de genes (DEGs)
 - 3.4. Análisis de correlación y construcción de redes de genes asociados
 - 3.5. Evaluación de la infiltración inmune y el microambiente tumoral
 - 3.6. Análisis de enriquecimiento de vías mediante GSEA
 4. Resultados
 - 4.1. Análisis de expresión diferencial de genes para determinar la posición de LTA4H, LTB4R y LTB4R2
 - 4.2. Diferencias de expresión de LTA4H, LTB4R y LTB4R2
 - 4.3. Enriquecimiento funcional de los genes correlacionados con LTA4H, LTB4R y LTB4R2
 - 4.4. Correlaciones entre LTA4H, LTB4R y LTB4R2 y las infiltraciones inmunes en hepatocarcinoma celular
 - 4.5. Impacto de la sobreexpresión de LTA4H en la proliferación tumoral, apoptosis y pronóstico del carcinoma hepatocelular
 5. Discusión y conclusiones
 6. Bibliografía
-

1. Introducción

1.1. Cáncer de hígado

El carcinoma hepatocelular (HCC) es el tipo más común de cáncer de hígado primario, y su incidencia ha aumentado globalmente en las últimas décadas(Sung et al., 2021).

El HCC se desarrolla principalmente en el contexto de enfermedades hepáticas crónicas como la cirrosis, que puede ser causada por infecciones virales (virus de la hepatitis B, VHB, y virus de la hepatitis C, VHC,), consumo excesivo de alcohol o enfermedades metabólicas como la esteatohepatitis no alcohólica (NASH)(Lozano et al., 2012).

Anteriormente, las infecciones virales eran las causas más comunes de HCC. Sin embargo, gracias a las campañas de vacunación contra la hepatitis B y los tratamientos antivirales efectivos para la hepatitis C, la incidencia de HCC de etiología viral ha disminuido(Zamor et al., 2017). Al mismo tiempo, se ha observado un aumento en los casos de HCC asociados con la enfermedad del hígado graso no alcohólico (NAFLD), vinculada al incremento de la obesidad y el síndrome metabólico en la población(Yang et al., 2019).

El HCC es el sexto cáncer más frecuente y la tercera causa principal de muerte relacionada con cáncer en el mundo(Sung et al., 2021). El HCC ocupa una posición más alta en el ranking de mortalidad por cáncer debido a que, en la mayoría de los casos, el diagnóstico se realiza en etapas avanzadas. Esto se debe a que las manifestaciones clínicas son sutiles e inespecíficas, lo que retrasa su detección. Las opciones terapéuticas más efectivas, especialmente los tratamientos quirúrgicos, solo son viables cuando el diagnóstico es temprano, lo que limita significativamente las posibilidades de tratamiento. Esto lleva a recurrir a terapias sistémicas que solo resultan efectivas en un 30% de los casos y generan resistencia a los fármacos en un lapso de seis meses(Chidambaranathan-Reghupaty et al., 2021).

Por esta razón, es crucial continuar investigando los mecanismos subyacentes del HCC, no solo para mejorar las terapias actuales, sino también para identificar nuevos enfoques terapéuticos y desarrollar marcadores específicos que permitan un diagnóstico temprano.

Estadios de cáncer

La estadificación del HCC según el sistema de la American Joint Committee on Cancer (AJCC)(FAG and T.A., 2010), basado en el sistema TNM (Tumor, Nodo, Metástasis), proporciona un marco estandarizado para evaluar la extensión anatómica del tumor. En la figura 1 se muestra esta estadificación. Este sistema clasifica la enfermedad en diferentes estadios según el tamaño y número de tumores, la presencia de invasión vascular, la afectación de ganglios linfáticos

regionales y la existencia de metástasis a distancia. La clasificación en estadios no solo permite estimar la supervivencia y la progresión de la enfermedad, sino que también desempeña un papel crucial en la toma de decisiones terapéuticas, facilitando la selección de tratamientos como la resección quirúrgica, el trasplante hepático, las terapias locorregionales o los enfoques sistémicos.

Además, la estadificación AJCC es una herramienta esencial para la comunicación entre profesionales de la salud, la comparación de resultados en investigación clínica y la evaluación de la respuesta a los tratamientos. El HCC tiene la particularidad de que su pronóstico está determinado no solo por las características del tumor, sino también por la función hepática. Por ello, la estadificación TNM suele complementarse con otros sistemas (Chun et al., 2011; Duseja, 2014), como el BCLC (Barcelona Clinic Liver Cancer), que integran tanto la función hepática como el estado general del paciente.

Etapa ACJJ	Agrupación por etapas	Descripción de la etapa
I	T1 N0 M0	Tumor solitario sin invasión vascular (T1) que no se ha propagado a los ganglios linfáticos adyacentes (N0), ni a sitios distantes (M0).
II	T2 N0 M0	Tumor solitario con invasión vascular o tumores multifocales, ninguno mayor a 5 cm que no se ha propagado a ganglios linfáticos adyacentes (N0), ni a sitios distantes (M0).
III A	T3a N0 M0	Tumores multifocales, al menos uno de los cuales mide más de 5 cm (T3a)
III B	T3b N0 M0	Tumor único o multifocales de cualquier tamaño que involucren una rama principal de la vena porta o la vena hepática (T3b)
III C	T4 N0 M0	Tumor con invasión directa de órganos adyacentes distintos de la vesícula biliar o con perforación del peritoneo visceral (T4)
IV A	Todo T N1 M0	Tumor único o multifocales con metástasis en ganglio linfático regional
IV B	Todo T Toda N M1	Tumor único o multifocales con o sin metástasis en ganglio linfático regional con metástasis a distancia

Figura 1. Clasificación TNM para HCC según la 7ma (2010) ed. del AJCC.

1.2. Microambiente tumoral

Un tumor no es solo un conjunto de células cancerosas, sino una estructura compleja formada por una combinación de células del propio tumor, células del huésped infiltradas o residentes, factores secretados y matriz extracelular. Las células tumorales generan cambios profundos a nivel molecular, celular y físico en los tejidos circundantes, creando un entorno que favorece su crecimiento y progresión (Sautès-Fridman et al., 2011).

El microambiente tumoral es dinámico y está en constante evolución. Su composición varía según el tipo de tumor, pero suele incluir células inmunitarias, células del estroma, vasos sanguíneos y matriz extracelular.

A medida que el tumor se desarrolla, es infiltrado por una variedad de células del sistema inmunológico, tanto del sistema innato como del adaptativo. Estas células inmunes pueden desempeñar funciones opuestas: en algunos casos, contribuyen a la eliminación del tumor, mientras que en otros pueden favorecer su crecimiento (Anderson and Simon, 2020).

1.3. Vía LTA4H

La vía de la LTA4H se esquematiza en la figura 2. El LTB4 es un mediador lipídico de familia de los eicosanoides derivados del ácido araquidónico (AA) que se produce a partir de la acción de dos enzimas: la 5-lipooxigenasa (5-LOX) que actúa sobre el AA para dar el producto intermedio leucotrieno A4 (LTA4), el cual es convertido en LTB4 por la acción de la enzima LTA4 hidrolasa (LTA4H) (W and R, 2008). Por otra parte, el metabolito intermedio LTA4 puede ser

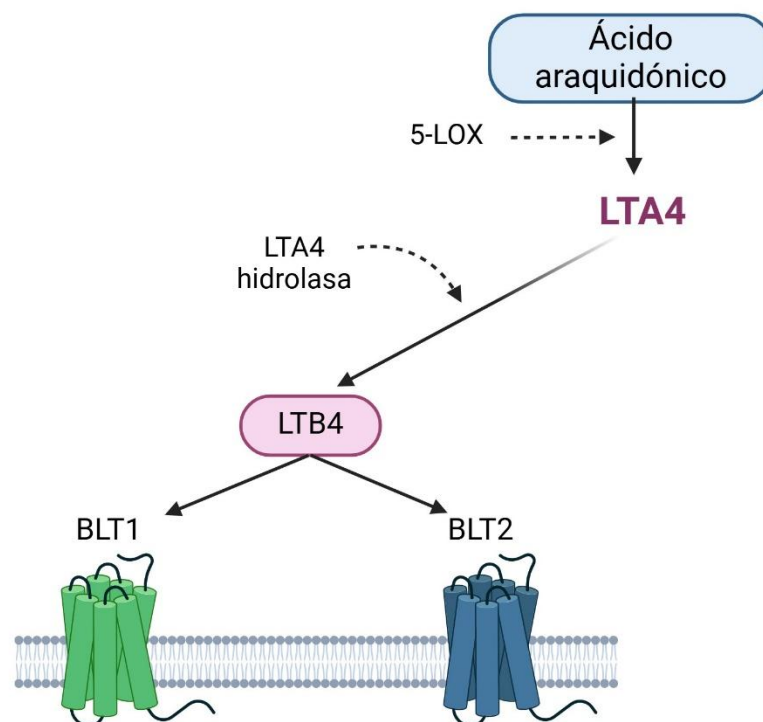


Figura 2. Diagrama de la vía de la LTA4H

metabolizado por distintas enzimas de la familia para producir otros eicosanoides con acciones diferentes al LTB4 (Rådmark and Samuelsson, 2010; Serhan et al., 2008; W and R, 2008). Se han identificado dos receptores para LTB4: el receptor con mayor afinidad, conocido como BLT1, y el receptor con menor afinidad, denominado BLT2.

La primera función descrita para la vía de la LTA4H fue la proinflamatoria (Serhan et al., 2008) generando la quimiotaxis de neutrófilos y

macrófagos. Luego se estableció su relación con enfermedades inflamatorias agudas y crónicas, como infecciones, alergias, enfermedades autoinmunes y trastornos metabólicos(He et al., 2020). Pero en estos últimos años ha habido estudios que la relacionan con distintos tipos de cánceres. Se ha reportado que hay una sobreproducción de LTB4 en varios tipos de cánceres humanos, y que el mismo estimula la proliferación de las células cancerosas(Hu et al., 2011; Oi et al., 2017; Yang et al., 2008; Zhao et al., 2019). Además, se ha documentado la relación entre diversos tipos de cánceres y los receptores BLT1 y BLT2(Park et al., 2019). Por lo tanto, la vía LTA4H/LTB4/BLT1-BLT2 podría jugar un papel importante en el proceso de carcinogénesis(Chen et al., 2005).

1.4. Desregulación del ciclo celular y muerte celular en el cáncer

En el ser humano el crecimiento, la proliferación y la muerte celular están altamente regulados para mantener el equilibrio del organismo. El ciclo celular está estrictamente regulado para garantizar la homeostasis, y la muerte celular programada desempeña un papel clave en un desarrollo adecuado.

