



UNIVERSIDAD AUTÓNOMA DE GUERRERO
FACULTAD DE CIENCIAS QUÍMICO-BIOLÓGICAS
MAESTRIA EN CIENCIAS BIOMEDICAS



**Análisis de la expresión de las proteínas no caracterizadas
del cromosoma 19 en cáncer cervicouterino, mama y ovario:
Proyecto Proteoma Humano (HUPO)**

T E S I S

**PARA OBTENER EL GRADO DE
MAESTRA EN CIENCIAS BIOMEDICAS**

P R E S E N T A

Q.B.P. Diana Laura Gonzalez Tolentino

Directora de tesis: Dra. Olga Lilia Garibay Cerdenares

Codirector: Dr. Sergio Manuel Encarnación Guevara

Chilpancingo Gro., diciembre de 2022



UNIVERSIDAD AUTÓNOMA DE GUERRERO
FACULTAD DE CIENCIAS QUÍMICO BIOLÓGICAS
FACULTAD DE MEDICINA
UNIDAD DE INVESTIGACIÓN ESPECIALIZADA EN MICROBIOLOGÍA
MAESTRÍA EN CIENCIAS BIOMÉDICAS

ACTA DE APROBACIÓN DE TESIS

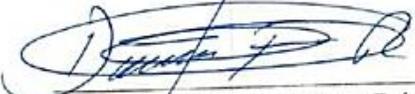
En la ciudad de Chilpancingo, Guerrero, siendo los 04 días del mes de julio del año dos mil veintidós, se reunieron los miembros del Comité Tutorial designado por la Academia de Posgrado de la Maestría en Ciencias Biomédicas, para examinar la tesis titulada **Análisis de la expresión de las proteínas no caracterizadas del cromosoma 19 en cáncer cervicouterino, mama y ovario: Proyecto Proteoma Humano (HUPO)**, presentada por la alumna Diana Laura Gonzalez Tolentino, para obtener el Grado de Maestría en Ciencias Biomédicas. Después del análisis correspondiente, los miembros del comité manifiestan su aprobación de la tesis, autorizan la impresión final de la misma y aceptan que, cuando se satisfagan los requisitos señalados en el Reglamento General de Estudios de Posgrado e Investigación Vigente, se proceda a la presentación del examen de grado.

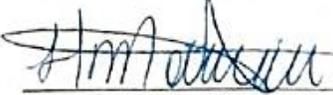
El Comité Tutorial


Dra. Olga Lilia Garibay Cerdanares
Dirección de tesis


Dr. Marco Antonio Leyva Vázquez

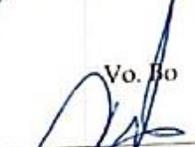

Dra. Luz del Carmen Alarcón Romero


Dr. Donaciano Flores Robles


M. en C. Mónica Virginia Saavedra Herrera

Vo. Bo




Coordinación de la Maestría en Ciencias Biomédicas
Dr. Daniel Hernández Sotelo
Coordinador de la Maestría en Ciencias Biomédicas



Facultad de Ciencias Químico Biológicas
Vo. Bo


Dr. Oscar del Moral Hernández
Director de la Facultad de Ciencias Químico Biológicas

Este trabajo se realizó en el Laboratorio de Biomedicina Molecular y el Laboratorio de Investigación en Citopatología e Histoquímica de la Facultad de Ciencias Químico Biológicas (FCQB) de la Universidad Autónoma de Guerrero (UAGro) en Chilpancingo, Guerrero.

Bajo la dirección de:

Dra. Olga Lilia Garibay Cerdenares

La Co-dirección de:

Dr. Sergio Manuel Encarnación Guevara

La asesoría de:

Dra. Luz del Carmen Alarcón Romero

Dr. Marco Antonio Leyva Vásquez

Dr. Donaciano Flores Robles

MenC. Mónica Virginia Saavedra Herrera

Durante el periodo en que curso la Maestría en Ciencias Biomédicas, la C. Diana Laura Gonzalez Tolentino recibió la beca (No.772011) del 01 de septiembre del 2020 al 01 de agosto de del 2022 otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACyT) a los programas del Sistema Nacional de Posgrados (SNP).

AGRADECIMIENTOS

Mi agradecimiento, cariño y respeto siempre a la **Dra. Olga Lilia Garibay Cerdaneres**, por aceptarme, para trabajar en este proyecto. Por motivarnos a continuar alcanzando más metas, por la disponibilidad y apoyo incondicional, para concluir este trabajo. Siempre admire su profesionalismo y trabajo en el laboratorio, gracias por enseñarnos tanto.

Al **Dr. Sergio Manuel Encarnación Guevara**, por confiar, en nosotros para ser partes de este proyecto tan importante. Por la disponibilidad para revisar y mejorar este trabajo.

A la **Dra. Luz del Carmen Alarcón Romero**, por su apoyo y disposición siempre, para mejorar este trabajo, por permitirme trabajar en su laboratorio, la técnica de inmunohistoquímica, también por siempre motivarnos en clases, recuerdo cuando nos decía que a pesar de las dificultades siempre tenemos que buscar la forma de salir adelante. Mi admiración y respeto siempre.

Al **Dr. Marco Antonio Leyva Vázquez** por su apoyo, para mejorar el trabajo, por aceptarnos para trabajar en el Laboratorio de Biomedicina molecular.

Al **Dr. Donaciano Flores Robles**, muchas gracias por su gran apoyo en las correcciones de este trabajo, por la disponibilidad en todos los seminarios, mi admiración y respeto por su profesionalismo y trabajo como investigador.

A la **MC. Mónica Virginia Saavedra Herrera**, le agradezco mucho su apoyo, para formar parte del comité tutorial, gracias por los comentarios y observaciones al trabajo, por la disponibilidad en los seminarios, muchas gracias.

A la **Dra. Eugenia Flores Alfaro**, coordinadora del seminario de investigación, por guiarnos en cada seminario, por sus comentarios y observaciones.

Al **Dr. Israel Francisco Torres Rojas**, por su apoyo y disponibilidad siempre para resolver nuestras dudas, por enseñarnos y apoyarnos en la estandarización de la técnica de RT- qPCR. Muchas gracias Dr. por ser también parte de este proyecto.

Al **Dr. Hugo Alberto Ruiz Alarcón**, por su apoyo y enseñanzas, en la técnica de RT-qPCR.

A la **Dra. Ana Elvira Zacapala Gómez** y el **Dr. Miguel Ángel Mendoza Catalán**, a quienes admiro y respeto mucho, por creer en mí y darme mi primera oportunidad en la investigación, lo valoro muchísimo, por seguirme apoyando y enseñarme con su ejemplo el enorme valor que tiene ser solidarios y ayudar a los demás.

A la química **Natividad Sales Linares** y **Josué Feliciano Ortíz** del Laboratorio de Biomedicina Molecular, por su enorme apoyo, y por el gran trabajo que hacen en el LBM.

