

Grado en ODONTOLOGÍA

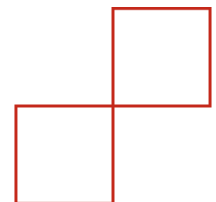
Trabajo Fin de Grado

Curso 2024-25

**Alteraciones genéticas asociadas a la invasión
angiolinfática en el carcinoma de células escamosas de
cabeza y cuello. Metaanálisis de expresión génica.**

Presentado por: Antonio Belli

**Tutoras: Marta Serna García
Nicla Flacco**



Agradecimientos

Llegar hasta aquí no ha sido un camino solitario. Este trabajo, y todo lo que representa, es también fruto del acompañamiento, la paciencia y el cariño de muchas personas que han estado a mi lado a lo largo de estos años.

Quiero agradecer a mi familia, mi mayor pilar. A mi padre, que ha sido mi compañero más fiel en este viaje. Nunca estuve solo: tus llamadas diarias, tus consejos a cualquier hora y tu apoyo incondicional, emocional y, sí, también económico, fueron el verdadero motor detrás de este logro. Gracias por ser mi mayor apoyo, mi modelo a seguir y por demostrarme cada día que la verdadera fuerza está en estar, sin hacer ruido.

Gracias, mamá, por tu infinita comprensión y por confiar en mí incluso en los momentos de duda. Tu apoyo silencioso, pero firme, ha sido una presencia esencial en mi vida y en este proceso.

A mi hermano, por venir hasta Valencia cuando hizo falta, por ayudarme con el piso, por estar ahí sin que tuviera que pedírtelo. Tu presencia, sencilla pero siempre oportuna, me dio alivio en momentos clave.

Y a mi abuela, por ser esa presencia constante que siempre estuvo.

A Margherita, porque sin ti esta historia probablemente habría sido distinta. Fuiste la persona que me animó a dar el salto, a salir de Italia y perseguir eso que durante tanto tiempo había soñado. Si estoy aquí, terminando esta etapa en Valencia, es en gran parte gracias a tu impulso, a tu confianza en mí cuando yo todavía no la tenía del todo. Compartimos mucho más que años: compartimos decisiones, sueños y caminos que hoy siguen, aunque en formas diferentes. Gracias por estar en un momento clave, por acompañarme.

Quiero agradecer A mis amigos de Alfedena, mi segunda familia. Crecimos juntos, compartimos veranos, inviernos, historias que no caben en ningún resumen, y, lo más importante, un vínculo que no ha cambiado, aunque hayan pasado los años o la distancia. Gracias por estar siempre y por recordarme que tengo un lugar al que siempre puedo volver.

A quienes compartieron conmigo los años en “Viale Ippocrate”, en Roma, donde la vida era caótica, ruidosa y, al mismo tiempo, increíblemente viva. Aquella casa fue mucho más que un lugar donde dormir. No éramos solo compañeros de piso: fuimos familia.

A Nicola, compañero de batallas, de estudio, de vida. Desde el primer día en Valencia fuiste más que un amigo: fuiste esa persona con la que se puede contar en todo momento, desde las cosas más simples hasta las más complicadas.

A Leo, que empezaste a estudiar con nosotros y, aunque dejaste la carrera, nuestra amistad siguió siendo la misma de siempre.

Y, por último, pero no por eso menos importante, quiero agradecer a todos los compañeros de universidad con los que compartí este camino.

A mis tutoras, gracias por haberme acompañado en este trabajo final con atención, claridad y paciencia. Vuestra guía fue fundamental para mantener el rumbo, incluso cuando las dudas eran muchas. Agradezco el tiempo, la dedicación y la confianza que depositaron en mí durante este proceso.

Y a la Universidad Europea, por haber sido el espacio donde pude formarme no solo como profesional, sino también como persona.

ÍNDICE

1. RESUMEN	7
2. ABSTRACT.....	9
3. PALABRAS CLAVES Y ABREVIATURAS.....	11
4. INTRODUCCIÓN	14
4.1. Generalidades.....	14
4.2. Invasión angiolinfática en el cáncer oral.....	17
4.3. Transcriptómica	20
4.4. Análisis de enriquecimiento	22
5. JUSTIFICACIÓN	25
6. HIPÓTESIS.....	26
7. OBJETIVOS.....	28
7.1. General.....	28
7.2. Específicos	28
8. MATERIALES Y MÉTODOS.....	30
8.1. Diseño del estudio	30
8.1.1. Selección de los datos	30
8.1.2. Análisis de expresión diferencial.....	32
8.1.3. Análisis de enriquecimiento	32
8.1.4. Análisis estadístico de las variables clínicas	33
9. RESULTADOS.....	35
9.1. Análisis de expresión diferencial.....	35
9.1. Análisis de enriquecimiento con DAVID software	37
9.2. Análisis de variables clínicas en la invasión angiolinfática en HNSCC.....	38
9.2.1. Tabaco	38
9.2.2. Alcohol	40
9.2.3. Edad	43
9.2.4. Sexo	45
10. DISCUSIÓN	50
11. CONCLUSIÓN.....	64
12. BIBLIOGRAFÍA	66
13. ANEXOS	71

1. RESUMEN

Introducción:

El carcinoma escamoso de cabeza y cuello (HNSCC) es una neoplasia maligna agresiva con alta incidencia y mortalidad a nivel mundial. Un factor clave en su progresión es la invasión angiolímfática (IAL), que facilita la diseminación metastásica y se asocia con un peor pronóstico. Este estudio realiza un metaanálisis de expresión génica diferencial en muestras de HNSCC con y sin IAL, utilizando datos transcriptómicos del consorcio TCGA. Se identifican firmas moleculares clave relacionadas con la invasión angiolímfática. Además, mediante análisis de enriquecimiento funcional, se exploran vías biológicas implicadas en la progresión tumoral y la invasión angiolímfática. Estos hallazgos pueden contribuir al desarrollo de biomarcadores pronósticos.

Materiales y Métodos:

Se analizaron 358 muestras transcriptómicas de HNSCC extraídas de la base de datos TCGA, clasificadas en 126 con invasión angiolímfática (IAL) y 232 sin ella. Se aplicó un análisis de expresión diferencial con los paquetes EdgeR y Limma ($FDR \leq 0,05$), seguido de un análisis de enriquecimiento funcional con DAVID. Además, se evaluaron variables clínicas mediante pruebas estadísticas (χ^2 , test exacto de Fisher, t de Student y Mann–Whitney U).

Resultados:

Se identificaron 375 genes diferencialmente expresados: 274 sobreexpresados y 101 infraexpresados en tumores con IAL. Los genes más relevantes incluyeron *KRT18*, *GADD45G*, *SIGIRR*, *S100A12*, *DSG1*, *GSDMA* y *WFDC12*. Los procesos más afectados fueron la exocitosis, la adhesión célula-célula y la respuesta inmune. El tabaquismo fue la única variable clínicamente asociada de forma significativa con la presencia de IAL ($p < 0,05$), mientras que el sexo fue estadísticamente significativo, aunque el resultado debe interpretarse con cautela debido al desequilibrio muestral. La edad mostró una relación débil y el alcohol no presentó asociación significativa.

Conclusión:

Los resultados del estudio evidenciaron diferencias claras en la expresión génica entre los tumores con y sin invasión angiolímfática, identificando genes relevantes y procesos biológicos implicados en la progresión tumoral. Se propusieron posibles biomarcadores asociados a la adhesión celular, exocitosis y respuesta inmune. Además, se observaron asociaciones clínicas significativas con el tabaquismo y con el sexo, aunque esta última requiere una interpretación cautelosa por el desequilibrio en la muestra.

2. ABSTRACT

Introduction:

Head and neck squamous cell carcinoma (HNSCC) is an aggressive malignant neoplasm with high incidence and mortality worldwide. A key factor in its progression is lymphovascular invasion (LVI), which facilitates metastatic dissemination and is associated with a worse prognosis. This study performs a meta-analysis of differential gene expression in HNSCC samples with and without LVI, using transcriptomic data from the TCGA consortium. Key molecular signatures related to lymphovascular invasion are identified. In addition, functional enrichment analysis is used to explore biological pathways involved in tumor progression and lymphovascular invasion. These findings may contribute to the development of prognostic biomarkers.

Materials and Methods:

A total of 358 transcriptomic HNSCC samples from the TCGA database were analyzed, classified into 126 with lymphovascular invasion (LVI) and 232 without. Differential expression analysis was performed using EdgeR and Limma packages ($FDR \leq 0.05$), followed by functional enrichment analysis with DAVID. Clinical variables were also evaluated using statistical tests (χ^2 , Fisher's exact test, Student's t-test, and Mann–Whitney U test).

Results:

A total of 375 differentially expressed genes were identified: 274 upregulated and 101 downregulated in tumors with LVI. The most relevant genes included *KRT18*, *GADD45G*, *SIGIRR*, *S100A12*, *DSG1*, *GSDMA*, and *WFDC12*. The most affected processes were exocytosis, cell-cell adhesion, and immune response. Tobacco use was the only clinical variable significantly associated with the presence of LVI ($p < 0.05$), while sex was statistically significant, although the result should be interpreted with caution due to sample imbalance. Age showed a weak relationship and alcohol did not present a significant association.

Conclusion:

The results of the study showed clear differences in gene expression between tumours with and without angiolymphatic invasion, identifying relevant genes and biological processes involved in tumour progression. Possible biomarkers associated with cell adhesion, exocytosis and immune response were proposed. In addition, significant clinical associations were observed with smoking and sex, although the latter requires cautious interpretation due to sample imbalance.

3. PALABRAS CLAVES Y ABREVIATURAS

- I. **Cáncer de cabeza y cuello**
- II. **Invasión angiolinfática**
- III. **Transcriptómica**
- IV. **DAVID**
- V. **Análisis de enriquecimiento**
- VI. **Biomarcadores de pronóstico**
- VII. **Genes**
- VIII. **Rutas biológicas**

ABREVIATURAS

HNSCC:	Carcinoma de células escamosas de cabeza y cuello (<i>Head and Neck Squamous Cell Carcinoma</i>)
IAL:	Invasión angiolinfática
TCGA:	<i>The Cancer Genome Atlas Project</i>
ACP:	Análisis de Componentes Principales
NCI:	<i>National Cancer Institute</i>
DEGs:	Genes diferencialmente expresados (<i>Differentially Expressed Genes</i>)
DAVID:	<i>Database for Annotation, Visualization, and Integrated Discovery</i>
TPM:	<i>Transcripts Per Million</i> (Transcripciones por millón, medida de normalización de datos de expresión génica)
FDR:	<i>False Discovery Rate</i> (Tasa de descubrimiento falso)
WGCNA:	<i>Weighted Gene Co-expression Network Analysis</i> (Análisis de redes de coexpresión génica ponderadas)
NGS:	<i>Next Generation Sequencing</i> (Secuenciación de nueva generación)
mRNA:	ARN mensajero (<i>Messenger Ribonucleic Acid</i>)
rRNA:	ARN ribosomal (<i>Ribosomal Ribonucleic Acid</i>)
tRNA:	ARN de transferencia (<i>Transfer Ribonucleic Acid</i>)
DNA:	Ácido desoxirribonucleico (<i>Deoxyribonucleic Acid</i>)
GO:	<i>Gene Ontology</i> (Ontología de genes)

KEGG: *Kyoto Encyclopedia of Genes and Genomes* (Enciclopedia de genes y genomas de Kyoto)

4. INTRODUCCIÓN

4.1. Generalidades

El cáncer es una enfermedad compleja que se puede describir como una proliferación incontrolada y anormal de células. La mayoría de las veces, estas células anormales pueden propagarse e invadir los tejidos circundantes, incluso migrando y llegando a otras partes del cuerpo utilizando el torrente sanguíneo o el sistema linfático.

El carcinoma de células escamosas de cabeza y cuello (HNSCC) es un conjunto heterogéneo de neoplasias malignas que se originan en el epitelio escamoso de la mucosa, incluyendo la cavidad oral, la orofaringe, la hipofaringe y la laringe, tiene una estructura anatómica única y un microambiente tumoral impredecible (1). La localización anatómica del HNSCC influye significativamente en el pronóstico. Según un estudio retrospectivo realizado en Brasil, las tasas de supervivencia global a cinco años varían según el sitio anatómico: 37.0% para cáncer de laringe, 35.8% para cavidad oral, 26.6% para orofaringe y 14.8% para hipofaringe (2). Se origina en las células escamosas planas, una fina capa de tejido en la superficie de las estructuras de la cabeza y el cuello. Si este cáncer se limita a la capa de células escamosas se llama carcinoma in situ y si las células cancerosas de HNSCC proliferan más allá de esta capa celular e invaden los tejidos más profundos, se clasifica como invasivo (1).

Se estima que el HNSCC constituye aproximadamente el 90% de los tumores malignos en esta región y, según datos de GLOBOCAN, se sitúa en el puesto 7 en incidencia a nivel mundial, con cerca de 890,000 casos nuevos al año y alrededor de 450,000 muertes. El HNSCC representa aproximadamente el 4.5% de todos los diagnósticos y muertes por cáncer a nivel mundial. Hay una gran variación geográfica en la incidencia y distribución anatómica del HNSCC en todo el mundo, En los países en desarrollo, la incidencia está aumentando debido al incremento en el consumo de tabaco (fumado o masticado), alcohol y nuez de areca (betel). En cambio, en los países desarrollados, el HNSCC relacionado con el VPH supera al asociado con el consumo de tabaco y alcohol (3).

A nivel mundial, el HNSCC afecta a los hombres entre dos y cuatro veces más que a las mujeres, con estimaciones que superan los 20 casos por cada 100,000 habitantes. Se proyecta que

la incidencia del HNSCC aumentará en un 30% para el año 2030, alcanzando aproximadamente 1.08 millones de nuevos casos anuales (4).

El diagnóstico y el manejo actuales del HNSCC son desafiantes. Actualmente no hay una sola herramienta moderna disponible que pueda detectar y predecir con precisión la progresión de este tipo de cáncer (1). La heterogeneidad molecular del HNSCC representa otro desafío significativo en su manejo. Esta heterogeneidad se manifiesta tanto entre tumores de diferentes pacientes (intertumoral) como dentro de un mismo tumor (intratumoral). Las variaciones genéticas y epigenéticas entre las células tumorales pueden influir en la respuesta al tratamiento y en la progresión de la enfermedad (5). Por ejemplo, se ha observado que diferentes regiones de un mismo tumor pueden albergar sub clones con distintas mutaciones, lo que complica la eficacia de las terapias dirigidas y contribuye a la resistencia al tratamiento (6).

La tasa general de supervivencia a 5 años no ha mostrado una mejoría significativa, y el número de individuos que sucumben al HNSCC sigue siendo alto. El problema radica en el hecho de que los HNSCC se detectan típicamente en etapas avanzadas, contribuyendo considerablemente a las altas tasas de mortalidad asociadas con esta enfermedad (1). Más de la mitad de los pacientes con carcinoma escamoso de cabeza y cuello (HNSCC) e son diagnosticados en etapas avanzadas (regional o metastásica), lo que limita las opciones terapéuticas y reduce significativamente la supervivencia global. En un estudio, que analizó datos poblacionales de cáncer en Europa, se observó que el 54% de los casos de HNSCC se diagnosticaban en estadios avanzados. La supervivencia relativa a cinco años para los tumores localizados oscilaba entre el 42% en los de hipofaringe y el 74% en los de laringe (7). Métodos para detectar la presencia de HNSCC en las primeras etapas y el uso de terapias personalizadas pueden ayudar a mejorar los resultados del tratamiento de HNSCC, y esto puede ser posible mediante el uso de nuevos biomarcadores para este tipo de cáncer.

El HNSCC se asocia a diversos factores de riesgo, siendo el consumo de tabaco y alcohol los más prevalentes, contribuyendo al desarrollo de aproximadamente el 80% de los casos a nivel mundial. El tabaquismo excesivo de productos de tabaco y sus derivados, junto con el consumo regular y fuerte de alcohol, especialmente cuando se combinan, aumentan significativamente el riesgo de desarrollar HNSCC. La combinación del alcohol y del tabaco tiene una mayor asociación con la aparición de cánceres de la laringe (89%) que con los cánceres orales (64%) (4). Además, la

infección por el virus del papiloma humano (VPH), comúnmente transmitida a través de la actividad sexual oral, y la exposición prolongada al sol son factores importantes, contribuyendo al desarrollo de aproximadamente el 80% de los casos a nivel mundial (3).

El tabaquismo excesivo de productos de tabaco y sus derivados, junto con el consumo regular y fuerte de alcohol, especialmente cuando se combinan, aumentan significativamente el riesgo de desarrollar HNSCC. Otros factores de riesgo incluyen la enfermedad por reflujo gastroesofágico y la enfermedad por reflujo laringofaríngeo, ya que el flujo hacia atrás de ácido estomacal hacia las vías respiratorias superiores puede contribuir al desarrollo de HNSCC (1). También está relacionado con la masticación de la nuez de areca, mayormente consumida en países en desarrollo por su efecto eufórico, psicoactivo y de supresión del apetito. Otros factores incluyen la exposición ocupacional a sustancias carcinógenas, la mala higiene oral, la predisposición genética y el estado socioeconómico bajo, que puede limitar el acceso a cuidados de salud y aumentar el riesgo de diagnósticos en estadios avanzados (3). Además, estudios recientes han señalado que la microbiota oral alterada podría actuar como un factor contribuyente en la carcinogénesis del HNSCC, promoviendo la inflamación crónica y alteraciones del microambiente epitelial (8). La mayoría de los factores de riesgo del cáncer oral son modificables, por lo cual juega un rol importante en la prevención la motivación al cambio de los estilos de vida y la vacunación contra el VPH (3).

Cuando hablamos de cáncer, es fundamental conocer el concepto de *tumor microenvironment (TME)*. Este entorno cercano al tumor primario juega un papel crucial en procesos como la recurrencia y la metástasis (9,10). Se compone de fibroblastos, células inmunes, componentes de la matriz extracelular, vasos sanguíneos, linfocitos, células inflamatorias y moléculas señalizadoras, controlando el desarrollo tumoral a través de diversas vías de señalización (11). Este entorno también participa activamente en la facilitación de la invasión angiolímfática (ALI), un proceso clave en la diseminación tumoral. Se ha demostrado que los fibroblastos asociados al cáncer (CAFs) promueven la migración e invasión de células tumorales mediante mecanismos como el eje *BDNF-TrkB*, favoreciendo la metástasis linfática y vascular (12).

Un TME alterado puede favorecer la infiltración tumoral y la metástasis, caracterizándose por hipoxia, pH reducido, alta presencia de especies reactivas de oxígeno (ROS), vascularización desorganizada y aumento de la presión osmótica (9). Investigaciones recientes han demostrado que

la modulación del TME influye en la respuesta a la terapia y en la progresión tumoral (11,13). Además, ciertos genes implicados en la respuesta inmune y la heterogeneidad intratumoral del HNSCC, influenciada por la diversidad del microambiente tumoral, contribuye significativamente a la progresión tumoral y a la resistencia terapéutica (14).

En particular, la hipoxia tumoral induce respuestas celulares que promueven la angiogénesis y la evasión inmune, contribuyendo a la resistencia a terapias convencionales como la radioterapia. Este fenómeno resalta la importancia de estrategias terapéuticas dirigidas al TME para mejorar la efectividad del tratamiento contra el cáncer (11).

Es común ver alteraciones a nivel de la matriz extracelular en los *TMEs*; se ha visto que los tumores provocan la formación de matriz extracelular y que en cambio la matriz extracelular puede regular la proliferación de las células tumorales. Las funciones generales de la matriz extracelular son la regulación del crecimiento, de la supervivencia, la diferenciación y de la movilidad celular. Estas funciones se llevan a cabo por medio de uniones a diferentes receptores como por ejemplo las integrinas. Además, la matriz extracelular se encarga de la correcta homeostasis celular y está adaptada según el órgano, para poder asegurar su función de una manera óptima (15,16).

Muchos estudios han demostrado que el brote tumoral, la profundidad de la invasión y la invasión angiolinfática (IAL) son predictores importantes del pronóstico, incluyendo la supervivencia a largo plazo y la probabilidad de metástasis ganglionar (*LNM*). Estos factores están estrechamente relacionados con la agresividad del tumor y la respuesta al tratamiento, por lo que su evaluación es fundamental para la planificación terapéutica (17). La invasión angiolinfática se ha identificado como un factor pronóstico negativo, asociado con una mayor tasa de metástasis linfática y una menor supervivencia global en diversos tipos de cáncer. En particular, la invasión angiolinfática (IAL) se ha relacionado con firmas de expresión génica específicas que reflejan una biología tumoral más agresiva, especialmente en el carcinoma escamoso de laringe (18). En el carcinoma escamoso, su evaluación sigue siendo un reto debido a la variabilidad en los criterios diagnósticos y la dificultad de detección en cortes histológicos de rutina (19).

4.2. Invasión angiolinfática en el cáncer oral

La invasión angiolinfática (IAL) es una característica histopatológica clave en la progresión tumoral y se define como la presencia de células tumorales dentro de espacios vasculares o linfáticos revestidos por endotelio (17,19). Este fenómeno patológico facilita la diseminación metastática a los ganglios linfáticos regionales y, en estadios avanzados, a órganos distantes, lo que contribuye a un peor pronóstico en los pacientes con carcinoma oral. Estudios recientes han identificado la IAL como un marcador pronóstico independiente de metástasis ganglionar y supervivencia global (20,21).

Desde el punto de vista biológico, la invasión angiolinfática implica la capacidad de las células tumorales para acceder a la circulación linfática y sanguínea mediante la degradación de la membrana basal y la disrupción del endotelio vascular. Este proceso es facilitado por proteínas clave como las metaloproteinasas de matriz (*MMPs*), que favorecen la degradación de la matriz extracelular, y por factores de crecimiento como VEGF-C y VEGF-D, que promueven la linfangiogénesis (19). Además, La remodelación temprana del colágeno en el microambiente tumoral también ha sido implicada en la progresión invasiva del HNSCC, facilitando la migración de células tumorales hacia vasos linfáticos y contribuyendo así a la aparición de IAL (22). Paralelamente, la infiltración de macrófagos tipo M2 en el estroma peritumoral ha mostrado una correlación significativa con la presencia de IAL, subrayando el papel de la inmunomodulación en la progresión tumoral (23).

La linfangiogénesis y la invasión angiolinfática están intrínsecamente relacionadas. Se ha demostrado que la proliferación de nuevos vasos linfáticos facilita la propagación de células tumorales hacia los ganglios linfáticos, aumentando el riesgo de metástasis regional (19). En particular, en el carcinoma de células escamosas de cabeza y cuello (HNSCC), se ha encontrado que los pacientes con presencia de IAL tienen una mayor incidencia de metástasis linfática y peor pronóstico (17,20).

Desde el punto de vista clínico, la presencia de IAL se ha asociado con una mayor agresividad tumoral y una menor tasa de supervivencia a cinco años (19). Por ello, su detección ha sido incluida en las guías de estadificación del cáncer como un factor pronóstico adicional que puede guiar la toma de decisiones terapéuticas (17). Los avances en el análisis genético han permitido identificar firmas moleculares asociadas a la invasión angiolinfática en HNSCC. En este contexto, CerS1

(*ceramide synthase 1*), también conocido como hLASS1 (*longevity assurance gene 1*), juega un papel clave en la biosíntesis de ceramida C18, un lípido bioactivo que regula procesos celulares como la apoptosis y la respuesta al estrés celular. Se ha observado que la disminución de los niveles de ceramida C18 está correlacionada con un mayor riesgo de invasión angiolinfática y metástasis linfática, lo que sugiere que CerS1 podría influir en la progresión tumoral y en la modulación del microambiente tumoral, lo que indica un papel relevante en la regulación de la diseminación tumoral (21).

Los avances en el análisis genético han permitido identificar firmas moleculares asociadas a la invasión angiolinfática en HNSCC. Un estudio reciente identificó módulos de coexpresión génica relacionados con la invasión linfovascular utilizando análisis de redes de coexpresión ponderadas (WGCNA), encontrando genes clave como *CNFN*, *KIF18B*, *KIF23*, *PRC1*, *CCNA2*, *DEPDC1* y *TTK*, que están implicados en la proliferación celular y la progresión tumoral (20).

Estudios transcriptómicos han demostrado que la activación de la vía *PI3K/AKT* está significativamente vinculada con el desarrollo y progresión del carcinoma escamoso de cabeza y cuello (HNSCC). Se ha identificado que aproximadamente 30.5% de los HNSCC presentan alteraciones en la vía *PI3K*, lo que contribuye a un aumento en la inestabilidad genómica y la reparación del daño en el DNA (*Deoxyribonucleic Acid*). En particular, las mutaciones activadoras en *PIK3CA* se han relacionado con una mayor sensibilidad a inhibidores específicos de *PI3K/mTOR* (24).

En particular, el uso de bases de datos como (TCGA *firehouse legacy*) ha permitido validar la expresión diferencial de estos genes en tumores con invasión angiolinfática, facilitando la identificación de biomarcadores pronósticos y el desarrollo de nuevas estrategias terapéuticas dirigidas (20). Estudios recientes han demostrado que la expresión diferencial de genes involucrados en la remodelación del microambiente tumoral y la adhesión celular, como *TP53*, VEGF-C, E-cadherina y MMP9, puede influir en la capacidad de diseminación tumoral (20,21).

Además, la aplicación de tecnologías de secuenciación masiva y estudios transcriptómicos han permitido detectar patrones de expresión génica específicos en tumores con IAL, lo que facilita la identificación de posibles biomarcadores para estratificar mejor a los pacientes y desarrollar estrategias terapéuticas más personalizadas (20).

Un análisis de expresión génica diferencial realizado en cohortes de HNSCC del TCGA identificó 88 genes significativamente asociados con la presencia de invasión angiolinfática, muchos de ellos implicados en rutas de adhesión celular, ciclo celular y regulación del citoesqueleto, lo que sugiere un papel clave de estos mecanismos en la diseminación tumoral (20).

Por ejemplo, se ha observado que la sobreexpresión de *CCNA2* y *DEPDC1* en tumores con IAL podría estar relacionada con una mayor proliferación celular y agresividad tumoral, dado su papel en la regulación del ciclo celular y la mitosis. *CCNA2* (ciclina A2) es un regulador clave de la transición G1/S y G2/M del ciclo celular, cuya sobreexpresión en HNSCC ha sido asociada con un mayor índice proliferativo y una peor diferenciación tumoral. Por otro lado, *DEPDC1* (*DEP domain containing 1*) ha sido identificado como un oncogén en diversos tipos de cáncer y se ha vinculado con la evasión de la apoptosis y una mayor invasión celular (20).

4.3. Transcriptómica

La genómica, por su parte, es una ciencia que se concentra en el estudio del genoma completo (ADN) de un organismo, mientras que disciplinas como la transcriptómica y la proteómica se encargan, respectivamente, del análisis del transcriptoma (ARN) y del proteoma (conjunto de proteínas), además del estudio de las interacciones entre estos componentes y en el entendimiento de vías metabólicas de los organismos, se dedica a documentar las secuencias génicas de los organismos. Es un área interdisciplinar que se ocupa de estudiar el genoma completo (25). La genómica puede dividirse en diferentes áreas, como la genómica estructural, que se enfoca en la organización física del material genético, la genómica funcional, que estudia la expresión y regulación de los genes, y la genómica comparativa, que analiza las similitudes y diferencias entre los genomas de distintas especies (26).