Dado su papel fundamental, la alteración de la proliferación o la muerte celular no solo contribuye, sino que en muchos casos define, el desarrollo de diversas enfermedades, como trastornos autoinmunes, enfermedades neurodegenerativas y cáncer(Loftus L V., 2022). En este último, las consecuencias de una proliferación descontrolada y una reducción en la muerte

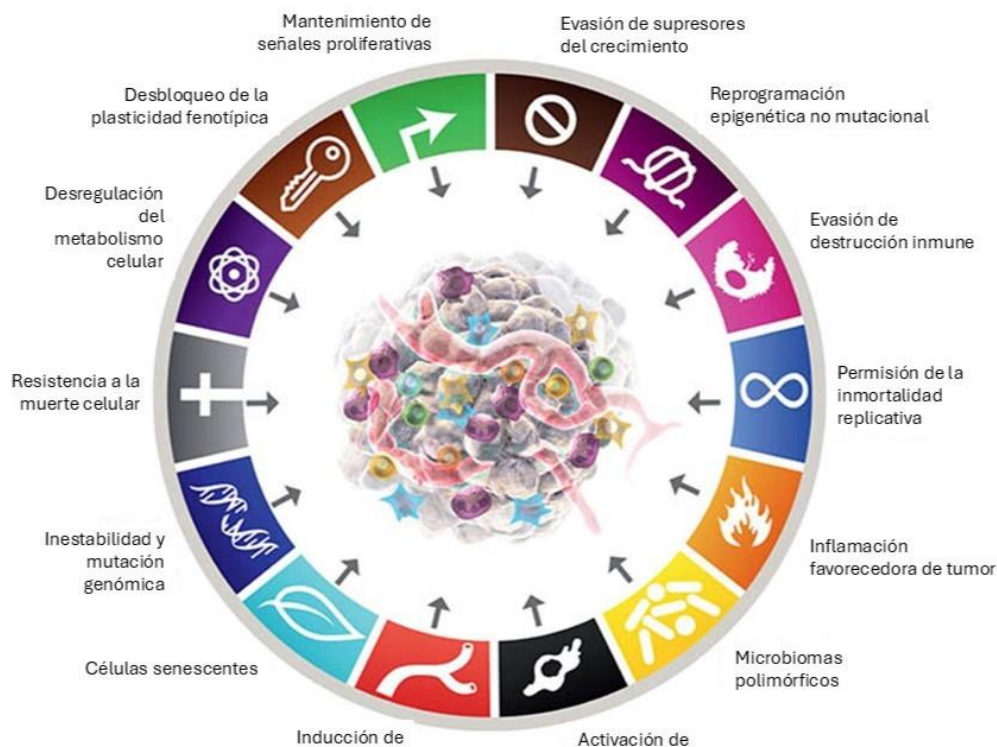


Figura 3. Representación de los Hallmarks del cáncer. Perspectiva actualizada (2022)(23).

celular son particularmente evidentes. El crecimiento desmedido del tumor se debe a la activación de señales proliferativas sostenidas y a la evasión de supresores del crecimiento, mientras que la resistencia a la muerte celular y la evasión del sistema inmunológico contribuyen a su progresión. Este desequilibrio fundamental se describe en los “Hallmarks of Cancer”(Hanahan et al., 2011)(ver figura 3) y ha sido ampliado con una característica adicional propuesta para los cánceres letales: la resistencia a los tratamientos.

A medida que la investigación en cáncer y muerte celular ha avanzado, también ha evolucionado la comprensión de su biología subyacente. Hoy se sabe que la muerte celular inducida por los puntos de control del ciclo celular ocurre principalmente a través de la apoptosis, un proceso altamente regulado de suicidio celular. Además, dado que la apoptosis es la principal forma de muerte celular durante el desarrollo, se ha esclarecido que la relación evolutiva entre la muerte celular y la proliferación radica específicamente en la conexión entre la apoptosis y el ciclo celular(Varghese et al., 2019).

1.5. Bioinformática

La bioinformática ha emergido como una disciplina esencial en la investigación biomédica, integrando principios de biología, informática y estadística para analizar y comprender datos biológicos complejos(Mount, 2004). Su aplicación es fundamental para desentrañar los mecanismos moleculares subyacentes en diversas enfermedades, incluyendo el carcinoma HCC, el tipo más común de cáncer de hígado. A través de herramientas bioinformáticas, los investigadores pueden identificar patrones en la expresión génica y las interacciones moleculares, facilitando la comprensión de la progresión y desarrollo de enfermedades como el HCC(Tomczak et al., 2015).

1.5.1. Obtención y cuantificación de datos de expresión génica

En la figura 4 se muestran los pasos de la secuenciación de ARN y análisis bioinformático. El análisis de expresión génica comienza con la recolección de muestras biológicas, que pueden provenir de tejidos, células en cultivo o fluidos biológicos(Conesa, 2016). La calidad y el procesamiento adecuado de estas muestras son fundamentales para garantizar la fiabilidad de los datos obtenidos.

Una vez obtenida la muestra, se extrae el ARN total utilizando métodos como la lisis celular y la purificación con kits comerciales basados en columnas de sílice o separación con fases orgánicas. Es crucial evitar la degradación del ARN, por lo que se utilizan inhibidores de ARNasa y condiciones controladas de temperatura. Posteriormente, la calidad e integridad del ARN se evalúan mediante electroforesis en gel de agarosa o sistemas automatizados como el Bioanalyzer, asegurando que las muestras sean aptas para su análisis(Wagner et al., 2012).

El siguiente paso es la conversión del ARN en ADN complementario (cDNA) mediante transcripción inversa. Este proceso es clave porque el ARN es una molécula inestable y difícil de manipular, mientras que el cDNA es más estable

y compatible con las técnicas de secuenciación. Una vez sintetizado el cDNA, se fragmenta en pequeños segmentos para facilitar su posterior secuenciación y se prepara una biblioteca de secuenciación mediante la adición de adaptadores específicos.

La secuenciación de ARN de alto rendimiento (RNA-Seq) es la tecnología más utilizada para obtener datos de expresión génica. Se basa en la lectura masiva de fragmentos de cDNA mediante plataformas como Illumina o Nanopore. Los datos crudos obtenidos son secuencias cortas (reads) que deben ser procesadas bioinformáticamente. Estas lecturas se alinean contra un genoma de referencia utilizando algoritmos como STAR(Dobin et al., 2013) o HISAT2(Kim et al., 2019), permitiendo cuantificar la abundancia de cada transcripto.

Para normalizar la expresión génica entre muestras, se utilizan métricas como los “Transcripts Per Million” (TPM), que corrigen el número de lecturas asignadas a un gen teniendo en cuenta la longitud del transcripto y el total de secuencias en la muestra. Esto permite realizar comparaciones precisas de expresión génica, eliminando sesgos derivados de diferencias en la profundidad de secuenciación.



Figura 4. Diagrama de flujo de la secuenciación de ARN y análisis bioinformático

1.5.2. Análisis de expresión génica diferencial

El análisis de expresión diferencial es una estrategia clave en bioinformática para identificar genes cuya expresión varía significativamente entre diferentes condiciones experimentales. Se aplica en estudios comparativos entre muestras

de pacientes y controles, tejidos sanos y tumorales, o células expuestas a distintos tratamientos.

Para determinar qué genes presentan cambios significativos en su expresión, se utilizan modelos estadísticos que comparan los niveles de expresión normalizados entre condiciones. Es fundamental aplicar métodos de corrección por comparaciones múltiples, como el ajuste del valor de p mediante el método de Benjamini-Hochberg, para controlar la tasa de falsos descubrimientos (Benjamini and Hochberg, 1995).

Este análisis permite identificar genes potencialmente relevantes en procesos biológicos específicos. Los genes diferencialmente expresados pueden ser marcadores de enfermedades, objetivos terapéuticos o elementos clave en la regulación de vías metabólicas. Además, el análisis de expresión diferencial es un paso previo esencial para estudios de enriquecimiento funcional y análisis de redes de interacción génica (Conesa, 2016).

1.5.3. Análisis de enriquecimiento funcional

El análisis de enriquecimiento funcional permite interpretar listas de genes identificando procesos biológicos, funciones moleculares o componentes celulares sobrerrepresentados. Este enfoque es fundamental para contextualizar la información obtenida del análisis de expresión diferencial y comprender cómo las alteraciones en la expresión génica impactan en las funciones celulares (Subramanian et al., 2005).

Los genes diferencialmente expresados se agrupan en categorías funcionales definidas en bases de datos especializadas, como ontologías génicas (Ashburner et al., 2000) o conjuntos de genes previamente asociados a funciones biológicas (Kanehisa et al., 2021). Este análisis revela qué procesos están afectados en una condición determinada y ayuda a priorizar genes para estudios experimentales posteriores.

El enriquecimiento funcional es ampliamente utilizado en estudios biomédicos, permitiendo, por ejemplo, identificar rutas metabólicas alteradas en enfermedades como el HCC. También es útil para descubrir interacciones moleculares relevantes (Jensen et al., 2009) y generar hipótesis sobre mecanismos biológicos subyacentes.

1.5.4. Bases de datos públicas

En los últimos años, el acceso a bases de datos públicas ha revolucionado la investigación en bioinformática. Plataformas como The Cancer Genome Atlas (TCGA) (Tomczak et al., 2015) y Gene Expression Omnibus (GEO) (Barrett et al., 2013) contienen información genómica, transcriptómica y epigenética de una amplia variedad de enfermedades.

Estos repositorios permiten a los investigadores acceder a grandes volúmenes de datos sin necesidad de generar sus propias muestras, lo que reduce costos y acelera los estudios. La disponibilidad de datos abiertos fomenta la reproducibilidad científica, ya que distintos grupos de investigación pueden

validar hallazgos previos en conjuntos de datos independientes(National Academies of Sciences, 2018).

A lo largo del tiempo, las bases de datos públicas han evolucionado para ofrecer información cada vez más detallada y estandarizada. Actualmente, incluyen metadatos clínicos, información de seguimiento de pacientes y datos multi-ómicos, lo que permite estudios integrados que combinan expresión génica, mutaciones genéticas y perfiles epigenéticos en un mismo análisis(Zhang et al., 2020).

1.5.5. Impacto de la bioinformática en la investigación biomédica

La integración de la bioinformática en la investigación biomédica ha transformado la manera en que se analizan y comprenden los datos biológicos. El uso de herramientas computacionales permite manejar grandes volúmenes de información de manera eficiente, facilitando la identificación de patrones en la expresión génica, la interacción molecular y la progresión de enfermedades como el HCC(Venter et al., 2001).