A la **Dra. Berenice Illades Aguiar**, muchas gracias, Dra. Bere, por apoyar a muchas jóvenes como yo, con el programa de incorporación de mujeres indígenas a posgrados, por confiar en nuestros sueños y habilidades, por motivarnos a proponernos nuestras metas y cumplirlas.

A mis amigos y compañeros de la maestría, **Esli Janaí Flores Palma**, **Ramiro Gonzalez Romero**, **Yesica Eulogio Metodio** y **José Eduardo García**, me enseñaron a trabajar en equipo, para reunirnos a leer artículos, discutir nuestras dudas, y también pasar momentos muy agradables y felices, gracias por permitirme ser parte del equipo de los doctores, mi cariño, admiración y buenos deseos siempre para ustedes.

A Dios por dirigir mi vida,
por su amor
por bendecirme con una vida bonita,
por enseñarme que las cosas más valiosas, son las risas y
los momentos de felicidad junto a mi familia,
y también por los momentos de dificultad
que nos hacen despertar y crecer.

Índice

Índice	7
Resumen	8
Abstract	9
Introducción	10
Materiales y métodos	12
Resultados	18
Discusión	43
Conclusiones	47
Perspectivas	48
Anexos	49
Referencias	52

Resumen

Introducción: el cáncer es una de las principales causas de muerte a nivel mundial, el origen de esta enfermedad es multifactorial e involucra la desregulación de procesos genéticos, epigenéticos y proteómicos, entre otros. The Human Proteome Organization (HUPO) es una organización científica internacional que promueve el análisis de las proteínas conocidas como “dark proteins o “proteínas desconocidas”, en las que se desconoce su nivel de expresión y su papel biológico en las células. Este proyecto tiene como objetivo el análisis de las proteínas C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160, que son parte del cromosoma 19 en los cánceres cérvicouterino, ovario y mama. **Metodología:** la expresión de C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160 se evaluó mediante RT-qPCR e inmunohistoquímica en líneas celulares y biopsias de pacientes en comparación con tejido normal, adyacente y líneas celulares no tumorales. **Resultados:** Los datos de expresión para CCDC61 y CCDC97 fueron disminuídos y no detectados a nivel de proteína en los tres tipos de cáncer. LENG8 una subexpresión en CaCu, CaMa y CaOv ($p \leq 0.05$); mientras que para TMEM160, una sobreexpresión en CaOv, respecto a CaCu y CaMa ($p < 0.05$). A nivel de proteína C19ORF47, TMEM160 y LENG8 se observa una sobreexpresión diferencial en cáncer cérvicouterino, mama y ovario, con localización subcelular en membrana y citoplasma. La línea celular HeLa sobreexpresa a C19ORF47 y TMEM160 en comparación con la línea no tumoral HaCaT ($p > 0.001$). En la línea celular C33A, LENG8 está sobreexpresado en comparación con la línea control HaCaT ($p \leq 0.05$). Las líneas celulares de CaMa se observó la sobreexpresión de C19ORF47 en MCF7, y TMEM160 en MCF7 y MDA-MB-231 en comparación con la línea control MCF10A ($p < 0.05$). **Conclusión:** C19ORF47, LENG8 y TMEM160 presentan un nivel de expresión diferencial en CaCu, CaMa y CaOv, probablemente asociado a cáncer.

Palabras clave: C19ORF47, CCDC61, CCDC97, LENG8, TMEM160, HUPO, cáncer.

Abstract

Introduction: Cancer is one of the leading causes of death worldwide, the origin of this disease is multifactorial and involves the deregulation of genetic, epigenetic, and proteomic processes, among others. The Human Proteome Organization (HUPO) is an international scientific organization that promotes the analysis of proteins known as "dark proteins" or "unknown proteins", whose level of expression and biological role in cells is unknown. This project aims to analyze the proteins C19ORF47, CCDC61, CCDC97, LENG8, and TMEM160, which are part of chromosome 19 in cervical, ovarian, and breast cancers. Methodology: the expression of C19ORF47, CCDC61, CCDC97, LENG8, and TMEM160 was assessed by RT-qPCR and immunohistochemistry in patient cell lines and biopsies compared to normal, adjacent tissue and non-tumor cell lines. Results: Expression data for CCDC61 and CCDC97 were decreased and not detected at the protein level in all three cancer types. LENG8 had an underexpression in CaCu, CaMa, and CaOv ($p \leq 0.05$); while for TMEM160, an overexpression in CaOv, concerning CaCu and CaMa ($p < 0.05$). At the C19ORF47, TMEM160, and LENG8 protein levels, differential overexpression was observed in cervical, breast, and ovarian cancer, with subcellular localization in the membrane and cytoplasm. The HeLa cell line overexpressed *C19ORF47* and *TMEM160* compared to the non-tumor line HaCaT ($p > 0.001$). In the C33A cell line, *LENG8* is overexpressed compared to the HaCaT control line ($p \leq 0.05$). CaMa cell lines, overexpression of *C19ORF47* was observed in MCF7 and *TMEM160* in MCF7 and MDA-MB-231 compared to the control line MCF10A ($p < 0.05$). Conclusion: C19ORF47, LENG8, and TMEM160 present a differential expression level in CaCu, CaMa, and CaOv, probably associated with cancer.

Keywords: C19ORF47, CCDC61, CCDC97, LENG8, TMEM160, HUPO, cancer.

Introducción

El cáncer es una de las principales enfermedades a nivel mundial. La Organización Mundial de la Salud considera al cáncer como la segunda causa de muerte en 2020. De acuerdo con Globocan, el cáncer de mama, el cáncer cervicouterino y al cáncer de ovario se ubican entre los primeros lugares de incidencia y mortalidad en mujeres en todo el mundo (Sung *et al.*, 2021).

Las células cancerosas durante el proceso de progresión tumoral utilizan diversos mecanismos genéticos y epigenéticos que les permiten evadir los procesos que regulan la proliferación y la supervivencia celular, y como consecuencia, las células tumorales proliferan de manera aberrante, siendo capaces de inducir metástasis (Chen *et al.*, 2017).

Una de las principales estrategias en el diagnóstico y tratamiento de esta enfermedad ha sido el análisis de proteínas mediante diversas herramientas proteómicas como posibles biomarcadores y/o blancos terapéuticos, sin embargo, durante el proceso se han detectado proteínas que han sido clasificadas en el “Dark Side” del proteoma, ya que en este grupo se consideran a aquellas proteínas que han sido identificadas mediante herramientas genómicas, transcriptómicas y proteómicas, pero que carecen de una función biológica asignada. En febrero del 2001 se inició oficialmente la Organización del Proteoma Humano (HUPO) (<https://www.hupo.org>), el cual tiene como objetivo principal caracterizar todas las proteínas codificadas por los genes en cada uno de los cromosomas humanos (Paik *et al.*, 2012), para ello se creó una red de colaboración entre diferentes equipos de investigación para el análisis de cromosomas individuales, como un recurso para identificar la función biológica y molecular que permita avanzar en el diagnóstico, tratamiento y prevención de enfermedades (Paik *et al.*, 2018).