La transcriptómica, por otro lado, es una rama de la genómica funcional que estudia el conjunto de RNA (ácido ribonucleico) transcritos en un organismo o tejido en un momento determinado. Esto incluye el mRNA (ARN mensajero), rRNA (ARN ribosomal), tRNA (ARN de transferencia) y diferentes tipos de ARN no codificantes (25). El estudio del transcriptoma es clave para comprender cómo y cuándo se expresan los genes, permitiendo analizar la regulación génica en respuesta a diferentes estímulos biológicos o patológicos (26).

Además, el análisis del transcriptoma nos ayuda a entender procesos moleculares involucrados en ciertos procesos biológicos y procesos patológicos. Hoy en día hay muchos métodos de secuenciación avanzada que ayudan en el estudio del transcriptoma y uno de ellos es el RNA-seq (*RNA sequencing*) (26).

El *RNA-Seq* es una herramienta que forma parte de las tecnologías NGS (*Next Generation Sequencing*), o también llamadas *high-throughput* o *deep sequencing*, que revolucionaron el ámbito de la biología molecular. El RNA-Seq nos permite determinar niveles de expresión, el *splicing* diferencial y la expresión específica de alelos con una mayor precisión y a un coste más eficiente (27).

El objetivo de un análisis diferencial es averiguar diferencias en los transcriptomas de tejidos y células, además de la sobreexpresión e infraexpresión de los genes de ellos. La sobreexpresión se refiere a cuando un gen tiene un nivel de expresión mayor al normal, lo que puede estar relacionado con procesos como proliferación celular descontrolada en el cáncer. La infraexpresión, por otro lado, ocurre cuando un gen se expresa en menor cantidad de lo esperado, lo que puede afectar funciones celulares esenciales, como la respuesta inmune o la reparación del ADN (28).

En el contexto de la invasión angiolinfática (IAL), el RNA-seq permite caracterizar la heterogeneidad celular e identificar subpoblaciones de células tumorales con mayor potencial invasivo (27). Además, ha revelado nuevos biomarcadores específicos de células tumorales con mayor potencial metastásico. Esta técnica ofrece una visión detallada del paisaje transcriptómico del tumor y permite identificar interacciones entre células del microambiente tumoral que favorecen la invasión linfovascular (26).

El análisis transcriptómico ha demostrado ser una herramienta valiosa para identificar firmas génicas asociadas a la invasión linfovascular en HNSCC, permitiendo una estratificación más precisa del riesgo de metástasis y facilitando la selección de tratamientos personalizados basados en el perfil molecular del tumor (20). Se han revelado perfiles de expresión génica específicos en tumores con IAL, destacando genes implicados en la angiogénesis, proliferación celular y remodelación de la matriz extracelular (27).

Estas herramientas han permitido detectar con alta sensibilidad patrones moleculares vinculados con riesgo de metástasis temprana y mal pronóstico, posicionando al RNA-seq como una herramienta clave en la caracterización molecular del HNSCC con invasión angiolinfática (20).

Los datos obtenidos mediante *RNA-seq* se procesan habitualmente con herramientas bioinformáticas que permiten identificar genes diferencialmente expresados entre condiciones biológicas, como la presencia o ausencia de invasión angiolinfática, mediante modelos estadísticos adecuados para datos de conteo (29).

4.4. Análisis de enriquecimiento

El análisis de enriquecimiento sirve para interpretar datos de expresión génica, permitiendo la identificación de categorías funcionales sobrerrepresentadas en un conjunto de genes en comparación con lo esperado por azar (30). Esta técnica es ampliamente utilizada en estudios transcriptómicos para asociar cambios en la expresión génica con procesos biológicos relevantes. Este método intenta averiguar si un conjunto de genes está asociado a una característica (31). Para ello, calcula un *enrichment score* (ES), el cual mide la sobre-representación de genes de un conjunto específico en una lista ordenada de genes expresados, asignando un puntaje según su relevancia biológica (30).

Para realizar un análisis de enriquecimiento, podemos utilizar diversas herramientas y programas bioinformáticos que faciliten la interpretación de datos como la expresión génica, las asociaciones genómicas u otros datos biológicos. Una de estas herramientas es DAVID (*Database for Annotation, Visualization, and Integrated Discovery*). El propósito del análisis de enriquecimiento es identificar categorías de términos funcionales (por ejemplo, vías metabólicas, funciones biológicas o procesos celulares) que están sobrerrepresentadas en un conjunto de genes de interés, en comparación con lo que se esperaría aleatoriamente. Estos análisis proporcionan información sobre funciones biológicas, rutas metabólicas o términos de ontología génica (*Gene Ontology, GO*) relevantes para un conjunto específico de genes.

DAVID es una de las herramientas más utilizadas para el análisis de enriquecimiento funcional y la anotación de listas de genes. Proporciona un conjunto de herramientas para analizar grandes

volúmenes de datos, permitiendo a los investigadores identificar asociaciones significativas entre conjuntos de genes y procesos biológicos específicos. Entre sus funcionalidades clave se incluyen:

- **Análisis de enriquecimiento funcional:** Permite evaluar la sobrerrepresentación de genes en distintas categorías biológicas como KEGG (*Kyoto Encyclopedia of Genes and Genomes*), *Reactome* y *Gene Ontology* (GO).
- **Clasificación funcional:** Agrupamiento de genes en módulos funcionales según similitud de anotaciones, facilitando la identificación de patrones en los datos.
- **Conversión de identificadores de genes:** Facilita la integración de datos de distintas plataformas a través de la normalización de identificadores de genes y proteínas.

Además, DAVID ha sido recientemente actualizado para mejorar la cobertura taxonómica y la integración de nuevas bases de datos, como *DisGeNET* para asociaciones gen-enfermedad y *WikiPathways* para análisis de rutas metabólicas. La continua evolución de esta herramienta permite una mejor interpretación de datos ómicos y su aplicación en estudios de biología molecular y genética (32).

A pesar de su utilidad, el análisis de enriquecimiento presenta limitaciones relacionadas con la redundancia funcional entre categorías, el tamaño de los conjuntos de genes y la dependencia del conjunto de referencia (*background*) utilizado. Estas limitaciones pueden afectar la interpretación de los resultados y aumentar la probabilidad de obtener falsos positivos. Por ello, se recomienda validar los resultados mediante herramientas complementarias y una revisión manual cuidadosa de los términos de ontología génica (GO) relevantes (33).

5. JUSTIFICACIÓN

El carcinoma de células escamosas de cabeza y cuello representa uno de los cánceres más comunes a nivel mundial, con una tasa de supervivencia a los 5 años de aproximadamente 50% y una incidencia creciente (3). Dentro de los factores que influyen en su progresión y pronóstico, la invasión angiolinfática (IAL) juega un papel fundamental, ya que se ha identificado como un marcador pronóstico de metástasis ganglionar y peor supervivencia en los pacientes con HNSCC (20).

El diagnóstico y tratamiento temprano de la IAL en el carcinoma de células escamosas de cabeza y cuello son esenciales para mejorar la supervivencia. Por ello, es fundamental la implementación de estrategias basadas en el análisis transcriptómico que permitan una mejor identificación de los mecanismos moleculares asociados a la IAL.

En este contexto, el estudio de la expresión génica en tumores con invasión angiolinfática representa una oportunidad para desarrollar estrategias diagnósticas y terapéuticas más efectivas. Un metaanálisis centrado en la expresión génica de estos tumores permitirá identificar patrones moleculares específicos que podrían diferenciar tumores con alto potencial metastásico de aquellos con menor agresividad, facilitando el desarrollo de biomarcadores pronósticos y terapias dirigidas.

Este trabajo de fin de grado se enmarca en el Objetivo de Desarrollo Sostenible (ODS) número 3, “Salud y Bienestar”. Al profundizar en la investigación del carcinoma oral y su asociación con la invasión angiolinfática, se busca contribuir al objetivo global de promover una vida saludable y bienestar para todos. Los resultados obtenidos podrían tener aplicaciones clínicas concretas, mejorando la detección temprana y el desarrollo de terapias personalizadas, alineándose así con la promoción de la salud y la prevención de enfermedades.

6. HIPÓTESIS

Se espera que existan diferencias significativas en la expresión génica entre muestras con invasión angiolinfática en comparación con muestras sin invasión angiolinfática en el carcinoma de células escamosas de cabeza y cuello. Estas diferencias en la expresión génica podrían estar asociadas a vías moleculares específicas que favorecen la diseminación linfovascular y la progresión tumoral, permitiendo la identificación de biomarcadores pronósticos y potenciales dianas terapéuticas.

7. OBJETIVOS

7.1. General

Identificar y caracterizar los genes diferencialmente expresados asociados a la invasión angiolímfática en el carcinoma de células escamosas de cabeza y cuello, con el fin de mejorar la comprensión de los mecanismos moleculares subyacentes y proponer posibles biomarcadores pronósticos y terapéuticos.

7.2. Específicos

1. Detectar genes diferencialmente expresados en muestras de carcinoma de células escamosas de cabeza y cuello con y sin invasión angiolímfática.
2. Analizar las funciones biológicas y las rutas moleculares asociadas a los genes identificados.
3. Identificar biomarcadores potenciales para la detección precoz de la invasión angiolímfática.

8. MATERIALES Y MÉTODOS

Realizamos un análisis bioinformático de expresión diferencial en muestras de pacientes con carcinoma de células escamosas de cabeza y cuello, utilizando datos de la base de datos TCGA (*The Cancer Genome Atlas Project*), con el propósito de identificar genes diferencialmente expresados en pacientes con o sin invasión angiolinfática.

8.1. Diseño del estudio

8.1.1. Selección de los datos

Para llevar a cabo este estudio se emplearon datos extraídos de la base de datos TCGA (*The Cancer Genome Atlas Project*, <http://tcga.cancer.gov>), un portal que facilita tanto el análisis como la descarga de una amplia gama de datos óhmicos relacionados con el cáncer. Este recurso, desarrollado por el NCI (*National Cancer Institute*) y el Instituto Nacional de Investigación del Genoma, ha acumulado más de 2.5 petabits de información genómica, epigenómica, transcriptómica y proteómica, contribuyendo significativamente a mejorar el diagnóstico, tratamiento y prevención del cáncer.

Dentro de la plataforma TCGA se filtraron las muestras según el tipo de cáncer de interés para nuestro estudio: cáncer de cabeza y cuello. Así, se obtuvieron 528 muestra de carcinoma de células escamosas de cabeza y cuello, luego se procedió utilizando un filtro que contuvieran datos de invasión y que no tuvieran invasión por ello se redujo a al final a 358 muestra, correspondientes al estudio *Firehose Legacy*.

Una vez seleccionadas las muestras, se llevó a cabo una observación previa de estas antes de analizarlas, mediante un Análisis de Componentes Principales (ACP) con el *Software R* (34) sobre las 358 muestras, de las cuales 126 presentaban invasión angiolinfática (Grupo 2) y 232 sin invasión (Grupo 1) con el objetivo de identificar posibles valores atípicos, así como su distribución clara entre ambos grupos (Figura 1).

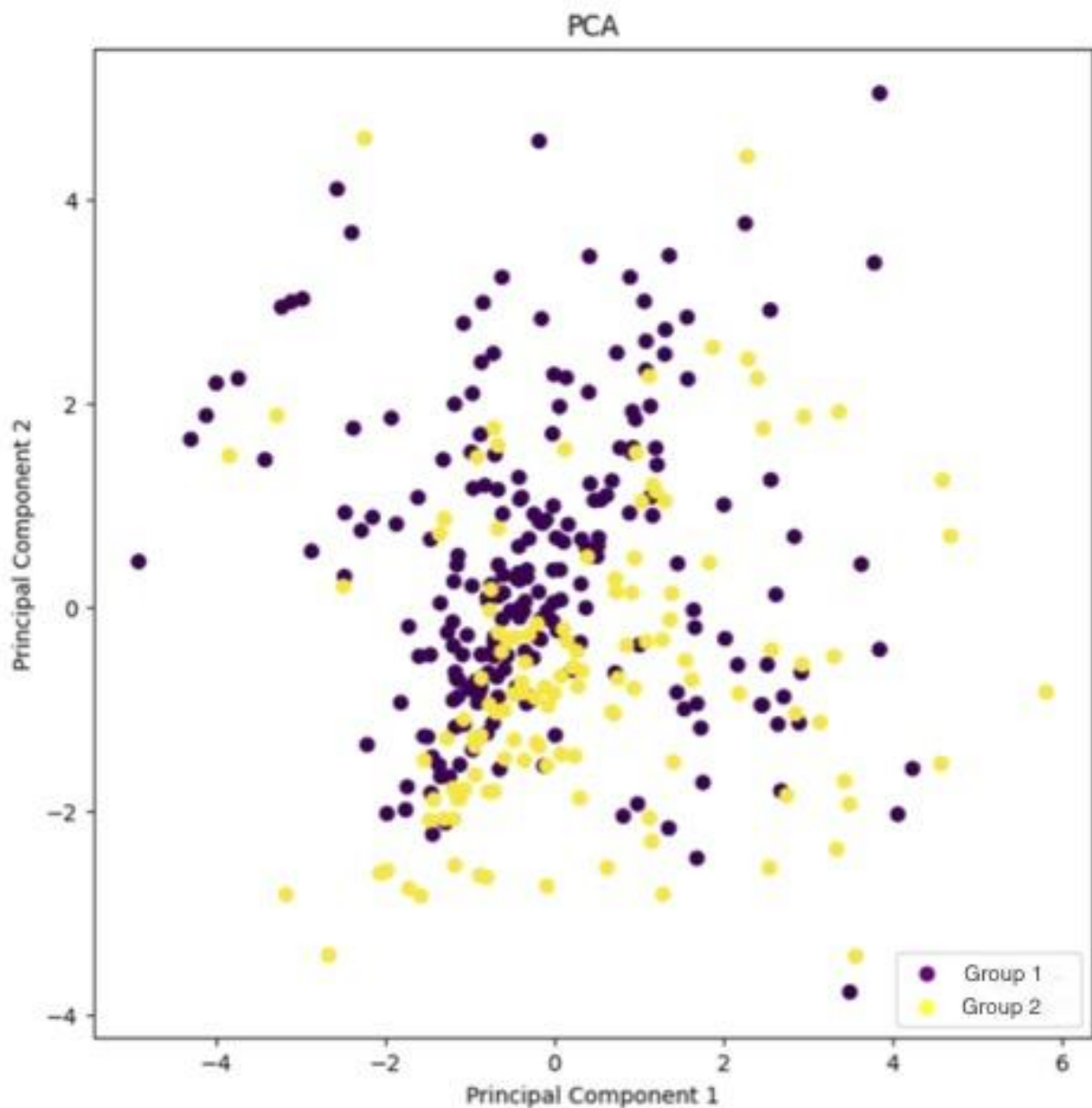


Figura 1. Análisis de Componentes Principales de 358 muestras, libre de la invasión (232 muestras: *Group 1*) e invasión angiolinfática (126 muestras: *Group 2*).

A través del portal cBioPortal (https://www.cbioportal.org/study/summary?id=hnscc_tcga), se aplicaron filtros a 358 muestras de pacientes con carcinoma escamoso de cabeza y cuello (HNSCC), obteniendo un total de 126 muestras correspondientes a 124 pacientes (34.8%) con invasión angiolinfática y 232 muestras/pacientes sin invasión (65.2%).

En cuanto a la distribución por sexo, 258 pacientes (72.5%) eran hombres y 98 (27.5%) mujeres, de los cuales 99 hombres (79.8%) y 25 mujeres (20.2%) presentaban invasión angiolinfática. La edad al momento del diagnóstico se clasificó en dos grupos: 193 muestras correspondían a pacientes menores de 61 años y 162 a mayores de esta edad.

Además, se evaluó la presencia de hábitos como el consumo de alcohol y tabaco. Respecto al consumo de alcohol, 241 muestras (67.7%) pertenecían a consumidores, de los cuales 90 (72.6%) presentaban invasión angiolinfática, 109 (30.6%) a personas que no presentaban este hábito de los cuales 33 (26.6%) presentaban invasión angiolinfática y 6 (1.7%) no contaban con datos disponibles. En relación con el consumo de tabaco, 119 pacientes (33.4%) eran fumadores activos, de los cuales 49 (39.5%) presentaban invasión angiolinfática, 84 (23.6%) nunca habían fumado de los cuales 23 (18.5%) presentaban invasión angiolinfática, 145 (40.8%) habían dejado el hábito, de los cuales 49 (39.8%) presentaban invasión angiolinfática y 8 (2.3%) no tenían información registrada.

8.1.2. Análisis de expresión diferencial

Para el análisis de expresión diferencial, se compararon dos grupos de muestras de pacientes con cáncer oral: 126 muestras correspondientes a pacientes con invasión angiolinfática frente a 232 muestras de pacientes sin invasión angiolinfática. El análisis se llevó a cabo estableciendo una comparación directa entre ambos grupos, con el objetivo de identificar genes diferencialmente expresados asociados a la presencia o ausencia de invasión angiolinfática.

Para realizar este análisis se utilizaron los *softwares* EdgeR y Limma (*Bioconductor*), donde se estimaron lecturas de TPMs normalizadas, se identificaron de manera precisa los genes diferencialmente expresados (*DEGs*) entre ambos grupos, a través de métodos estadísticos (i.e., *lineal modeled and empirical Bayes*) según la metodología de Law y cols (35). Se consideraron como *DEGs* aquellos genes que mostraron un *false discovery rate* (FDR) ≤ 0.05 como expresados diferencialmente. lo que aseguró la relevancia estadística de los hallazgos y aportó información valiosa sobre los posibles mecanismos moleculares asociados a la invasión angiolinfática en el cáncer oral.

8.1.3. Análisis de enriquecimiento

Para el estudio de procesos biológicos de los genes mediante análisis de enriquecimiento se utilizó el programa DAVID. Este análisis permitió identificar las rutas biológicas y funciones moleculares más representativas asociados a los genes diferencialmente expresados (*DEGs*) de interés (30,32). Aquellos procesos biológicos con un Pvalor ≤ 0.05 e consideraron cómo estadísticamente significativos en el DAVID.

8.1.4. Análisis estadístico de las variables clínicas

Para analizar la relación entre variables clínicas y la presencia de invasión angiolinfática en pacientes con carcinoma escamoso de cabeza y cuello, se empleó el *software* Jamovi (36), una herramienta estadística de libre acceso basada en R.

Las variables categóricas, como el historial de consumo de tabaco, alcohol y el sexo, se evaluaron mediante la construcción de tablas de contingencia. Para contrastar la independencia entre estas variables y la invasión angiolinfática, se utilizó la prueba de Chi-cuadrado (χ^2). En aquellos casos en los que las frecuencias esperadas en alguna celda fueron inferiores a 5, se aplicó también el *test* exacto de *Fisher*, que ofrece mayor precisión en muestras pequeñas o distribuciones asimétricas.

Para la variable edad, al tratarse de una variable continua, se evaluó inicialmente la distribución de los datos mediante la prueba de normalidad de *Shapiro–Wilk*. Si los datos cumplían con el criterio de normalidad, se aplicó la prueba t de *Student* para muestras independientes. En caso contrario, se optó por la prueba no paramétrica de *Mann–Whitney U*, que no requiere el supuesto de normalidad. En todos los análisis, se consideró un valor de $p < 0.05$ como umbral de significación estadística.

9. RESULTADOS

9.1. Análisis de expresión diferencial

En este estudio se identificaron. 375 genes diferencialmente expresados. en muestras de tejido oral de pacientes con HNSCC e invasión angiolinfática, en comparación con aquellos sin esta característica, utilizando un criterio de FDR inferior a 0.05. De estos genes, 274 presentaron sobreexpresión (*Fold Change* positivo), mientras que 101 mostraron infraexpresión (*Fold Change* negativo) en los casos con invasión angiolinfática.

En la Tabla 1, se presentan los 20 genes con mayor sobreexpresión, mientras que en la Tabla 2. se listan los 20 genes con mayor infraexpresión en los tumores con invasión angiolinfática. Además, en el anexo 1 se encuentra la lista completa de los 274 genes sobreexpresados y en el anexo 2 la lista completa de los 101 genes infraexpresados, junto con sus respectivos valores estadísticos.

Tabla 1. Los 20 genes diferenciales con FDR <0.05 con mayor sobreexpresión (*Fold Change* positivo)

GENES	Nombre completo	fold change	Pvalue	FDR
FAM174B	family with sequence similarity 174 member B	0.462	5.1×10^{-8}	0.002
ARNILA	androgen receptor negatively regulated lncRNA	0.550	1.4×10^{-7}	0.002
NPDC1	neural proliferation, differentiation and control 1	0.490	2.0×10^{-7}	0.002
RHPN1	rhophilin Rho GTPase binding protein 1	0.673	3.7×10^{-7}	0.003
DNASE2	deoxyribonuclease 2, lysosomal	0.268	6.5×10^{-7}	0.005
COX14	cytochrome c oxidase assembly factor COX14	0.264	8.1×10^{-7}	0.005
FAM50A	family with sequence similarity 50 member A	0.316	1.0×10^{-6}	0.006
PKN1	protein kinase N1	0.412	1.6×10^{-6}	0.007
P4HTM	prolyl 4-hydroxylase, transmembrane	0.449	2.1×10^{-6}	0.007

PCBD1	pterin-4 alpha-carbinolamine dehydratase 1	0.284	2.7×10^{-6}	0.008
PRR15	proline rich 15	0.762	3.5×10^{-6}	0.009
PRKAG2-AS1	PRKAG2 antisense RNA 1	0.626	3.8×10^{-6}	0.009
CATSPERG	cation channel sperm associated auxiliary subunit gamma	0.393	3.7×10^{-6}	0.009
CFAP43	cilia and flagella associated protein 43	0.523	5.9×10^{-6}	0.009
KRT18	keratin 18	0.481	6.1×10^{-6}	0.009
GADD45G	growth arrest and DNA damage inducible gamma	0.473	6.6×10^{-6}	0.009
CCNO	cyclin O	0.467	6.1×10^{-6}	0.009
TSC22D3	TSC22 domain family member 3	0.353	7.4×10^{-6}	0.009
SIGIRR	single Ig and TIR domain containing	0.322	7.3×10^{-6}	0.009
ARRDC2	arrestin domain containing 2	0.288	6.7×10^{-6}	0.009

*Fold Change se refiere al grupo con invasión en comparación con sin invasión angioblástica.

Tabla 2. Los 20 genes diferenciales con FDR <0.05 con menor infraexpresión (*Fold Change* negativo)

GENES	Nombre completo	fold change	Pvalue	FDR
SPRR2G	small proline rich protein 2G	-1.312	1.3×10^{-6}	0.006
EPGN	epithelial mitogen	-1.050	2.1×10^{-6}	0.007
WFDC12	WAP four-disulfide core domain 12	-1.154	2.7×10^{-6}	0.008
LCE3E	late cornified envelope 3E	-1.149	3.6×10^{-6}	0.009
SPRR5	small proline rich protein 5	-0.840	4.0×10^{-6}	0.009
SLEAR	STAT1 regulated ILF2 complex interacting lncRNA	-0.634	6.3×10^{-6}	0.009
GJB2	gap junction protein beta 2	-0.647	7.3×10^{-6}	0.009
SPRR2F	small proline rich protein 2F	-1.214	6.0×10^{-6}	0.009
DSG1	desmoglein 1	-1.198	7.6×10^{-6}	0.009
KRT17P1	keratin 17 pseudogene 1	-0.472	9.7×10^{-6}	0.011
GSDMA	gasdermin A	-0.816	1.1×10^{-5}	0.012
CRCT1	cysteine rich C-terminal 1	-1.030	1.6×10^{-5}	0.015
PRR9	proline rich 9	-1.095	1.8×10^{-5}	0.016

GNAI1	G protein subunit alpha i1	-0.299	2.8×10^{-5}	0.019
TUBB2A	tubulin beta 2A class IIa	-0.370	3.0×10^{-5}	0.020
LINC01740	long intergenic non-protein coding RNA 1740	-0.313	3.7×10^{-5}	0.020
ELL2P1	elongation factor for RNA polymerase II 2 pseudogene 1	-0.410	3.3×10^{-5}	0.020
PNPLA1	patatin like phospholipase domain containing 1	-0.735	3.7×10^{-5}	0.020
S100A12	S100 calcium binding protein A12	-0.793	3.5×10^{-5}	0.020
UBQLN1P1	ubiquilin 1 pseudogene 1	-0.294	4.2×10^{-5}	0.022

*Fold Change se refiere al grupo con invasión en comparación con sin invasión angirolinfática.

9.1. Análisis de enriquecimiento con DAVID software

Se llevó a cabo un análisis de enriquecimiento funcional utilizando el *software* DAVID, incorporando los 274 genes sobreexpresados y los 101 genes infraexpresados con un FDR inferior a 0.05.

A partir de este análisis, se identificaron un total de 10 procesos biológicos significativamente enriquecidos asociados a los genes sobreexpresados, mientras que 9 procesos biológicos mostraron una relación con los genes infraexpresados (p-valor < 0.05).

A continuación, encontramos las tablas completas de los procesos biológicos. En la Tabla 3, se presentan los procesos biológicos que se ven impactados por los genes sobreexpresados, proporcionando una visión global de las funciones celulares más afectadas en los casos de HNSCC con invasión angirolinfática, y en la Tabla 4, los procesos biológicos relacionados con los genes infraexpresados.

Tabla 3. Procesos biológicos significativamente enriquecidos (p-valor < 0.05) entre los genes sobreexpresados, según el análisis DAVID.

Term	Proceso biológico	Count	p-valor	Genes
GO:0017157	regulation of exocytosis	4	0.004	RABL2A, RABL2B, HAP1, VAMP2
GO:0003341	cilium movement	4	0.011	DNAAF1, ZMYND10, CFAP43, PIERCE1

GO:0008637	apoptotic mitochondrial changes	3	0.013	AIFM2, JTB, SLC25A4
GO:0042733	embryonic digit morphogenesis	4	0.020	FBXW4, SHH, FLVCR1, TMEM231
GO:0006783	heme biosynthetic process	3	0.025	SLC25A38, FLVCR1, PPOX
GO:0006406	mRNA export from nucleus	4	0.027	CHTOP, HHEX, DDX39A, RBM8A
GO:0033617	mitochondrial cytochrome c oxidase assembly	3	0.034	UQCC6, COX14, COA5
GO:1903008	organelle disassembly	2	0.041	KIF9, PLAAT3
GO:0008089	anterograde axonal transport	3	0.041	DTNBP1, KIF1A, HAP1
GO:0035082	axoneme assembly	3	0.046	RSPH1, DNAAF1, PIERCE1

* Los procesos biológicos resaltados en verde fueron seleccionados para su análisis detallado y la discusión en el presente estudio.

Tabla 4. Los Procesos biológicos significativamente enriquecidos (p-valor < 0.05) entre los genes infraexpresados, según el análisis DAVID.