Antes del desarrollo de la bioinformática, el análisis de expresión génica se realizaba mediante técnicas de biología molecular que requerían largos tiempos de procesamiento y permitían estudiar solo un número limitado de genes. En contraste, los enfoques actuales permiten analizar el transcriptoma completo en cuestión de días, obteniendo información detallada sobre miles de genes simultáneamente(Wang et al., 2009).

Además, la bioinformática permite la integración de datos de diferentes fuentes, lo que amplía la capacidad de interpretación y la identificación de nuevas asociaciones biológicas(Ge et al., 2020). Al automatizar el análisis de datos, se reduce la posibilidad de errores humanos y se optimiza el tiempo de investigación, permitiendo que los científicos se enfoquen en la interpretación de los resultados y el diseño de nuevas estrategias experimentales.

En conjunto, estas herramientas han impulsado el desarrollo de estrategias diagnósticas y terapéuticas más efectivas, facilitando el descubrimiento de nuevos biomarcadores y el diseño de terapias personalizadas basadas en el perfil molecular de cada paciente(Schilsky and Kim, 2021).

2. Objetivos

2.1. Objetivo general: explorar la participación de la vía de la LTA4H en el desarrollo y progresión del HCC a partir de un análisis *in silico*.

2.2. Objetivos específicos:

→ 1: Comparar la expresión de LTA4H y sus receptores (BLT1 y BLT2) en tejido tumoral en relación con el tejido normal, teniendo en cuenta los estadios de HCC.

→ 2: Analizar la coexpresión génica y enriquecimiento funcional de dichos genes, para evidenciar potenciales funciones nuevas de la vía de señalización en estudio.

→ 3: Estudiar la influencia de la expresión de LTA4H, LTB4R y LTB4R2 en la respuesta inmune del microambiente tumoral del HCC.

→ 4: Investigar la correlación entre la sobreexpresión de LTA4H y la activación de rutas moleculares implicadas en la proliferación, la apoptosis y la supervivencia.

3. Metodología:

3.1. Obtención y preprocesamiento de datos

Los datos de expresión génica fueron descargados del repositorio The Cancer Genome Atlas (TCGA) mediante el paquete R/Bioconductor 'TCGAbiolinks' versión 2.30.3 (Colaprico, A, et al., 2016). Se incluyeron muestras únicas de tumores primarios de carcinoma hepatocelular (HCC) (n = 374), con una matriz inicial de 60.659 genes por paciente.

Variable	HCC Caso, n (%)	Hígado sano Caso, n (%)
<u>Edad al diagnóstico</u>		
<65	222 (59.3 %)	
>=65	151 (40.3 %)	
ND	1 (0.2 %)	
<u>Sexo</u>		
Varón	253 (67.6 %)	28 (56 %)
Mujer	121 (32.3 %)	22 (44 %)
<u>Raza</u>		
Blanco	185 (49.4 %)	34 (68 %)
Negra	17 (4.5 %)	7 (14 %)
Asiático	160 (42.7 %)	6 (12%)
ND	1 (0.2 %)	3 (6 %)
<u>AJCC TNM</u>		
I	173 (46 %)	
II	87 (23.2 %)	
III	85 (22.7 %)	
IV	5 (1.3 %)	
ND	24 (6.4 %)	
<u>Estado Vital</u>		
Vivo	243 (64.7 %)	
Muerto	130 (34.7 %)	

ND	1 (0.2 %)	
TOTAL (N)	374 (100 %)	50 (100 %)
AJCC TNM, American Joint Committee on Cancer ND, no disponible		

Tabla 1. Las características clínico-patológicas de los 374 pacientes con HCC de la base de datos TCGA.

El análisis se realizó en R versión 4.3.0, utilizando los paquetes 'data.table' versión 1.14.8 (Dowle, M.; Srinivasan, A., 2023) y 'tidyverse' versión 2.0.0 (Wickham, H. et al., 2019) para la manipulación y limpieza de los datos. Las cuentas crudas (raw counts) fueron filtradas para excluir genes con baja expresión (valores cercanos a cero en >80% de las muestras), reduciendo el ruido técnico y mejorando la potencia estadística.

La normalización se efectuó mediante el método de transcritos por millón (TPM), que ajusta la abundancia de cada gen según su longitud y la profundidad total de secuenciación. Este enfoque permite comparar niveles de expresión entre genes y entre muestras, evitando sesgos por longitud génica y sin asumir que la mayoría de los genes no están diferencialmente expresados.

Posteriormente, se seleccionaron los genes de interés LTA4H, LTB4R y LTB4R2.

La información clínica (edad, sexo, etnia y estado vital) se obtuvo del Genomic Data Commons (GDC), eliminando registros con valores faltantes. Para caracterizar la cohorte, se calcularon proporciones de sexo, raza y estado vital según la expresión de LTA4H. Las comparaciones se realizaron mediante pruebas de chi-cuadrado, asumiendo independencia entre categorías y tamaño muestral adecuado.

Las características de la cohorte se resumen en la Tabla 1.

3.2. Comparación de la expresión génica según estadio tumoral

Para evitar la fragmentación de los datos en múltiples subcategorías dentro de un mismo estadio, se homogenizó la nomenclatura de los estadios III y IV, agrupando sus subdivisiones en una única categoría.

La expresión génica cuantificada en Transcripts Per Million (TPM) fue analizada mediante pruebas estadísticas implementadas en R versión 4.3.0. Para comparar la expresión entre los distintos estadios tumorales, se utilizó un modelo lineal clásico de un factor (ANOVA). Las diferencias específicas entre grupos se evaluaron con comparaciones post hoc de Tukey (TukeyHSD), que ajusta los valores p para múltiples comparaciones bajo el supuesto de varianzas iguales.

Las visualizaciones se generaron con 'tidyverse' versión 2.0.0 (Wickham, H. et al., 2019). Se emplearon boxplots para mostrar la distribución de expresión por estadio tumoral.

3.3. Análisis de expresión diferencial de genes (DEGS)

El análisis de expresión diferencial se realizó utilizando el paquete DESeq2 versión 1.38.3 (Love et al., 2014), empleando como referencia la condición de tejido sano para las comparaciones. Se utilizaron datos "unstranded" crudos (raw counts) en lugar de TPM, lo que permite que DESeq2 aplique su modelo generalizado de Poisson-Negativa Binomial (NB-GLM), diseñado para manejar la sobre-dispersión típica de los datos de RNA-seq.

Antes del análisis, se filtraron los genes con menos de 10 lecturas totales para reducir el sesgo asociado a baja expresión y mejorar la potencia estadística, evitando falsos positivos derivados de valores muy cercanos a cero.

DESeq2 estima el log₂FoldChange ajustando los tamaños de biblioteca y corrigiendo por efectos de dispersión, y realiza pruebas estadísticas de Wald para cada gen, generando valores p que luego son ajustados mediante el método Benjamini-Hochberg (Benjamini and Hochberg, 1995) para controlar la tasa de falsos descubrimientos (False Discovery Rate, FDR).

Los genes diferencialmente expresados (DEGs) se definieron mediante un cribado bifactorial que combinó la significancia estadística y la magnitud del cambio: aquellos con p-ajustada < 0.01 (filtro de confianza estadística riguroso) y log₂FoldChange > 1 o < -1 (filtro de magnitud). Este enfoque garantiza la selección de genes con un efecto biológico sustancial (cambio de al menos el doble) y alta certeza estadística, clasificándolos como sobreexpresados o subexpresados según corresponda.

Para visualizar la distribución de los DEGs, se generó un Volcano plot utilizando el paquete EnhancedVolcano versión 1.18.0 (Blighe, et al., 2021), en el que los genes sobreexpresados se resaltaron en rojo y los subexpresados en azul. Los genes de interés LTA4H, LTB4R y LTB4R2 fueron destacados para facilitar su interpretación biológica.

3.4. Análisis de correlación y construcción de redes de genes asociados

Se analizaron los genes con las 20 correlaciones más altas y más bajas para cada uno de los tres genes de interés (LTA4H, LTB4R y LTB4R2) utilizando correlación de Spearman, que evalúa asociaciones monotónicas entre variables y no requiere distribución normal, aunque asume independencia entre observaciones. Esta elección permite capturar relaciones no lineales entre genes y es robusta frente a valores atípicos.

Para visualizar las relaciones, se construyeron redes interactivas mediante los paquetes `networkD3` y `visNetwork` (Allaire, 2017; B.V. and `visNetwork` contributors, 2023). Se generaron matrices de adyacencia a partir de los coeficientes de correlación, asignando valores de conexión entre los genes de interés y sus genes correlacionados. Cada red incluyó nodos correspondientes a los genes seleccionados, donde el tamaño del nodo refleja la magnitud de la correlación y las escalas de color diferencian correlaciones positivas y negativas.

Todas las visualizaciones fueron exportadas en formato HTML a “rPubs” para facilitar la exploración interactiva fuera del entorno de análisis.

Para el análisis de enriquecimiento funcional, se empleó Metascape (Zhou et al., 2019), considerando procesos biológicos de Gene Ontology (GO), vías de señalización de KEGG y rutas de Reactome. El método utiliza pruebas de hipergeométricas para evaluar la sobre-representación de genes en cada categoría, aplicando corrección de múltiples comparaciones mediante FDR.

3.5. Evaluación de la infiltración inmune y el microambiente tumoral

Las puntuaciones de infiltración inmune (immune score) y de microambiente tumoral (microenvironment score) se estimaron para cada muestra mediante el método de deconvolución `xCell` (Aran et al., 2017), implementado en el paquete homónimo de R. `xCell` combina técnicas de regresión de enriquecimiento por firmas génicas con ajustes de spillover entre tipos celulares, permitiendo cuantificar la contribución relativa de 64 subpoblaciones inmunes y estromales.

Las correlaciones entre las puntuaciones de `xCell` y la expresión génica de LTA4H, LTB4R y LTB4R2 (convertida a valores numéricos y validados para ausencia de NA) se calcularon con el coeficiente de Spearman (ρ) usando `cor.test` del paquete `stats` (R. C. Team, 2023). Este enfoque no paramétrico evalúa asociaciones monotónicas y es robusto frente a valores atípicos, aunque la potencia estadística depende del tamaño muestral y de la magnitud de la correlación. Se consideraron asociaciones fuertes ($|\rho| \geq 0.6$), moderadas ($0.3 \leq |\rho| < 0.6$) y débiles ($|\rho| < 0.3$), con significancia $p < 0.05$.

Los resultados se organizaron en tablas formateadas con `kableExtra` (Zhu, 2021) y se visualizaron mediante “heatmaps” generados con `pheatmap` (pheatmap, 2019), utilizando los colores: azul-blanco-rojo y sin agrupamiento jerárquico para preservar la estructura original del dataset. Todos los análisis se ejecutaron en RStudio (Rs. Team, 2023) bajo condiciones controladas de reproducibilidad, empleando `tidyverse` (Wickham, 2023) para manipulación de datos.