En México, el grupo científico “Consortio Cromosoma 19” es el responsable de indagar sobre las proteínas de este cromosoma, el segundo cromosoma más grande en humanos, que se ha asociado a enfermedades como el cáncer. El cromosoma 19, forma parte de los 22 cromosomas autosómicos, constituido de 64

millones de pares de bases aproximadamente, y representan el 2% del DNA total en las células humanas (Nilsson *et al.*, 2013). En promedio contiene 26 loci que abarcan 28.1 Mb, sus transcritos maduros contiene en promedio 58% de regiones codificantes y 42% de no traducción, con un alto contenido de G+C (Grimwood *et al.*, 2004).

A través de un análisis en bases de datos como neXProt y UNIPROT, fueron identificadas 1429 proteínas codificadas en el cromosoma 19, sin embargo, sólo en el 60% de ellas existe evidencia traduccional a proteína considerando los métodos de identificación internacionales como la secuenciación Edman, espectrometría de masas, estructura por cristalografía de rayos X o resonancia magnética (RMN), interacción proteína-proteína y su detección a través de anticuerpos, es decir, la presencia de evidencia experimental validada sobre su existencia, mientras que en el 36% sólo se presenta evidencia experimental a nivel del transcrito (RT-PCR o Northern Blot), indicando la presencia de una región codificante, pero que no ha sido estrictamente comprobada. Finalmente, el 4% restante, corresponden a proteínas conocidas como inferidas (identificadas como ortólogas en especies estrechamente relacionadas), predichas (sin evidencia experimental de transcrito o proteína) o proteínas inciertas (sin evidencia experimental) (Taylor *et al.*, 2008).

En un estudio realizado por el Consorcio Cromosoma 19, utilizando espectrometría de masas, se identificaron los péptidos de las proteínas: Cromosoma 19 Marco de lectura abierto 47 (C19ORF47), la proteína dominio en espiral 61 (CCDC61), la proteína dominio en espiral 97 (CCDC97), la proteína miembro del grupo de receptores de leucocitos 8 (LENG8) y la proteína transmembranal 160 (TMEM160) (Nilsson *et al.*, 2012). Dichas proteínas también fueron identificadas en células no tumorales HaCaT y en células CaLo y SiHa derivadas de cáncer cervicouterino mediante LC-MS/MS, sin embargo, estas proteínas carecen de función asignada (Gil *et al.*, 2017).

El gen *C19ORF47* (marco de lectura abierto 47 del cromosoma 19) se encuentra localizado en el cromosoma 19q13.2, y la proteína que codifica consta de 422 aminoácidos y tiene un peso molecular de 45 kDa. Las proteínas que contienen

dominios en espiral (CCDC) pertenecen a una familia de proteínas que participan en diversas funciones biológicas, las cuales se han asociado al cáncer (Modjtahedi *et al.*, 2016). *CCDC61* es un gen codificante, localizado en el cromosoma 19q13.32, mientras que la proteína que codifica consta de 512 aminoácidos y tiene un peso molecular de 57.36 kDa. *CCDC61* es una proteína centrosomal de unión a microtúbulos necesaria para la cohesión del centriolo. Por su parte, *CCDC97* es codificada por el gen con el mismo nombre, ubicado en el cromosoma 19q13.2, su proteína consta de 343 aminoácidos y tiene un peso aproximado de 55 kDa. (GeneCards; Rigbolt *et al.*, 2011). *CCDC97* fue identificada en un análisis en tejidos pulmonares de rata y su expresión fue validada por RT-qPCR (Liu *et al.*, 2020).

Por otra parte, el gen *LENG8* (miembro del grupo de receptores de leucocitos 8) se localiza en el cromosoma 19q13.42, y su proteína que codifica consta de 800 aminoácidos y tiene un peso molecular de 90 kDa. Entre las proteínas no caracterizadas de la familia TMEM, se encuentra TMEM160, la cual consta de 188 aminoácidos y un peso molecular de 19.65 kDa, es codificada por el gen con el mismo nombre ubicado en el cromosoma 19q13.32, sin embargo, TMEM 160 no ha sido caracterizada experimentalmente a nivel funcional (Marx *et al.*, 2020a).

Debido a que no se conoce la función biológica y el nivel de expresión basal y en cáncer de las proteínas C19ORF47, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160*. El objetivo de este estudio fue analizar la expresión de estas cuatro proteínas, a nivel de RNAm y proteína, en cáncer de mama (CaMa), ovario (CaOv) y cervicouterino (CaCu).

Materiales y métodos

Selección de las muestras

Biopsias de cáncer cervicouterino, mama y ovario obtenidas del Biobanco del Laboratorio de Biomedicina Molecular de la Facultad de Ciencias Químico-Biológicas de la Universidad Autónoma de Guerrero través de proyectos aprobados en la Subdirección de Enseñanza e Investigación del Instituto Estatal de Cancerología (IECAN) Dr. Arturo Beltrán Ortega, en Acapulco Guerrero, México.

Para el análisis de expresión se consideraron 20 biopsias de cáncer de pacientes de recién diagnóstico cáncer de mama, con sus respectivos tejidos adyacentes (Anexos, tabla 2), 19 biopsias de cáncer de ovario: 3 de ovario libre de cáncer (OLC), 4 de teratomas benignos (COG) y 12 de carcinomas epiteliales (CEO) (Anexos, tabla 3) finalmente 15 biopsias de cáncer cervicouterino: 4 biopsias sin lesión negativas al virus del papiloma humano (VPH) (Sin LEI, VPH-), 2 sin lesión positivas a VPH (Sin LEI, VPH+) y 9 biopsias de cáncer cervicouterino (CaCu) (sin considerar el genotipo de VPH). Para el análisis de la expresión de proteínas, se utilizaron cortes histológicos montados en parafina: 20 laminillas provenientes de tejidos de cáncer de mama (Anexo, tabla 4), 15 laminillas de cáncer cervical y 7 laminillas de cáncer de ovario (sin información). Se obtuvo acceso a los expedientes clínicos de algunas de las pacientes de las muestras colectadas. Se realizó una base de datos con la información clínica relevante, características histopatológicas y demográficas de los pacientes.

Líneas Celulares

Las células HaCaT (queratinocitos inmortalizados), C33A, HeLa y SiHa derivadas de cáncer cervicouterino se obtuvieron del Laboratorio de Biomedicina, y se cultivaron en medio DMEM suplementado con suero bovino fetal al 10 % (Gibco, Life Technologies, Gaithersburg, MD, EE. UU.), penicilina 100 U/ml y estreptomycin 100 µg/ml, a 37°C en presencia de CO₂ al 5 %. Las líneas celulares de cáncer de mama MCF7, MDA-MB-231 y MCF10A (línea celular epitelial no tumorigénica) se obtuvieron del Laboratorio de Biología Celular del Cáncer, y se cultivaron en medio DMEM/F12 (Sigma-Aldrich) suplementado con suero fetal bovino (FBS) al 5 % y antibióticos al 1 %. penicilina G/Estreptomycin, Gibco) en una atmósfera humidificada que contiene 5% de CO₂ a 37°C.