Term	Proceso biológico	Court	p-valor	Genes
GO:0009395	phospholipid catabolic process	3	0.001	PLA2G4E, LIPG, PLD2
GO:0019731	antibacterial humoral response	4	0.002	WFDC12, PI3, SPRR2A, WFDC5
GO:0002437	inflammatory response to antigenic stimulus	3	0.004	IL36B, IL36RN, IL1F10
GO:0008366	axon ensheathment	2	0.021	POU3F1, PPARD
GO:0036149	phosphatidylinositol acyl-chain remodeling	2	0.021	PLA2G4D, PLA2G4E
GO:0042742	defense response to bacterium	4	0.022	WFDC12, GSDMA, S100A12, DEFB4A
GO:0071222	cellular response to lipopolysaccharide	4	0.036	IL36B, IL36RN, PPARD, IL1F10
GO:0098609	cell-cell adhesion	4	0.037	TENM2, FBLIM1, DSG1, DSC2
GO:0031640	killing of cells of another organism	3	0.045	LCE3A, S100A12, DEFB4A

* Los procesos biológicos resaltados en verde fueron seleccionados para su análisis detallado y la discusión en el presente estudio.

9.2. Análisis de variables clínicas en la invasión angiolinfática en HNSCC

9.2.1. Tabaco

Se evaluó la relación entre el historial de tabaquismo y la presencia de invasión angiolinfática en pacientes diagnosticados con carcinoma escamoso de cabeza y cuello (HNSCC). Para ello, se analizaron un total de 351 pacientes. Se construyeron gráficos para comparar la distribución de la invasión angiolinfática entre los distintos grupos de tabaquismo.

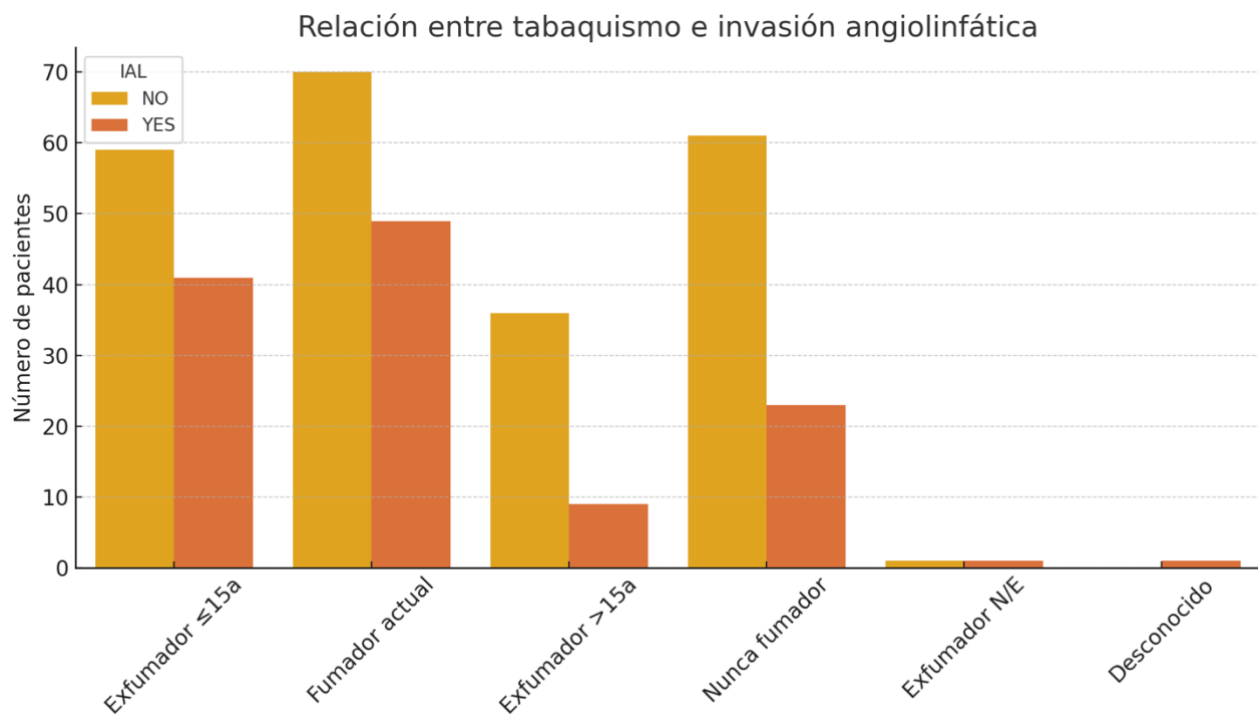


Figura 1. Distribución del número de pacientes con (YES) y sin (NO) invasión angiolinfática en función del historial de tabaquismo.

En la figura se representa la relación entre el historial de tabaquismo y la presencia de invasión angiolinfática en pacientes con carcinoma escamoso de cabeza y cuello (HNSCC). El gráfico compara el número de pacientes con (YES) y sin (NO) invasión angiolinfática en cada una de las categorías de tabaquismo: fumador actual, exfumador ≤15 años, exfumador >15 años, nunca fumador, exfumador N/E (no especificado) y desconocido.

Se observa que en los grupos de fumadores actuales y exfumadores con abstinencia ≤15 años, la proporción de pacientes con invasión angiolinfática (barras naranjas) es notablemente mayor en comparación con los no fumadores o aquellos con abstinencia prolongada. En particular, los nunca fumadores presentan una proporción más baja de invasión, lo que sugiere una posible relación entre el tabaquismo activo o reciente y una mayor frecuencia de invasión angiolinfática.

Este análisis gráfico tiene un carácter exploratorio y permite visualizar patrones iniciales, que posteriormente se pondrán a prueba mediante métodos estadísticos.

Para determinar la asociación entre el tabaquismo y la invasión angiolinfática, se aplicaron pruebas estadísticas de contraste de hipótesis.

A continuación, se aplicó la prueba de Chi-cuadrado (χ^2) para determinar si existía una asociación estadísticamente significativa entre el historial de tabaquismo y la presencia de invasión angiolinfática. El resultado fue $\chi^2 = 12.2$ con 5 grados de libertad y un valor $p = 0.033$, lo cual indica que existe una asociación significativa entre ambas variables ($p < 0.05$).

Dado que algunas celdas presentaban frecuencias esperadas bajas, se utilizó también el *test* exacto de *Fisher* como prueba complementaria, el cual mostró un valor $p = 0.015$. Este resultado refuerza la evidencia estadística de asociación entre el tabaquismo y la invasión angiolinfática.

Tabla 5. Tabla de la prueba de chi-cuadrado complementada con el Test exacto de Fischer.

Test χ^2

	Valor	gdl	p
χ^2	12.2	5	0.033
Test exacto de Fisher.			0.015
N	351		

En conjunto, estos resultados permiten concluir que existe una relación significativa entre el historial de tabaquismo y la presencia de invasión angiolinfática en pacientes con HNSCC. Específicamente, los pacientes con tabaquismo activo o con abstinencia reciente muestran una mayor frecuencia de invasión en comparación con los nunca fumadores o aquellos con abstinencia prolongada.

9.2.2. Alcohol

Se evaluó la asociación entre el consumo de alcohol y la presencia de invasión angiolinfática en pacientes con HNSCC. Para ello, se analizaron un total de 352 pacientes, divididos en dos grupos

según su historial de consumo de alcohol: consumidores y no consumidores. Se construyeron gráficos para comparar la distribución de la invasión angiolinfática entre ambos grupos.

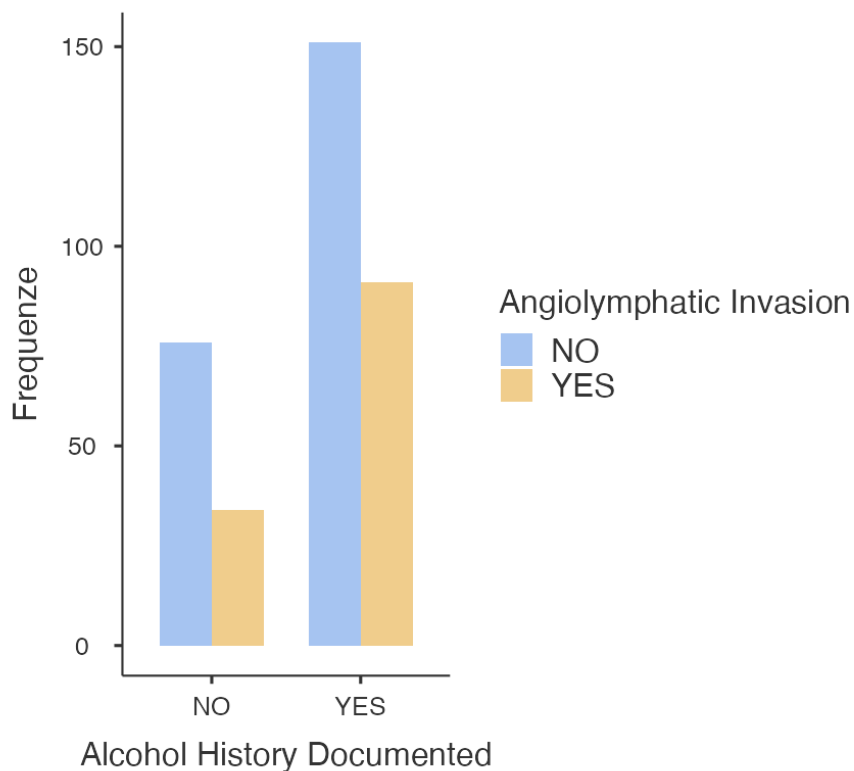


Figura 2. Distribución del número de pacientes con (YES) y sin (NO) invasión angiolinfática en función de si presentan o no historial documentado de consumo de alcohol.

En la figura 2 se observa que, entre los pacientes sin historial de consumo de alcohol, el número de casos sin invasión (barras azules) es considerablemente mayor que los casos con invasión (barras naranjas). Por otro lado, en los pacientes con historial de consumo de alcohol, también hay un mayor número de pacientes sin invasión, pero la proporción de pacientes con invasión angiolinfática es notablemente más alta en comparación con el grupo que no consume alcohol.

Este patrón sugiere una posible asociación entre el consumo de alcohol y una mayor frecuencia de invasión angiolinfática, que será evaluada mediante pruebas estadísticas específicas. A continuación, se aplicó la prueba de Chi-cuadrado (χ^2) para evaluar la asociación entre el consumo de alcohol y la presencia de invasión angiolinfática.

Tabla 6. Tabla de la prueba de chi-cuadrado complementada con el Test exacto de Fischer.

Test χ^2

	Valor	gdl	p
χ^2	1.48	1	0.224
Test exacto de Fisher.			0.232
N	352		

El resultado fue $\chi^2 = 1.48$ con 1 grado de libertad y un valor $p = 0.224$, lo cual indica que no existe una asociación estadísticamente significativa entre ambas variables. Este resultado fue complementado con el *test* exacto de *Fisher*, que obtuvo un valor $p = 0.232$, confirmando la ausencia de significación estadística.

Asimismo, se calculó la medida de asociación *Odds Ratio* (OR), obteniéndose un valor de OR = 1.35 con un intervalo de confianza del 95% entre 0.833 y 2.18.

Tabla 7. Tabla de mediadas comparativas utilizando el Odds-Ratio.

Medidas comparativas

	Valor	95% Intervallo de confianza	
		Inferiore	Superior
Odds-Ratio	1.35	0.833	2.18

Aunque este valor sugiere un aumento en las probabilidades de invasión angiolinfática entre los pacientes con historial de consumo de alcohol, el intervalo de confianza incluye el valor 1, lo que indica que la diferencia no es estadísticamente significativa.

En conjunto, estos resultados muestran que, si bien se observa una tendencia visual hacia una mayor frecuencia de invasión en los consumidores de alcohol, no se encontraron diferencias estadísticamente significativas en esta cohorte.

9.2.3. Edad

Se evaluó si existía una diferencia en la edad de diagnóstico entre pacientes con y sin invasión angiolinfática en el carcinoma escamoso de cabeza y cuello (HNSCC). Para ello, se realizaron pruebas estadísticas para determinar la distribución de los datos y la validez de las comparaciones entre los grupos.

Para analizar la relación entre la edad al diagnóstico y la presencia de invasión angiolinfática, se presentan en la Figura 3 dos representaciones gráficas complementarias. La Figura 3 muestra un histograma de densidad estratificado por grupo, comparando la distribución de la edad entre los pacientes con y sin invasión angiolinfática. La Figura 3B representa un gráfico de violín (*violin plot*), que permite visualizar la distribución y dispersión de la edad en ambos grupos de forma comparativa.

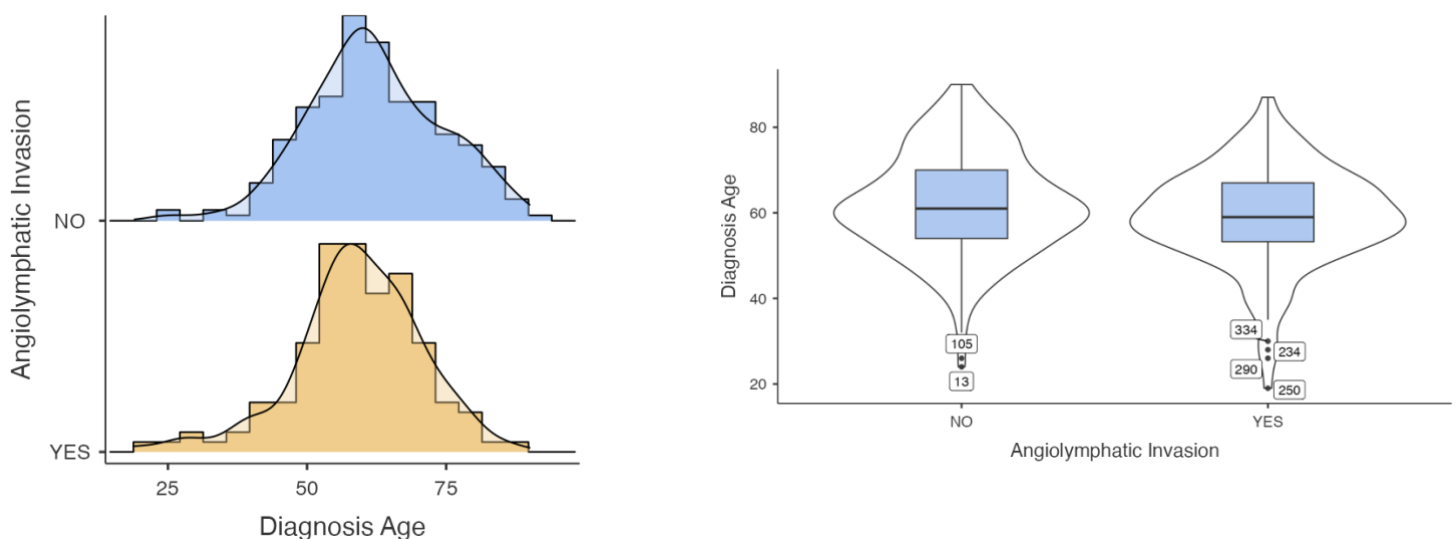


Figura 3.

A. Histograma de densidad estratificado por grupo (NO = sin invasión angiolinfática; YES = con invasión angiolinfática), que permite observar la forma y distribución de la variable edad en cada grupo, resaltando posibles diferencias en tendencia central y dispersión.

B. Gráfico de violín que representa la distribución de la edad en ambos grupos, incluyendo la mediana, los intervalos intercuartílicos y los valores atípicos. Se observan diferencias en la densidad de edad entre los grupos con y sin invasión angiolinfática

El histograma de densidad muestra la distribución de edades en pacientes con (grupo YES) y sin (grupo NO) invasión angiolinfática. Se observa que los pacientes sin invasión presentan una

distribución más simétrica, centrada alrededor de los 65 años. En cambio, los pacientes con invasión muestran una mayor concentración de casos en edades ligeramente inferiores, con un pico alrededor de los 55–60 años. Esta diferencia sugiere una posible tendencia hacia una edad más joven en el grupo con invasión.

El gráfico de violín, por su parte, combina la información del *boxplot* con la densidad de probabilidad. Se observa que la mediana de edad es ligeramente inferior en el grupo con invasión, y que existe una mayor concentración de pacientes en el rango de 50–60 años. Además, se identifican varios valores atípicos en ambos grupos, especialmente en edades jóvenes, lo que indica cierta heterogeneidad en la distribución. La forma del violín refuerza visualmente la idea de que los pacientes con invasión podrían ser más jóvenes en comparación con los que no presentan invasión.

Estas diferencias observadas gráficamente se analizaron estadísticamente mediante la prueba t de *Student* y la prueba no paramétrica U de *Mann-Whitney*. Inicialmente, se aplicó la prueba t de *Student* para muestras independientes para comparar la edad media entre los grupos. Dado que esta prueba asume normalidad en la distribución de los datos, se realizó una prueba de normalidad de *Shapiro-Wilk* para verificar este supuesto. Adicionalmente, debido a que los datos no seguían una distribución normal, se complementó el análisis con la prueba no paramétrica U de *Mann-Whitney*, la cual no requiere el supuesto de normalidad.

Tabla 8. Tabla con Estadísticas descriptivas del grupo según la variable de edad. (SD=Desviación Estándar. SE=Error estándar)

Estadísticas descriptivas del grupo						
	Grupo	N	Media	Mediana	SD	SE
Edad del diagnóstico	NO	231	62.1	61.0	12.0	0.791
	YES	126	59.1	59.0	11.6	1.03

Tabla 9. Tabla del *Test* de Normalidad (*Shapiro-Wilk*)

Test de Normalidad (Shapiro-Wilk)

	W	p
Edad del diagnóstico	0.989	0.011

Tabla 10. Tabla del *Test* t para muestras independientes complementada con el *Test* U de *Mann-Whitney*.

Test t para muestras independientes

		Estadística	gdl	p
Edad del diagnóstico	t de Student	2.28	355	0.023
	U de Mann-Whitney	12794		0.059

* $H_a \mu_{NO} \neq \mu_{YES}$

La prueba de normalidad de *Shapiro-Wilk* mostró un valor $p = 0.011$, Un valor de $p < 0.05$ indica que los datos no siguen una distribución normal, lo que evidencia una violación del supuesto de normalidad en la variable edad. Por esta razón, además de la prueba t de *Student* para muestras independientes, se aplicó también el *test* no paramétrico de *Mann-Whitney*, más adecuado en este contexto. Aunque la prueba t sugiere una diferencia significativa en la edad al diagnóstico entre los pacientes con y sin invasión angiolinfática ($p = 0.023$), esta diferencia no se confirma mediante el *test* U de *Mann-Whitney* ($p = 0.059$). Por tanto, no puede afirmarse con certeza estadística que exista una diferencia significativa en la edad entre ambos grupos, aunque se observa una tendencia visual y numérica hacia una edad ligeramente menor en los pacientes con invasión.

9.2.4. Sexo

Se evaluó la relación entre el sexo y la presencia de invasión angiolinfática en pacientes con HNSCC. Para ello, se analizaron un total de 358 pacientes, de los cuales 99 eran mujeres y 259 eran hombres.

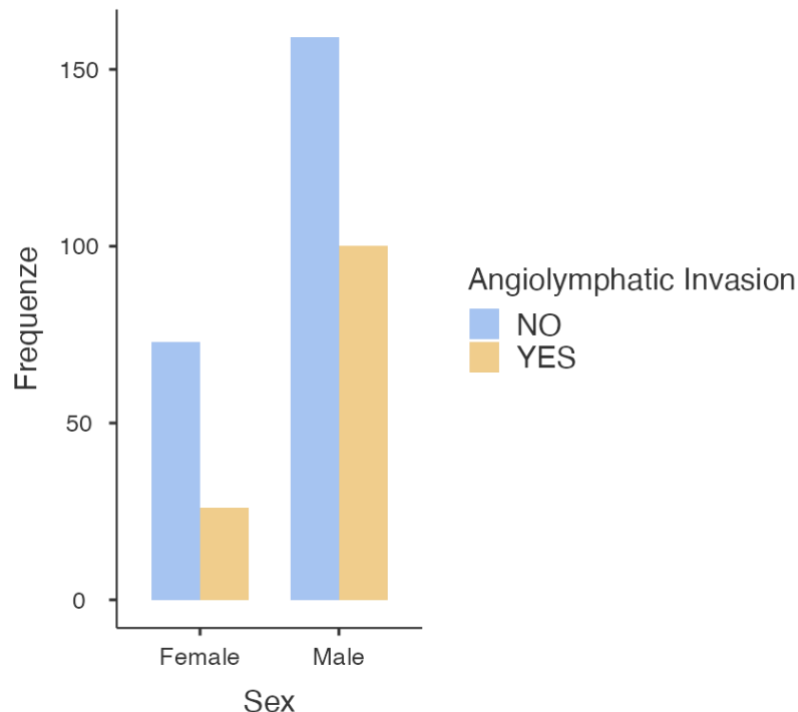


Figura 4. Frecuencia de pacientes con (YES) y sin (NO) invasión angiolinfática según el sexo.

En la Figura 4, se representa la distribución de pacientes con y sin invasión angiolinfática según el sexo. Se observa que, en ambos grupos, el número de pacientes sin invasión (barras azules) es superior al número de pacientes con invasión (barras naranjas). No obstante, los hombres presentan una mayor frecuencia absoluta de invasión angiolinfática en comparación con las mujeres. Este patrón visual sugiere una posible relación entre el sexo masculino y una mayor prevalencia de invasión, la cual será evaluada mediante pruebas estadísticas específicas. No obstante, dicha relación debe interpretarse con cautela, ya que se trata de una asociación no causal y podría estar influenciada por el desequilibrio en el número de hombres y mujeres en la muestra.

Para determinar si dicha diferencia es estadísticamente significativa, se aplicaron pruebas de contraste de hipótesis. Inicialmente, se utilizó la prueba de Chi-cuadrado (χ^2) de Pearson, adecuada para analizar la independencia entre variables categóricas. Debido a la posibilidad de frecuencias bajas en alguna de las combinaciones, se complementó el análisis con el *test* exacto de *Fisher*, más robusto en este tipo de situaciones.

Tabla 11. Tabla de la prueba de chi-cuadrado complementada con el *Test* exacto de *Fisher*.

Test χ^2

	Valor	gdl	p
χ^2	4.79	1	0.029
Test exacto de Fisher.			0.035
N	358		

Los resultados mostraron un valor de $\chi^2 = 4.79$ con 1 grado de libertad y un valor $p = 0.029$, lo que indica una asociación estadísticamente significativa entre el sexo y la presencia de invasión angiolinfática ($p < 0.05$). El *test* exacto de *Fisher* confirmó este hallazgo, mostrando un valor $p = 0.035$. Por tanto, se puede concluir que existe una diferencia significativa en la proporción de invasión angiolinfática entre hombres y mujeres.

Finalmente, se estimó el *odds ratio* (OR) con su intervalo de confianza del 95% como medida de asociación entre el sexo y la presencia de invasión angiolinfática.

Tabla 12. Tabla de mediadas comparativas utilizando el *Odds-Ratio*.

Medidas comparativas

	Valor	95% Intervalo de confianza	
		Inferior	Superior
Odds-Ratio	1.77	1.06	2.95

El valor obtenido fue $OR = 1.77$, con un intervalo de confianza del 95 % entre 1.06 y 2.95. Esto indica que los hombres presentaron un 77 % más de probabilidad de desarrollar invasión angiolinfática en comparación con las mujeres. Dado que el intervalo de confianza no incluye el valor 1, la asociación observada se considera estadísticamente significativa.

Los resultados sugieren que existe una asociación moderada y significativa entre el sexo y la presencia de invasión angiolinfática en pacientes con HNSCC. Sin embargo, esta asociación no implica causalidad directa, y debe interpretarse con cautela. Además, es importante tener en cuenta que la muestra analizada presentaba una desproporción entre el número de hombres y mujeres, lo que podría influir en la robustez clínica del resultado, a pesar de la significación estadística. Esta limitación puede afectar la confiabilidad del OR , especialmente en el subgrupo femenino con menor tamaño muestral, aun cuando se hayan aplicado pruebas estadísticas apropiadas.

10. DISCUSIÓN

Diversos estudios han señalado que la invasión angiolinfática (IAL) representa un evento histopatológico de relevancia clínica en el carcinoma escamoso de cabeza y cuello (HNSCC), ya que se ha asociado a una mayor tasa de recurrencia locorregional y a peor supervivencia (17,19). No obstante, los mecanismos moleculares que regulan este proceso aún no han sido completamente definidos, y existe interés creciente en identificar perfiles de expresión génica que permitan comprender mejor su biología subyacente.

En este trabajo, se identificaron siete genes con expresión diferencial estadísticamente significativa entre los grupos con y sin IAL: *KRT18*, *GADD45G*, *SIGIRR*, *S100A12*, *DSG1*, *GSDMA* y *WFDC12*. Estos genes fueron seleccionados por su posible implicación funcional en procesos relacionados con la diseminación tumoral, como la adhesión intercelular, la regulación de la inflamación y la homeostasis epitelial. Cuatro de ellos (*S100A12*, *DSG1*, *GSDMA* y *WFDC12*) presentaron una menor expresión en los tumores con IAL, mientras que *KRT18*, *GADD45G* y *SIGIRR* mostraron sobreexpresión en ese mismo grupo.

El análisis de enriquecimiento funcional mediante DAVID identificó que los genes diferencialmente expresados en tumores con IAL se agrupan en procesos biológicos como la respuesta inflamatoria, la defensa frente a patógenos y la adhesión célula-célula. Estos procesos han sido relacionados con la progresión invasiva del carcinoma escamoso de cabeza y cuello, dado su papel en la disrupción de la barrera epitelial y la modulación del microambiente inmunológico (13,20).

Cabe resaltar que varios de los genes identificados también han sido implicados en otros tipos de tumores epiteliales, lo que sugiere que podrían participar en mecanismos comunes de progresión tumoral. Sin embargo, es importante interpretar estos hallazgos con cautela, ya que los patrones de expresión pueden estar influenciados por múltiples factores, incluyendo el tipo histológico, el grado de diferenciación y el microambiente tumoral.

A continuación, se explora de forma individual el posible papel de cada uno de estos genes en el contexto de la invasión angiolinfática en HNSCC, contrastando nuestros resultados con la literatura científica disponible

KRT18

KRT18 (queratina 18) es una proteína de filamentos intermedios tipo I que forma heterodímeros con la queratina 8 en epitelios simples, constituyendo un componente esencial del citoesqueleto epitelial y brindando soporte mecánico a las células epiteliales (37). Más allá de su función estructural, la pareja *K8/K18* participa en la regulación de procesos celulares clave, incluyendo la resistencia a la apoptosis y la progresión del ciclo celular (38).

Esta capacidad de *KRT18* para conferir mayor supervivencia y estabilidad estructural podría proporcionar a las células cancerosas ventajas durante la invasión y diseminación metastásica. La expresión de *KRT18* suele estar alterada en las neoplasias epiteliales malignas. Numerosos estudios han documentado una sobreexpresión de *KRT18* en carcinomas humanos comparados con los tejidos epiteliales normales correspondientes. Esta sobreexpresión frecuentemente se asocia con un fenotipo tumoral más agresivo y con peor pronóstico clínico. Por ejemplo, en carcinoma colorrectal los niveles elevados de *KRT18* se correlacionan significativamente con invasión tumoral profunda, presencia de metástasis ganglionares, enfermedad metastásica a distancia y menor supervivencia (39).

De manera consistente, en carcinoma pulmonar de tipo no microcítico (adenocarcinoma), la inhibición experimental de *KRT18* disminuye la capacidad migratoria e invasiva de las células tumorales e incrementa su sensibilidad a agentes quimioterápicos, lo que indica que *KRT18* favorece la migración tumoral y la resistencia al estrés citotóxico (40).