3.6. Análisis de enriquecimiento de vías mediante GSEA

Se evaluó el enriquecimiento de grupos de genes entre los fenotipos de alta y baja expresión de LTA4H a partir de la matriz de expresión génica. Para ello se utilizó GSEA (Gene Set Enrichment Analysis) (Subramanian et al., 2021), que

implementa un algoritmo basado en permutaciones de fenotipo para determinar si un conjunto de genes predefinido está sesgadamente distribuido hacia los extremos de la lista de genes ordenados por expresión.

Se analizaron firmas moleculares de bases de datos como Gene Ontology (GO), KEGG, Reactome y colecciones seleccionadas de MSigDB 3.0 (Liberzon et al., 2011). Se aplicaron 1000 permutaciones de fenotipo para estimar la significancia de cada vía, calculando un Normalized Enrichment Score (NES) que refleja el grado de sobre- o subregulación de la vía. Se consideraron vías sobrerreguladas aquellas con $NES > 0$ y subreguladas con $NES < 0$, tomando como significativas las que tenían $FDR-Q < 0.01$.

Para visualizar los resultados, se generó un heatmap del contraste fenotípico de LTA4H (alta vs baja), ordenando las muestras según su expresión y los genes de acuerdo con su contribución al enriquecimiento.

4. Resultados

4.1. Análisis de expresión diferencial de genes para determinar la posición de LTA4H, LTB4R y LTB4R2.

Las tres proteínas de interés en nuestro estudio eran LTA4H, BLT1 y BLT2, ya que son clave en la vía investigada. Identificamos que los genes que las codifican son los siguientes: LTA4H está codificada por el gen LTA4H, BLT1 por el gen LTB4R y BLT2 por el gen LTB4R2. Nos propusimos determinar si estos tres genes se encontraban dentro de los genes diferencialmente expresados (DEGs).

Como se puede ver en la figura 5, el análisis inicial incluyó un total de 60.660 genes. Se consideraron DEGs aquellos con un valor absoluto del cambio de expresión $Fold\ Change \geq 1$ y un FDR (False Discovery Rate) ≤ 0.01 . De esta forma, detectamos 8.066 genes sobreexpresados y 1.968 subexpresados. Este análisis permitió identificar cambios significativos en la expresión génica entre las condiciones evaluadas.

Sin embargo, al examinar específicamente los genes LTA4H, LTB4R y LTB4R2, descubrimos que ninguno de estos genes cumple con los criterios para ser considerado como diferencialmente expresado.

LTA4H presenta un valor de $p\ adj > 0.01$, lo cual indica que la diferencia en su expresión entre las condiciones comparadas no es estadísticamente significativa.

Por otro lado, los genes LTB4R y LTB4R2 muestran una expresión diferencial que es estadísticamente significativa, con valores de $p\ adj$ menores a 0.01. Esto sugiere que hay una alta probabilidad de que las diferencias observadas en la expresión de estos genes sean reales y no debidas al azar. Además, ambos genes exhiben una tendencia a estar sobreexpresados en la condición

experimental. No obstante, aunque la diferencia en la expresión es significativa, el $\log_2\text{FoldChange}$ de estos genes es menor a 1. Esto indica que la magnitud de la diferencia en su expresión no es muy grande, es decir, los cambios en la expresión de LTB4R y LTB4R2, aunque significativos, no son muy pronunciados.

Aunque los genes LTA4H, LTB4R y LTB4R2 no cumplieron los criterios globales de significancia estadística, la evidencia bibliográfica que respalda su implicación en procesos inflamatorios y tumorales motivó su análisis más detallado.

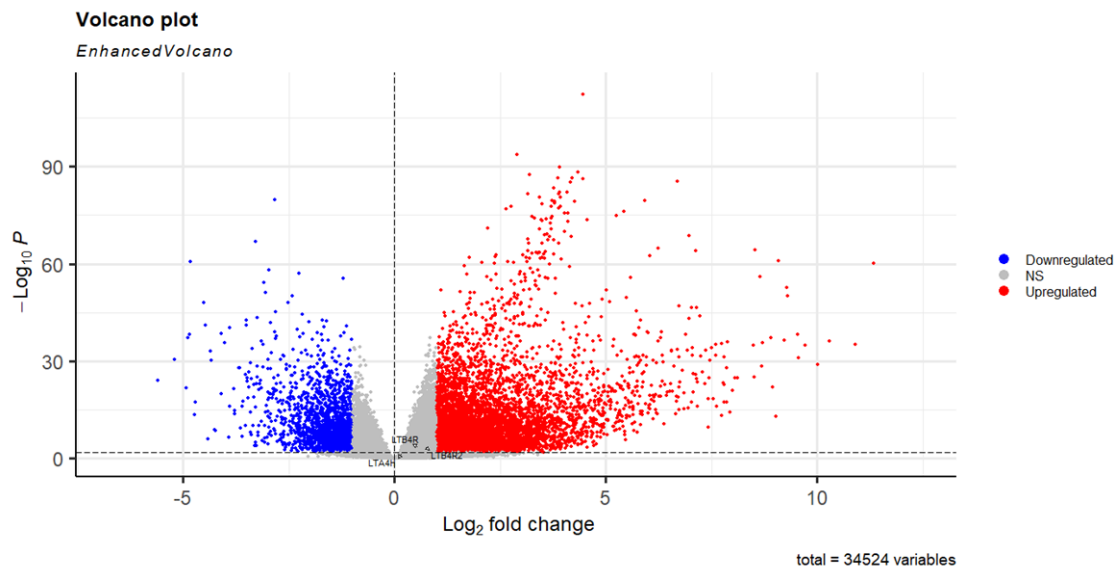


Figura 5. Volcano plot de los genes diferencialmente expresados (DEGs) en HCC, considerando un cambio de expresión (Fold Change) ≥ 1 y un FDR (False Discovery Rate) ≤ 0.01 . Los subexpresados se muestran en azul y los sobreexpresados en rojo. Las ubicaciones de los tres de interés (LTA4H, LTB4R, LTB4R2) están resaltadas y etiquetadas en el gráfico.

4.2. Diferencias de expresión de LTA4H, LTB4R y LTB4R2.

En nuestro análisis de expresión génica, como se muestra en la figura 6, observamos una significativa sobreexpresión en el carcinoma HCC en comparación con el tejido hepático sano para los tres genes clave de nuestra investigación: LTA4H, LTB4R y LTB4R2 ($p < 2.2e-16$ en todos los casos). Al desglosar la expresión por etapa de HCC según AJCC, observamos un patrón consistente en los tres genes. Hubo un aumento significativo en la expresión en las primeras tres etapas en comparación con el control, seguido de una disminución en la etapa 4, como se puede ver en la figura 7.

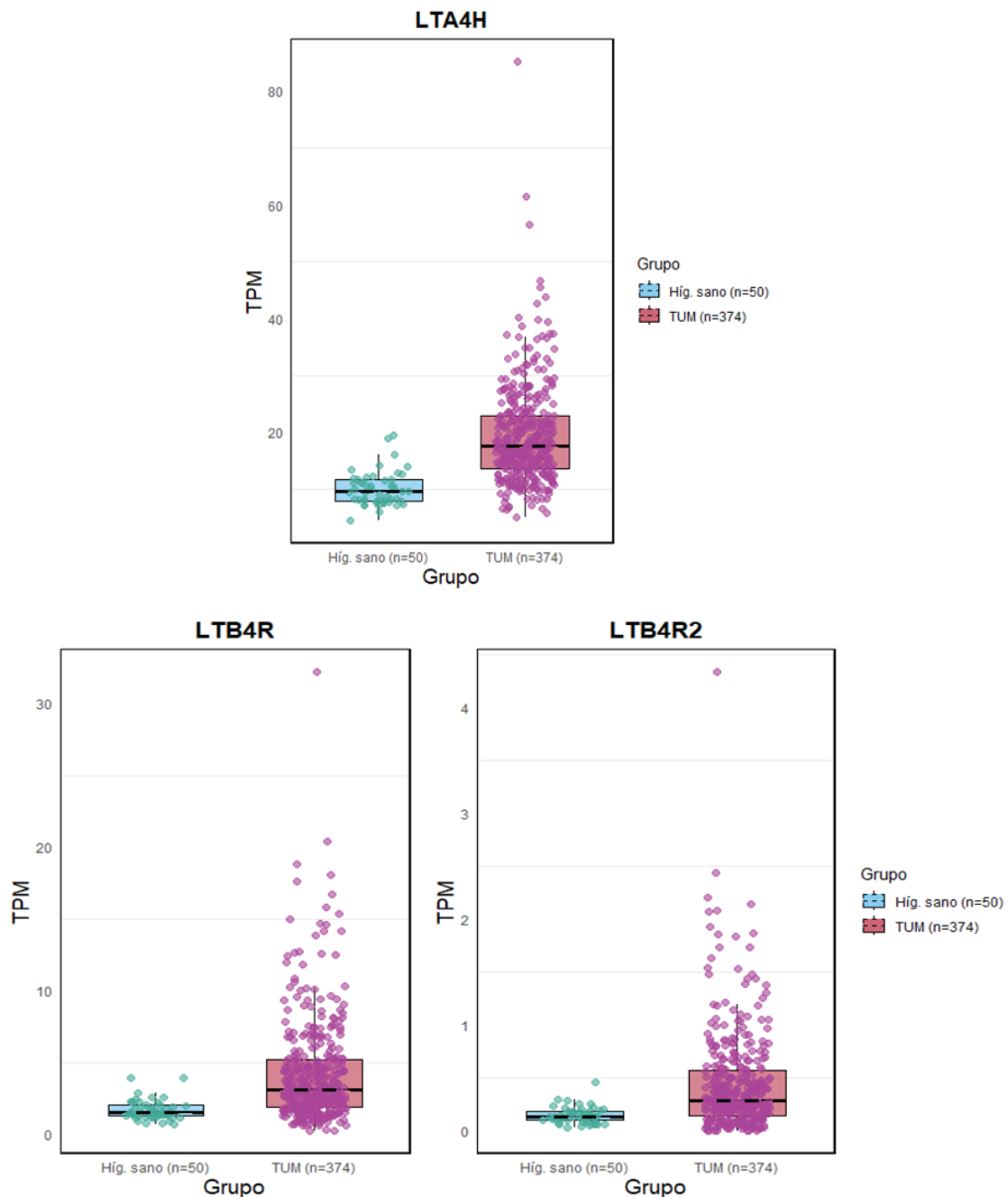


Figura 6. Comparación de la expresión de LTA4H, LTB4R y LTB4R2 entre todas las muestras de HCC (n=374) e hígado sano (n=50) en el conjunto de datos TCGA-LIHC. $p < 0.05$

Para LTA4H, un análisis de varianza (ANOVA) reveló un efecto significativo del estadio patológica AJCC en la expresión ($p < 0.001$). Las pruebas post-hoc de Tukey indicaron que los niveles de expresión fueron significativamente más altos en las etapas I, II y III en comparación con el tejido sano ($p < 0.001$), con una diferencia no significativa en la etapa IV ($p = 0.072$). De manera similar, para LTB4R y LTB4R2, las pruebas ANOVA mostraron efectos significativos de la etapa patológica AJCC en la expresión (LTB4R: $p < 0.001$; LTB4R2: $p < 0.001$). Las pruebas post-hoc de Tukey revelaron aumentos significativos en los niveles de expresión para las etapas I, II y III en comparación con el tejido sano ($p < 0.001$), con la etapa IV mostrando diferencias no significativas (LTB4R: $p = 1.000$; LTB4R2: $p = 0.841$).