Extracción de RNA

El RNA total se obtuvo de los cultivos celulares y biopsias usando el reactivo TRIzol (Reagent® Invitrogen™ Life Technologies, Carlsbad, EUA) (Chomczynski y Sacchi, 1987). Para las biopsias se trituraron ~2g de tejido en un mortero, agregando 100 µL de PBS 1x, posteriormente se transfirió a un tubo Eppendorf de 1.5 mL, después

se agregó 600 μL de TRIzol aplicando vórtex durante 15 s y se dejó reposar en hielo por 5 min, (para las muestras con alta cantidad de grasa, se centrifugaron a 12,000 rpm por 5 min y se transfirió el sobrenadante a otro tubo, se dejó incubar por 5 min), se agregaron 60 μL de cloroformo y se mezcló en vórtex por 15 s, y se incubó por 10 min en hielo, se centrifugaron a 13,000 rpm por 15 min para obtener las tres fases (acuosa, intermedia y orgánica), después se recuperó la fase acuosa (RNA), y se agregaron 300 μL de isopropanol, y se incubaron a -80°C toda la noche. Después se realizó una serie de lavados con etanol y centrifugaciones a 4°C , el RNA se resuspendió en agua libre de nucleasas. La pureza fue determinada mediante espectrofotometría utilizando el Nanodrop 2000C (Thermo Scientific Massachusetts, EUA), una pureza de 1.7–2.0 fue considerada como adecuada para la realización del análisis de la expresión génica. Las muestras se utilizaron inmediatamente o se almacenaron a -80°C hasta su uso.

Síntesis de cDNA

Se utilizó 400 ng de RNA total, para la reacción de transcripción reversa, 1 μL del Oligo dT 15 (Promega Corporation, Wisconsin, EUA) en un volumen final de 13 μL con agua estéril. La mezcla se incubó a 70°C por 8 min en un termociclador SelectCycler (Applied Biosystems, Foster City, EUA), seguido de una incubación en hielo por 1 min. Se adicionaron 4 μL del buffer de reacción M-MLV 5X (Promega Corporation), 1 μL de una solución 10 mM de desoxinucleótidos trifosfatados (dNTP's) (Promega Corporation), 1 μL (10 U) de inhibidor de RNAsas (Promega Corporation), 1 μL (200 U) de la enzima transcriptasa (RT) reversa M-MLV (Promega Corporation). La mezcla de reacción se ajustó a un volumen final de 20 μL con agua estéril y se incubó a 42°C por 1 h. Se adicionaron 0.5 μL (1.5 U/ μL) de ribonucleasa H a la mezcla y se incubó a 37°C por 20 min. Por último, la muestra se utilizó inmediatamente o se almacenó a -20°C hasta su uso.

RT-qPCR

La expresión del mRNA se determinó mediante PCR cuantitativa (qPCR) utilizando un sistema de tiempo real CFX96 Touch™ (BioRad Laboratories, Hercules, EUA), con SYBR Green (Applied Biosystems, Carlsbad, CA, EE. UU.) Las reacciones de

qPCR se realizaron por separado para cada gen y se normalizaron con la expresión del mRNA de β -actina (Hs99999903_m1). Los ensayos de PCR se realizaron por duplicado. La reacción se llevó a cabo en un volumen final de 20 μ L que contiene 6 μ L de SYBR Green (Applied Biosystems), 12 μ L de Máster Mix (Applied Biosystems) y 2 μ L de cDNA (equivalente a 200 ng). Se realizaron amplificaciones durante 40 ciclos a 94 °C durante 15 s y 63 °C durante 1 min. Los cambios en la expresión génica relativa se determinaron utilizando la fórmula $2^{-\Delta CT}$, $2^{-\Delta\Delta CT}$ y Log 2.

Tabla1. Secuencias de los iniciadores para los genes *C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160* analizados por RT-qPCR.

Nombre del gen	Secuencia del Primer usado en RT-qPCR (forward y reverse)
<i>C19ORF47</i>	5' -CGATGGTCTCCGTGACTATGG 3' 5' -CGGCTGAGGACTTACTCTTGG 3'
<i>CCDC61</i>	5' -GCGAGTTCGATGCTGGCTTCAT 3' 5' -ACTCCAGGTCTGTGTAGGTCAG 3'
<i>CCDC97</i>	5' -ATCCTGCGAGAGGAGTTCACCA 3' 5' CAGGTTCTTCCTCATCAAAGTACC 3'
<i>LENG8</i>	5' -CCAAGTTTGGCAACCGCAACGT 3' 5' -AGTAGGAGCTGTCCGAGTCTGA 3'
<i>TMEM160</i>	5' -CTTGCCCGTCTTCGCTTCC 3' 5' -CGCCTCATCAGGGCCTTCC 3'

Inmunohistoquímica

Los bloques de tejidos de cáncer fijados en parafina se cortaron en secciones de 3 μ m, y fueron transferidos en portaobjetos electrocargados (BIOCare Medical, Concord, CA, USA), colocándolos sobre agua a 40°C, posteriormente fueron desparafinados en xilol y deshidratados en concentraciones decrecientes de alcohol (OH-Abs., OH-96% y OH-70%). La recuperación antigénica se realizó durante 10 min a 120°C en tampón de citrato pH 6.0 en la olla Pascal. El bloqueo se realizó con

BIOcare's peroxidized 1 por 5 min para la peroxidasa endógena. Se usaron los anticuerpos policlonales anti-C19ORF47 (abcam, No. Cat. Ab236288), anti-CCDC61 (MyBioSource, No. Cat. MBS9202921), anti-CCDC97(MyBioSource, No. Cat. MBS9202950), anti-LENG8 (MyBioSource, No. Cat. MBS416086) y anti-TMEM160 (MyBioSource, No. Cat. MBS8556351). Se usó el polímero MACH 1 Universal HRP-Polymer (No. Cat. MRH538L10) como anticuerpo secundario y el cromógeno diaminobenzidina (BIOcare's Betazoid DAB, No. Cat. BDB900G). Finalmente, se realizó la contra tinción con hematoxilina de Harris. La expresión de las proteínas se evaluó de acuerdo con la distribución y localización de la reacción positiva.

Análisis densitométrico (Inmunohistoquímica)

Se obtuvieron capturas de 5 campos representativos de cada laminilla para cada proteína, mediante un microscopio óptico (Leica DM1000 LED) acoplado a una cámara digital (Leica EC3). Las imágenes se analizaron en el Software Image-Pro Plus 6.0® (IPP6) (Media Cybernetics, Silver Spring, MD. USA). Los resultados se expresaron como el promedio de la densidad óptica relativa multiplicado por $1.00 \pm EE$ (medida en unidades arbitrarias de intensidad neta). Los datos fueron exportados al programa Excel para calcular los promedios de la densidad media de los 5 campos.