En conjunto, estos hallazgos respaldan un rol pro-oncogénico de *KRT18* en la progresión tumoral, facilitando procesos de invasión y metástasis. Aunque el epitelio escamoso de la cabeza y cuello normalmente no expresa *K8/K18*, la detección anómala de estas queratinas en carcinomas escamosos de esta región se ha vinculado con mayor agresividad. Estudios inmunohistoquímicos han demostrado que la coexpresión ectópica de *KRT8/KRT18* en carcinomas escamosos ausente en el epitelio escamoso sano aparece en los tumores más invasivos y con propensión metastásica (41).

De hecho, en carcinomas escamosos de la cavidad oral (un subtipo de HNSCC), la presencia de *KRT8/18* se ha identificado como un marcador pronóstico independiente de mal pronóstico, asociado a una disminución de la supervivencia global y libre de progresión (42).

En el carcinoma nasofaríngeo otro tumor de cabeza y cuello se ha reportado que la sobreexpresión de *KRT18*, tanto en el tejido tumoral como en células tumorales circulantes detectadas en sangre, se asocia con mayor riesgo de metástasis y peor evolución clínica (43).

Todos estos datos sugieren que, en el contexto de carcinomas escamosos de cabeza y cuello, la reactivación o sobreexpresión de *KRT18* podría conferir a las células tumorales propiedades proinvasivas, favoreciendo especialmente la invasión angiolímfática y la diseminación metastásica inicial.

Este gen, identificado como sobreexpresado en los tumores de HNSCC con invasión angiolímfática, emerge, así como un posible mediador de la progresión metastásica, respaldado por hallazgos en HNSCC y en otros modelos de cáncer invasivo (39,41,42).

Cada uno de estos estudios refuerza la noción de que *KRT18* contribuye a un fenotipo más invasivo y agresivo, ya sea manteniendo ciertas características epiteliales que facilitan la invasión colectiva o protegiendo a las células tumorales del estrés apoptótico durante su tránsito por vasos sanguíneos y linfáticos. La integración de estos datos apoya la relevancia de *KRT18* en la biología invasiva del carcinoma escamoso de cabeza y cuello, posicionándolo como un posible marcador molecular de comportamiento agresivo (39,42).

GADD45G

Uno de los genes sobreexpresados más relevantes en nuestro análisis fue *GADD45G* (*Growth Arrest and DNA Damage-Inducible Gamma*), miembro de la familia *GADD45*, conocida por su papel en la respuesta al daño del ADN, el control del ciclo celular y la supresión tumoral. Este gen se activa bajo condiciones de estrés celular y participa en vías que conducen a la detención del ciclo y a la apoptosis (44).

En el contexto oncológico, se ha descrito que *GADD45G* actúa como supresor de la migración e invasión celular. En modelos de carcinoma escamoso de esófago, se ha demostrado que la expresión de *GADD45G* reduce significativamente la capacidad invasiva de las células tumorales al modular positivamente la expresión de E-cadherina, una proteína clave en la adhesión intercelular y la supresión de la transición epitelio-mesenquimal (EMT) (45). Dado que el EMT representa uno de los mecanismos fundamentales que permiten a las células epiteliales adquirir un fenotipo

mesenquimal y migratorio, este hallazgo sugiere un posible efecto protector de *GADD45G* frente a la invasión vascular.

En nuestro estudio, la sobreexpresión de *GADD45G* en pacientes con invasión angiolinfática podría representar un intento del tumor por compensar el aumento en agresividad biológica. No obstante, también es posible que el nivel de sobreexpresión refleje heterogeneidad tumoral o diferencias en el microambiente, dado que este gen puede estar regulado por estímulos como la hipoxia, la replicación aberrante o el estrés genotóxico.

En conjunto, aunque *GADD45G* ha sido tradicionalmente considerado un supresor tumoral, su sobreexpresión en los casos con invasión angiolinfática en nuestro estudio plantea una posible función no convencional, cuya relevancia biológica aún no ha sido aclarada. Esta hipótesis debe ser explorada en estudios funcionales adicionales para determinar si *GADD45G* contribuye directamente a la progresión invasiva o si actúa como marcador secundario de agresividad tumoral.

SIGIRR

Otro gen que ha surgido como relevante en nuestro análisis es *SIGIRR* (*Single Immunoglobulin Interleukin-1 Receptor-Related molecule*), también conocido como TIR8. Este gen pertenece a la familia de los receptores tipo Toll (TLR) e interleucina-1 (IL-1R), pero se diferencia funcionalmente por su papel inhibidor en las respuestas inflamatorias mediadas por estas vías. *SIGIRR* actúa como un regulador negativo de la señalización inflamatoria, particularmente bloqueando la activación de NF- κ B y JNK inducida por TLRs e IL-1, y desempeña un rol importante en la homeostasis del sistema inmune (46).

En el contexto tumoral, *SIGIRR* ha demostrado modular el microambiente tumoral al disminuir la inflamación crónica y la inmunosupresión asociada al cáncer, especialmente en tumores donde la inflamación juega un papel en la progresión. En estudios recientes sobre HNSCC positivo para *HPV*, se ha observado que una mayor expresión de *SIGIRR* se asocia con un perfil inmunológico más favorable, caracterizado por una mejor infiltración de linfocitos T CD8+ y una expresión reducida de genes inmunosupresores (47). Este patrón sugiere un posible efecto antitumoral mediado por la restauración de la vigilancia inmunitaria.

En nuestro análisis, se observó una sobreexpresión de *SIGIRR* en muestras con invasión angiolinfática. Aunque este gen ha sido descrito como regulador negativo de la inflamación y supresor de la progresión tumoral en varios contextos, su expresión aumentada en tumores más agresivos podría reflejar un mecanismo compensatorio del propio tumor o del microambiente inmunitario para limitar la activación excesiva de señales inflamatorias. Esta interpretación es coherente con el papel conocido de *SIGIRR* como modulador negativo de las vías IL-1/TLR, frecuentemente activadas en contextos tumorales inflamados (48,49). No obstante, aún no se ha establecido con claridad si esta sobreexpresión representa una respuesta reguladora funcional o una consecuencia secundaria de un entorno proinflamatorio.

Por otro lado, se ha propuesto que la pérdida o infraexpresión de *SIGIRR* en otros tumores (como colon) puede favorecer la progresión tumoral a través de una mayor activación inflamatoria, mientras que su preservación podría limitar la remodelación del microambiente y la permeabilidad vascular (50). Si bien faltan estudios directos que conecten *SIGIRR* con la invasión angiolinfática en HNSCC, sus funciones biológicas apoyan una implicación indirecta mediante la modulación del entorno tumoral, especialmente de las células endoteliales y del infiltrado inmune.

S100A12

S100A12 (*S100 calcium-binding protein A12*), también conocido como EN-RAGE, forma parte de la familia de proteínas S100, ampliamente implicadas en la regulación de la respuesta inmune, inflamación, migración celular y remodelación tisular. Se trata de una proteína secretada que actúa como ligando de RAGE (*receptor for advanced glycation end products*) y de TLR4, activando rutas proinflamatorias como NF- κ B y promoviendo la producción de citoquinas y especies reactivas de oxígeno (51).

En el contexto oncológico, la expresión de *S100A12* ha sido descrita principalmente en relación con procesos inflamatorios del microambiente tumoral, y su activación puede favorecer indirectamente fenómenos como la angiogénesis, la degradación de matriz extracelular y la diseminación celular. En estudios sobre carcinoma papilar de tiroides, se observó que el silenciamiento de *S100A12* inhibía significativamente la proliferación, migración e invasión de las células tumorales, lo cual indica que su expresión podría estar asociada con un fenotipo más agresivo (52).

De forma más directa, en cáncer de mama se ha reportado que niveles elevados de *S100A12* se correlacionan con invasión vascular y recurrencia. En este estudio, la expresión de *S100A12* se asoció con una mayor incidencia de células tumorales dentro de vasos linfáticos, sugiriendo un papel funcional en la invasión angiolinfática (53). Este tipo de evidencia es especialmente relevante dado el objetivo de nuestro trabajo, centrado precisamente en este tipo de diseminación.

En nuestro estudio, *S100A12* se encontró infraexpresado en los pacientes con invasión angiolinfática, lo que podría representar un hallazgo contradictorio con la literatura previa. No obstante, la función de *S100A12* en cáncer no es universalmente promotora; puede variar según el tipo de tumor y el microambiente. En el caso particular del carcinoma escamoso de cabeza y cuello (HNSCC), la evidencia disponible sugiere un rol complejo. Por un lado, estudios de expresión génica indican que varios miembros de la familia *S100*, incluyendo *S100A12*, están *down-regulated* (infraexpresados) en HNSCC en comparación con tejidos normales. En un análisis bioinformático integral, Bai et al. encontraron que la expresión de *S100A12* se correlacionaba significativamente con parámetros clínicos de agresividad en HNSCC, en particular con el grado histológico y con la presencia de invasión linfovascular y metástasis ganglionares (54).

Es decir, los casos de HNSCC con compromiso linfático o vascular tendían a mostrar niveles más bajos de *S100A12*. Este dato concuerda plenamente con nuestros hallazgos: en el metaanálisis de expresión génica realizado, *S100A12* resultó uno de los genes significativamente infraexpresados en los tumores con invasión angiolinfática (ALI) en HNSCC. La disminución de *S100A12* en tales tumores invasivos sugiere que la pérdida de su función podría contribuir a un entorno tumoral más permisivo para la diseminación.

Además, la vía *S100A12*–*RAGE* ha sido implicada en la activación del endotelio vascular, favoreciendo la extravasación celular (55). Una disminución de esta señalización en ciertos tipos de HNSCC podría indicar una menor participación del sistema inmune o una pérdida de control inflamatorio local, lo cual paradójicamente puede facilitar una invasión silenciosa hacia vasos linfáticos sin generar respuesta inmunológica fuerte.

S100A12 es una proteína inmunológica multifuncional vinculada a la respuesta de defensa contra infecciones y a la regulación de la inflamación. Su papel en el cáncer parece depender del equilibrio

entre sus efectos proinflamatorios y sus efectos de estimulación de la inmunidad antitumoral. La infraexpresión de S100A12 observada en los carcinomas de cabeza y cuello con invasión angiolinfática sugiere una pérdida de la respuesta inflamatoria defensiva local, lo cual podría facilitar la invasión y metástasis linfática. Este hallazgo es coherente con estudios previos que vinculan la disminución de S100A12 con mayor agresividad en HNSCC (56). No obstante, resulta necesario profundizar en cómo S100A12 interactúa con el microambiente tumoral inflamatorio de HNSCC.

Nuestros resultados aportan evidencia de que S100A12, al estar infraexpresado en tumores con invasión angiolinfática, podría servir como indicador molecular de un microambiente menos vigilado inmunológicamente y más propenso a la diseminación metastásica.

DSG1

DSG1 (*Desmoglein 1*) codifica una proteína transmembrana perteneciente a la familia de las cadherinas desmosomales, responsables de la adhesión celular intercelular en epitelios escamosos estratificados. Se localiza en los desmosomas, uniones intercelulares especializadas que anclan filamentos intermedios entre células adyacentes, proporcionando una fuerte adhesión mecánica y manteniendo la integridad estructural del tejido. Su función es esencial para mantener la cohesión y la arquitectura del epitelio, particularmente en tejidos como la epidermis y las mucosas orales (57).

La pérdida de función o la infraexpresión de *DSG1* ha sido asociada en diversos tumores con disrupción de la barrera epitelial, aumento de la motilidad celular y activación de vías oncogénicas como EGFR/ERK, que favorecen la proliferación y la invasión. En el contexto del carcinoma escamoso, incluyendo el HNSCC, esta pérdida de adhesión estructural puede facilitar la transición epitelio-mesénquima (EMT), promoviendo la migración celular y la penetración en estructuras vasculares(58).

En un estudio realizado por Valenzuela y colaboradores, se demostró que la pérdida de *DSG1* induce la formación de invadopodios, estructuras citoplasmáticas especializadas en degradar la matriz extracelular y facilitar la invasión tumoral y activa la señalización de EGFR, lo que sugiere un rol directo en la promoción de la diseminación celular a través de mecanismos vasculares. Asimismo,

la reintroducción de *DSG1* en células tumorales resultó en una reducción de la invasividad, confirmando su función como supresor de la invasión(58).

En nuestro estudio, *DSG1* aparece como uno de los genes significativamente infraexpresados en las muestras de HNSCC con invasión angiolinfática, lo que refuerza su papel funcional como barrera estructural frente a la diseminación tumoral.

un estudio en carcinoma escamoso de cabeza y cuello mostró que la expresión de *DSG1* se correlaciona positivamente con la de E-cadherina, y que la pérdida de *DSG1* es frecuente en tumores poco cohesionados (59). De hecho, cultivos de células de carcinoma escamoso (incluyendo líneas de carcinoma oral y cutáneo) presentan niveles muy reducidos o ausencia total de *DSG1* en comparación con queratinocitos normales (60).

Concordantemente, dichas células tumorales exhiben una drástica disminución en el número de desmosomas observados por microscopía electrónica, reflejando una pérdida estructural de las uniones adhesivas (60). Este patrón indica una alteración de las moléculas de adhesión epitelial en el cáncer escamoso y sugiere un cambio en las isoformas desmosómicas expresadas: mientras el epitelio escamoso sano expresa abundantemente *DSG1*, las células carcinomatosas tienden a perder *DSG1* y a depender de otras desmogleínas más basales (como *DSG2/DSG3*) (61). En conjunto, la infraexpresión de *DSG1* es una característica recurrente de los carcinomas escamosos, especialmente marcada en tumores agresivos con menor cohesión celular.

Cabe señalar que, además de su función mecánica, *DSG1* participa en la regulación de la diferenciación celular epitelial. Su infraexpresión ha sido asociada a tumores pobremente diferenciados, lo cual se correlaciona con un peor pronóstico y una mayor frecuencia de invasión linfovascular en múltiples tipos de carcinoma escamoso (59).

la evidencia disponible apoya que *DSG1* actúa como un supresor de la progresión invasiva en los carcinomas escamosos: su infraexpresión conlleva pérdida de cohesión epitelial, facilita cambios EMT-like y se asocia a mayor invasividad tumoral. Cada uno de estos aspectos refuerza la relevancia de *DSG1* como marcador de diferenciación epitelial y potencial factor protector contra la diseminación angiolinfática en el carcinoma escamoso de cabeza y cuello (18,59).

GSDMA

GSDMA (*Gasdermin A*) forma parte de la familia de las gasderminas, un conjunto de proteínas que han adquirido gran relevancia en los últimos años por su papel en la piroptosis, un tipo de muerte celular inflamatoria dependiente de caspasas. Esta proteína también se ha asociado a funciones en la homeostasis epitelial, diferenciación celular y apoptosis, actuando como un supresor potencial de la progresión tumoral en ciertos contextos celulares (62).

Estas proteínas son expresadas principalmente en epitelios escamosos y mucosos. En tejidos sanos, *GSDMA* se encuentra abundantemente en células epiteliales diferenciadas de la piel, lengua, esófago, estómago y otros epitelios mucosos (63).

En condiciones fisiológicas, *GSDMA* se expresa de forma preferencial en epitelios escamosos, incluyendo piel, tracto digestivo y mucosas. Estudios previos han sugerido que la pérdida de *GSDMA* puede estar relacionada con un fenotipo celular más agresivo, ya que su expresión es capaz de inducir apoptosis mediada por TGF- β , un proceso esencial para mantener la homeostasis de tejidos epiteliales y prevenir la acumulación de mutaciones oncogénicas (64).

En el contexto del carcinoma escamoso de cabeza y cuello (HNSCC), nuestro metaanálisis de expresión génica identificó a *GSDMA* significativamente infraexpresado en los tumores que presentan invasión angiolinfática. Aunque la literatura específica en HNSCC es limitada, este hallazgo concuerda con estudios en otros carcinomas epiteliales. Por ejemplo, en carcinomas escamosos del tracto digestivo superior, *GSDMA* sufre silenciamiento frecuente por mecanismos epigenéticos (p.ej., metilación del promotor) y su expresión desaparece en las células tumorales (64).

En cáncer gástrico, la pérdida de *GSDMA* es un hecho bien documentado la proteína prácticamente no se detecta en muchas líneas celulares y muestras de tumor gástrico en comparación con el epitelio normal (64).

Asimismo, en el carcinoma escamoso de esófago se ha observado el apagado transcripcional de *GSDMA*, sugiriendo que la pérdida de esta gasdermina es un rasgo común en la transformación maligna de epitelios escamosos (64). Dado que *GSDMA* normalmente funciona como guardiana de la integridad epitelial promoviendo la eliminación de células diferenciadas dañadas o infectadas, su ausencia podría conferir una ventaja a las células neoplásicas. La infraexpresión o inactivación de

GSDMA en tumores agresivos podría facilitar la evasión de la muerte celular regulada (evitando tanto la apoptosis como la piroptosis) y la evitación de la respuesta inmune local, favoreciendo así la supervivencia y diseminación de las células tumorales.

Si bien hacen falta más estudios específicos en HNSCC, la evidencia extrapolada de otros carcinomas epiteliales respalda la relevancia de GSDMA: su infraexpresión en tumores invasivos sugiere la pérdida de una barrera inmunológica y pro-apoptótica, lo cual podría contribuir al comportamiento agresivo y la capacidad de invasión angiolinfática de estas neoplasias.

WFDC12

WFDC12 (*WAP Four-Disulfide Core Domain 12*) pertenece a la familia de proteínas WFDC, caracterizadas por contener dominios tipo WAP (*whey acidic protein*) que actúan como inhibidores de proteasas y moduladores de procesos inflamatorios. Estas proteínas tienen un papel relevante en la respuesta inmunitaria innata y en la regulación de la infiltración celular en epitelios y tejidos linfáticos, y su disfunción se ha vinculado con alteraciones en la homeostasis tisular y la progresión tumoral (65).

WFDC12 ha sido implicado en procesos inflamatorios crónicos y enfermedades inmunomediadas. Su expresión en epitelios y su participación en la regulación de la respuesta inmune sugieren que podría tener relevancia en el microambiente tumoral (65).

En el contexto oncológico, la pérdida de expresión de *WFDC12* podría favorecer un microambiente pro-invasivo al eliminar un factor protector antiproteásico e inmunorregulador. Si bien no existen hasta el momento estudios específicos de *WFDC12* en carcinoma escamoso de cabeza y cuello, la evidencia en carcinomas escamosos de localizaciones análogas sugiere un rol tumor-supresor de este gen. En carcinoma escamoso de esófago (ESCC), *WFDC12* se encuentra significativamente infraexpresado en los tejidos tumorales en comparación con el epitelio normal. Más aún, experimentos funcionales demostraron que la supresión adicional de *WFDC12* (mediante *knockdown*) en líneas celulares de ESCC produce un aumento de la proliferación y migración de las células tumorales (66).

Este comportamiento concuerda con nuestros hallazgos de metanálisis en HNSCC, donde los tumores con invasión angiolinfática mostraron niveles reducidos de *WFDC12*. Es posible que los carcinomas más agresivos pierdan características de diferenciación escamosa como la expresión de *WFDC12*, un marcador de queratinocitos tardíos lo cual conlleva menor control sobre enzimas proteolíticas y señales inflamatorias en el entorno tumoral (65).

Además, el análisis de enriquecimiento funcional mostró que *WFDC12* participa en rutas relacionadas con la respuesta antimicrobiana y la inflamación, incluyendo el proceso GO:0002437 ("*inflammatory response to antigenic stimulus*") y GO:0002742 ("*defense response to bacterium*"), junto a otros genes como *S100A12* y *GSDMA*. Esta coincidencia sugiere un posible eje de regulación inflamatoria común en los tumores con menor expresión de estos genes, que podría afectar la estabilidad del endotelio linfático y la inmunidad local.

Relación entre el Invasión Angiolinfática y el HNSCC

Los resultados de nuestro análisis clínico revelan asociaciones significativas entre la presencia de invasión angiolinfática (IAL) y ciertas variables epidemiológicas clásicas del carcinoma escamoso de cabeza y cuello (HNSCC). En particular, se observó una mayor frecuencia de IAL en varones y en pacientes con antecedentes de tabaquismo activo o reciente. Estos hallazgos refuerzan la hipótesis de que algunos perfiles clínicos, además de predisponer al desarrollo tumoral, podrían también modular su comportamiento biológico invasivo. Aunque la edad mostró una tendencia sugestiva, y el alcohol no se asoció significativamente, estas variables podrían estar influidas por otros factores no controlados y requieren mayor exploración en cohortes más amplias.

Es importante señalar que, si bien la asociación entre el sexo masculino y la IAL fue estadísticamente significativa, nuestra cohorte presentó un desequilibrio de género con una mayor proporción de varones, lo que podría influir en la interpretación de este resultado. Este hallazgo sugiere la necesidad de estudios adicionales con una distribución de género más equilibrada para evaluar con mayor precisión la relación entre el sexo y la IAL en HNSCC. Diversos estudios, incluido uno de cohorte en Corea del Sur, han confirmado que los hombres presentan una incidencia significativamente mayor de HNSCC que las mujeres, incluso en ausencia de tabaquismo o consumo de alcohol (67). Además, se ha descrito que el sexo masculino también se asocia con una peor

supervivencia en pacientes con HNSCC, particularmente en aquellos con tumores negativos para VPH, lo que podría reflejar una mayor susceptibilidad biológica más allá de los factores conductuales clásicos (68).

La asociación significativa entre el tabaquismo y la invasión angiolinfática (IAL) en nuestro análisis clínico encuentra respaldo en la literatura reciente sobre HNSCC. Se ha observado que el tabaco no solo actúa como agente carcinógeno, sino que también puede inducir cambios moleculares que favorecen la progresión tumoral. En particular, estudios recientes han evidenciado que la exposición al tabaco altera la expresión de ciertos microARNs y genes reguladores asociados con procesos de migración, adhesión celular e invasión, todos ellos implicados en la diseminación linfovascular (69).

Estos efectos podrían estar relacionados con la alteración de genes identificados en nuestro análisis como *DSG1*, *GSDMA* y *WFDC12*, que se encontraron infraexpresados en pacientes con IAL y están implicados en funciones epiteliales, inmunológicas y de barrera estructural. La pérdida de *DSG1* ha sido asociada con una menor cohesión celular y mayor capacidad invasiva en carcinomas escamosos, mientras que *GSDMA* y miembros de la familia *WFDC* participan en la regulación de la inmunidad local y la integridad epitelial. Aunque no se ha demostrado de forma directa que el tabaco reduzca la expresión de estos genes en HNSCC, Por ejemplo, estudios han demostrado que la exposición al humo del cigarrillo altera la expresión génica en el epitelio respiratorio, afectando funciones clave como la adhesión celular y la respuesta inmune. Aunque estos estudios no se centran específicamente en los genes mencionados ni en el HNSCC, proporcionan evidencia de que el tabaquismo puede influir en la expresión génica en tejidos epiteliales (70).

En cuanto a la variable edad, aunque no se alcanzó significación estadística, la tendencia observada podría indicar una mayor susceptibilidad a la IAL en pacientes de mayor edad, posiblemente debido a una disminución en la capacidad de respuesta inmunitaria o a la acumulación de exposiciones a factores de riesgo a lo largo del tiempo. Sin embargo, se requieren estudios adicionales para confirmar esta hipótesis.

Estos hallazgos sugieren que los factores clínicos y conductuales podrían estar reflejando un trasfondo molecular más agresivo, detectable mediante el análisis de expresión génica. Por tanto, la integración de datos clínicos y moleculares es fundamental para entender la biología de la invasión angiolinfática y para avanzar hacia una medicina de precisión en HNSCC.

Limitaciones

Este estudio se basa exclusivamente en análisis computacionales, utilizando datos transcriptómicos secundarios extraídos de la base pública TCGA. Al no haber sido generados específicamente para responder a la hipótesis de investigación planteada, el diseño retrospectivo conlleva limitaciones inherentes, entre ellas la imposibilidad de controlar las condiciones originales de recolección de las muestras y de asegurar la inclusión sistemática de todas las variables clínicas o moleculares relevantes. Aunque la base de datos utilizada proporciona una cantidad considerable de información clínica, incluidos factores como el consumo de tabaco y alcohol, edad, estadio tumoral y estado del tumor, ciertos campos presentan valores ausentes o categorizaciones amplias que dificultan un análisis más fino. Por ejemplo, las categorías de tabaquismo o alcohol no permiten una cuantificación exacta de la exposición, lo que limita el ajuste por intensidad o duración del hábito. Esta situación se repite en numerosos estudios que utilizan cohortes públicas como TCGA, generando una cierta heterogeneidad metodológica y dificultando la comparación directa entre investigaciones.

Adicionalmente, la composición de la cohorte presenta un claro desequilibrio entre los grupos comparados, tanto en número con una mayoría de pacientes sin invasión angiolinfática frente a aquellos con ella como en variables demográficas, especialmente el sexo, donde existe un predominio masculino. Aunque se aplicaron análisis estadísticos para explorar la relación de variables como sexo, edad y hábitos tóxicos con la presencia de invasión, no puede descartarse la influencia de factores de confusión no contemplados o interacciones complejas entre variables. Esta asimetría entre grupos puede afectar la robustez de los resultados y la potencia estadística para detectar asociaciones significativas. Por tanto, los hallazgos obtenidos deben ser interpretados como preliminares y exploratorios, a la espera de estudios adicionales que evalúen de forma funcional los genes identificados.

Además, cabe señalar que actualmente existen pocos estudios centrados específicamente en el análisis de expresión génica diferencial asociado a la invasión angiolinfática en el carcinoma escamoso de cabeza y cuello. Esta escasez de literatura limita la posibilidad de contrastar y validar los resultados con trabajos previos, pero al mismo tiempo subraya la relevancia y la novedad de este estudio, que puede contribuir al avance del conocimiento en un área poco explorada, pero de gran interés clínico.

11. CONCLUSIÓN

La invasión angiolinfática en el carcinoma escamoso de cabeza y cuello se asoció a una firma génica diferencial compuesta por 375 genes, entre los que destacan *KRT18*, *GADD45G*, *SIGIRR*, *S100A12*, *DSG1*, *GSDMA* y *WFDC12*. El análisis funcional mostró una implicación significativa de procesos como la adhesión célula-célula, la exocitosis y la respuesta inmune, lo que refuerza su papel en la progresión tumoral. Estos resultados permitieron proponer dichos genes como posibles biomarcadores para la detección precoz de la invasión angiolinfática. Además, se identificaron asociaciones clínicas con el tabaquismo y el sexo, si bien esta última debe interpretarse con cautela debido al desequilibrio en la distribución de la muestra.