Dado que la vía del enzima LTA4H está principalmente involucrada en la proliferación, se espera que tenga menos actividad en la etapa indicativa de metástasis.

Es así que la segmentación por estadio tumoral (I–III vs. tejido sano) aumentó la significancia de los resultados al reducir la heterogeneidad biológica y resaltar el contexto proliferativo temprano, donde la expresión de LTA4H, LTB4R y LTB4R2 mostró su mayor actividad diferencial.

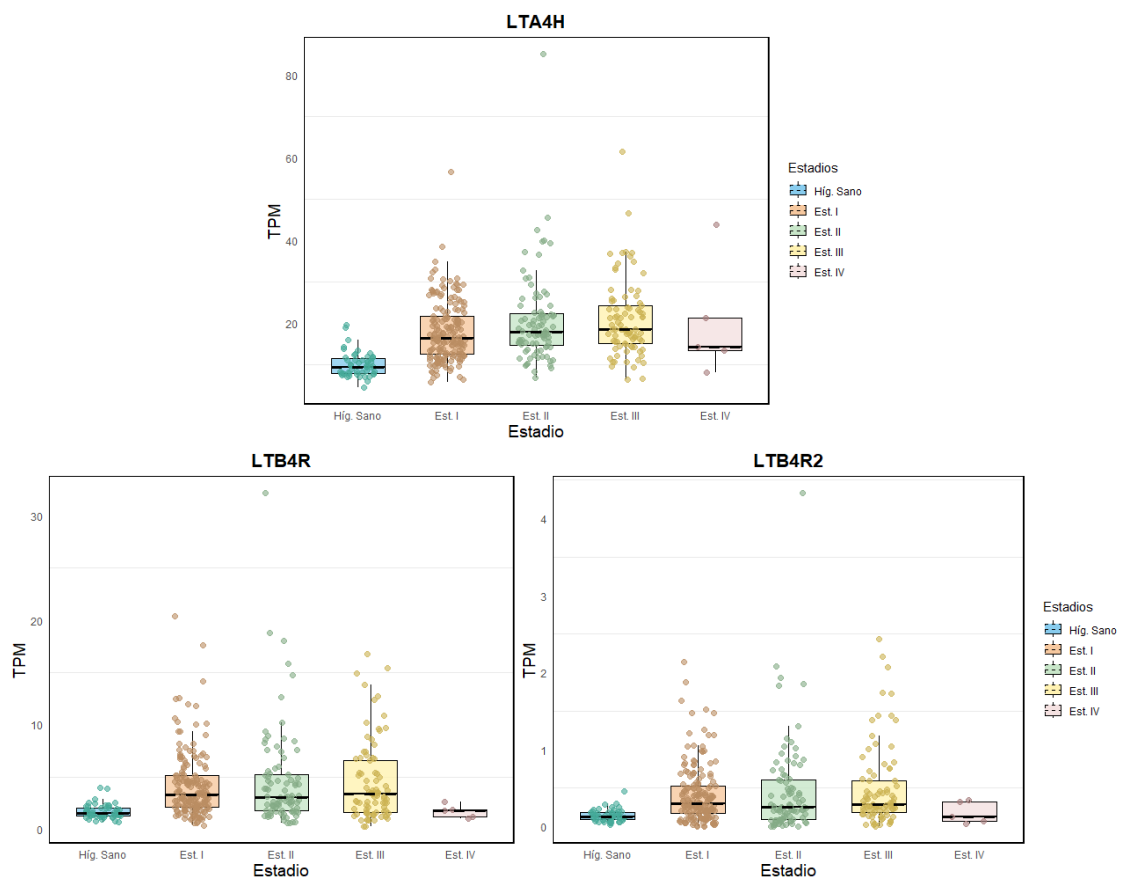


Figura 7. Comparación de la expresión de LTA4H, LTB4R y LTB4R2 entre los distintos estadios del HCC, analizada mediante la prueba post-hoc de Tukey. $p < 0.05$

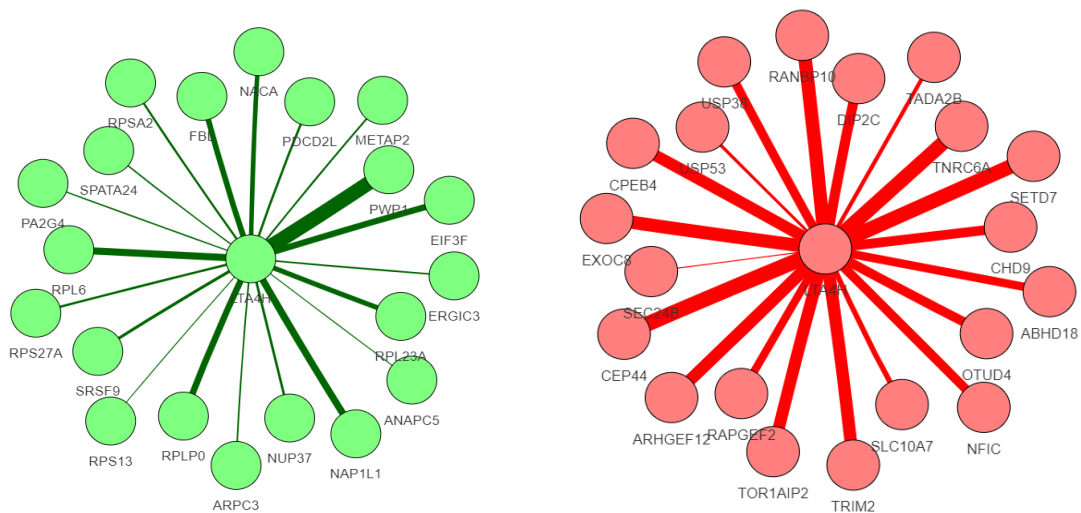


Figura 8. LTA4H y sus genes correlacionados en HCC. Los nodos representan los 20 genes con mayor correlación positiva (nodos verdes) y negativa (nodos rojos) con LTA4H. El grosor de las líneas indica la magnitud de la correlación.

4.3. Enriquecimiento funcional de los genes correlacionados con LTA4H, LTB4R y LTB4R2.

Se utilizó la base de datos cBioPortal para identificar los genes correlacionados con LTA4H, LTB4R y LTB4R2 en el conjunto de datos TCGA-LIHC. Se seleccionaron los 20 genes con mayor correlación positiva y negativa para cada uno de ellos, utilizando el coeficiente de correlación de Spearman. Los genes con mayor correlación positiva mostraron valores de Rho entre 0.493 y 0.393, mientras que los genes con mayor correlación negativa presentaron valores de Rho entre -0.376 y -0.318. En ambos casos, la fuerza de la correlación se considera moderada. Los figuras 8, 9 y 10 se muestran como imágenes estáticas, pero son interactivos. Para una mejor visualización, incluyo un [enlace](https://rpubs.com/luciaoviedobustos) a "Rpubs" donde se pueden explorar dinámicamente. <https://rpubs.com/luciaoviedobustos>

Luego, se realizó un análisis de enriquecimiento funcional con Metascape para cada uno de estos genes de manera individual (figura 11). Se presentaron los términos de enriquecimiento más significativos asociados a los genes correlacionados con los tres genes. Sin embargo, los genes relacionados positivamente con LTB4R2 no se incluyeron en los resultados, ya que ninguno de ellos cumplió con los criterios establecidos: $p < 0.01$, un mínimo de tres genes por término y un factor de enriquecimiento superior a 1.5. Si bien los términos no se repiten, están interrelacionados principalmente en torno a procesos de traducción, biogénesis de ribonucleoproteínas, y regulación genética a través de la remodelación de la cromatina y la señalización. También incluyen funciones

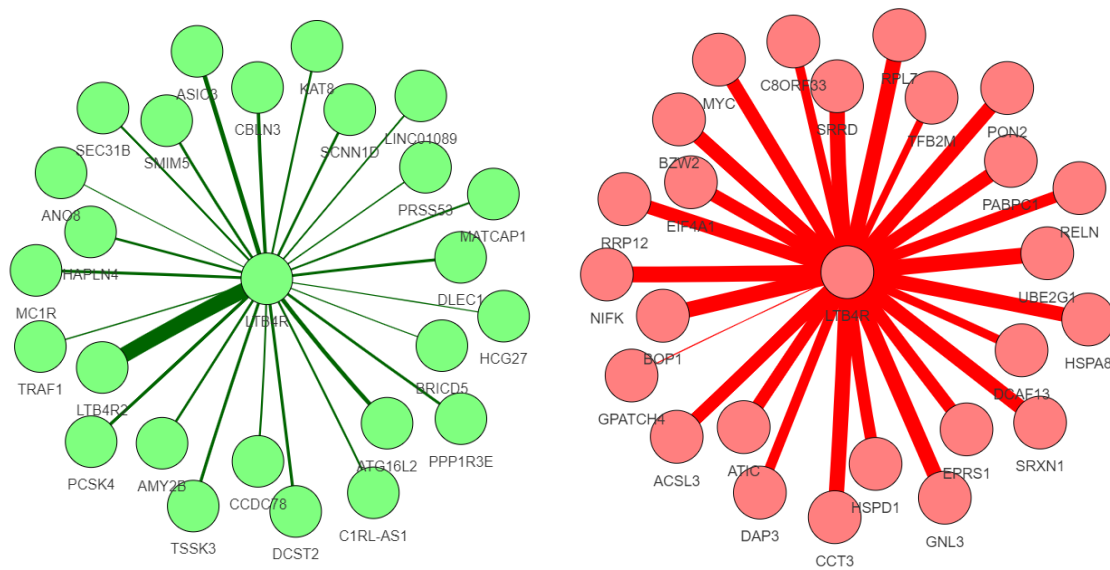


Figura 9. LTB4R y sus genes correlacionados en HCC. Los nodos representan los 20 genes con mayor correlación positiva (nodos verdes) y negativa (nodos rojos) con LTB4R. El grosor de las líneas indica la magnitud de la correlación.