Análisis de expresión en GEPIA

A partir de la base de datos GEPIA la cual es una herramienta basada en la web que proporciona funciones interactivas y personalizables basadas en datos TCGA (Atlas del genoma del cáncer) y GTEx (Expresión de tejido del genotipo), se evaluó la expresión génica de *C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160* en los tejidos de carcinomas cervicales, de mama y ovario y se compararon contra tejido normal, paratumoral o adyacente ((<http://gepia.cancer-pku.cn/index.html>);Tang, Z. et al. 2017).

Análisis de expresión en The Human Protein Atlas

Para conocer en nivel de expresión a nivel de proteína por la técnica de inmunohistoquímica se obtuvieron imágenes representativas de la base de datos

The human Protein Atlas, en tejido normal y tumoral de mama, ovario y cervical. Cada imagen representa el nivel de tinción DAB, en una escala de bajo (<25% de las células teñidas) moderado (25%- 75% de las células teñidas) e intenso (>75% de las células teñidas). Las cuales se utilizaron como referencia sobre el nivel de expresión de cada proteína de este estudio <https://www.proteinatlas.org/>.

Análisis estadístico

Para la evaluación de la expresión de los genes *C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160* en pacientes y líneas celulares cáncer cervicouterino, mama y ovario, se utilizó la prueba estadística de ANOVA de una vía y la prueba de Tukey.

Para la evaluación de la expresión de las proteínas *C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160* en pacientes, se utilizó la prueba estadística de ANOVA de una vía y la prueba de Bonferroni. Los resultados con un valor de ($p \leq 0.05$), se consideraron estadísticamente significativos.

Resultados

Para determinar el perfil de expresión de los genes analizados (*C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160*) en los tumores de CaCu, CaMa y CaOv, se utilizó la base de datos GEPIA (<http://gepia.cancer-pku.cn/>) la cual proporciona funciones interactivas de la expresión diferencial entre tejidos normales y tumorales, en diferentes etapas patológicas. El análisis mostró que *C19ORF47* tiene una tendencia a la expresión alta en CaCu, CaMa y CaOv, aunque no significativa (Figura 1, a-c).

En este estudio se utilizaron 15 muestras de CaCu (9 correspondían a tejido tumoral, 4 Sin LEI y sin VPH y 2 Sin LEI- VPH+), 40 muestras de CaMa de las cuales 20 corresponden a tejido tumoral y 20 a tejido adyacente y 19 muestras de CaOv, del total 3 corresponden a OLC, 5 a COG y 11 a CEO. De acuerdo con la expresión de *C19ORF47*, en pacientes presenta una tendencia a la expresión alta en CaCu y CaOv, aunque no fue significativo (Figura 1 d-f), mientras que en las líneas celulares de CaCu, se observó la sobreexpresión en la línea celular HeLa en comparación con la línea control HaCaT ($p > 0.001$). En las líneas celulares de CaMa se observó sobreexpresión en la línea celular MCF7 en comparación con la línea control MCF10A ($p < 0.05$) (Figura 1j-k).

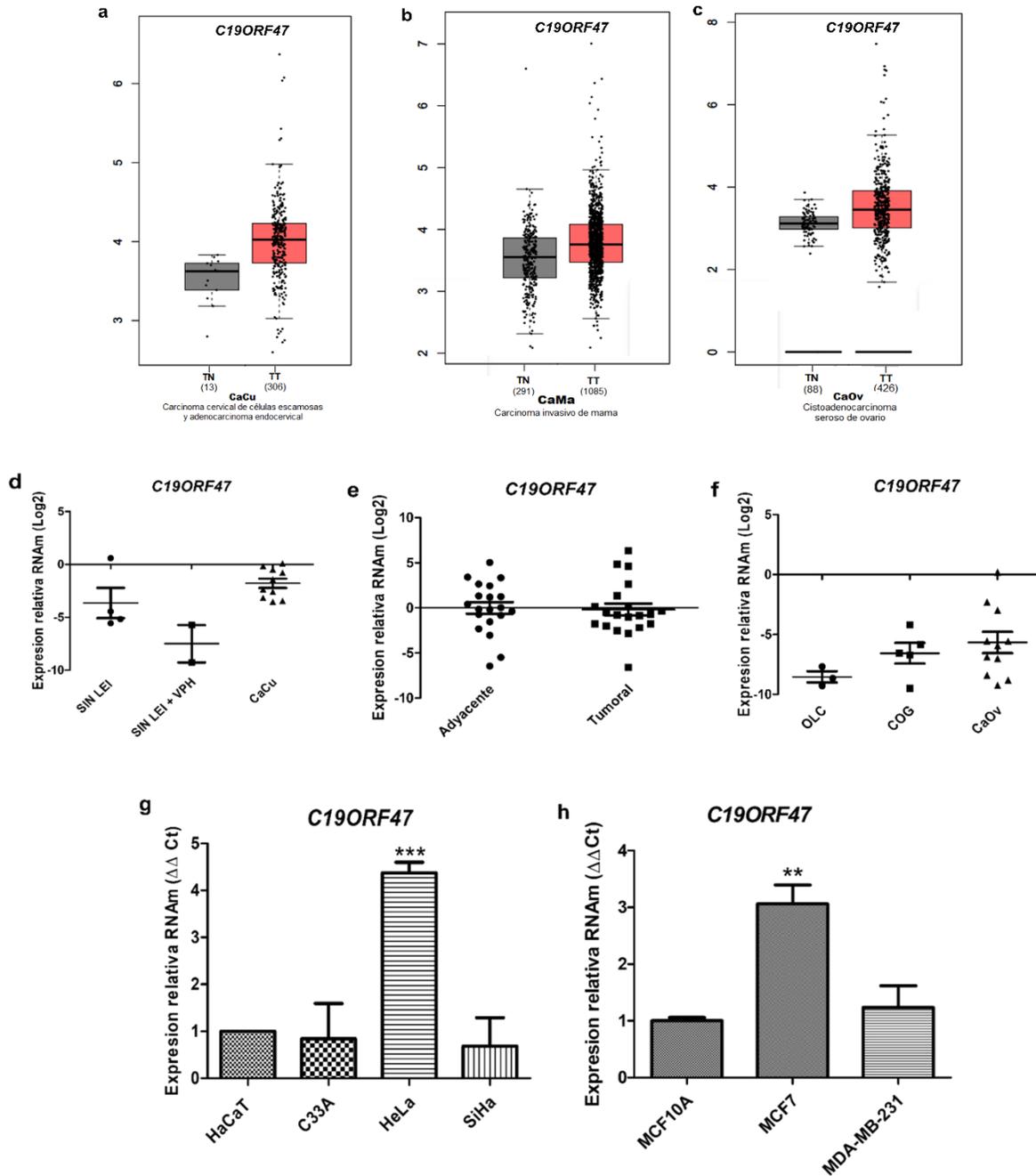


Figura 1. Expresión del gen *C19orf47* en líneas celulares y pacientes con cáncer cervicouterino, mama y ovario. a-c) Datos de expresión en base de datos GEPIA. d-f) Datos de expresión en pacientes con CaCu, CaMa y CaOv mediante Log2. g-h) Datos de expresión en líneas celulares de CaCu (C33A, HeLa y SiHa) y CaMa (MCF7 y MDA-MB-231) mediante $2^{-\Delta\Delta Ct}$. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

Para *CCDC61*, la base de datos GEPIA mostró en CaOv una expresión baja, en CaCu y CaMa en comparación con tejido normal, aunque no significativa (Figura 2, a-c), por otra parte, en las biopsias de pacientes analizados en este trabajo, con CaCu, CaMa y CaOv, no se observaron diferencias en el nivel de expresión en tejido tumoral y tejido normal (Figura 2, d-f). En las líneas celulares de CaCu, *CCDC61* tiene una subexpresión, en las células C33A, HeLa y SiHa respecto a HaCaT ($p < 0.001$); mientras que en las líneas celulares MCF7 y MDA-MB-231 presenta una expresión alta, aunque no significativa (Figura 2, g-h).