12. BIBLIOGRAFÍA

1. Shirima CA, Bleotu C, Spandidos DA, El-Naggar AK, Pircalabioru GG, Michalopoulos I. Epithelial-derived head and neck squamous tumourigenesis (Review). Vol. 52, *Oncology Reports*. Spandidos Publications; 2024.
2. Mores AL, Bonfim-Alves CG, López RVM, Rodrigues-Oliveira L, Palmier NR, Mariz BALA, et al. Prognostic Factors in Head and Neck Cancer: A Retrospective Cohort Study of 3052 Patients in Brazil. *Oral Dis*. 2024 Apr 1;
3. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Medical Sciences*. 2023 Jun 13;11(2):42.
4. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J*. 2022 Nov 11;233(9):780–6.
5. Mroz EA, Rocco JW. Intra-tumor heterogeneity in head and neck cancer and its clinical implications. Vol. 2, *World Journal of Otorhinolaryngology - Head and Neck Surgery*. KeAi Communications Co.; 2016. p. 60–7.
6. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015 Jan 29;517(7536):576–82.
7. Zigon G, Berrino F, Gatta G, Sánchez MJ, van Dijk B, Van Eycken E, et al. Prognoses for head and neck cancers in Europe diagnosed in 1995-1999: A population-based study. *Annals of Oncology*. 2011 Jan 1;22(1):165–74.
8. Perera M, Al-Hebshi NN, Speicher DJ, Perera I, Johnson NW. Emerging role of bacteria in oral carcinogenesis: A review with special reference to perio-pathogenic bacteria. Vol. 8, *Journal of Oral Microbiology*. Co-Action Publishing; 2016.
9. Liu C, Wang M, Zhang H, Li C, Zhang T, Liu H, et al. Tumor microenvironment and immunotherapy of oral cancer. Vol. 27, *European Journal of Medical Research*. BioMed Central Ltd; 2022.
10. Jarosz-Biej M, Smolarczyk R, Cichoń T, Kułach N. Tumor microenvironment as a “game changer” in cancer radiotherapy. Vol. 20, *International Journal of Molecular Sciences*. MDPI AG; 2019.
11. Arneth B. Tumor microenvironment. Vol. 56, *Medicina (Lithuania)*. MDPI AG; 2020.
12. Jiffar Tilahun, Turker Yilmaz, Junegoo Lee, Yair Miller, Lei Feng, Adel El-Naggar, et al. Brain derived neutrophic factor (BDNF) coordinates lymphovascular metastasis through a fibroblast-governed paracrine axis in the tumor microenvironment. *Cancer Cell Microenviron*. 2017 Jul 10;
13. Nathan CA, Khandelwal AR, Wolf GT, Rodrigo JP, Mäkitie AA, Saba NF, et al. TP53 mutations in head and neck cancer. *Mol Carcinog*. 2022 Apr 1;61(4):385–91.
14. Van den bossche V, Zaryouh H, Vara-Messler M, Vignau J, Machiels JP, Wouters A, et al. Microenvironment-driven intratumoral heterogeneity in head and neck cancers: clinical challenges and opportunities for precision medicine. *Drug Resistance Updates*. 2022 Jan 1;60.
15. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep*. 2014 Dec;15(12):1243–53.
16. Karamanos NK, Theocharis AD, Piperigkou Z, Manou D, Passi A, Skandalis SS, et al. A guide to the composition and functions of the extracellular matrix. *FEBS Journal*. 2021 Dec 1;288(24):6850–912.
17. Huang S, Zhu Y, Cai H, Zhang Y, Hou J. Impact of lymphovascular invasion in oral squamous cell carcinoma: A meta-analysis. Vol. 131, *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*. Mosby Inc.; 2021. p. 319-328.e1.
18. Metzger K, Moratin J, Freier K, Hoffmann J, Zaoui K, Plath M, et al. A six-gene expression signature related to angiolymphatic invasion is associated with poor survival in laryngeal squamous cell carcinoma. *European Archives of Oto-Rhino-Laryngology*. 2021 Apr 1;278(4):1199–207.

19. Mascitti M, Togni L, Caponio VCA, Zhurakivska K, Bizzoca ME, Contaldo M, et al. Lymphovascular invasion as a prognostic tool for oral squamous cell carcinoma: a comprehensive review. Vol. 51, *International Journal of Oral and Maxillofacial Surgery*. Churchill Livingstone; 2022. p. 1–9.
20. Zhang J, Lin H, Jiang H, Jiang H, Xie T, Wang B, et al. A key genomic signature associated with lymphovascular invasion in head and neck squamous cell carcinoma. *BMC Cancer*. 2020 Mar 30;20(1).
21. Karahatay S, Thomas K, Koybasi S, Senkal CE, Elojeimy S, Liu X, et al. Clinical Relevance of Ceramide Metabolism in the Pathogenesis of Human Head and Neck Squamous Cell Carcinoma (HNSCC): Attenuation of C 18-ceramide in HNSCC Tumors Correlates with Lymphovascular Invasion and Nodal Metastasis. Vol. 256, *Cancer Lett*. 2007.
22. Chen YQ, Kuo JC, Wei MT, Chen YC, Yang MH, Chiou A. Early stage mechanical remodeling of collagen surrounding head and neck squamous cell carcinoma spheroids correlates strongly with their invasion capability. *Acta Biomater*. 2019 Jan 15;84:280–92.
23. Kumar AT, Knops A, Swendseid B, Martinez-Outschoom U, Harshyne L, Philp N, et al. Prognostic Significance of Tumor-Associated Macrophage Content in Head and Neck Squamous Cell Carcinoma: A Meta-Analysis. *Front Oncol*. 2019 Jul 23;9.
24. Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. Vol. 61, *Seminars in Cancer Biology*. Academic Press; 2020. p. 71–83.
25. Solanke AU. Genomics: An Integrative Approach for Molecular Biology [Internet]. Available from: <https://www.researchgate.net/publication/283084372>
26. Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. *PLoS Comput Biol*. 2017 May 1;13(5).
27. Costa V, Angelini C, De Feis I, Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-Seq. Vol. 2010, *Journal of Biomedicine and Biotechnology*. 2010.
28. Khodadadian A, Darzi S, Haghi-Daredeh S, Eshaghi FS, Babakhanzadeh E, Mirabutalebi SH, et al. Genomics and transcriptomics: The powerful technologies in precision medicine. Vol. 13, *International Journal of General Medicine*. Dove Medical Press Ltd; 2020. p. 627–40.
29. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec 5;15(12).
30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles [Internet]. 2005. Available from: www.pnas.org/doi/10.1073/pnas.0506580102
31. Lin R, Dai S, Irwin RD, Heinloth AN, Boorman GA, Li L. Gene set enrichment analysis for non-monotone association and multiple experimental categories. *BMC Bioinformatics*. 2008 Nov 14;9.
32. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022 Jul 5;50(W1):W216–21.
33. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
34. Citation del software R es . R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/> .
35. Law CW, Alhamdoosh M, Su S, Smyth GK, Ritchie ME. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res*. 2016 Jun 17;5:1408.
36. The jamovi project. Jamovi, The jamovi project. (2024). jamovi (2.6.26). <https://www.jamovi.org>. 2024.
37. Toivola DM, Ku NO, Resurreccion EZ, Nelson DR, Wright TL, Omary MB. Keratin 8 and 18 hyperphosphorylation is a marker of progression of human liver disease. *Hepatology*. 2004 Aug;40(2):459–66.

38. Chen B, Xu X, Lin DD, Chen X, Xu YT, Liu X, et al. KRT18 Modulates Alternative Splicing of Genes Involved in Proliferation and Apoptosis Processes in Both Gastric Cancer Cells and Clinical Samples. *Front Genet.* 2021 Jul 5;12.
39. Zhang J, Hu S, Li Y. KRT18 is correlated with the malignant status and acts as an oncogene in colorectal cancer. *Biosci Rep.* 2019 Aug 13;39(8).
40. Zhang B, Wang J, Liu W, Yin Y, Qian D, Zhang H, et al. Cytokeratin 18 knockdown decreases cell migration and increases chemosensitivity in non-small cell lung cancer. *J Cancer Res Clin Oncol.* 2016 Dec 1;142(12):2479–87.
41. Yamashiro Y, Takei K, Umikawa M, Asato T, Oshiro M, Uechi Y, et al. Ectopic coexpression of keratin 8 and 18 promotes invasion of transformed keratinocytes and is induced in patients with cutaneous squamous cell carcinoma. *Biochem Biophys Res Commun.* 2010 Aug;399(3):365–72.
42. Fillies T, Werkmeister R, Packeisen J, Brandt B, Morin P, Weingart D, et al. Cytokeratin 8/18 expression indicates a poor prognosis in squamous cell carcinomas of the oral cavity. *BMC Cancer.* 2006 Jan 13;6.
43. Lin M, Zhang XL, You R, Liu YP, Cai HM, Liu LZ, et al. Evolutionary route of nasopharyngeal carcinoma metastasis and its clinical significance. *Nat Commun.* 2023 Dec 1;14(1).
44. Zhan Q, Antinore MJ, Wang XW, Carrier F, Smith ML, Harris CC, et al. Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene.* 1999 May 6;18(18):2892–900.
45. Li T, Xu L, Teng J, Ma Y, Liu W, Wang Y, et al. GADD45G Interacts with E-cadherin to Suppress the Migration and Invasion of Esophageal Squamous Cell Carcinoma. *Dig Dis Sci.* 2020 Apr 1;65(4):1032–41.
46. Garlanda C, Anders HJ, Mantovani A. TIR8/SIGIRR: an IL-1R/TLR family member with regulatory functions in inflammation and T cell polarization. Vol. 30, *Trends in Immunology.* 2009. p. 439–46.
47. Yin G, Guo W, Wang R, Li N, Chen X, Zhang Y, et al. Analysis of the role of IL-1 family and related genes in head and neck squamous cell carcinoma. *Braz J Otorhinolaryngol.* 2025 Jan 1;91(1).
48. Watari K, Shibata T, Kawahara A, Sata KI, Nabeshima H, Shinoda A, et al. Tumor-derived interleukin-1 promotes lymphangiogenesis and lymph node metastasis through M2-type macrophages. *PLoS One.* 2014 Jun 12;9(6).
49. Mantovani A, Barajon I, Garlanda C. IL-1 and IL-1 regulatory pathways in cancer progression and therapy. Vol. 281, *Immunological Reviews.* Blackwell Publishing Ltd; 2018. p. 57–61.
50. Zhao J, Zepp J, Bulek K, Li X. SIGIRR, a negative regulator of colon tumorigenesis. Vol. 8, *Drug Discovery Today: Disease Mechanisms.* Elsevier Ltd; 2011.
51. Foell D, Wittkowski H, Vogl T, Roth J. S100 proteins expressed in phagocytes: a novel group of damage-associated molecular pattern molecules. *J Leukoc Biol.* 2007 Jan 1;81(1):28–37.
52. Wang X, Sun Z, Tian W, Piao C, Xie X, Zang J, et al. S100A12 is a promising biomarker in papillary thyroid cancer. *Sci Rep.* 2020 Dec 1;10(1).
53. Gunaldi M, Okuturlar Y, Gedikbasi A, Akarsu C, Karabulut M, Kural A. Diagnostic importance of S100A9 and S100A12 in breast cancer. *Biomedicine and Pharmacotherapy.* 2015 Dec 1;76:52–6.
54. Bai G, Yue S, You Y, Zhao J, Tian Z, Wang Y, et al. An integrated bioinformatics analysis of the S100 in head and neck squamous cell carcinoma. *Transl Cancer Res.* 2023;12(4):717–31.
55. Sionov RV. Leveling up the controversial role of neutrophils in cancer: When the complexity becomes entangled. Vol. 10, *Cells.* MDPI; 2021.
56. Lu C, Liu J, Yao M, Li L, Li G. Downregulation of S100 calcium binding protein A12 inhibits the growth of glioma cells. *BMC Cancer.* 2020 Mar 30;20(1).
57. Garrod D, Chidgeyt M, North A. Desmosomes: differentiation, development, dynamics and disease Abbreviations AJ adherens junction CAR cell adhesion recognition Dsc desmocollin Dsg

- desmoglein IgCAM Ig-like cell adhesion molecule PKC protein kinase C TPA 12-O-tetradecanoylphorbol 13-acetate. *Curr Opin Cell Biol.* 1996;8:6–6.
58. Valenzuela-Iglesias A, Burks HE, Arnette CR, Yalamanchili A, Nekrasova O, Godsel LM, et al. Desmoglein 1 regulates invadopodia by suppressing EGFR/ERK signaling in an ERBIN-dependent manner. *Molecular Cancer Research.* 2019 May 1;17(5):1195–206.
 59. Wong MP, Cheang M, Yorlida E, Coldman A, Gilks CB, Huntsman D, et al. Loss of desmoglein 1 expression associated with worse prognosis in head and neck squamous cell carcinoma patients. *Pathology.* 2008;40(6):611–6.
 60. Han BA, Yang XP, Hosseini DK, Zhang P, Zhang Y, Yu JT, et al. Identification of candidate aberrantly methylated and differentially expressed genes in Esophageal squamous cell carcinoma. *Sci Rep.* 2020 Dec 1;10(1).
 61. Harada H, Iwatsuki K, Ohtsuka M, Han G, Kaneko F. Abnormal desmoglein expression by squamous cell carcinoma cells. *Acta Derm Venereol* [Internet]. 1996 Nov 1;76(6):417–20. Available from: <https://medicaljournalssweden.se/actadv/article/view/15041>
 62. Rogers C, Erkes DA, Nardone A, Aplin AE, Fernandes-Alnemri T, Alnemri ES. Gasdermin pores permeabilize mitochondria to augment caspase-3 activation during apoptosis and inflammasome activation. *Nat Commun.* 2019 Dec 1;10(1).
 63. Zou J, Zheng Y, Huang Y, Tang D, Kang R, Chen R. The Versatile Gasdermin Family: Their Function and Roles in Diseases. Vol. 12, *Frontiers in Immunology.* Frontiers Media S.A.; 2021.
 64. Saeki N, Usui T, Aoyagi K, Kim DH, Sato M, Mabuchi T, et al. Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. *Genes Chromosomes Cancer.* 2009 Mar;48(3):261–71.
 65. Kalinina P, Vorstandlechner V, Buchberger M, Eckhart L, Lengauer B, Golabi B, et al. The Whey Acidic Protein WFDC12 Is Specifically Expressed in Terminally Differentiated Keratinocytes and Regulates Epidermal Serine Protease Activity. *Journal of Investigative Dermatology.* 2021 May 1;141(5):1198-1206.e13.
 66. Zhuang Y, Wang L, Ji C, Sun Y, Shao F. Construction of a novel pyroptosis-related prognostic model of esophageal square cell carcinoma and determination of the anti-tumor effect of WFDC12. *Funct Integr Genomics.* 2023 Jun 1;23(2).
 67. Park JO, Nam IC, Kim CS, Park SJ, Lee DH, Kim HB, et al. Sex Differences in the Prevalence of Head and Neck Cancers: A 10-Year Follow-Up Study of 10 Million Healthy People. *Cancers (Basel).* 2022 May 1;14(10).
 68. Mazul AL, Naik AN, Zhan KY, Stepan KO, Old MO, Kang SY, et al. Gender and race interact to influence survival disparities in head and neck cancer. *Oral Oncol.* 2021 Jan 1;112.
 69. Shen Y jie, Ji M you, Huang Q, Hsueh CY, Du H dong, Zhang M, et al. Nicotine downregulates miR-375–3p via neurotrophic tyrosine receptor kinase 2 to enhance the malignant behaviors of laryngopharyngeal squamous epithelial cells. *Ecotoxicol Environ Saf.* 2024 Apr 1;274.
 70. Floreani AA, Wyatt TA, Stoner J, Sanderson SD, Thompson EG, Allen-Gipson D, et al. Smoke and C5a induce airway epithelial intercellular adhesion molecule-1 and cell adhesion. *Am J Respir Cell Mol Biol.* 2003 Oct 1;29(4):472–82.

13. ANEXOS

Tabla 1 Anexo. Los 274 genes diferenciales sobreexpresados con FDR < 0.05.

GENES	fold change	pvalue	FDR
FAM174B	0.462	5.1×10^{-08}	0.002
ARNILA	0.55	1.4×10^{-07}	0.003
NPDC1	0.49	2.0×10^{-07}	0.003
RHPN1	0.673	3.7×10^{-07}	0.004
DNASE2	0.268	6.5×10^{-07}	0.005
COX14	0.264	8.1×10^{-07}	0.006
FAM50A	0.316	1.0×10^{-06}	0.006
PKN1	0.413	1.6×10^{-06}	0.007
P4HTM	0.449	2.1×10^{-06}	0.008
PCBD1	0.285	2.7×10^{-06}	0.009
PRR15	0.763	3.5×10^{-06}	0.009
PRKAG2-AS1	0.626	3.8×10^{-06}	0.009
CATSPERG	0.394	3.7×10^{-06}	0.009
CFAP43	0.523	5.9×10^{-06}	0.010
KRT18	0.481	6.1×10^{-06}	0.010
GADD45G	0.473	6.6×10^{-06}	0.010
CCNO	0.467	6.1×10^{-06}	0.010
TSC22D3	0.354	7.4×10^{-06}	0.010
SIGIRR	0.322	7.3×10^{-06}	0.010
ARRDC2	0.289	6.7×10^{-06}	0.010
DPP7	0.28	6.0×10^{-06}	0.010
TRAPPC9	0.226	6.3×10^{-06}	0.010
ZFYVE27	0.218	6.6×10^{-06}	0.010
NAT8L	0.663	9.1×10^{-06}	0.011
FAM186A	0.527	1.1×10^{-05}	0.012
FN3K	0.476	1.0×10^{-05}	0.012
FKBP4	0.272	1.1×10^{-05}	0.012
TTC19	0.229	1.1×10^{-05}	0.012
LONRF2	0.895	1.3×10^{-05}	0.013

CHKA	0.355	1.6×10^{-05}	0.015
PYCR2	0.-259	1.6×10^{-05}	0.015
TOR3A	0.241	1.5×10^{-05}	0.015
COQ10A	0.277	2.0×10^{-05}	0.018
AIFM2	0.309	2.1×10^{-05}	0.019
HAGH	0.199	2.2×10^{-05}	0.019
RPL39L	0.663	2.3×10^{-05}	0.019
RNF32-DT	0.567	2.5×10^{-05}	0.019
MELTF	0.517	2.5×10^{-05}	0.019
GKAP1	0.472	2.5×10^{-05}	0.019
FAM241B	0.47	2.4×10^{-05}	0.019
RHEBL1	0.385	2.5×10^{-05}	0.019
LINC02367	0.568	2.6×10^{-05}	0.020
SLC25A38	0.207	2.7×10^{-05}	0.020
CD63-AS1	0.289	2.8×10^{-05}	0.020
CUEDC2	0.208	3.0×10^{-05}	0.020
VAMP2	0.204	2.9×10^{-05}	0.020
CHTOP	0.155	2.9×10^{-05}	0.020
CRB2	0.747	3.4×10^{-05}	0.021
GFRA3	0.733	3.7×10^{-05}	0.021
PRR15L	0.689	3.4×10^{-05}	0.021
KLF15	0.626	3.3×10^{-05}	0.021
TG	0.576	3.6×10^{-05}	0.021
ASRGL1	0.54	3.7×10^{-05}	0.021
LAMB2P1	0.425	3.2×10^{-05}	0.021
FBXW4	0.193	3.8×10^{-05}	0.021
ANAPC16	0.176	3.9×10^{-05}	0.021
KRT8	0.804	4.3×10^{-05}	0.022
PIERCE1	0.378	4.1×10^{-05}	0.022
TRIM11	0.166	4.3×10^{-05}	0.022
ACBD7	0.587	4.6×10^{-05}	0.022
MYEF2	0.555	4.6×10^{-05}	0.022
BICDL3P	0.512	5.0×10^{-05}	0.022

HES6	0.429	5.0×10^{-05}	0.022
SLC29A2	0.422	4.9×10^{-05}	0.022
NR1I3	0.356	4.9×10^{-05}	0.022
INTS6-AS1	0.331	4.5×10^{-05}	0.022
UBL4A	0.171	4.9×10^{-05}	0.022
SUPV3L1	0.165	5.0×10^{-05}	0.022
SHH	0.87	5.2×10^{-05}	0.023
SEPSECS-AS1	0.347	5.2×10^{-05}	0.023
RNF157	0.449	5.3×10^{-05}	0.023
GOT2P3	0.454	5.4×10^{-05}	0.023
PLEKHB1	0.497	5.8×10^{-05}	0.023
ABCA2	0.388	5.9×10^{-05}	0.023
TCEA2	0.309	5.8×10^{-05}	0.023
PPP1R16A	0.257	6.0×10^{-05}	0.023
CDKN1B	0.228	5.9×10^{-05}	0.023
UQCC6	0.218	5.7×10^{-05}	0.023
ASCL5	0.528	7.0×10^{-05}	0.024
USP44	0.48	7.0×10^{-05}	0.024
CENPV	0.462	6.3×10^{-05}	0.024
LAMTOR5-AS1	0.298	6.7×10^{-05}	0.024
UBE2T	0.288	6.9×10^{-05}	0.024
PRRT3	0.287	6.2×10^{-05}	0.024
MPC2	0.194	6.4×10^{-05}	0.024
IRF2BP1	0.199	7.2×10^{-05}	0.024
MAPKAPK5-AS1	0.231	7.4×10^{-05}	0.025
B9D2	0.222	7.4×10^{-05}	0.025
FLVCR1	0.334	7.6×10^{-05}	0.025
SUZ12P1	0.296	7.7×10^{-05}	0.025
IFITM10	0.448	7.7×10^{-05}	0.025
SLC45A1	0.435	7.8×10^{-05}	0.025
C12orf57	0.305	8.1×10^{-05}	0.026
MYB	0.558	8.6×10^{-05}	0.026
TMEM163	0.528	8.4×10^{-05}	0.026

HNRNPD-DT	0.404	8.6×10^{-05}	0.026
NEIL1	0.426	8.9×10^{-05}	0.027
EFCAB2	0.335	9.3×10^{-05}	0.027
SLC16A4-AS1	0.45	9.6×10^{-05}	0.028
ARFGEF3	0.624	1.0×10^{-04}	0.028
DNAJC27-AS1	0.336	1.0×10^{-04}	0.028
LRRC38	0.685	1.0×10^{-04}	0.028
CACNA2D2	0.376	1.0×10^{-04}	0.028
TMEM109-DT	0.349	1.0×10^{-04}	0.028
CFAP36	0.216	1.0×10^{-04}	0.028
RBMX2	0.154	1.0×10^{-04}	0.028
CD63	0.183	1.0×10^{-04}	0.029
ZNF768	0.208	1.1×10^{-04}	0.030
ATP6V0E2	0.446	1.1×10^{-04}	0.030
CBX8	0.23	1.1×10^{-04}	0.030
SLC25A14	0.198	1.2×10^{-04}	0.030
CLDN3	0.866	1.2×10^{-04}	0.030
NEURL1	0.554	1.2×10^{-04}	0.030
KCNC3	0.468	1.2×10^{-04}	0.030
MACROD1	0.323	1.2×10^{-04}	0.030
TSPYL2	0.295	1.2×10^{-04}	0.030
TMEM205	0.265	1.2×10^{-04}	0.030
SSBP4	0.201	1.2×10^{-04}	0.030
PIM3	0.192	1.2×10^{-04}	0.030
MELTF-AS1	0.496	1.3×10^{-04}	0.030
KIF1A	0.859	1.3×10^{-04}	0.030
RNF157-AS1	0.635	1.3×10^{-04}	0.030
ABCA3	0.571	1.3×10^{-04}	0.030
DSTNP2	0.26	1.3×10^{-04}	0.030
PEMT	0.216	1.4×10^{-04}	0.030
NDUFA9	0.21	1.3×10^{-04}	0.030
MED9	0.173	1.4×10^{-04}	0.030
RNF113A	0.167	1.3×10^{-04}	0.030

SCAMP3	0.141	1.3×10^{-04}	0.030
BEND7	0.528	1.4×10^{-04}	0.031
REM2	0.392	1.4×10^{-04}	0.031
DDIT3	0.339	1.5×10^{-04}	0.031
DLGAP1-AS1	0.303	1.5×10^{-04}	0.032
HAP1	0.588	1.6×10^{-04}	0.032
ITPKA	0.433	1.5×10^{-04}	0.032
BEGAIN	0.431	1.5×10^{-04}	0.032
MCRIP2	0.291	1.5×10^{-04}	0.032
ABHD11	0.287	1.5×10^{-04}	0.032
RAD23A	0.187	1.5×10^{-04}	0.032
PDCD6IPP2	0.392	1.6×10^{-04}	0.033
PGPEP1	0.331	1.6×10^{-04}	0.033
TUSC2	0.141	1.6×10^{-04}	0.033
CRACR2B	0.414	1.6×10^{-04}	0.033
UBXN11	0.295	1.6×10^{-04}	0.033
TSNARE1	0.251	1.7×10^{-04}	0.034
C7	0.777	1.9×10^{-04}	0.034
PODXL2	0.721	1.8×10^{-04}	0.034
PLCH1	0.608	1.8×10^{-04}	0.034
PLCB4	0.597	1.9×10^{-04}	0.034
EPCAM	0.56	1.7×10^{-04}	0.034
OR7E14P	0.519	1.9×10^{-04}	0.034
AK8	0.502	1.9×10^{-04}	0.034
NT5M	0.488	1.8×10^{-04}	0.034
UPK3A	0.448	1.9×10^{-04}	0.034
SYNM	0.411	1.7×10^{-04}	0.034
MINCR	0.361	1.9×10^{-04}	0.034
PPP1R3E	0.267	1.8×10^{-04}	0.034
ASPSCR1	0.23	1.8×10^{-04}	0.034
RABL2B	0.224	1.8×10^{-04}	0.034
MRPS26	0.209	1.8×10^{-04}	0.034
EXOSC5	0.186	1.9×10^{-04}	0.034

MRPL24	0.174	1.9×10^{-04}	0.034
CEBPG	0.151	1.9×10^{-04}	0.034
PSMD8	0.12	1.8×10^{-04}	0.034
ZNF653	0.188	1.9×10^{-04}	0.034
KCTD13	0.237	2.0×10^{-04}	0.035
NPTXR	0.593	2.0×10^{-04}	0.035
GPT	0.463	2.0×10^{-04}	0.035
ZFHX2	0.418	2.0×10^{-04}	0.035
ARHGAP39	0.261	2.0×10^{-04}	0.035
RHPN1-AS1	0.433	2.1×10^{-04}	0.036
KIF9	0.215	2.3×10^{-04}	0.038
NRBF2P6	0.349	2.3×10^{-04}	0.039
KRT8P45	0.548	2.3×10^{-04}	0.039
DNAAF1	0.445	2.3×10^{-04}	0.039
PLAAT3	0.44	2.3×10^{-04}	0.039
CD27-AS1	0.305	2.3×10^{-04}	0.039
SPTBN4	0.414	2.4×10^{-04}	0.039
ZMYND10	0.453	2.4×10^{-04}	0.040
RPSAP69	0.484	2.5×10^{-04}	0.040
LINC01063	0.426	2.5×10^{-04}	0.040
GABRP	0.949	2.6×10^{-04}	0.041
CORO6	0.616	2.6×10^{-04}	0.041
SLC12A7	0.39	2.7×10^{-04}	0.041
PRR19	0.324	2.6×10^{-04}	0.041
TM9SF5P	0.32	2.6×10^{-04}	0.041
P2RX4	0.257	2.6×10^{-04}	0.041
LRRC27	0.242	2.6×10^{-04}	0.041
NCBP2AS2	0.224	2.6×10^{-04}	0.041
EMC10	0.205	2.6×10^{-04}	0.041
WDR83	0.205	2.7×10^{-04}	0.041
JMJD4	0.185	2.5×10^{-04}	0.041
ZNF691	0.184	2.7×10^{-04}	0.041
CEP89	0.172	2.7×10^{-04}	0.041