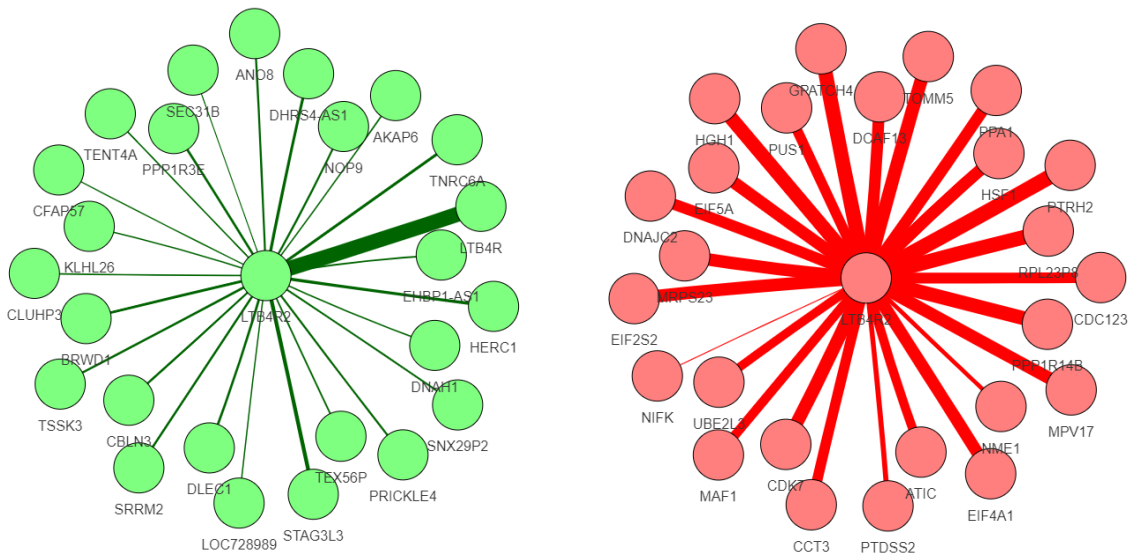


Figura 10. LTB4R2 y sus genes correlacionados en HCC. Los nodos representan los 20 genes con mayor correlación positiva (nodos verdes) y negativa (nodos rojos) con LTB4R2. El grosor de las líneas indica la magnitud de la correlación

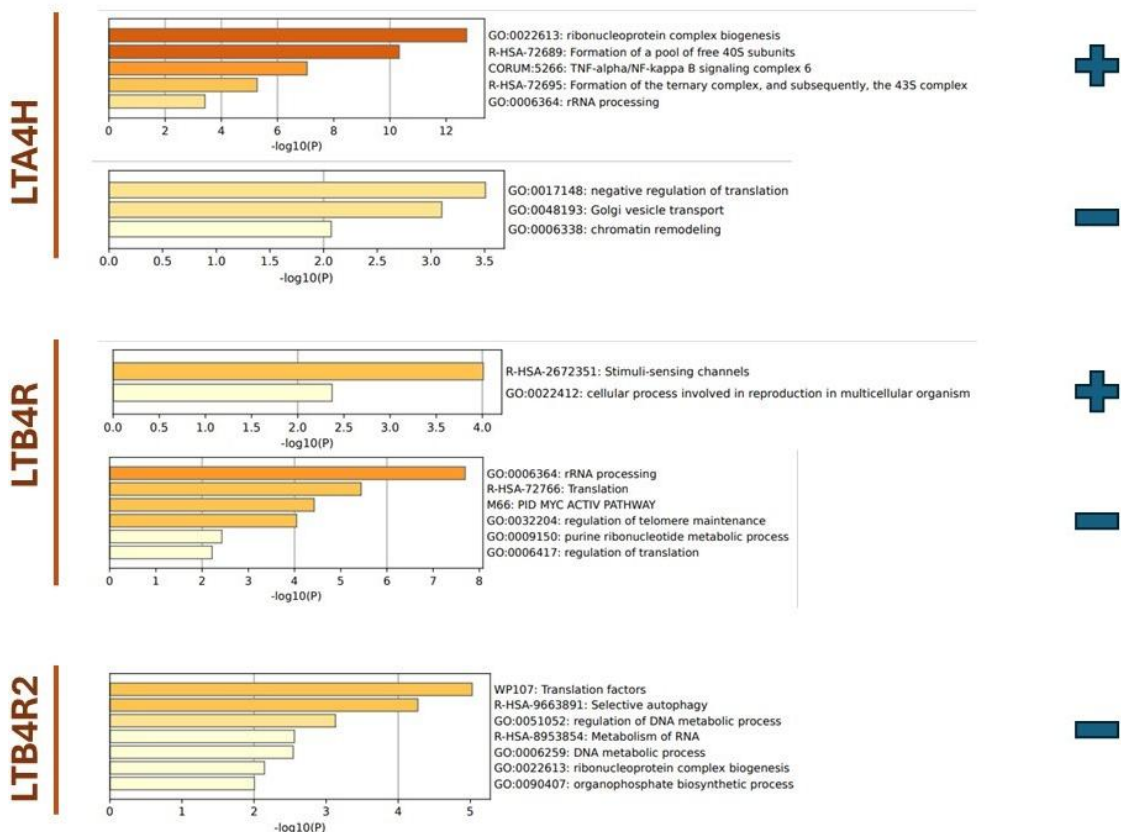


Figura 11. Enriquecimiento funcional de los genes correlacionados con LTA4H, LTB4R y LTB4R2. Se muestran los principales procesos enriquecidos por los 20 genes con mayor correlación positiva, así como aquellos asociados a los genes con mayor correlación negativa. Solo se incluyen los términos con $p < 0.01$ en los gráficos.

metabólicas esenciales y procesos de transporte celular, que están relacionados con la producción y regulación de proteínas.

4.4. Correlaciones entre LTA4H, LTB4R y LTB4R2 y las infiltraciones inmunes en hepatocarcinoma celular.

Como la vía de la LTA4H es típicamente descripta como inflamatoria, nos dispusimos a evaluar la correlación entre los infiltrados inmunológicos y los niveles de expresión de los 3 genes en estudio: LTA4H, LTB4R y LTB4R2.

Para ello, realizamos análisis de correlación utilizando distintas metodologías de inferencia de infiltración inmune, entre las que se incluyeron TIMER y xCell. En el análisis con TIMER se evaluaron las correlaciones entre los genes de interés y los niveles de infiltración de seis tipos de células inmunitarias, entre ellas linfocitos T (CD8+ y CD4+), células B, macrófagos, neutrófilos y células dendríticas. Se utilizó una correlación parcial ajustada según la pureza tumoral, lo que ayudó a reducir el efecto de la variabilidad entre muestras y a obtener estimaciones más precisas de la abundancia celular.

Los resultados obtenidos con TIMER revelaron patrones diferenciados en función de cada tipo celular. Si bien algunas correlaciones fueron estadísticamente significativas, al observar las gráficas de la figura 12 y analizar los valores de Rho obtenidos fueron bajos en ciertos casos, lo que sugiere que la relación entre la expresión génica y la infiltración inmune podría no ser

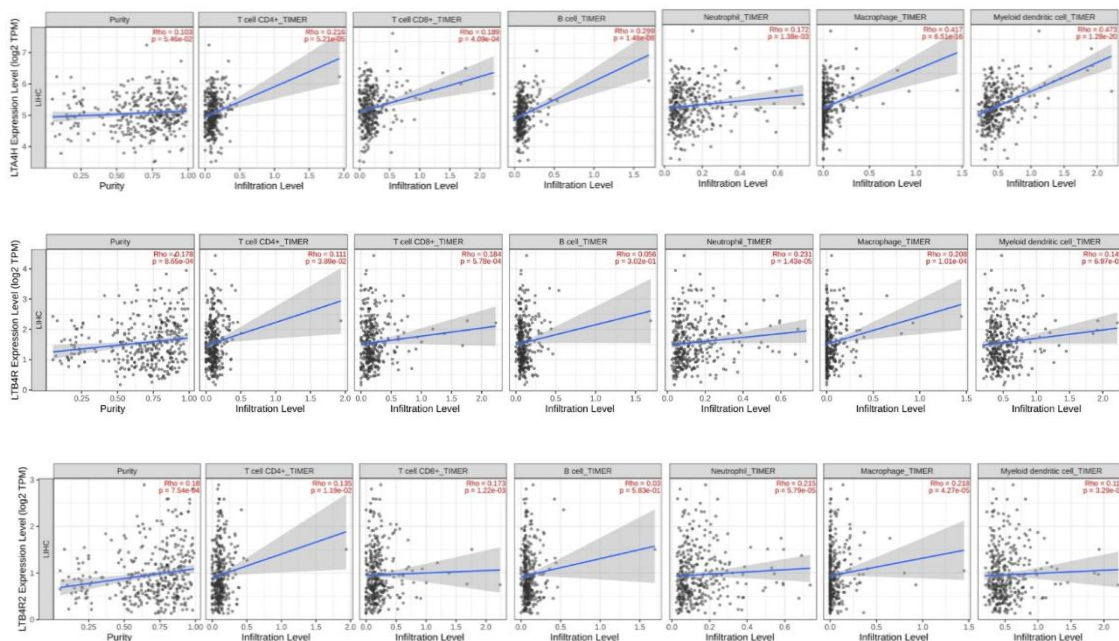


Figura 12. Análisis de correlación entre LTA4H, LTB4R y LTB4R2 y la infiltración inmune. Se presentan las correlaciones, corregidas por pureza, entre LTA4H, LTB4R y LTB4R2 y la infiltración de células B, células T CD8+, células T CD4+, macrófagos, neutrófilos y células dendríticas. El análisis se realizó mediante correlación parcial de Spearman, considerando $p < 0.01$ como significativo.

estrictamente lineal. Se observó que la expresión de LTA4H se correlacionaba de manera más marcada con la presencia de macrófagos y células dendríticas, sugiriendo que esta proteína podría estar implicada en la regulación de células vinculadas a funciones fagocíticas y de presentación antigénica. También, se detectaron asociaciones particulares en el contexto de los linfocitos T: mientras

que LTB4R mostró una tendencia a relacionarse con poblaciones de linfocitos T, la correlación no alcanzó significancia estadística, lo que apunta a una posible variabilidad en la forma en que este receptor modula la respuesta inmunitaria adaptativa. En el caso de LTB4R2, TIMER indicó una correlación significativa con ciertos tipos celulares, aunque la relación negativa esperada no fue claramente evidente en las gráficas. Esto sugiere que podrían intervenir otros factores en la regulación de su expresión y función en el microambiente tumoral.