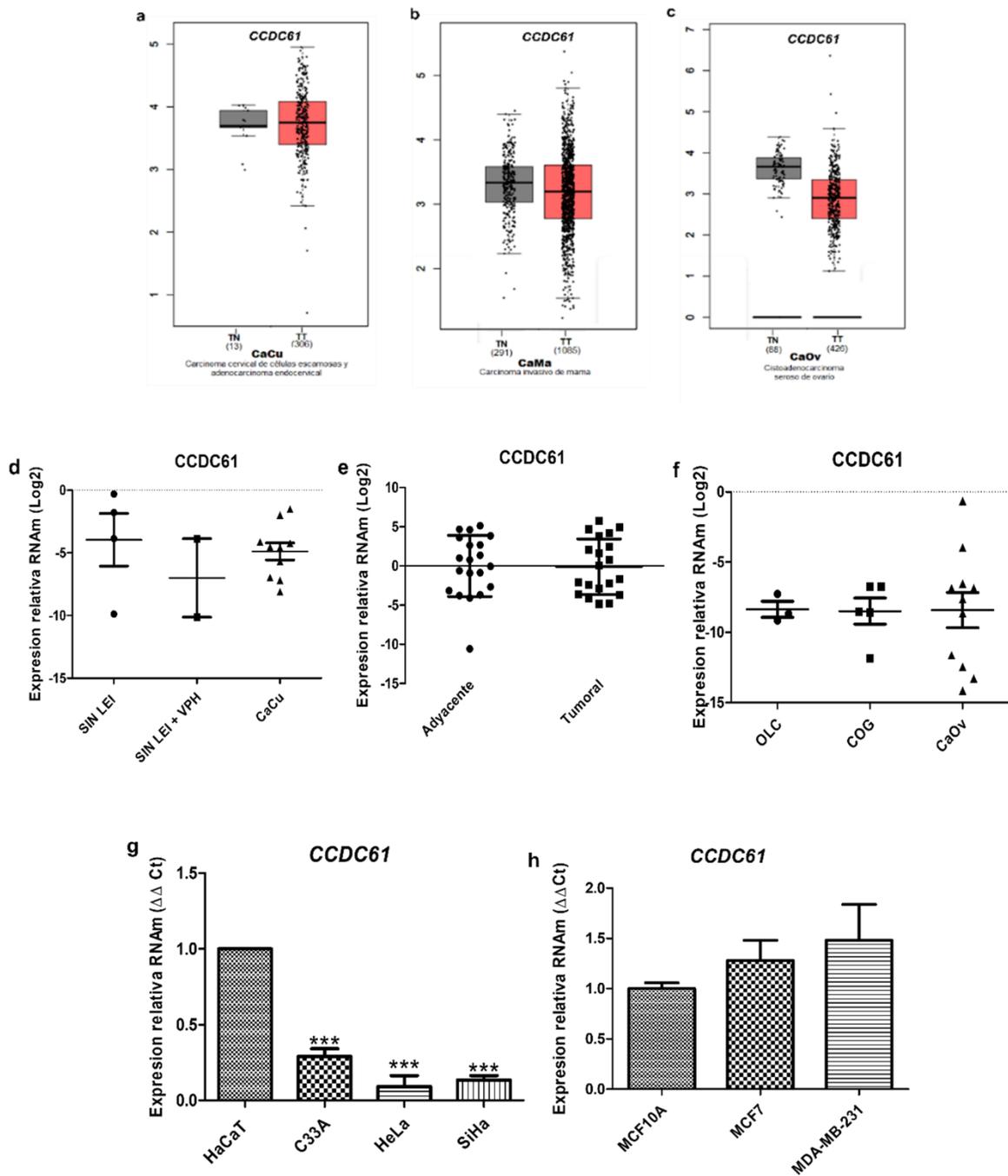


Figura 2. Expresión del gen *CCDC61* en pacientes y líneas celulares de cáncer cervicouterino, cáncer de mama y cáncer de ovario. a-c) análisis de expresión en la base de datos GEPIA. **d-f)** Análisis *in vivo* en pacientes con CaCu, CaMa y CaOv mediante Log2. **g-h)** Análisis *in vitro* en líneas celulares de CaCu (C33A, HeLa y SiHa) y CaMa (MCF7 y MDA-MB-231) mediante $2^{-\Delta\Delta Ct}$. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

Para el gen *CCDC97*, la base de datos GEPIA muestra una tendencia a la expresión alta, aunque no significativa en CaMa ($p>0.05$), en comparación con CaCu y CaOv (Figura 3, a-c), sin embargo, en las biopsias analizadas en este trabajo en CaCu, CaMa y CaOv no se encontraron diferencias significativas entre el nivel de expresión en tejido tumoral y tejido normal (Figura 3, d-f). En las líneas celulares C33A y SiHa, el nivel de expresión de *CCDC97* es bajo, mientras que en HeLa se observa una sobreexpresión ($p<0.001$). En las líneas celulares de CaMa se observó un nivel de expresión alto en la línea celular MCF7 en comparación con MCF10A (Figura 3, g-h).