MCAT	0.157	2.7×10^{-04}	0.041
CRTC2	0.138	2.6×10^{-04}	0.041
AIP	0.195	2.7×10^{-04}	0.041
ZNF48	0.257	2.7×10^{-04}	0.041
CLDN7	0.477	2.8×10^{-04}	0.041
TRIM46	0.328	2.7×10^{-04}	0.041
RBM8A	0.146	2.8×10^{-04}	0.041
CYSTM1	0.266	2.9×10^{-04}	0.042
DTNBP1	0.199	2.9×10^{-04}	0.042
CA11	0.34	2.9×10^{-04}	0.042
C12orf76	0.224	3.0×10^{-04}	0.043
NUP210	0.51	3.1×10^{-04}	0.043
TMEM52	0.456	3.1×10^{-04}	0.043
H1-10-AS1	0.382	3.1×10^{-04}	0.043
TIPARP-AS1	0.336	3.1×10^{-04}	0.043
PRAF2	0.276	3.1×10^{-04}	0.043
LRRC23	0.25	3.1×10^{-04}	0.043
NR2C2AP	0.215	3.1×10^{-04}	0.043
RPIA	0.204	3.1×10^{-04}	0.043
PPOX	0.2	3.2×10^{-04}	0.043
GIT1	0.179	3.2×10^{-04}	0.043
DDX39A	0.171	3.1×10^{-04}	0.043
NFKBIB	0.149	3.2×10^{-04}	0.043
PNCK	0.773	3.2×10^{-04}	0.043
GOLGA2P5	0.416	3.2×10^{-04}	0.043
RPL31P34	0.372	3.2×10^{-04}	0.044
SYNE4	0.687	3.4×10^{-04}	0.044
SCAT2	0.441	3.3×10^{-04}	0.044
TMEM231	0.357	3.4×10^{-04}	0.044
LRRC75A	0.356	3.3×10^{-04}	0.044
PPP1R3F	0.318	3.4×10^{-04}	0.044
RABL2A	0.312	3.4×10^{-04}	0.044
OLMALINC	0.31	3.4×10^{-04}	0.044

SNHG30	0.271	3.3×10^{-04}	0.044
MZF1	0.232	3.4×10^{-04}	0.044
SYNGR2	0.183	3.3×10^{-04}	0.044
TBCB	0.145	3.4×10^{-04}	0.044
ABHD12	0.166	3.4×10^{-04}	0.044
CST2	0.516	3.6×10^{-04}	0.045
HOXB6	0.488	3.5×10^{-04}	0.045
SYCE2	0.469	3.5×10^{-04}	0.045
NEBL-AS1	0.381	3.6×10^{-04}	0.045
GABARAPL1	0.282	3.6×10^{-04}	0.045
TROAP	0.27	3.6×10^{-04}	0.045
LINC00847	0.232	3.6×10^{-04}	0.045
GPR89B	0.193	3.6×10^{-04}	0.045
COA5	0.186	3.5×10^{-04}	0.045
SAMD1	0.158	3.6×10^{-04}	0.045
CC2D1A	0.154	3.6×10^{-04}	0.045
USF1	0.149	3.6×10^{-04}	0.045
AIFM1	0.146	3.5×10^{-04}	0.045
USP20	0.143	3.6×10^{-04}	0.045
LINC01290	0.298	3.7×10^{-04}	0.045
PLD3	0.229	3.7×10^{-04}	0.045
SLC25A4	0.351	3.7×10^{-04}	0.045
TRAPPC6A	0.302	3.9×10^{-04}	0.047
ZNF672	0.162	3.9×10^{-04}	0.047
PPP1R26-AS1	0.386	4.0×10^{-04}	0.047
DYNLT2B	0.365	4.0×10^{-04}	0.047
HHEX	0.336	4.0×10^{-04}	0.047
LINC02019	0.285	4.0×10^{-04}	0.047
CEBPB-AS1	0.322	4.0×10^{-04}	0.048
FUNDC2	0.147	4.0×10^{-04}	0.048
RSPH1	0.493	4.1×10^{-04}	0.048
SEPTIN3	0.504	4.2×10^{-04}	0.048
REXO5	0.266	4.1×10^{-04}	0.048

H1-10	0.228	4.1×10^{-04}	0.048
AJM1	0.312	4.3×10^{-04}	0.049
JTB	0.244	4.3×10^{-04}	0.049
PRKAA2	0.75	4.5×10^{-04}	0.050
MYEOV	0.652	4.4×10^{-04}	0.050
FMNL1-AS1	0.443	4.4×10^{-04}	0.050
MISP3	0.378	4.4×10^{-04}	0.050
LAGE3	0.242	4.4×10^{-04}	0.050
TRIM52-AS1	0.239	4.4×10^{-04}	0.050
CKS1B	0.227	4.4×10^{-04}	0.050
SMDT1	0.174	4.5×10^{-04}	0.050
STK33	0.595	4.5×10^{-04}	0.050
PIMREG	0.317	4.5×10^{-04}	0.050
ISL2	0.366	4.5×10^{-04}	0.050

Tabla 2 Anexo. Los 101 genes diferenciales infraexpresados con FDR < 0.05.

GENES	fold change	pvalue	FDR
SPRR2G	-1.312	1.3×10^{-6}	0.006
EPGN	-1.050	2.1×10^{-6}	0.007
WFDC12	-1.154	2.7×10^{-6}	0.008
LCE3E	-1.149	3.6×10^{-6}	0.009
SPRR5	-0.840	4.0×10^{-6}	0.009
SLEAR	-0.634	6.3×10^{-6}	0.009
GJB2	-0.647	7.3×10^{-6}	0.009
SPRR2F	-1.214	6.0×10^{-6}	0.009
DSG1	-1.198	7.6×10^{-6}	0.009
KRT17P1	-0.472	9.7×10^{-6}	0.011
GSDMA	-0.816	1.1×10^{-5}	0.012
CRCT1	-1.030	1.6×10^{-5}	0.015
PRR9	-1.095	1.8×10^{-5}	0.016
GNAI1	-0.299	2.8×10^{-5}	0.019
TUBB2A	-0.370	3.0×10^{-5}	0.020
LINC01740	-0.313	3.7×10^{-5}	0.020

ELL2P1	-0.410	3.3×10^{-5}	0.020
PNPLA1	-0.735	3.7×10^{-5}	0.020
S100A12	-0.793	3.5×10^{-5}	0.020
UBQLN1P1	-0.294	4.2×10^{-5}	0.022
DSC2	-0.464	4.3×10^{-5}	0.022
HEPHL1	-0.889	4.3×10^{-5}	0.022
RAB38	-0.410	4.4×10^{-5}	0.022
POU3F1	-0.703	5.0×10^{-5}	0.022
SPRR2B	-1.104	4.6×10^{-5}	0.022
PPARD	-0.245	5.5×10^{-5}	0.022
SPRR2E	-1.007	5.5×10^{-5}	0.022
ADAP2	-0.300	5.9×10^{-5}	0.023
TCAF2	-0.302	6.2×10^{-5}	0.023
LINC02323	-0.518	6.6×10^{-5}	0.023
FABP5P1	-0.570	6.5×10^{-5}	0.023
PLA2G4E	-0.783	7.0×10^{-5}	0.023
RNF222	-0.793	6.9×10^{-5}	0.023
LCE3A	-0.952	7.0×10^{-5}	0.023
KRT6C	-0.959	6.5×10^{-5}	0.023
LCE3D	-1.030	6.6×10^{-5}	0.023
RN7SKP112	-0.163	7.5×10^{-5}	0.024
FABP5P2	-0.495	8.3×10^{-5}	0.026
LINC01280	-0.458	8.5×10^{-5}	0.026
KRT14	-0.664	8.9×10^{-5}	0.026
RN7SL252P	-0.131	9.0×10^{-5}	0.026
MTCYBP17	-0.169	9.1×10^{-5}	0.026
PI3	-0.873	9.1×10^{-5}	0.026
HES5	-0.536	9.2×10^{-5}	0.026
SPRR2D	-0.825	9.8×10^{-5}	0.028
DNAJB8	-0.563	1.0×10^{-4}	0.028
TGM1	-0.707	1.0×10^{-4}	0.028
S100A7A	-1.052	1.0×10^{-4}	0.028
ELL2	-0.256	1.1×10^{-4}	0.029

WFDC5	-0.639	1.1×10^{-4}	0.029
PLA2G4D	-0.772	1.1×10^{-4}	0.029
SLC37A2	-0.344	1.2×10^{-4}	0.030
RDH12	-0.791	1.2×10^{-4}	0.030
KRTDAP	-1.027	1.2×10^{-4}	0.030
IL1F10	-0.665	1.2×10^{-4}	0.030
EFNA5	-0.415	1.3×10^{-4}	0.030
LINC01797	-0.492	1.3×10^{-4}	0.030
SCNN1D	-0.531	1.4×10^{-4}	0.030
F11	-0.726	1.4×10^{-4}	0.030
SPRR2A	-0.873	1.3×10^{-4}	0.030
SPRR2C	-0.906	1.3×10^{-4}	0.030
KRT16P3	-0.593	1.4×10^{-4}	0.030
FTMT	-0.241	1.5×10^{-4}	0.032
PON1	-0.683	1.7×10^{-4}	0.033
OTOP3	-0.906	1.7×10^{-4}	0.033
CTNNBIP1	-0.175	1.9×10^{-4}	0.033
KRT17	-0.441	1.8×10^{-4}	0.033
DYDC1	-0.479	1.7×10^{-4}	0.033
SPINK7	-0.978	1.8×10^{-4}	0.033
DEFB4A	-1.001	1.8×10^{-4}	0.033
KRT1	-1.225	1.8×10^{-4}	0.033
IL36RN	-0.740	2.1×10^{-4}	0.036
MICALL1	-0.193	2.2×10^{-4}	0.038
KRT82	-0.499	2.3×10^{-4}	0.038
CADM2-AS2	-0.132	2.7×10^{-4}	0.040
LIPG	-0.523	2.7×10^{-4}	0.040
KHSRPP1	-0.542	2.6×10^{-4}	0.040
NIPAL1	-0.368	2.7×10^{-4}	0.040
DAAM1	-0.253	2.8×10^{-4}	0.041
IVL	-0.689	2.8×10^{-4}	0.041
PLD2	-0.203	2.8×10^{-4}	0.041
ZNF101P1	-0.408	2.8×10^{-4}	0.041

SPRR4	-0.855	2.8×10^{-4}	0.041
MINDY4B	-0.423	2.9×10^{-4}	0.041
TENM2	-0.631	2.9×10^{-4}	0.041
GPR78	-0.655	3.2×10^{-4}	0.043
OR51A5P	-0.140	3.4×10^{-4}	0.044
SLC22A23	-0.274	3.3×10^{-4}	0.044
CLCA2	-0.473	3.4×10^{-4}	0.044
FBLIM1	-0.274	3.5×10^{-4}	0.044
DMKN	-0.518	3.6×10^{-4}	0.044
IL36B	-0.675	3.7×10^{-4}	0.045
LNCBRM	-0.598	3.8×10^{-4}	0.046
DKKL1P1	-0.243	4.1×10^{-4}	0.047
IFFO2	-0.317	4.2×10^{-4}	0.048
CFAP100-DT	-0.412	4.2×10^{-4}	0.048
CNFN	-0.659	4.3×10^{-4}	0.049
FAM83G	-0.262	4.4×10^{-4}	0.049
MKI67P1	-0.566	4.4×10^{-4}	0.049
SLURP2	-0.611	4.4×10^{-4}	0.049
CPA4	-0.689	4.5×10^{-4}	0.049

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	PORTADA
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	6-8
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	13-22
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	27
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	29-30
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	30
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	31-32
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	32
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	31-32
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	31
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	31-32
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	31-32

Section and Topic	Item #	Checklist item	Location where item is reported
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	31-32
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	34-37 39-46
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	34-37
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	39-46
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	49
	23b	Discuss any limitations of the evidence included in the review.	59-60
	23c	Discuss any limitations of the review processes used.	61
	23d	Discuss implications of the results for practice, policy, and future research.	50-60
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	

Section and Topic	Item #	Checklist item	Location where item is reported
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. This work is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

Declaración de uso de Inteligencia Artificial (IA) en la elaboración del TFG

Herramienta:

ChatGPT 4.0 (OpenAI)

Funciones:

La herramienta se utilizó para traducir textos entre español e inglés, manteniendo un estilo académico adecuado, así como para mejorar la redacción mediante el uso de un vocabulario más amplio y la ortografía.

Prompts utilizados:

- “Traduce este texto del inglés al español”
- “Reescribe esta parte con diferentes sinónimos y mantenga el lenguaje profesional y adecuado para un trabajo científico.”
- “Pásame esta introducción al español sin perder el tono formal del original en inglés.”
- “Puedes mejorar el estilo de esta frase en español para que suene más académica?”
- “Cuidar la ortografía”

ANEXO 2

" Alteraciones genéticas asociadas a la invasión angiolinfática en el carcinoma de células escamosas de cabeza y cuello. Metaanálisis de expresión génica"

Antonio Belli¹, Marta Serna García¹, Nicla Flacco¹

¹Universidad Europea de Valencia, Facultad de Ciencias de la Salud, Departamento de Odontología, Valencia, España

Correspondence:
Universidad Europea de Valencia
Paseo Alameda, 7
46010 – Valencia, Spain
B3ellianyonio@gmail.com

Introducción:

La invasión angiolinfática (IAL) representa un marcador pronóstico clave en el carcinoma escamoso de cabeza y cuello (HNSCC), favoreciendo la diseminación metastásica y asociándose con peor supervivencia. La identificación de firmas moleculares asociadas a IAL puede contribuir al desarrollo de nuevos biomarcadores clínicos.

Métodos:

Se analizaron datos transcriptómicos de 358 muestras de HNSCC extraídas del consorcio TCGA (Firehose Legacy), clasificadas en función de la presencia (n=126) o ausencia (n=232) de IAL. Se realizó un análisis de expresión diferencial mediante los paquetes edgeR y limma ($FDR \leq 0,05$), seguido de un análisis de enriquecimiento funcional con DAVID. Paralelamente, se evaluaron asociaciones clínicas (edad, sexo, tabaco, alcohol) mediante pruebas χ^2 , test exacto de Fisher, t de Student y U de Mann–Whitney.

Resultados:

Se identificaron 375 genes diferencialmente expresados entre tumores con y sin IAL (274 sobreexpresados, 101 infraexpresados). Entre los genes más relevantes se encuentran KRT18, GADD45G, SIGIRR (sobreexpresados) y S100A12, DSG1, GSDMA, WFDC12 (infraexpresados). Los procesos biológicos enriquecidos se relacionaron con la exocitosis, adhesión celular, inmunidad epitelial y respuesta inflamatoria. El tabaquismo fue la única variable clínica significativamente asociada a la presencia de IAL ($p < 0,05$).

Conclusiones:

La invasión angiolinfática en HNSCC se asocia a una firma molecular específica y a la desregulación de rutas funcionales implicadas en la progresión tumoral. El perfil transcriptómico diferencial hallado podría ser útil para el desarrollo de biomarcadores pronósticos y estrategias terapéuticas dirigidas en este subtipo de cáncer. El tabaquismo podría influir en estos mecanismos moleculares, reforzando su impacto clínico más allá del rol etiológico.

Keywords: Carcinoma escamoso de cabeza y cuello, Invasión angiolinfática, Análisis transcriptómico, Genes diferencialmente expresados, Enriquecimiento funcional, Biomarcadores, TCGA

Introducción:

El carcinoma escamoso de cabeza y cuello (HNSCC) es un grupo heterogéneo de neoplasias malignas que se originan en el epitelio escamoso de la mucosa oral, faríngea y laríngea, y se caracteriza por su agresividad clínica y elevada mortalidad (1). La localización anatómica del tumor influye notablemente en el pronóstico, con tasas de supervivencia a cinco años que oscilan entre el 14.8 % en hipofaringe y el 74.0 % en laringe (2).

A nivel mundial, el HNSCC representa aproximadamente el 90 % de los tumores malignos en esta región, con una incidencia estimada de 890.000 casos nuevos y 450.000 muertes por año (3). En países en desarrollo, su incidencia ha aumentado debido al consumo de tabaco, alcohol y nuez de areca (4). En contraste, en países desarrollados, predomina la forma asociada al virus del papiloma humano (VPH) (3).

El diagnóstico de HNSCC en estadios avanzados limita las opciones terapéuticas y reduce la supervivencia global (4). La heterogeneidad molecular tumoral dificulta la respuesta uniforme al tratamiento y contribuye a la resistencia terapéutica (5). Esta heterogeneidad se presenta tanto entre tumores de diferentes pacientes como dentro de un mismo tumor (6).

Uno de los factores histopatológicos con mayor impacto pronóstico es la invasión angiolinfática (IAL), definida como la presencia de células tumorales dentro de vasos linfáticos o sanguíneos revestidos por endotelio (7). La IAL favorece la diseminación metastásica y se asocia con una menor supervivencia global (8). Además, se ha identificado como marcador independiente de metástasis ganglionar (9).

Desde el punto de vista molecular, la IAL se ha relacionado con alteraciones en vías como PI3K/AKT, remodelación de la matriz extracelular y linfangiogénesis (10). Estas vías contribuyen a la capacidad invasiva del tumor mediante la interacción con el microambiente tumoral (11). En este entorno, participan fibroblastos, células inmunes, componentes vasculares y elementos de señalización que regulan el desarrollo tumoral. (12).

El análisis transcriptómico mediante RNA-Seq ha permitido identificar genes diferencialmente expresados (DEGs) implicados en estos procesos, facilitando el descubrimiento de biomarcadores (13). Herramientas como DAVID permiten relacionar dichos genes con rutas funcionales clave para la progresión tumoral (14).

En este contexto, el presente estudio tiene como objetivo identificar genes diferencialmente expresados entre tumores de HNSCC con y sin IAL, analizar las rutas biológicas asociadas y explorar posibles asociaciones clínicas, con el fin de aportar nuevos biomarcadores

Material y métodos:

Se realizó un estudio observacional y retrospectivo basado en el análisis de datos transcriptómicos públicos disponibles en el consorcio The Cancer Genome Atlas (TCGA), específicamente del conjunto de datos Firehose Legacy correspondiente a pacientes diagnosticados con carcinoma escamoso de cabeza y cuello (HNSCC) (The Cancer Genome Atlas, <http://tcga.cancer.gov>). Se seleccionaron un total de 358 muestras tumorales, previamente clasificadas según la presencia (n=126) o ausencia (n=232) de invasión angiolinfática (IAL), conforme a los datos clínicos disponibles en la plataforma.

Los datos de expresión génica, previamente normalizados a TPM (Transcripts Per Million), fueron procesados con los paquetes edgeR y limma del entorno Bioconductor para la detección de genes diferencialmente expresados (DEGs). Se utilizó un enfoque estadístico basado en modelos lineales moderados y estimación empírica bayesiana, lo que permitió controlar la dispersión de los datos y aumentar la robustez del análisis estadístico (15).

Se consideraron significativos aquellos genes con un valor de FDR ≤ 0.05 , tras corrección por pruebas múltiples mediante el método de Benjamini-Hochberg (13). Los DEGs identificados fueron posteriormente analizados mediante la herramienta DAVID (Database for Annotation, Visualization and Integrated Discovery), con el objetivo de identificar rutas funcionales y procesos biológicos sobrerrepresentados, en base a las ontologías GO (Gene Ontology) y las rutas KEGG (Kyoto Encyclopedia of Genes and Genomes), considerando un umbral de significación estadística de $p < 0.05$ (14).

En paralelo, se evaluaron posibles asociaciones clínicas entre la presencia de IAL y variables epidemiológicas como sexo, edad, consumo de tabaco y consumo de alcohol. Para ello, se empleó el software estadístico Jamovi, basado en R (The jamovi project). Las variables categóricas se analizaron mediante pruebas de independencia χ^2 o, en su defecto, el test exacto de Fisher, aplicado cuando las frecuencias esperadas en alguna celda eran inferiores a 5. Para la variable continua edad, se aplicó inicialmente la prueba de normalidad de Shapiro–Wilk; en función del resultado, se utilizó la prueba t de Student para muestras independientes o la prueba no paramétrica de Mann–Whitney U. En todos los análisis estadísticos se adoptó un nivel de significación de $p < 0.05$.

Resultados:

El análisis de expresión génica diferencial entre los tumores con y sin invasión angiolímfática permitió identificar un total de 375 genes significativamente alterados. De estos, 274 genes se encontraban sobreexpresados en los tumores con invasión y 101 infraexpresados, aplicando un umbral de significación de FDR ≤ 0.05 . La mayoría de los genes diferencialmente expresados mostró cambios moderados en el nivel de expresión, aunque con alta significación estadística. Entre los genes con mayor sobreexpresión se destacaron KRT18, GADD45G, SIGIRR, PRR15 y CFAP43, con valores de fold change positivos y FDRs inferiores a 0.01. En contraste, los genes con mayor infraexpresión incluyeron S100A12, DSG1, GSDMA, WFDC12 y SPRR2G, todos ellos implicados en funciones epiteliales, inmunológicas o estructurales. Las Tablas 1 y 2 presentan los 20 genes más significativamente sobreexpresados e infraexpresados, respectivamente, ordenados por su fold change.

Tabla 1. Los 20 genes diferenciales con FDR < 0.05 con mayor sobreexpresión (*Fold Change* positivo)

GENES	Nombre completo	fold change	Pvalue	FDR
FAM174B	family with sequence similarity 174 member B	0.462	5.1×10^{-8}	0.002
ARNILA	androgen receptor negatively regulated lncRNA	0.550	1.4×10^{-7}	0.002

NPDC1	neural proliferation, differentiation and control 1	0.490	2.0×10^{-7}	0.002
RHPN1	rhophilin Rho GTPase binding protein 1	0.673	3.7×10^{-7}	0.003
DNASE2	deoxyribonuclease 2, lysosomal	0.268	6.5×10^{-7}	0.005
COX14	cytochrome c oxidase assembly factor COX14	0.264	8.1×10^{-7}	0.005
FAM50A	family with sequence similarity 50 member A	0.316	1.0×10^{-6}	0.006
PKN1	protein kinase N1	0.412	1.6×10^{-6}	0.007
P4HTM	prolyl 4-hydroxylase, transmembrane	0.449	2.1×10^{-6}	0.007
PCBD1	pterin-4 alpha-carbinolamine dehydratase 1	0.284	2.7×10^{-6}	0.008
PRR15	proline rich 15	0.762	3.5×10^{-6}	0.009
PRKAG2-AS1	PRKAG2 antisense RNA 1	0.626	3.8×10^{-6}	0.009
CATSPERG	cation channel sperm associated auxiliary subunit gamma	0.393	3.7×10^{-6}	0.009
CFAP43	cilia and flagella associated protein 43	0.523	5.9×10^{-6}	0.009
KRT18	keratin 18	0.481	6.1×10^{-6}	0.009
GADD45G	growth arrest and DNA damage inducible gamma	0.473	6.6×10^{-6}	0.009
CCNO	cyclin O	0.467	6.1×10^{-6}	0.009
TSC22D3	TSC22 domain family member 3	0.353	7.4×10^{-6}	0.009
SIGIRR	single Ig and TIR domain containing	0.322	7.3×10^{-6}	0.009
ARRDC2	arrestin domain containing 2	0.288	6.7×10^{-6}	0.009

*Fold Change se refiere al grupo con invasión en comparación con sin invasión angirolinfática.

Tabla 2. Los 20 genes diferenciales con FDR <0,05 con menor infraexpresión (*Fold Change* negativo)

GENES	Nombre completo	fold change	Pvalue	FDR
SPRR2G	small proline rich protein 2G	-1.312	1.3×10^{-6}	0.006
EPGN	epithelial mitogen	-1.050	2.1×10^{-6}	0.007

WFDC12	WAP four-disulfide core domain 12	-1.154	2.7×10^{-6}	0.008
LCE3E	late cornified envelope 3E	-1.149	3.6×10^{-6}	0.009
SPRR5	small proline rich protein 5	-0.840	4.0×10^{-6}	0.009
SLEAR	STAT1 regulated ILF2 complex interacting lncRNA	-0.634	6.3×10^{-6}	0.009
GJB2	gap junction protein beta 2	-0.647	7.3×10^{-6}	0.009
SPRR2F	small proline rich protein 2F	-1.214	6.0×10^{-6}	0.009
DSG1	desmoglein 1	-1.198	7.6×10^{-6}	0.009
KRT17P1	keratin 17 pseudogene 1	-0.472	9.7×10^{-6}	0.011
GSDMA	gasdermin A	-0.816	1.1×10^{-5}	0.012
CRCT1	cysteine rich C-terminal 1	-1.030	1.6×10^{-5}	0.015
PRR9	proline rich 9	-1.095	1.8×10^{-5}	0.016
GNAI1	G protein subunit alpha i1	-0.299	2.8×10^{-5}	0.019
TUBB2A	tubulin beta 2A class IIa	-0.370	3.0×10^{-5}	0.020
LINC01740	long intergenic non-protein coding RNA 1740	-0.313	3.7×10^{-5}	0.020
ELL2P1	elongation factor for RNA polymerase II 2 pseudogene 1	-0.410	3.3×10^{-5}	0.020
PNPLA1	patatin like phospholipase domain containing 1	-0.735	3.7×10^{-5}	0.020
S100A12	S100 calcium binding protein A12	-0.793	3.5×10^{-5}	0.020
UBQLN1P1	ubiquilin 1 pseudogene 1	-0.294	4.2×10^{-5}	0.022

*Fold Change se refiere al grupo con invasión en comparación con sin invasión angirolinfática.

Posteriormente, se realizó un análisis de enriquecimiento funcional utilizando el software DAVID (Database for Annotation, Visualization and Integrated Discovery), aplicado a los genes diferencialmente expresados con un $FDR \leq 0.05$. En el conjunto de genes sobreexpresados se identificaron diez procesos biológicos significativamente enriquecidos ($p < 0.05$), principalmente relacionados con la regulación de la exocitosis, el movimiento ciliar, los cambios mitocondriales apoptóticos, el ensamblaje del axonema y la exportación de mRNA desde el núcleo. Estos hallazgos sugieren un incremento en la capacidad de secreción, remodelación intracelular y reorganización del citoesqueleto, potencialmente asociados a una mayor agresividad tumoral. En el conjunto de genes infraexpresados, el análisis con DAVID reveló nueve procesos biológicos enriquecidos, como la adhesión célula-célula, la respuesta antimicrobiana, la activación de la respuesta a lipopolisacáridos y la catálisis de fosfolípidos. La pérdida de expresión de estos genes puede reflejar una desregulación de las funciones epiteliales defensivas y una mayor fragilidad de las uniones celulares, favoreciendo

la penetración tumoral en vasos linfáticos. Las Tablas 3 y 4 recogen los procesos significativamente enriquecidos junto con el número de genes implicados y sus identificadores.

Tabla 3. Los procesos biológicos con menor Pvalor < 0.05 enriquecidos según el análisis DAVID.

Term	Proceso biológico	Count	PValue	Genes
GO:0017157	regulation of exocytosis	4	0.004	RABL2A, RABL2B, HAP1, VAMP2
GO:0003341	cilium movement	4	0.011	DNAAF1, ZMYND10, CFAP43, PIERCE1
GO:0008637	apoptotic mitochondrial changes	3	0.013	AIFM2, JTB, SLC25A4
GO:0042733	embryonic digit morphogenesis	4	0.020	FBXW4, SHH, FLVCR1, TMEM231
GO:0006783	heme biosynthetic process	3	0.025	SLC25A38, FLVCR1, PPOX
GO:0006406	mRNA export from nucleus	4	0.027	CHTOP, HHEX, DDX39A, RBM8A
GO:0033617	mitochondrial cytochrome c oxidase assembly	3	0.034	UQCC6, COX14, COA5
GO:1903008	organelle disassembly	2	0.041	KIF9, PLAAT3
GO:0008089	anterograde axonal transport	3	0.041	DTNBP1, KIF1A, HAP1
GO:0035082	axoneme assembly	3	0.046	RSPH1, DNAAF1, PIERCE1

* Los procesos biológicos resaltados en verde fueron seleccionados para su análisis detallado y la discusión en el presente estudio.