	Puntuación	Método	LTB4R_coef	LTB4R_p	LTB4R2_coef	LTB4R2_p	LTA4H_coef	LTA4H_p
Inmune_xCell	Inmune	xCell	-0.0872129	0.0934622	-0.1729122	0.0008243	-0.0371595	0.4754906
Microambiente_xCell	Microambiente	xCell	-0.0808458	0.1200335	-0.1274259	0.0140440	-0.1650785	0.0014187

Tabla 2. Coeficiente de correlación entre los niveles de expresión génica y las puntuaciones de infiltración inmune y microambiente tumoral, junto con su valor de p. Se informa el coeficiente de correlación de Spearman (rho).

Además, los resultados obtenidos con xCell (tabla 2) mostraron correlaciones negativas entre la expresión de los genes LTA4H, LTB4R y LTB4R2 y las puntuaciones inmunes y del microambiente tumoral. En particular, LTB4R2 presentó la correlación negativa más pronunciada en el compartimento inmune ($p = -0.1729$, $p = 0.0008$), mientras que LTA4H mostró una correlación negativa significativa con la puntuación del microambiente tumoral ($p = -0.1650$, $p = 0.0014$). Por su parte, LTB4R exhibió una correlación negativa más débil, sin alcanzar la significación estadística en el contexto inmune ($p = 0.0934$) y presentando una tendencia similar en el microambiente tumoral ($p = 0.1200$).

Para profundizar en estos hallazgos, se generó un heatmap para visualizar la relación entre la expresión génica y la presencia de distintos tipos celulares en el tumor. Este análisis evidenció que la relación entre la expresión de LTA4H, LTB4R y LTB4R2 y el infiltrado inmunológico no es uniforme. En algunos casos, se observaron asociaciones particularmente marcadas con linajes específicos; por ejemplo, se destacó una correlación más fuerte en aquellas muestras con mayor abundancia de macrófagos M2 y células dendríticas activadas, situación que suele estar implicados en la creación de un microambiente inmunosupresor y en la promoción de la tolerancia tumoral. De igual modo, las diferencias en la magnitud de las correlaciones con linfocitos T (citotóxicos o reguladores) sugieren que la modulación de la respuesta inmunitaria por esta vía podría variar en función del estado de activación o de la polarización celular.

En conjunto, estos resultados resaltan la complejidad de la interacción entre la vía de la LTA4H y el microambiente tumoral. Tanto los datos de TIMER como los de xCell apuntan a que la influencia de LTA4H, LTB4R y LTB4R2 sobre la infiltración y la composición del microambiente no se comporta de manera lineal y uniforme, sino que depende en gran medida del tipo celular y del contexto inmunológico presente en cada muestra, si bien la relación entre LTA4H y las células dendríticas y los macrófagos, es llamativa. Esta variabilidad subraya la importancia de utilizar múltiples enfoques metodológicos para la estimación de la infiltración inmune y de considerar las particularidades de cada linaje celular al interpretar el papel de estos genes en la respuesta inmunológica tumoral.

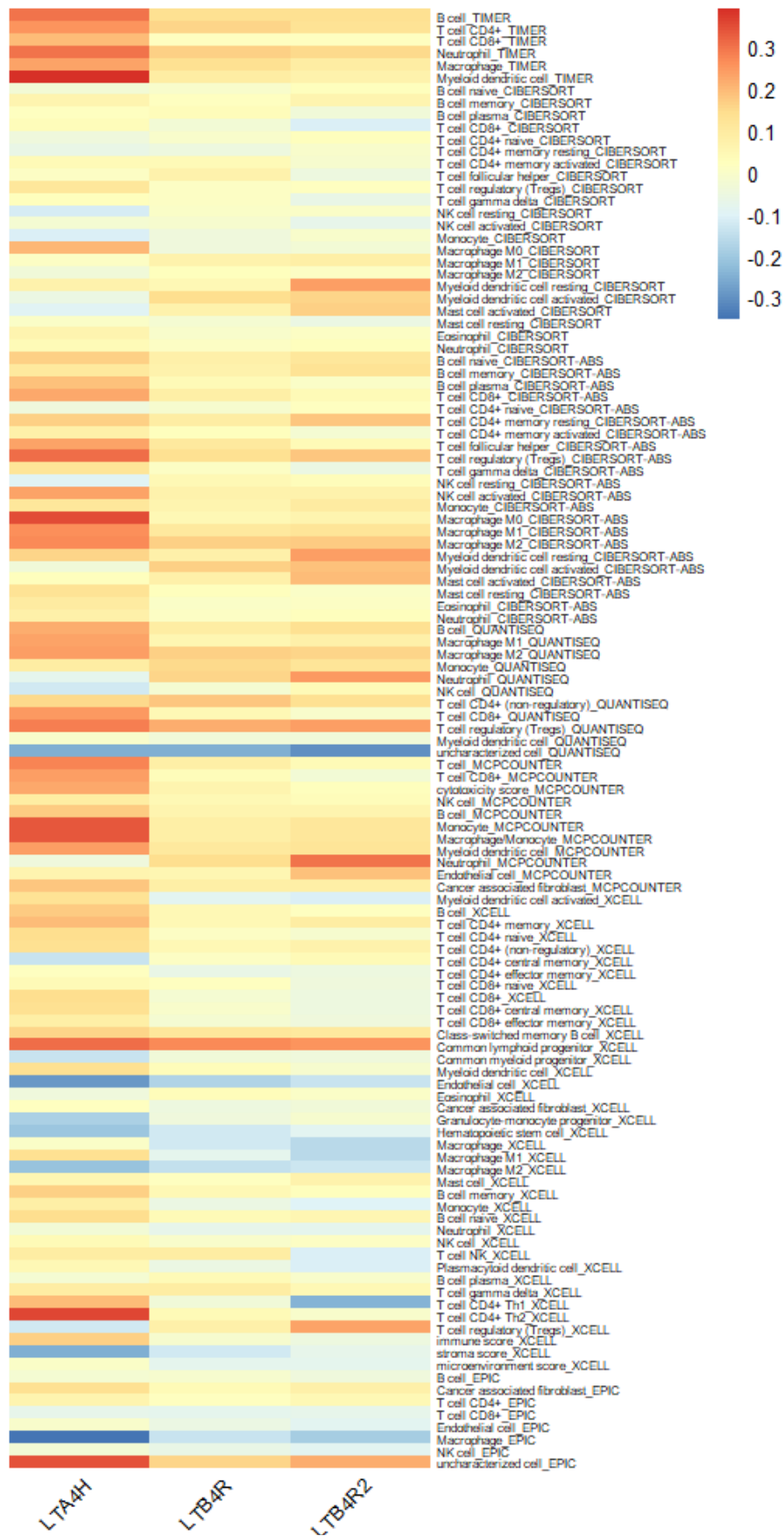


Figura 13. Heatmap que relaciona los perfiles celulares con la expresión de LTA4H, LTB4R y LTB4R2. La escala de colores a la derecha indica la intensidad y la dirección de la asociación.

4.5. Impacto de la sobreexpresión de LTA4H en la proliferación tumoral, apoptosis y pronóstico del carcinoma hepatocelular

El análisis GSEA reveló perfiles de expresión distintos para dos conjuntos de genes, categorizados como 'baja LTA4H' y 'alta LTA4H'. Utilizando el conjunto de datos de Reactome para el análisis de enriquecimiento funcional, el "heatmap" resultante destaca claras diferencias en los patrones de expresión génica entre los dos grupos de expresión de LTA4H.

Para investigar el impacto de los niveles de expresión de LTA4H, en los procesos más comunes involucrados en la progresión del HCC, realizamos un análisis GSEA utilizando varios conjuntos de genes. Nuestros hallazgos revelaron que el conjunto de genes "CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP", que

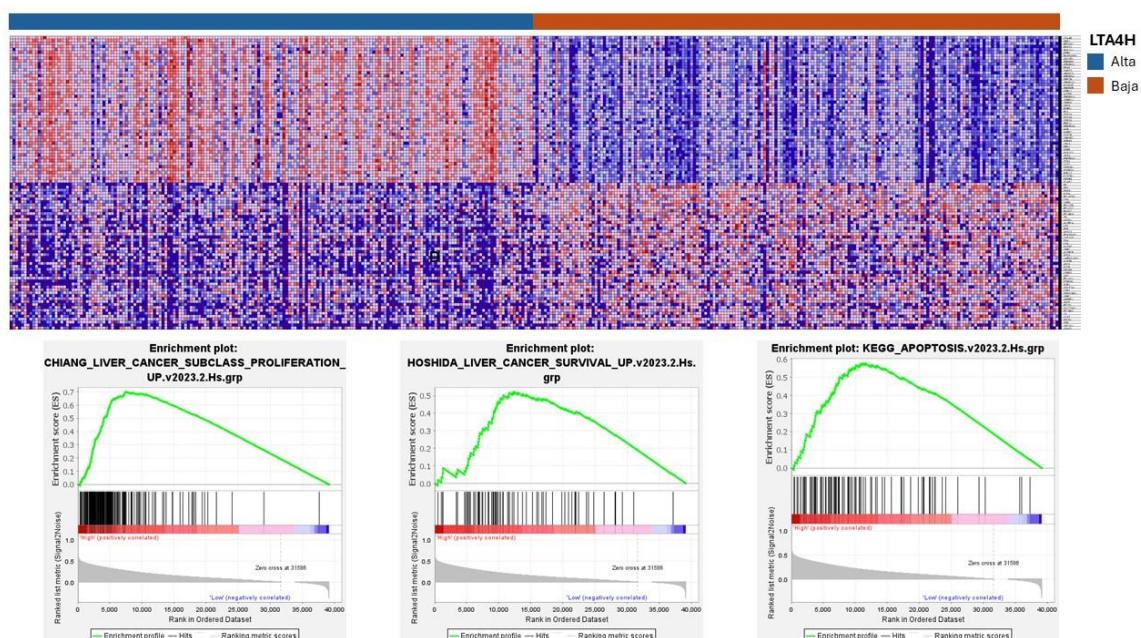


Figura 14. Asociación entre la expresión de LTA4H, distintas vías metabólicas y el análisis de supervivencia. Heatmap de los niveles de expresión de genes relacionados con el HCC en los dos grupos definidos por alta y baja expresión de LTA4H. Debajo están los gráficos de enriquecimiento que muestran la relación entre la expresión de LTA4H y proliferación, supervivencia y apoptosis en el cáncer de hígado.

comprende los 200 genes marcadores principales que están sobreexpresados en la subclase de 'proliferación' del HCC, mostró una asociación positiva en el grupo con alta expresión de LTA4H (ES 0.700, NES 1.821, $p = 0.0$). Esta subclase se caracteriza por un aumento en la proliferación, niveles elevados de alfa-fetoproteína sérica y una inestabilidad cromosómica.