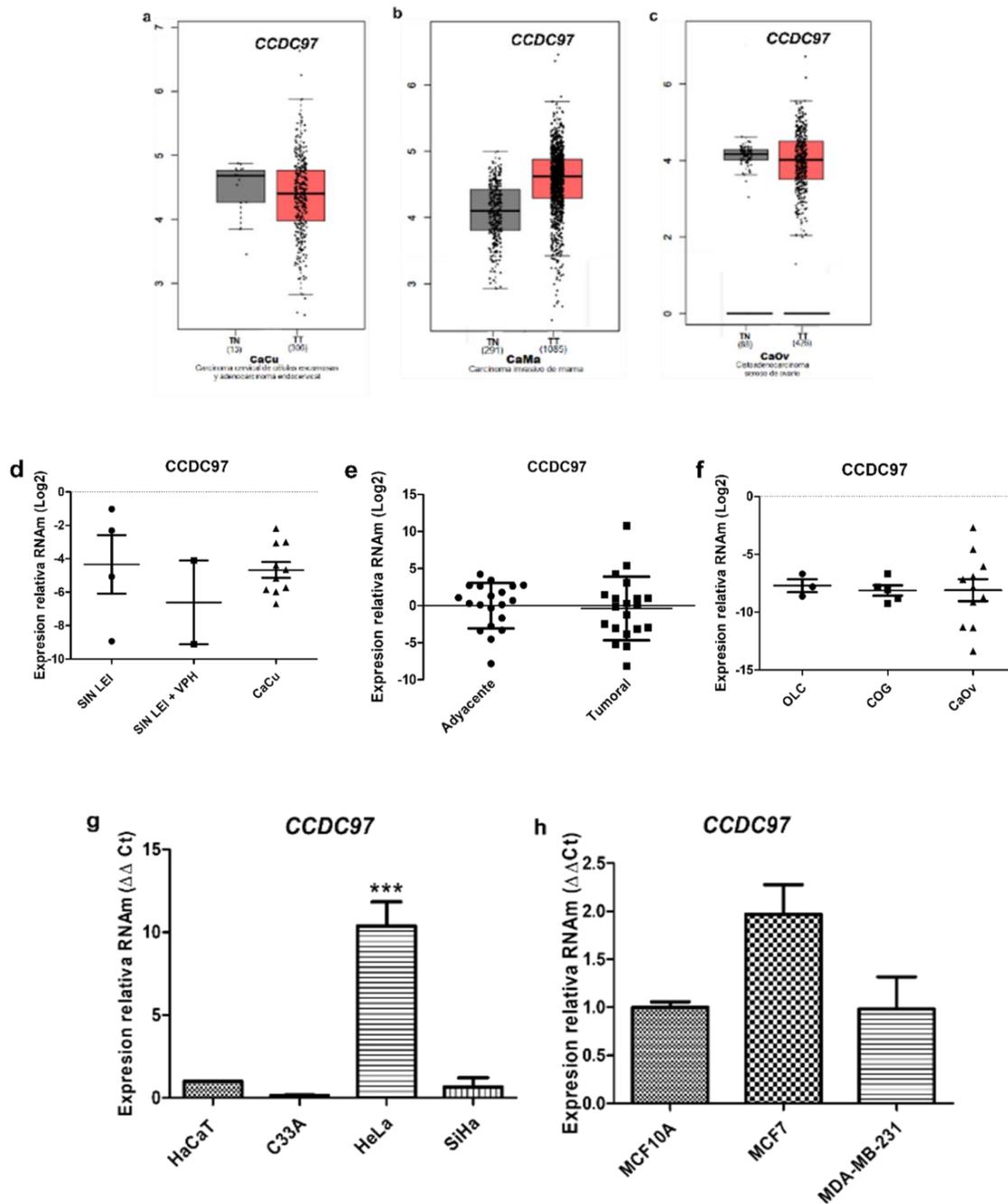


Figura 3. Expresión del gen *CCDC97* en pacientes y líneas celulares de cáncer cervicouterino, cáncer de mama y cáncer de ovario. a-c) Análisis de expresión en la base de datos GEPIA. d-f) Análisis *in vivo* en pacientes con CaCu, CaMa y CaOv mediante Log2. g-h) Análisis *in vitro* en líneas celulares de CaCu (C33A, HeLa y SiHa) y CaMa (MCF7 y MDA-MB-231) mediante 2- $\Delta\Delta$ Ct. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

En el análisis de datos de expresión para el gen *LENG8* en la base de datos GEPIA, se identificó una subexpresión en CaCu, CaMa y CaOv ($p \leq 0.05$), en comparación con el tejido normal o adyacente (Figura 4, a-c). Sin embargo, en las muestras de cáncer analizadas, se observó en los tres tipos de cáncer una expresión disminuida, en CaCu y CaMa, y una expresión alta en CaOv (Figura, 4 d-f), sin embargo, no significativa. En la línea celular C33A, *LENG8* está sobreexpresado, en comparación con la línea control HaCaT ($p \leq 0.05$) y una expresión alta en las líneas celulares MCF7 y MDA-MB-231 en comparación con la línea control MCF10A, aunque no significativa (Figura 4, g-h).