Tabla 4. Los procesos biológicos con pvalor <0.05inhibidos según el análisis DAVID.

Term	Proceso biológico	Court	PValue	Genes
GO:0009395	phospholipid catabolic process	3	0.001	PLA2G4E, LIPG, PLD2
GO:0019731	antibacterial humoral response	4	0.002	WFDC12, PI3, SPRR2A, WFDC5
GO:0002437	inflammatory response to antigenic stimulus	3	0.004	IL36B, IL36RN, IL1F10
GO:0008366	axon ensheathment	2	0.021	POU3F1, PPARD
GO:0036149	phosphatidylinositol acyl-chain remodeling	2	0.021	PLA2G4D, PLA2G4E
GO:0042742	defense response to bacterium	4	0.022	WFDC12, GSDMA, S100A12, DEFB4A
GO:0071222	cellular response to lipopolysaccharide	4	0.036	IL36B, IL36RN, PPARD, IL1F10

GO:0098609	cell-cell adhesion	4	0.037	TENM2, FBLIM1, DSG1, DSC2
GO:0031640	killing of cells of another organism	3	0.045	LCE3A, S100A12, DEFB4A

* Los procesos biológicos resaltados en verde fueron seleccionados para su análisis detallado y la discusión en el presente estudio.

En el análisis de variables clínicas, se evaluó la relación entre el historial de tabaquismo y la presencia de invasión angiolinfática. Se analizaron 351 pacientes divididos en categorías de fumador actual, exfumador reciente, exfumador de más de 15 años, nunca fumador y casos sin datos. La mayor proporción de casos con invasión angiolinfática se encontró en los grupos de fumadores activos y exfumadores recientes. La prueba de Chi-cuadrado mostró una asociación estadísticamente significativa ($p = 0.033$), que fue confirmada por el test exacto de Fisher ($p = 0.015$). Estos resultados indican que el tabaquismo reciente o persistente se asocia con una mayor probabilidad de presentar invasión vascular, sugiriendo un efecto modulador del hábito tabáquico sobre los mecanismos tumorales de diseminación.

Respecto al consumo de alcohol, se evaluaron 352 pacientes con historial documentado. Aunque se observó una tendencia hacia una mayor frecuencia de invasión angiolinfática en consumidores, las diferencias no alcanzaron significación estadística ($p = 0.224$, prueba de Chi-cuadrado; $p = 0.232$, test exacto de Fisher). El odds ratio estimado fue de 1.35, con un intervalo de confianza del 95 % entre 0.83 y 2.18, lo cual sugiere un posible aumento del riesgo no concluyente en esta cohorte.

La variable edad se analizó como variable continua. La edad media en el grupo sin invasión fue de 62.1 años, mientras que en el grupo con invasión fue de 59.1 años. La diferencia fue significativa según la prueba t de Student ($p = 0.023$), pero no fue confirmada por la prueba no paramétrica de Mann–Whitney U ($p = 0.059$), lo que indica una posible tendencia, aunque no robusta desde el punto de vista estadístico. Este patrón podría reflejar una mayor susceptibilidad a la invasión en pacientes más jóvenes, aunque los datos deben interpretarse con cautela.

En cuanto al sexo, se observó una mayor proporción de invasión angiolinfática en varones respecto a mujeres. Esta diferencia fue significativa en la prueba de Chi-cuadrado ($p = 0.029$) y en el test exacto de Fisher ($p = 0.035$). El odds ratio calculado fue de 1.77, con un intervalo de confianza del 95 % entre 1.06 y 2.95, lo que indica que los pacientes varones presentan un riesgo 77 % mayor de presentar invasión angiolinfática. Este hallazgo podría estar vinculado a diferencias biológicas, moleculares o de exposición a factores de riesgo, aunque el desequilibrio de género en la muestra debe considerarse como posible factor de confusión.

Table 5: Association between clinical variables and lymphovascular invasion

Variab	Category	N (no LVI)	N (LVI)	p-val. (χ^2/Fisher)	OR (95% CI)
Sex	Male	160	99	0.029 / 0.035	1.77 (1.06–2.95)

	Female	99	27		
Age	Mean ($\pm$ SD)	62.1 $\pm$ 12.0	59.1 $\pm$ 11.6	0.023 / 0.059	
Tobacco	Current smoker	70	49	0.033 / 0.015	
	Former ≤ 15 years	21	21		
	Former > 15 years	76	49		
	Never smoker	61	23		
Alcohol	Consumer	152	90	0.224 / 0.232	1.35 (0.83–2.18)
	Non-consumer	80	33		

Discusión:

KRT18

KRT18 codifica una queratina tipo I presente en epitelios simples, cuya función principal es mantener la integridad estructural celular y la estabilidad frente a estrés mecánico o químico (16). Su implicación en procesos de señalización, organización del citoesqueleto y polaridad celular le otorgan un papel activo en la regulación del fenotipo epitelial. En nuestro estudio, KRT18 se encontró significativamente sobreexpresado en tumores con invasión angiolinfática, lo que sugiere su participación en la transición hacia un estado más migratorio y desorganizado, favoreciendo la penetración tumoral en vasos linfáticos. En otros tipos tumorales, como el hepatocarcinoma y el cáncer de mama, su sobreexpresión se ha vinculado con mayor capacidad de invasión y mal pronóstico, posiblemente por su efecto sobre la adhesión intercelular y la plasticidad estructural de las células tumorales (17,18). Estos hallazgos apoyan la hipótesis de que KRT18 no es un mero marcador de diferenciación, sino un componente funcional que puede facilitar la progresión tumoral en contextos con alta agresividad biológica, como los casos con invasión angiolinfática en HNSCC.

GADD45G

GADD45G (growth arrest and DNA damage-inducible gamma) forma parte de una familia de genes reguladores del ciclo celular implicados en la respuesta al estrés genotóxico, la detención del ciclo en G2/M y la modulación de la apoptosis. En modelos tumorales, su reexpresión ha demostrado suprimir la progresión maligna mediante la activación de vías de señalización asociadas a la estabilidad epitelial, como el aumento de E-cadherina y la inhibición de la transición epitelio-mesénquima (19). En cánceres escamosos, esta actividad se ha vinculado a una menor capacidad de migración e

invasión. En nuestro análisis, GADD45G se encontró significativamente sobreexpresado en tumores con invasión angiolinfática, un hallazgo que puede parecer contradictorio con su rol supresor, pero que podría interpretarse como una respuesta compensatoria frente a condiciones de alta agresividad biológica. Algunos autores han propuesto que esta sobreexpresión podría reflejar un intento de contención celular ante un microambiente pro-invasivo o un marcador de inestabilidad genética reactiva (20). Estos resultados sugieren un papel más complejo para GADD45G en la dinámica tumoral del HNSCC, posiblemente dependiente del contexto celular y del grado de diferenciación del tumor.

SIGIRR

SIGIRR (Single Immunoglobulin IL-1 Receptor-Related molecule), también conocido como TIR8, pertenece a la familia de receptores tipo Toll (TLR) e interleucina-1 (IL-1R), y actúa como un inhibidor negativo de estas vías inflamatorias, especialmente a través de la modulación de NF- κ B y JNK (21). En el contexto del carcinoma escamoso de cabeza y cuello (HNSCC), se ha observado que SIGIRR puede contribuir a limitar la inflamación crónica del microambiente tumoral, promoviendo una respuesta inmunitaria más equilibrada y favoreciendo la infiltración de linfocitos T citotóxicos. Además, en tumores positivos para VPH, la expresión elevada de SIGIRR se ha asociado con una mejor activación del sistema inmune adaptativo y una menor expresión de genes inmunosupresores (22). En nuestro estudio, la sobreexpresión de SIGIRR en tumores con invasión angiolinfática podría interpretarse como una respuesta compensatoria frente a un entorno proinflamatorio y altamente activo a nivel inmunológico. Este patrón sugiere que SIGIRR podría participar indirectamente en la regulación de procesos de remodelado vascular y linfangiogénesis, al limitar la activación excesiva de rutas inflamatorias que favorecen la progresión tumoral. Aunque su papel como supresor está documentado, la sobreexpresión observada en casos con IAL plantea la hipótesis de un rol dual dependiente del contexto celular e inmunológico, que debería ser explorado en futuros estudios funcionales.

DSG1

DSG1 (desmogleína 1) es una proteína de la familia de las cadherinas desmosomales que forma parte de los complejos de adhesión intercelular en epitelios escamosos estratificados, incluyendo la mucosa oral y la epidermis (23). Su función es crítica para la cohesión epitelial y la integridad estructural del tejido. En distintos modelos de cáncer, se ha demostrado que la pérdida de DSG1 favorece un fenotipo más agresivo, caracterizado por la pérdida de diferenciación celular, la formación de estructuras de invasión como invadopodios, y la activación de rutas como EGFR/ERK, asociadas con la migración e infiltración tumoral (24). En nuestro análisis, la infraexpresión de DSG1 en tumores con invasión angiolinfática refuerza su posible papel como marcador de integridad epitelial. Su pérdida podría facilitar la desorganización tisular y la transición hacia un fenotipo más móvil y menos cohesivo, características necesarias para la diseminación tumoral a través de vasos linfáticos. Además, el patrón de expresión observado sugiere que DSG1 podría desempeñar un papel clave en la regulación del comportamiento invasivo del HNSCC, no solo como elemento estructural, sino también como modulador de señalización tumoral.

GSDMA

GSDMA (Gasdermin A) pertenece a la familia de las gasderminas, un conjunto de proteínas asociadas a la piroptosis, un tipo de muerte celular inflamatoria dependiente de caspasas. Además de esta función, GSDMA interviene en la homeostasis epitelial, la diferenciación celular y la regulación de la apoptosis, actuando como un supresor potencial de la progresión tumoral en epitelios escamosos (25). En condiciones fisiológicas, se expresa abundantemente en piel, mucosa oral y tracto digestivo. Diversos estudios han demostrado que su pérdida de expresión se asocia con un fenotipo más agresivo, resistencia a señales apoptóticas y mayor capacidad de invasión. En nuestro análisis, GSDMA aparece significativamente infraexpresado en los tumores con invasión angiolinfática. Este hallazgo sugiere que su ausencia podría facilitar la evasión de la muerte celular durante la migración y entrada a vasos linfáticos. Además, el análisis de enriquecimiento funcional indicó su implicación en procesos inmunes frente a bacterias y lipopolisacáridos, lo cual refuerza su papel como modulador del microambiente inmunológico. La alteración de estas funciones podría comprometer la inmunovigilancia local y facilitar la entrada de células tumorales al sistema linfático (26).

WFDC12

WFDC12 (WAP four-disulfide core domain 12) pertenece a una familia de proteínas caracterizadas por contener dominios WAP con funciones antiproteásicas e inmunomoduladoras. Su expresión se ha descrito en epitelios con funciones de defensa innata, y su alteración se ha vinculado a enfermedades inflamatorias e inmunomediadas (27). Aunque la evidencia directa en HNSCC es limitada, estudios en carcinomas escamosos de esófago han demostrado que la supresión de WFDC12 favorece la proliferación y migración celular, lo que sugiere un papel tumor-supresor conservado (28). En nuestro análisis, se observó una infraexpresión significativa de WFDC12 en tumores con invasión angiolinfática. Este hallazgo es coherente con su posible función como regulador del microambiente inmunológico y como barrera frente a la degradación de matriz extracelular. Además, el análisis de enriquecimiento funcional mostró su implicación en rutas relacionadas con la respuesta antimicrobiana y la inflamación, en conjunto con GSDMA y S100A12. Esta convergencia funcional sugiere la existencia de un eje de regulación epitelial común que podría estar comprometido en los tumores más agresivos.

S100A12

S100A12 (S100 calcium-binding protein A12) es una proteína secretada de la familia S100, implicada en la regulación de la inflamación, la respuesta inmune innata y la activación endotelial. Actúa como ligando de los receptores RAGE y TLR4, activando rutas proinflamatorias como NF- κ B y promoviendo la producción de citoquinas y especies reactivas de oxígeno (29). Aunque en ciertos tumores se ha asociado a mayor agresividad, su papel varía según el tipo y el entorno tumoral. En nuestro estudio, S100A12 resultó infraexpresado en tumores con invasión angiolinfática, lo cual sugiere una pérdida de funciones protectoras del microambiente inflamatorio y una posible facilitación de la diseminación linfática. Este hallazgo coincide con análisis bioinformáticos previos en HNSCC que vinculan la baja expresión de S100A12 con mayor grado histológico, invasión linfovascular y metástasis ganglionares (30). Por tanto, su infraexpresión podría reflejar un microambiente menos inmunovigilado, más permisivo para la progresión y el escape tumoral.

Relación entre la invasión angiolinfática y las variables clínicas

Nuestro análisis clínico reveló asociaciones significativas entre la presencia de invasión angiolinfática (IAL) y ciertas variables epidemiológicas del HNSCC. Se observó una mayor frecuencia de IAL en pacientes varones y en aquellos con antecedentes de tabaquismo activo o reciente, lo que refuerza la hipótesis de que algunos perfiles clínicos, además de predisponer al desarrollo tumoral, podrían también modular su comportamiento biológico invasivo. Aunque la edad mostró una tendencia hacia la significación y el consumo de alcohol no se asoció de forma estadísticamente significativa, estas variables podrían estar influidas por factores no controlados y requerirían mayor exploración en cohortes más amplias.

En particular, la asociación entre el sexo masculino y la IAL, aunque estadísticamente significativa, debe interpretarse con cautela debido al desequilibrio de género presente en la muestra. Un estudio de cohorte realizado en Corea del Sur confirmó que los varones presentan una incidencia significativamente mayor de HNSCC que las mujeres, incluso sin antecedentes de tabaco o alcohol (31). Asimismo, el sexo masculino se ha asociado con una peor supervivencia en tumores negativos para VPH, lo que podría reflejar una mayor susceptibilidad biológica más allá de los factores conductuales clásicos (32).

Por otro lado, el tabaquismo mostró una asociación sólida con la IAL. Se ha descrito que la exposición al tabaco puede inducir alteraciones moleculares que afectan genes reguladores de migración, adhesión e inflamación, lo que podría facilitar la diseminación linfovascular (31). En nuestro estudio, esta hipótesis se refuerza por la infraexpresión de genes como *DSG1*, *GSDMA* y *WFDC12* en pacientes con IAL, todos ellos implicados en la organización epitelial, la cohesión celular y la inmunidad local.

Conclusión:

La invasión angiolinfática en el carcinoma escamoso de cabeza y cuello se asoció a una firma génica diferencial compuesta por 375 genes, entre los que destacan *KRT18*, *GADD45G*, *SIGIRR*, *S100A12*, *DSG1*, *GSDMA* y *WFDC12*. El análisis funcional mostró una implicación significativa de procesos como la adhesión célula-célula, la exocitosis y la respuesta inmune, lo que refuerza su papel en la progresión tumoral. Estos resultados permitieron proponer dichos genes como posibles biomarcadores para la detección precoz de la invasión angiolinfática. Además, se identificaron asociaciones clínicas con el tabaquismo y el sexo, si bien esta última debe interpretarse con cautela debido al desequilibrio en la distribución de la muestra.

Bibliografia:

1. Shirima CA, Bleotu C, Spandidos DA, El-Naggar AK, Pircalabioru GG, Michalopoulos I. Epithelial-derived head and neck squamous tumourigenesis (Review). Vol. 52, *Oncology Reports*. Spandidos Publications; 2024.
2. Mores AL, Bonfim-Alves CG, López RVM, Rodrigues-Oliveira L, Palmier NR, Mariz BALA, et al. Prognostic Factors in Head and Neck Cancer: A Retrospective Cohort Study of 3052 Patients in Brazil. *Oral Dis*. 2024 Apr 1;
3. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Medical Sciences*. 2023 Jun 13;11(2):42.
4. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J*. 2022 Nov 11;233(9):780–6.
5. Mroz EA, Rocco JW. Intra-tumor heterogeneity in head and neck cancer and its clinical implications. Vol. 2, *World Journal of Otorhinolaryngology - Head and Neck Surgery*. KeAi Communications Co.; 2016. p. 60–7.
6. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015 Jan 29;517(7536):576–82.
7. Mascitti M, Togni L, Caponio VCA, Zhurakivska K, Bizzoca ME, Contaldo M, et al. Lymphovascular invasion as a prognostic tool for oral squamous cell carcinoma: a comprehensive review. Vol. 51, *International Journal of Oral and Maxillofacial Surgery*. Churchill Livingstone; 2022. p. 1–9.
8. Zhang J, Lin H, Jiang H, Jiang H, Xie T, Wang B, et al. A key genomic signature associated with lymphovascular invasion in head and neck squamous cell carcinoma. *BMC Cancer*. 2020 Mar 30;20(1).
9. Huang S, Zhu Y, Cai H, Zhang Y, Hou J. Impact of lymphovascular invasion in oral squamous cell carcinoma: A meta-analysis. Vol. 131, *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*. Mosby Inc.; 2021. p. 319-328.e1.
10. Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. Vol. 61, *Seminars in Cancer Biology*. Academic Press; 2020. p. 71–83.
11. Chen YQ, Kuo JC, Wei MT, Chen YC, Yang MH, Chiou A. Early stage mechanical remodeling of collagen surrounding head and neck squamous cell carcinoma spheroids correlates strongly with their invasion capability. *Acta Biomater*. 2019 Jan 15;84:280–92.
12. Arneth B. Tumor microenvironment. Vol. 56, *Medicina (Lithuania)*. MDPI AG; 2020.
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec 5;15(12).
14. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022 Jul 5;50(W1):W216–21.
15. Law CW, Alhamdoosh M, Su S, Smyth GK, Ritchie ME. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res*. 2016 Jun 17;5:1408.
16. Yan J, Yang A, Tu S. The relationship between keratin 18 and epithelial-derived tumors: as a diagnostic marker, prognostic marker, and its role in tumorigenesis. Vol. 14, *Frontiers in Oncology*. Frontiers Media SA; 2024.
17. Golob-Schwarzl N, Bettermann K, Mehta AK, Kessler SM, Unterluggauer J, Krassnig S, et al. High Keratin 8/18 Ratio Predicts Aggressive Hepatocellular Cancer Phenotype. *Transl Oncol*. 2019 Feb 1;12(2):256–68.
18. Ha SA, Lee YS, Kim HK, Yoo J, Kim S, Gong GH, et al. The prognostic potential of keratin 18 in breast cancer associated with tumor dedifferentiation, and the loss of estrogen and progesterone receptors. *Cancer Biomarkers*. 2011;10(5):219–31.

19. Zhan Q, Antinore MJ, Wang XW, Carrier F, Smith ML, Harris CC, et al. Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene*. 1999 May 6;18(18):2892–900.
20. Li T, Xu L, Teng J, Ma Y, Liu W, Wang Y, et al. GADD45G Interacts with E-cadherin to Suppress the Migration and Invasion of Esophageal Squamous Cell Carcinoma. *Dig Dis Sci*. 2020 Apr 1;65(4):1032–41.
21. Garlanda C, Anders HJ, Mantovani A. TIR8/SIGIRR: an IL-1R/TLR family member with regulatory functions in inflammation and T cell polarization. Vol. 30, *Trends in Immunology*. 2009. p. 439–46.
22. Yin G, Guo W, Wang R, Li N, Chen X, Zhang Y, et al. Analysis of the role of IL-1 family and related genes in head and neck squamous cell carcinoma. *Braz J Otorhinolaryngol*. 2025 Jan 1;91(1).
23. Wong MP, Cheang M, Yorlida E, Coldman A, Gilks CB, Huntsman D, et al. Loss of desmoglein 1 expression associated with worse prognosis in head and neck squamous cell carcinoma patients. *Pathology*. 2008;40(6):611–6.
24. Valenzuela-Iglesias A, Burks HE, Arnette CR, Yalamanchili A, Nekrasova O, Godsel LM, et al. Desmoglein 1 regulates invadopodia by suppressing EGFR/ERK signaling in an ERBIN-dependent manner. *Molecular Cancer Research*. 2019 May 1;17(5):1195–206.
25. Zou J, Zheng Y, Huang Y, Tang D, Kang R, Chen R. The Versatile Gasdermin Family: Their Function and Roles in Diseases. Vol. 12, *Frontiers in Immunology*. Frontiers Media S.A.; 2021.
26. Saeki N, Usui T, Aoyagi K, Kim DH, Sato M, Mabuchi T, et al. Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. *Genes Chromosomes Cancer*. 2009 Mar;48(3):261–71.
27. Kalinina P, Vorstandlechener V, Buchberger M, Eckhart L, Lengauer B, Golabi B, et al. The Whey Acidic Protein WFDC12 Is Specifically Expressed in Terminally Differentiated Keratinocytes and Regulates Epidermal Serine Protease Activity. *Journal of Investigative Dermatology*. 2021 May 1;141(5):1198-1206.e13.
28. Zhuang Y, Wang L, Ji C, Sun Y, Shao F. Construction of a novel pyroptosis-related prognostic model of esophageal square cell carcinoma and determination of the anti-tumor effect of WFDC12. *Funct Integr Genomics*. 2023 Jun 1;23(2).
29. Sionov RV. Leveling up the controversial role of neutrophils in cancer: When the complexity becomes entangled. Vol. 10, *Cells*. MDPI; 2021.
30. Bai G, Yue S, You Y, Zhao J, Tian Z, Wang Y, et al. An integrated bioinformatics analysis of the S100 in head and neck squamous cell carcinoma. *Transl Cancer Res*. 2023;12(4):717–31.
31. Park JO, Nam IC, Kim CS, Park SJ, Lee DH, Kim HB, et al. Sex Differences in the Prevalence of Head and Neck Cancers: A 10-Year Follow-Up Study of 10 Million Healthy People. *Cancers (Basel)*. 2022 May 1;14(10).
32. Mazul AL, Naik AN, Zhan KY, Stepan KO, Old MO, Kang SY, et al. Gender and race interact to influence survival disparities in head and neck cancer. *Oral Oncol*. 2021 Jan 1;112.

ANEXO 3

" Genetic alterations associated with angiolymphatic invasion in head and neck squamous cell carcinoma. Meta-analysis of gene expression"

Antonio Belli¹, Marta Serna García¹, Nicla Flacco¹

¹Universidad Europea de Valencia, Faculty of Health Sciences, Department of Dentistry, Valencia, Spain

Correspondence:
Universidad Europea de Valencia
Paseo Alameda, 7
46010 – Valencia, Spain
B3ellianyonio@gmail.com

Introduction:

Angiolymphatic invasion (ALI) represents a key prognostic marker in head and neck squamous cell carcinoma (HNSCC), favoring metastatic spread and being associated with worse survival. The identification of molecular signatures associated with IAL may contribute to the development of new clinical biomarkers.

Methods:

Transcriptomic data from 358 HNSCC samples extracted from the TCGA (Firehose Legacy) consortium, classified according to the presence (n=126) or absence (n=232) of IAL, were analyzed. Differential expression analysis was performed using edgeR and limma packages ($FDR \leq 0.05$), followed by functional enrichment analysis with DAVID. In parallel, clinical associations (age, sex, smoking, alcohol) were evaluated using χ^2 tests, Fisher's exact test, Student's t test and Mann-Whitney U test.

Results:

375 differentially expressed genes were identified between tumors with and without ALI (274 overexpressed, 101 underexpressed). Among the most relevant genes were KRT18, GADD45G, SIGIRR (overexpressed) and S100A12, DSG1, GSDMA, WFDC12 (underexpressed). Enriched biological processes were related to exocytosis, cell adhesion, epithelial immunity and inflammatory response. Smoking was the only clinical variable significantly associated with the presence of ALI ($p < 0.05$).

Conclusions:

Angiolymphatic invasion in HNSCC is associated with a specific molecular signature and dysregulation of functional pathways involved in tumor progression. The differential transcriptomic profile found could be useful for the development of prognostic biomarkers and targeted therapeutic strategies in this cancer subtype. Smoking could influence these molecular mechanisms, reinforcing its clinical impact beyond its etiological role.

Keywords: Head and neck squamous cell carcinoma, Angiolymphatic invasion, Transcriptomic analysis, Differentially expressed genes, Functional enrichment, Biomarkers, TCGA.

Introduction:

Head and neck squamous cell carcinoma (HNSCC) is a heterogeneous group of malignant neoplasms originating from the squamous epithelium of the oral, pharyngeal and laryngeal mucosa, and is characterised by its clinical aggressiveness and high mortality (1). The anatomical location of the tumour significantly influences the prognosis, with five-year survival rates ranging from 14.8 % in hypopharynx to 74.0 % in larynx (2).

Worldwide, HNSCC accounts for approximately 90 % of malignant tumours in this region, with an estimated incidence of 890,000 new cases and 450,000 deaths per year (3). In developing countries, its incidence has increased due to tobacco, alcohol and areca nut consumption (4). In contrast, in developed countries, the human papillomavirus (HPV)-associated form predominates (3).

Diagnosis of HNSCC at advanced stages limits therapeutic options and reduces overall survival (4). Tumour molecular heterogeneity hinders uniform response to treatment and contributes to therapeutic resistance (5). This heterogeneity occurs both between tumours from different patients and within the same tumour (6).

One of the histopathological factors with the greatest prognostic impact is angiolymphatic invasion (ALI), defined as the presence of tumour cells within endothelium-lined blood or lymphatic vessels (7). ALI promotes metastatic spread and is associated with poorer overall survival (8). In addition, it has been identified as an independent marker of lymph node metastasis (9).

Molecularly, ALI has been linked to alterations in pathways such as PI3K/AKT, extracellular matrix remodelling and lymphangiogenesis (10). These pathways contribute to the invasiveness of the tumour through interaction with the tumour microenvironment (11). This environment involves fibroblasts, immune cells, vascular components and signalling elements that regulate tumour development (12).

Transcriptomic analysis using RNA-Seq has made it possible to identify differentially expressed genes (DEGs) involved in these processes, facilitating the discovery of biomarkers (13). Tools such as DAVID allow these genes to be linked to key functional pathways for tumour progression (14).

In this context, the present study aims to identify differentially expressed genes between HNSCC tumours with and without IAL, analyse the associated biological pathways and explore possible clinical associations, in order to provide new biomarkers.

Material and methods:

An observational, retrospective study was performed based on the analysis of publicly available transcriptomic data from The Cancer Genome Atlas (TCGA) consortium, specifically from the Firehose Legacy dataset corresponding to patients diagnosed with head and neck squamous cell carcinoma (HNSCC) (The Cancer Genome Atlas, <http://tcga.cancer.gov>). A total of 358 tumour samples, previously classified according to the presence ($n = 126$) or absence ($n = 232$) of angiolymphatic invasion (AI), were selected according to the clinical data available on the platform. Gene expression data, previously normalised to TPM (Transcripts Per Million), were processed with the edgeR and limma packages of the Bioconductor environment for the detection of differentially expressed genes (DEGs). A statistical approach based on moderated linear models and empirical Bayesian estimation was used to control for data dispersion and increase the robustness of the statistical analysis (15).