Además, el conjunto de genes "HOSHIDA_LIVER_CANCER_SURVIVAL_UP", asociado con malos resultados de supervivencia en pacientes con HCC, también mostró una estar enriquecidos en el grupo de alta LTA4H (ES 0.521, NES 1.299, valor p 0.003). De manera similar, el conjunto de genes "KEGG_APOPTOSIS"

también se encontró enriquecido en el grupo con alta expresión de LTA4H (ES 0.577, NES 1.451, $p = 0.0$).

Estos resultados sugieren una correlación entre niveles elevados de LTA4H y un aumento en la proliferación, una mayor apoptosis y una menor supervivencia en HCC.

5. Discusión y conclusiones

En este estudio, hemos examinado la expresión diferencial y la relevancia biológica de los genes LTA4H, LTB4R y LTB4R2 en el contexto del carcinoma HCC. A pesar de que se identificaron miles de genes con expresión diferencial, ni LTA4H ni LTB4R ni LTB4R2, mostraron cambios estadísticamente significativos en el análisis inicial de expresión diferencial, lo que sugiere que cualquier variación observada en estos genes podría deberse al azar.

El análisis más detallado reveló una sobreexpresión de los tres genes en las primeras etapas del HCC en comparación con el tejido hepático sano. Curiosamente, esta sobreexpresión disminuyó en la etapa IV de la enfermedad, lo que podría reflejar un cambio en la dinámica biológica del tumor a medida que avanza hacia la metástasis. Este patrón fue consistente entre los tres genes, y los análisis estadísticos corroboraron la validez de estas observaciones. Estos resultados apuntan a un rol relevante de LTA4H, LTB4R y LTB4R2 en las etapas iniciales del HCC, probablemente vinculado con la proliferación celular, como lo indica la reducción en su expresión en las etapas más avanzadas.

El análisis de correlación entre la expresión de estos genes y la infiltración de células inmunes también ofrece información clave. Se observó una relación positiva entre la sobreexpresión de estos genes y la infiltración inmune, lo que sugiere que la vía de LTA4H podría estar involucrada en la modulación del microambiente tumoral. Sin embargo, la distribución asimétrica de los datos en los gráficos de dispersión sugiere que hay factores adicionales que podrían estar influyendo en estas correlaciones, lo que requiere un estudio más detallado.

Finalmente, el análisis de enriquecimiento funcional mediante GSEA demostró que los altos niveles de expresión de LTA4H están asociados con vías biológicas relacionadas con la proliferación celular, la apoptosis y la supervivencia deficiente en pacientes con HCC. Estos hallazgos son consistentes con la hipótesis de que LTA4H podría estar desempeñando un papel clave en la progresión del HCC, particularmente en la subclase proliferativa del cáncer de hígado.

Si bien, este es un estudio exploratorio donde no todos los resultados son concluyentes; en conjunto los hallazgos sugieren que LTA4H, LTB4R y LTB4R2

podrían ser objetivos prometedores para futuras investigaciones sobre el tratamiento y pronóstico del HCC.

6. Bibliografía

- Allaire, J.J., 2017. networkD3: D3 JavaScript Network Graphs from R.
- Anderson, N.M., Simon, M.C., 2020. The tumor microenvironment. *Current Biology*.
- Aran, D., HZ, xCell, B.A.J., 2017. Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Bioconductor*.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., GO Consortium, 2000. Gene ontology: tool for the unification of biology. *Nat Genet* 25, 25–29.
- Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., et al., 2013. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 41, D991–D995.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.
- Blighe, et al., 2021. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.
- B.V., A., visNetwork contributors, 2023. visNetwork: Network Visualization using “vis.js” Library.
- Chen, X., Wang, S., Wu, N., Yang, C., 2005. Leukotriene A4 Hydrolase as a Target for Cancer Prevention and Therapy. *Curr Cancer Drug Targets* 18;4(3):267–83.
- Chidambaranathan-Reghupaty, S., Fisher, P.B., Sarkar, D., 2021. Hepatocellular carcinoma (HCC): Epidemiology, etiology and molecular classification, in: *Advances in Cancer Research*. Academic Press Inc, pp. 1–61.
- Chun, Y.H., Kim, S.U., Park, J.Y., Kim, D.Y., Han, K.H., Chon, C.Y., 2011. Prognostic value of the 7th edition of the AJCC staging system as a clinical staging system in patients with hepatocellular carcinoma. *Eur J Cancer*.
- Colaprico, A., et al., 2016. TCGAAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44, e71. <https://doi.org/10.1093/nar/gkv1507>
- Conesa, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17, 13.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Dowle, M.; Srinivasan, A., 2023. data.table: Extension of ‘data.frame.’
- Duseja, A., 2014. Staging of Hepatocellular Carcinoma. *J. Clin. Exp. Hepatol.* Elsevier BV 4, 74–9.
- FAG, G.F.L., T.A. (Eds.), 2010. , in: *AJCC Cancer Staging Manual*. 7a. Springer, Edge SB, BDR, CCC.
- Ge, Y., Huang, H., Lu, J., et al., 2020. Integrated analysis of omics data to identify molecular networks in complex diseases. *Brief Bioinform* 21, 1753–1768.
- Hanahan, D., Weinberg, R.A., *Cell*, 2011. Hallmarks of cancer: The next generation.
- He, R., Chen, Y., Cai, Q., 2020. The role of the LTB4-BLT1 axis in health and disease. *Pharmacological Research*. Academic Press.
- Hu, N., Li, Y., Zhao, Y., Wang, Q., You, J.C., Zhang, X.D., 2011. A novel positive feedback loop involving FASN/p-ERK1/2/5-LOX/LTB4/FASN sustains high growth of breast cancer cells. *Acta Pharmacol Sin.*
- Jensen, L.J., Kuhn, M., Stark, M., Chr. S, A., Wernstedt, A., et al., 2009. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* 37, D412–D416.
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., Tanabe, M., 2021. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res* 49, D514–D521.
- Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L., 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915.
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., Mesirov, J.P., 2011. The Molecular Signatures Database (MSigDB) 3.0.

- Loftus L V., P.K.J., Amend SR, 2022. Interplay between Cell Death and Cell Proliferation Reveals New Strategies for Cancer Therapy. *Int. J. Mol. Sci. MDPI* 23.
- Love, M.I., HW, A.S., 2014. DESeq2: Differential gene expression analysis based on the negative binomial distribution. *Bioconductor*.
- Lozano, R., Naghavi, M., Lim, S.S., MPH, A., SY, A., MB, A.M.P.H., K.G., 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Internet* 380.
- Mount, D.W., 2004. *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press.
- National Academies of Sciences, E., and Medicine, 2018. *Reproducibility and Replicability in Science*. The National Academies Press, Washington, DC.
- Oi, N., Yamamoto, H., Langfald, A., Bai, R., Lee, M.H., Bode, A.M., 2017. LTA4H regulates cell cycle and skin carcinogenesis. *Carcinogenesis* 1;38(7):728–37.
- Park, J.I., Jang, J.H., Kim, J.H., 2019. Mediatory role of BLT2 in the proliferation of KRAS mutant colorectal cancer cells. *Biochim Biophys Acta Mol Cell Res* 1;1866(3):329–36.
- pheatmap, K.R., 2019. *Pretty Heatmaps*. CRAN.
- Rådmark, O., Samuelsson, B., 2010. Regulation of the activity of 5-lipoxygenase, a key enzyme in leukotriene biosynthesis. *Biochem Biophys Res Commun*.
- Sautès-Fridman, C., Cherfils-Vicini, J., Damotte, D., Fisson, S., Fridman, W.H., Cremer, I., 2011. Tumor microenvironment is multifaceted. *Cancer Metastasis Rev*.
- Schilsky, R.L., Kim, E.S., 2021. The future of precision oncology: current realities and challenges. *Nat Rev Clin Oncol* 18, 653–661.
- Serhan, C.N., Chiang, N., Dyke, T.E., 2008. Resolving inflammation: Dual anti-inflammatory and pro-resolution lipid mediators. *Nat. Rev. Immunol.* 8, 349–61.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U A* 102, 15545–15550.
- 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., 2021. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* 102, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>
- Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* May;71(3):209–49.
- Team, R.C., 2023. *stats: R Statistical Functions*. R Found. Stat. Comput.
- Team, R., 2023. *RStudio: Integrated Development Environment for R*. Posit Software, PBC.
- Tomczak, K., Czerwińska, P., Wiznerowicz, M., 2015. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol.* 19, A68–A77.
- Varghese, E., Samuel, S.M., Sadiq, Z., Kubatka, P., Liskova, A., Benacka, J., 2019. Anti-cancer agents in proliferation and cell death: The calcium connection. *Int J Mol Sci*.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., et al., 2001. The sequence of the human genome. *Science* 291, 1304–51.
- W, S., R, M., 2008. The eicosanoids: cyclooxygenase, lipoxygenase, and epoxygenase pathways, in: *Biochemistry of Lipids, Lipoproteins and Membranes*. pp. 331–362.
- Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131, 281–285.
- Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10, 57–63.
- Wickham, H., 2023. *Tidyverse: Easily Install and Load the “Tidyverse.”*
- Wickham, H. et al., 2019. Welcome to the tidyverse. *J. Open Source Softw.* 4, 1686. <https://doi.org/10.21105/joss.01686>
- Yang, J.D., Hainaut, P., Gores, G.J., Amadou, A., Plymoth, A., Roberts, L.R., 2019. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol. Nat. Publ. Group* 16, 589–604.
- Yang, P., Sun, Z., Chan, D., Cartwright, C.A., Vijjeswarapu, M., Ding, J., 2008. Zyflamend® reduces LTB4 formation and prevents oral carcinogenesis in a 7,12-

- dimethylbenz[α]anthracene (DMBA)-induced hamster cheek pouch model. *Carcinogenesis* 29, 2182–9.
- Zamor, P.J., AS, Russo, M.W., 2017. Viral hepatitis and hepatocellular carcinoma: etiology and management. *Journal of Gastrointestinal Oncology*. AME Publishing Company.
- Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., 2020. Integrated multi-omics analysis of cancer. *J Exp Clin Cancer Res* 39, 198–198.
- Zhao, S., Yao, K., Li, D., Liu, K., Jin, G., Yan, M., 2019. Inhibition of LTA4H by bestatin in human and mouse colorectal cancer. *EBioMedicine*.
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., Chanda, S.K., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523. <https://doi.org/10.1038/s41467-019-09234-6>
- Zhu, H., 2021. kableExtra: Construct Complex Table with “kable” and Pipe Syntax.