Genes with an FDR value ≤ 0.05 were considered significant, after correction for multiple testing using the Benjamini-Hochberg method (13). The identified DEGs were subsequently analysed using the DAVID (Database for Annotation, Visualization and Integrated Discovery) tool, with the aim of identifying functional pathways and overrepresented biological processes, based on the GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) ontologies, considering a statistical significance threshold of $p < 0.05$ (14).

In parallel, possible clinical associations between the presence of ALI and epidemiological variables such as sex, age, smoking and alcohol consumption were evaluated. For this purpose, the R-based statistical software Jamovi (The jamovi project) was used. Categorical variables were analysed using χ^2 tests of independence or, failing that, Fisher's exact test, applied when the expected frequencies in any cell were less than 5. For the continuous variable age, the Shapiro-Wilk normality test was initially applied; depending on the result, Student's t-test for independent samples or the non-parametric Mann-Whitney U test was used. A significance level of $p < 0.05$ was adopted in all statistical analyses.

Results:

Differential gene expression analysis between tumours with and without angiolymphatic invasion identified a total of 375 significantly altered genes. Of these, 274 genes were overexpressed in tumours with invasion and 101 were underexpressed, applying a significance threshold of $\text{FDR} \leq 0.05$. Most of the differentially expressed genes showed moderate changes in expression level, albeit with high statistical significance. Among the most overexpressed genes, KRT18, GADD45G, SIGIRR, PRR15 and CFAP43 stood out, with positive fold change values and FDRs below 0.01. In contrast, the most under-expressed genes included S100A12, DSG1, GSDMA, WFDC12 and SPRR2G, all of which are involved in epithelial, immunological or structural functions. Tables 1 and 2 present the 20 most significantly over- and under-expressed genes, respectively, ordered by fold change.

Table 1. The 20 differential genes with $\text{FDR} < 0.05$ with the highest overexpression (*Fold Change* positive)

GENES	Full name	fold change	Pvalue	FDR
FAM174B	family with sequence similarity 174 member B	0.462	5.1×10^{-8}	0.002
ARNILA	androgen receptor negatively regulated lncRNA	0.550	1.4×10^{-7}	0.002
NPDC1	neural proliferation, differentiation and control 1	0.490	2.0×10^{-7}	0.002
RHPN1	rhophilin Rho GTPase binding protein 1	0.673	3.7×10^{-7}	0.003
DNASE2	deoxyribonuclease 2, lysosomal	0.268	6.5×10^{-7}	0.005

COX14	cytochrome c oxidase assembly factor COX14	0.264	8.1×10^{-7}	0.005
FAM50A	family with sequence similarity 50 member A	0.316	1.0×10^{-6}	0.006
PKN1	protein kinase N1	0.412	1.6×10^{-6}	0.007
P4HTM	prolyl 4-hydroxylase, transmembrane	0.449	2.1×10^{-6}	0.007
PCBD1	pterin-4 alpha-carbinolamine dehydratase 1	0.284	2.7×10^{-6}	0.008
PRR15	proline rich 15	0.762	3.5×10^{-6}	0.009
PRKAG2-AS1	PRKAG2 antisense RNA 1	0.626	3.8×10^{-6}	0.009
CATSPERG	cation channel sperm associated auxiliary subunit gamma	0.393	3.7×10^{-6}	0.009
CFAP43	cilia and flagella associated protein 43	0.523	5.9×10^{-6}	0.009
KRT18	keratin 18	0.481	6.1×10^{-6}	0.009
GADD45G	growth arrest and DNA damage inducible gamma	0.473	6.6×10^{-6}	0.009
CCNO	cyclin O	0.467	6.1×10^{-6}	0.009
TSC22D3	TSC22 domain family member 3	0.353	7.4×10^{-6}	0.009
SIGIRR	single Ig and TIR domain containing	0.322	7.3×10^{-6}	0.009
ARRDC2	arrestin domain containing 2	0.288	6.7×10^{-6}	0.009

*Fold Change refers to the group with invasion compared to without angiolymphatic invasion.

Table 2. The 20 differential genes with FDR <0.05 with the lowest underexpression (*Fold Change* negative)

GENES	Full name	fold change	Pvalue	FDR
SPRR2G	small proline rich protein 2G	-1.312	1.3×10^{-6}	0.006
EPGN	epithelial mitogen	-1.050	2.1×10^{-6}	0.007
WFDC12	WAP four-disulfide core domain 12	-1.154	2.7×10^{-6}	0.008
LCE3E	late cornified envelope 3E	-1.149	3.6×10^{-6}	0.009
SPRR5	small proline rich protein 5	-0.840	4.0×10^{-6}	0.009

SLEAR	STAT1 regulated ILF2 complex interacting lncRNA	-0.634	6.3×10^{-6}	0.009
GJB2	gap junction protein beta 2	-0.647	7.3×10^{-6}	0.009
SPRR2F	small proline rich protein 2F	-1.214	6.0×10^{-6}	0.009
DSG1	desmoglein 1	-1.198	7.6×10^{-6}	0.009
KRT17P1	Keratin 17 pseudogene 1	-0.472	9.7×10^{-6}	0.011
GSDMA	gasdermin A	-0.816	1.1×10^{-5}	0.012
CRCT1	cysteine rich C-terminal 1	-1.030	1.6×10^{-5}	0.015
PRR9	proline rich 9	-1.095	1.8×10^{-5}	0.016
GNAI1	G protein subunit alpha i1	-0.299	2.8×10^{-5}	0.019
TUBB2A	tubulin beta 2A class IIa	-0.370	3.0×10^{-5}	0.020
LINC01740	long intergenic non-protein coding RNA 1740	-0.313	3.7×10^{-5}	0.020
ELL2P1	elongation factor for RNA polymerase II 2 pseudogene 1	-0.410	3.3×10^{-5}	0.020
PNPLA1	patatin like phospholipase domain containing 1	-0.735	3.7×10^{-5}	0.020
S100A12	S100 calcium binding protein A12	-0.793	3.5×10^{-5}	0.020
UBQLN1P1	Ubiquiline 1 Pseudogene 1	-0.294	4.2×10^{-5}	0.022

*Fold Change refers to the group with invasion compared to without angiolymphatic invasion.

Subsequently, functional enrichment analysis was performed using DAVID (Database for Annotation, Visualization and Integrated Discovery) software, applied to differentially expressed genes with an FDR ≤ 0.05 . Ten significantly enriched ($p < 0.05$) biological processes were identified in the set of overexpressed genes, mainly related to the regulation of exocytosis, ciliary movement, apoptotic mitochondrial changes, axoneme assembly and mRNA export from the nucleus. These findings suggest an increased capacity for secretion, intracellular remodelling and cytoskeletal reorganisation, potentially associated with increased tumour aggressiveness. In the set of under-expressed genes, DAVID analysis revealed nine enriched biological processes, such as cell-cell adhesion, antimicrobial response, activation of lipopolysaccharide response and phospholipid catalysis. Loss of expression of these genes may reflect dysregulation of defensive epithelial functions and increased fragility of cell-cell junctions, favouring tumour penetration into lymphatic vessels. Tables 3 and 4 list the significantly enriched processes together with the number of genes involved and their identifiers.

Table 3. Biological processes with lower Pvalue < 0.05 enriched according to the DAVID analysis.

Term	Biological process	Count	PValue	Genes
------	--------------------	-------	--------	-------

GO:0017157	regulation of exocytosis	4	0.004	RABL2A, RABL2B, HAP1, VAMP2
GO:0003341	cilium movement	4	0.011	DNAAF1, ZMYND10, CFAP43, PIERCE1
GO:0008637	apoptotic mitochondrial changes	3	0.013	AIFM2, JTB, SLC25A4
GO:0042733	embryonic digit morphogenesis	4	0.020	FBXW4, SHH, FLVCR1, TMEM231
GO:0006783	heme biosynthetic process	3	0.025	SLC25A38, FLVCR1, PPOX
GO:0006406	mRNA export from nucleus	4	0.027	CHTOP, HHEX, DDX39A, RBM8A
GO:0033617	mitochondrial cytochrome c oxidase assembly	3	0.034	UQCC6, COX14, COA5
GO:1903008	organelle disassembly	2	0.041	KIF9, PLATE3
GO:0008089	anterograde axonal transport	3	0.041	DTNBP1, KIF1A, HAP1
GO:0035082	axoneme assembly	3	0.046	RSPH1, DNAAF1, PIERCE1

* The biological processes highlighted in green were selected for detailed analysis and discussion in this study.

Table 4. Biological processes with a pvalue <0.05 inhibited according to the DAVID analysis.

Term	Biological process	Court	PValue	Genes
GO:0009395	phospholipid catabolic process	3	0.001	PLA2G4E, LIPG, PLD2
GO:0019731	antibacterial humoral response	4	0.002	WFDC12, PI3, SPRR2A, WFDC5
GO:0002437	inflammatory response to antigenic stimulus	3	0.004	IL36B, IL36RN, IL1F10
GO:0008366	axon ensheathment	2	0.021	POU3F1, PPARD
GO:0036149	phosphatidylinositol acyl-chain remodeling	2	0.021	PLA2G4D, PLA2G4E
GO:0042742	defense response to bacterium	4	0.022	WFDC12, GSDMA, S100A12, DEFB4A
GO:0071222	cellular response to lipopolysaccharide	4	0.036	IL36B, IL36RN, PPARD, IL1F10
GO:0098609	cell-cell adhesion	4	0.037	TENM2, FBLIM1, DSG1, DSC2
GO:0031640	killing of cells of another organism	3	0.045	LCE3A, S100A12, DEFB4A

* The biological processes highlighted in green were selected for detailed analysis and discussion in this study.

In the analysis of clinical variables, the relationship between smoking history and the presence of angiolymphatic invasion was assessed. We analysed 351 patients divided into categories of current smoker, recent ex-smoker, ex-smoker more than 15 years, never smoker and cases with no data. The highest proportion of cases with angiolymphatic invasion was found in the groups of current smokers and recent ex-smokers. The Chi-square test showed a statistically significant association ($p = 0.033$), which was confirmed by Fisher's exact test ($p = 0.015$). These results indicate that recent or persistent smoking is associated with an increased likelihood of vascular invasion, suggesting a modulatory effect of smoking on tumour dissemination mechanisms.

Regarding alcohol consumption, 352 patients with a documented history were evaluated. Although a trend towards a higher frequency of angiolymphatic invasion was observed in consumers, the differences did not reach statistical significance ($p = 0.224$, Chi-square test; $p = 0.232$, Fisher's exact test). The estimated odds ratio was 1.35, with a 95 % confidence interval between 0.83 and 2.18, suggesting a possible inconclusive increased risk in this cohort.

Age was analysed as a continuous variable. The mean age in the non-invasive group was 62.1 years, while in the invasive group it was 59.1 years. The difference was significant by Student's t-test ($p = 0.023$), but was not confirmed by the non-parametric Mann-Whitney U test ($p = 0.059$), indicating a possible trend, although not statistically robust. This pattern could reflect a greater susceptibility to invasion in younger patients, although the data should be interpreted with caution.

In terms of gender, a higher proportion of angiolymphatic invasion was observed in males than in females. This difference was significant in the Chi-square test ($p = 0.029$) and Fisher's exact test ($p = 0.035$). The calculated odds ratio was 1.77, with a 95 % confidence interval between 1.06 and 2.95, indicating that male patients have a 77 % higher risk of angiolymphatic invasion. This finding could be linked to biological, molecular or risk factor exposure differences, although the gender imbalance in the sample should be considered as a possible confounding factor.

Table 5: Association between clinical variables and lymphovascular invasion

Variable	Categor y	N (no LVI)	N (LVI)	p-val. (x2/Fisher)	OR (95% CI)
Sex	Male	160	99	0.029 / 0.035	1.77 (1.06– 2.95)
	Female	99	27		
Age	Mean ($\pm$ SD)	62.1 $\pm$ 12.0	59.1 $\pm$ 11.6	0.023 / 0.059	
Tobacco	Current smoker	70	49	0.033 / 0.015	
	Former ≤ 15 years	21	21		
	Former > 15 years	76	49		
	Never smoker	61	23		

Alcohol	Consumer	152	90	0.224 / 0.232	1.35 (0.83– 2.18)
	Non- consumer	80	33		

Discussion:

KRT18

KRT18 encodes a type I keratin present in simple epithelia, whose main function is to maintain cellular structural integrity and stability against mechanical or chemical stress (16). Its involvement in signalling processes, cytoskeleton organisation and cell polarity give it an active role in the regulation of epithelial phenotype. In our study, KRT18 was found to be significantly overexpressed in tumours with angiolymphatic invasion, suggesting its involvement in the transition to a more migratory and disorganised state, favouring tumour penetration into lymphatic vessels. In other tumour types, such as hepatocarcinoma and breast cancer, its overexpression has been linked to increased invasiveness and poor prognosis, possibly due to its effect on intercellular adhesion and structural plasticity of tumour cells (17,18). These findings support the hypothesis that KRT18 is not merely a marker of differentiation, but a functional component that may facilitate tumour progression in biologically aggressive settings, such as cases with angiolymphatic invasion in HNSCC.

GADD45G

GADD45G (growth arrest and DNA damage-inducible gamma) is part of a family of cell cycle regulatory genes involved in the response to genotoxic stress, G2/M cycle arrest and modulation of apoptosis. In tumour models, its re-expression has been shown to suppress malignant progression by activating signalling pathways associated with epithelial stability, such as E-cadherin upregulation and inhibition of epithelial-mesenchymal transition (19). In squamous cancers, this activity has been linked to a reduced ability to migrate and invade. In our analysis, GADD45G was found to be significantly overexpressed in tumours with angiolymphatic invasion, a finding that may seem contradictory to its suppressive role, but could be interpreted as a compensatory response to conditions of high biological aggressiveness. Some authors have proposed that this overexpression may reflect an attempt at cellular containment in the face of a pro-invasive microenvironment or a marker of reactive genetic instability (20). These results suggest a more complex role for GADD45G in HNSCC tumour dynamics, possibly dependent on the cellular context and the degree of tumour differentiation.

SIGIRR

SIGIRR (Single Immunoglobulin IL-1 Receptor-Related molecule), also known as TIR8, belongs to the Toll-like receptor (TLR) and interleukin-1 receptor (IL-1R) family, and acts as a negative inhibitor of these inflammatory pathways, especially through modulation of NF- κ B and JNK (21). In the context of head and neck squamous cell carcinoma (HNSCC), it has been observed that SIGIRR may contribute to limiting chronic inflammation in the tumour microenvironment, promoting a more balanced immune response and favouring the infiltration of cytotoxic T-lymphocytes. Furthermore, in HPV-positive tumours, elevated SIGIRR expression has been associated with enhanced activation of the adaptive immune system and reduced expression of immunosuppressive genes (22). In our study, overexpression of SIGIRR in tumours with angiolymphatic invasion could be interpreted as a compensatory response to a proinflammatory and highly immune-active environment. This pattern suggests that SIGIRR may indirectly participate in the regulation of vascular remodelling and lymphangiogenesis processes by limiting the excessive activation of inflammatory pathways that favour tumour progression. Although its role as a suppressor is documented, the overexpression observed in cases with IAL raises the hypothesis of a dual cellular and immunological context-dependent role, which should be explored in future functional studies.

DSG1

DSG1 (desmoglein 1) is a protein of the desmosomal cadherin family that is part of intercellular adhesion complexes in stratified squamous epithelia, including the oral mucosa and epidermis (23). Its function is critical for epithelial cohesion and tissue structural integrity. In different cancer models, loss of DSG1 has been shown to favour a more aggressive phenotype, characterised by loss of cell differentiation, formation of invasion structures such as invadopodia, and activation of pathways such as EGFR/ERK, associated with tumour migration and infiltration (24). In our analysis, the under-expression of DSG1 in tumours with angiolymphatic invasion reinforces its possible role as a marker of epithelial integrity. Its loss could facilitate tissue disorganisation and transition to a more mobile and less cohesive phenotype, characteristics necessary for tumour dissemination through lymphatic vessels. Furthermore, the observed expression pattern suggests that DSG1 may play a key role in the regulation of HNSCC invasive behaviour, not only as a structural element, but also as a modulator of tumour signalling.

GSDMA

GSDMA (Gasdermin A) belongs to the gasdermin family, a set of proteins associated with pyroptosis, a type of caspase-dependent inflammatory cell death. In addition to this function, GSDMA is involved in epithelial homeostasis, cell differentiation and regulation of apoptosis, acting as a potential suppressor of tumour progression in squamous epithelia (25). Under physiological conditions, it is abundantly expressed in skin, oral mucosa and digestive tract. Several studies have shown that its loss of expression is associated with a more aggressive phenotype, resistance to apoptotic signals and increased invasiveness. In our analysis, GSDMA is significantly under-expressed in tumours with angiolymphatic invasion. This finding suggests that its absence may facilitate evasion of cell death during migration and entry into lymphatic vessels. In addition, functional enrichment analysis indicated its involvement in immune processes against bacteria and lipopolysaccharides, which reinforces its role as a modulator of the immune microenvironment. Disruption of these functions

could compromise local immunosurveillance and facilitate the entry of tumour cells into the lymphatic system (26).

WFDC12

WFDC12 (WAP four-disulfide core domain 12) belongs to a family of proteins characterised by WAP domains with anti-protease and immunomodulatory functions. Its expression has been described in epithelia with innate defence functions, and its alteration has been linked to inflammatory and immune-mediated diseases (27). Although direct evidence in HNSCC is limited, studies in oesophageal squamous cell carcinomas have shown that suppression of WFDC12 promotes cell proliferation and migration, suggesting a conserved tumour-suppressor role (28). In our analysis, significant under-expression of WFDC12 was observed in tumours with angiolymphatic invasion. This finding is consistent with its possible role as a regulator of the immune microenvironment and as a barrier to extracellular matrix degradation. In addition, functional enrichment analysis showed its involvement in pathways related to antimicrobial response and inflammation, in conjunction with GSDMA and S100A12. This functional convergence suggests the existence of a common epithelial regulatory axis that may be involved in the most aggressive tumours.

S100A12

S100A12 (S100 calcium-binding protein A12) is a secreted protein of the S100 family, involved in the regulation of inflammation, innate immune response and endothelial activation. It acts as a ligand for RAGE and TLR4 receptors, activating proinflammatory pathways such as NF- κ B and promoting the production of cytokines and reactive oxygen species (29). Although it has been associated with increased aggressiveness in certain tumours, its role varies according to tumour type and setting. In our study, S100A12 was found to be under-expressed in tumours with angiolymphatic invasion, suggesting a loss of protective functions of the inflammatory microenvironment and a possible facilitation of lymphatic spread. This finding is consistent with previous bioinformatics analyses in HNSCC linking low S100A12 expression with higher histological grade, lymphovascular invasion and lymph node metastasis (30). Therefore, its under-expression could reflect a less immunosurveilled microenvironment, more permissive for tumour progression and escape.

Relationship between angiolymphatic invasion and clinical variables

Our clinical analysis revealed significant associations between the presence of angiolymphatic invasion (ALI) and certain epidemiological variables of HNSCC. A higher frequency of ALI was observed in male patients and in those with a history of active or recent smoking, reinforcing the hypothesis that some clinical profiles, in addition to predisposing to tumour development, might also modulate its biological invasive behaviour. Although age showed a trend towards significance and alcohol consumption was not statistically significantly associated, these variables could be influenced by uncontrolled factors and would require further exploration in larger cohorts.

In particular, the association between male sex and ALI, although statistically significant, should be interpreted with caution due to the gender imbalance present in the sample. A cohort study in South Korea confirmed that males have a significantly higher incidence of HNSCC than females, even without a history of smoking or alcohol (31). Also, male sex has been associated with worse survival

in HPV-negative tumours, which may reflect a greater biological susceptibility beyond classical behavioural factors (32).

On the other hand, smoking showed a strong association with ALI. It has been described that tobacco exposure may induce molecular alterations affecting genes regulating migration, adhesion and inflammation, which could facilitate lymphovascular dissemination (31). In our study, this hypothesis is reinforced by the under-expression of genes such as DSG1, GSDMA and WFDC12 in patients with ALI, all of which are involved in epithelial organisation, cellular cohesion and local immunity.

Conclusion:

Angiolymphatic invasion in head and neck squamous cell carcinoma was associated with a differential gene signature composed of 375 genes, including KRT18, GADD45G, SIGIRR, S100A12, DSG1, GSDMA and WFDC12. Functional analysis showed significant involvement of processes such as cell-cell adhesion, exocytosis and immune response, which reinforces their role in tumour progression. These results allowed us to propose these genes as potential biomarkers for the early detection of angiolymphatic invasion. In addition, clinical associations with smoking and gender were identified, although the latter should be interpreted with caution due to the imbalance in the sample distribution.

Bibliography:

1. Shirima CA, Bleotu C, Spandidos DA, El-Naggar AK, Pircalabioru GG, Michalopoulos I. Epithelial-derived head and neck squamous tumourigenesis (Review). Vol. 52, *Oncology Reports*. Spandidos Publications; 2024.
2. Mores AL, Bonfim-Alves CG, López RVM, Rodrigues-Oliveira L, Palmier NR, Mariz BALA, et al. Prognostic Factors in Head and Neck Cancer: A Retrospective Cohort Study of 3052 Patients in Brazil. *Oral Dis*. 2024 Apr 1;
3. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Medical Sciences*. 2023 Jun 13;11(2):42.
4. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J*. 2022 Nov 11;233(9):780–6.
5. Mroz EA, Rocco JW. Intra-tumor heterogeneity in head and neck cancer and its clinical implications. Vol. 2, *World Journal of Otorhinolaryngology - Head and Neck Surgery*. KeAi Communications Co.; 2016. p. 60–7.
6. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015 Jan 29;517(7536):576–82.
7. Mascitti M, Togni L, Caponio VCA, Zhurakivska K, Bizzoca ME, Contaldo M, et al. Lymphovascular invasion as a prognostic tool for oral squamous cell carcinoma: a comprehensive review. Vol. 51, *International Journal of Oral and Maxillofacial Surgery*. Churchill Livingstone; 2022. p. 1–9.
8. Zhang J, Lin H, Jiang H, Jiang H, Xie T, Wang B, et al. A key genomic signature associated with lymphovascular invasion in head and neck squamous cell carcinoma. *BMC Cancer*. 2020 Mar 30;20(1).
9. Huang S, Zhu Y, Cai H, Zhang Y, Hou J. Impact of lymphovascular invasion in oral squamous cell carcinoma: A meta-analysis. Vol. 131, *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*. Mosby Inc.; 2021. p. 319-328.e1.
10. Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. Vol. 61, *Seminars in Cancer Biology*. Academic Press; 2020. p. 71–83.
11. Chen YQ, Kuo JC, Wei MT, Chen YC, Yang MH, Chiou A. Early stage mechanical remodeling of collagen surrounding head and neck squamous cell carcinoma spheroids correlates strongly with their invasion capability. *Acta Biomater*. 2019 Jan 15;84:280–92.
12. Arneth B. Tumor microenvironment. Vol. 56, *Medicina (Lithuania)*. MDPI AG; 2020.
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014 Dec 5;15(12).
14. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022 Jul 5;50(W1):W216–21.
15. Law CW, Alhamdoosh M, Su S, Smyth GK, Ritchie ME. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res*. 2016 Jun 17;5:1408.
16. Yan J, Yang A, Tu S. The relationship between keratin 18 and epithelial-derived tumors: as a diagnostic marker, prognostic marker, and its role in tumorigenesis. Vol. 14, *Frontiers in Oncology*. Frontiers Media SA; 2024.
17. Golob-Schwarzl N, Bettermann K, Mehta AK, Kessler SM, Unterluggauer J, Krassnig S, et al. High Keratin 8/18 Ratio Predicts Aggressive Hepatocellular Cancer Phenotype. *Transl Oncol*. 2019 Feb 1;12(2):256–68.

18. Ha SA, Lee YS, Kim HK, Yoo J, Kim S, Gong GH, et al. The prognostic potential of keratin 18 in breast cancer associated with tumor dedifferentiation, and the loss of estrogen and progesterone receptors. *Cancer Biomarkers*. 2011;10(5):219–31.
19. Zhan Q, Antinore MJ, Wang XW, Carrier F, Smith ML, Harris CC, et al. Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene*. 1999 May 6;18(18):2892–900.
20. Li T, Xu L, Teng J, Ma Y, Liu W, Wang Y, et al. GADD45G Interacts with E-cadherin to Suppress the Migration and Invasion of Esophageal Squamous Cell Carcinoma. *Dig Dis Sci*. 2020 Apr 1;65(4):1032–41.
21. Garlanda C, Anders HJ, Mantovani A. TIR8/SIGIRR: an IL-1R/TLR family member with regulatory functions in inflammation and T cell polarization. Vol. 30, *Trends in Immunology*. 2009. p. 439–46.
22. Yin G, Guo W, Wang R, Li N, Chen X, Zhang Y, et al. Analysis of the role of IL-1 family and related genes in head and neck squamous cell carcinoma. *Braz J Otorhinolaryngol*. 2025 Jan 1;91(1).
23. Wong MP, Cheang M, Yorlida E, Coldman A, Gilks CB, Huntsman D, et al. Loss of desmoglein 1 expression associated with worse prognosis in head and neck squamous cell carcinoma patients. *Pathology*. 2008;40(6):611–6.
24. Valenzuela-Iglesias A, Burks HE, Arnette CR, Yalamanchili A, Nekrasova O, Godsel LM, et al. Desmoglein 1 regulates invadopodia by suppressing EGFR/ERK signaling in an ERBIN-dependent manner. *Molecular Cancer Research*. 2019 May 1;17(5):1195–206.
25. Zou J, Zheng Y, Huang Y, Tang D, Kang R, Chen R. The Versatile Gasdermin Family: Their Function and Roles in Diseases. Vol. 12, *Frontiers in Immunology*. Frontiers Media S.A.; 2021.
26. Saeki N, Usui T, Aoyagi K, Kim DH, Sato M, Mabuchi T, et al. Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. *Genes Chromosomes Cancer*. 2009 Mar;48(3):261–71.
27. Kalinina P, Vorstandlechner V, Buchberger M, Eckhart L, Lengauer B, Golabi B, et al. The Whey Acidic Protein WFDC12 Is Specifically Expressed in Terminally Differentiated Keratinocytes and Regulates Epidermal Serine Protease Activity. *Journal of Investigative Dermatology*. 2021 May 1;141(5):1198–1206.e13.
28. Zhuang Y, Wang L, Ji C, Sun Y, Shao F. Construction of a novel pyroptosis-related prognostic model of esophageal square cell carcinoma and determination of the anti-tumor effect of WFDC12. *Funct Integr Genomics*. 2023 Jun 1;23(2).
29. Sionov RV. Leveling up the controversial role of neutrophils in cancer: When the complexity becomes entangled. Vol. 10, *Cells*. MDPI; 2021.
30. Bai G, Yue S, You Y, Zhao J, Tian Z, Wang Y, et al. An integrated bioinformatics analysis of the S100 in head and neck squamous cell carcinoma. *Transl Cancer Res*. 2023;12(4):717–31.
31. Park JO, Nam IC, Kim CS, Park SJ, Lee DH, Kim HB, et al. Sex Differences in the Prevalence of Head and Neck Cancers: A 10-Year Follow-Up Study of 10 Million Healthy People. *Cancers (Basel)*. 2022 May 1;14(10).
32. Mazul AL, Naik AN, Zhan KY, Stepan KO, Old MO, Kang SY, et al. Gender and race interact to influence survival disparities in head and neck cancer. *Oral Oncol*. 2021 Jan 1;112.