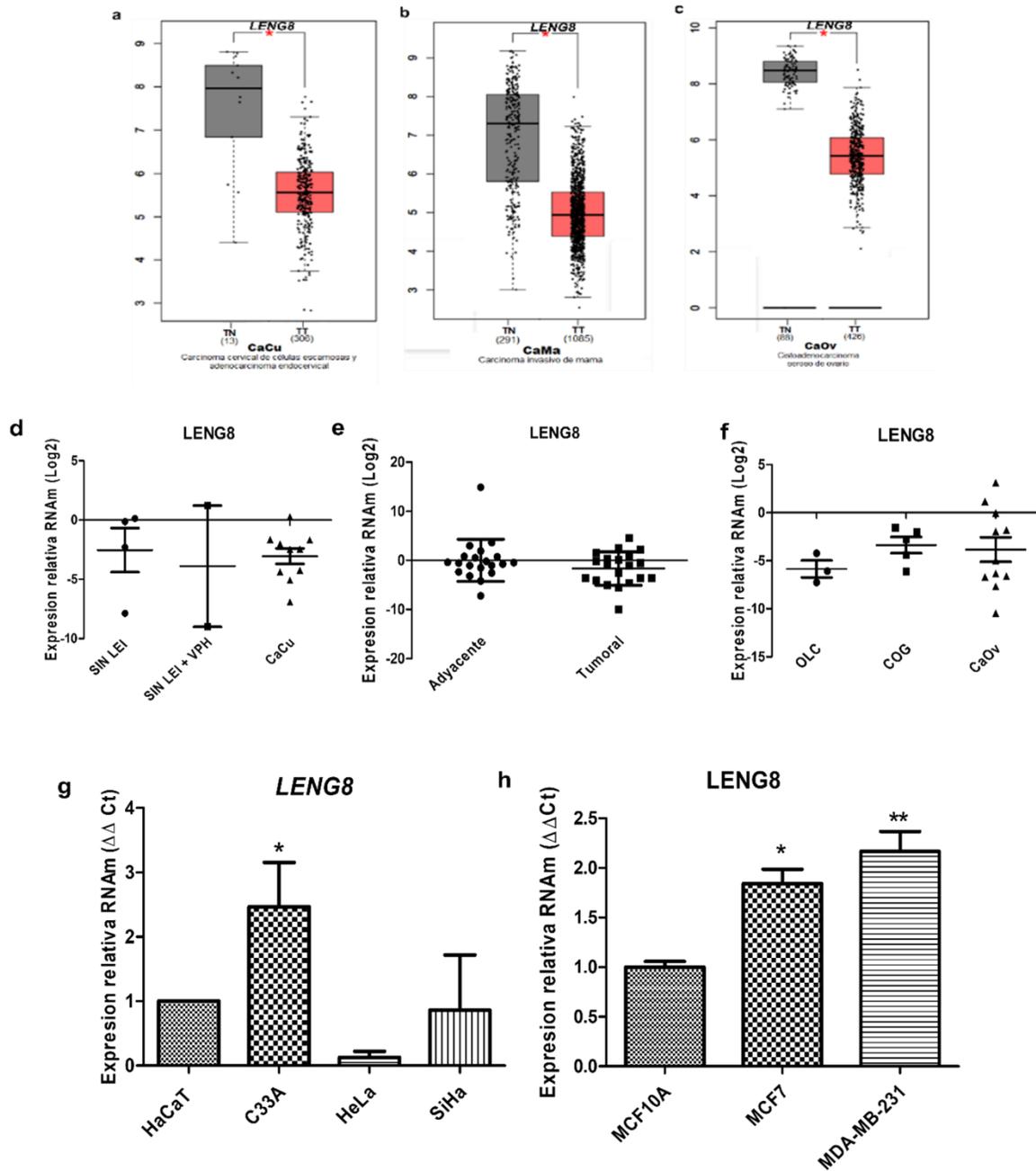


Figura 4. Expresión del gen *LENG8* en pacientes y líneas celulares de cáncer cervicouterino, cáncer de mama y cáncer de ovario. a-c) Análisis de expresión en la base de datos GEPIA. d-f) Análisis *in vivo* en pacientes con CaCu, CaMa y CaOv mediante Log2. g-h) Análisis *in vitro* en líneas celulares de CaCu (C33A, HeLa y SiHa) y CaMa (MCF7 y MDA-MB-231) mediante $2^{-\Delta\Delta Ct}$. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

Para *TMEM160*, se observa una sobreexpresión en CaOv, respecto a CaCu y CaMa en GEPIA ($p < 0.05$) (Figura 5, a-c), mientras que, en las biopsias analizadas, se observó una expresión alta en CaCu y CaOv, aunque no significativa (Figura 5 d-f). En las líneas celulares de CaCu se observó un mayor nivel de expresión en HeLa ($p < 0.001$) en comparación con la línea control HaCaT. En las líneas celulares derivadas de cáncer de mama, *TMEM160* presentó una sobreexpresión en las líneas celulares MCF7 y MDA-MB-231 en comparación con la línea control MCF10A ($p < 0.05$) (Figura 5, g-h).

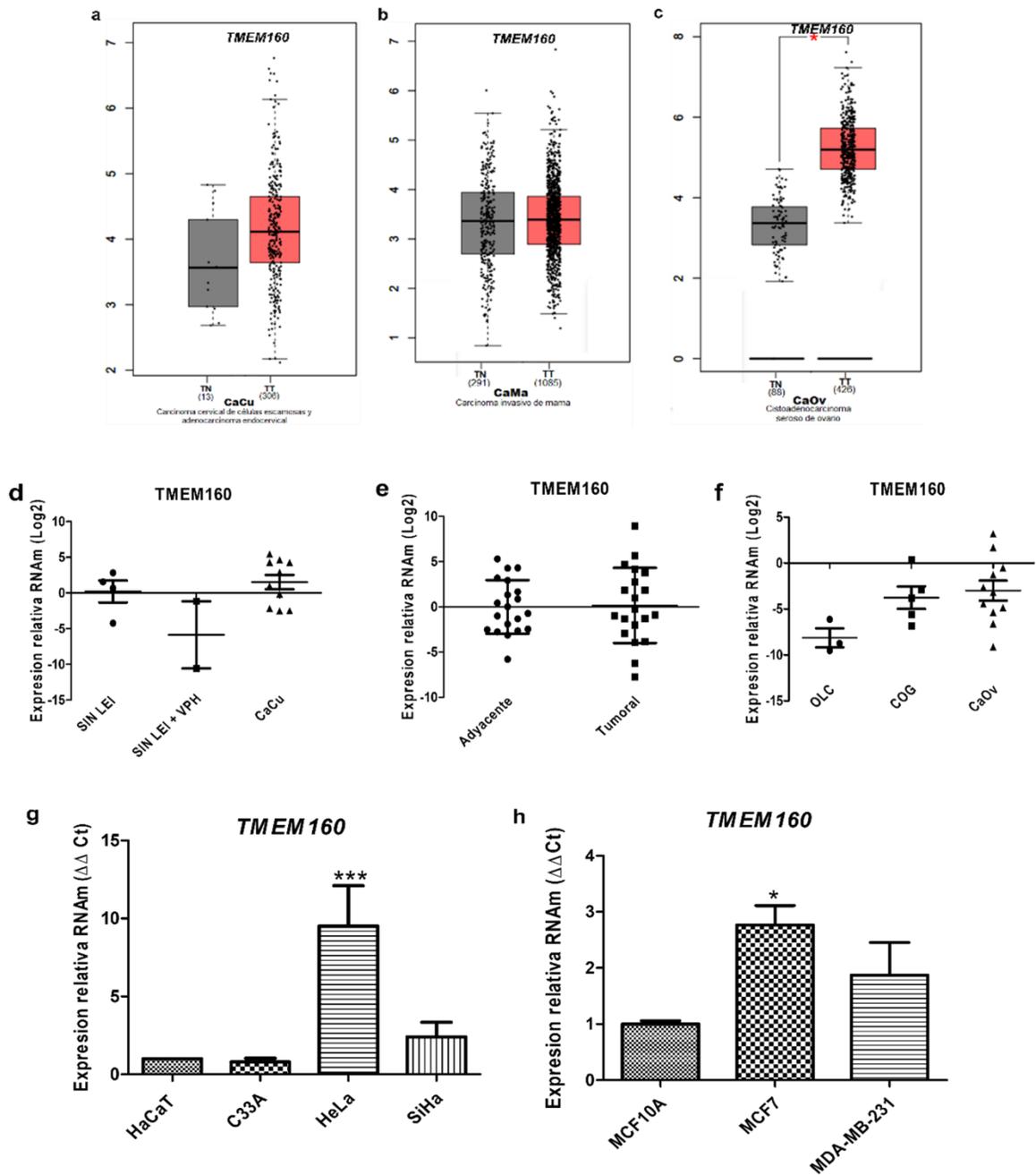


Figura 5. Expresión del gen *TMEM160* en pacientes y líneas celulares de cáncer cervicouterino, cáncer de mama y cáncer de ovario. a-c) Análisis de expresión en la base de datos GEPIA. d-f) Análisis *in vivo* en pacientes con CaCu, CaMa y CaOv mediante Log2. g-h) Análisis *in vitro* en líneas celulares de CaCu (C33A, HeLa y SiHa) y CaMa (MCF7 y MDA-MB-231) mediante $2^{-\Delta\Delta Ct}$. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

Una vez obtenidos el análisis de expresión individual por gen, se realizó el análisis grupal, mediante $2^{-\Delta\Delta CT}$. En CaCu se observó que los genes *C19ORF47* y *TMEM160* están sobreexpresados ($p < 0.05$), en CaOv sólo *TMEM160* está sobreexpresado ($p < 0.05$) (figura 6).

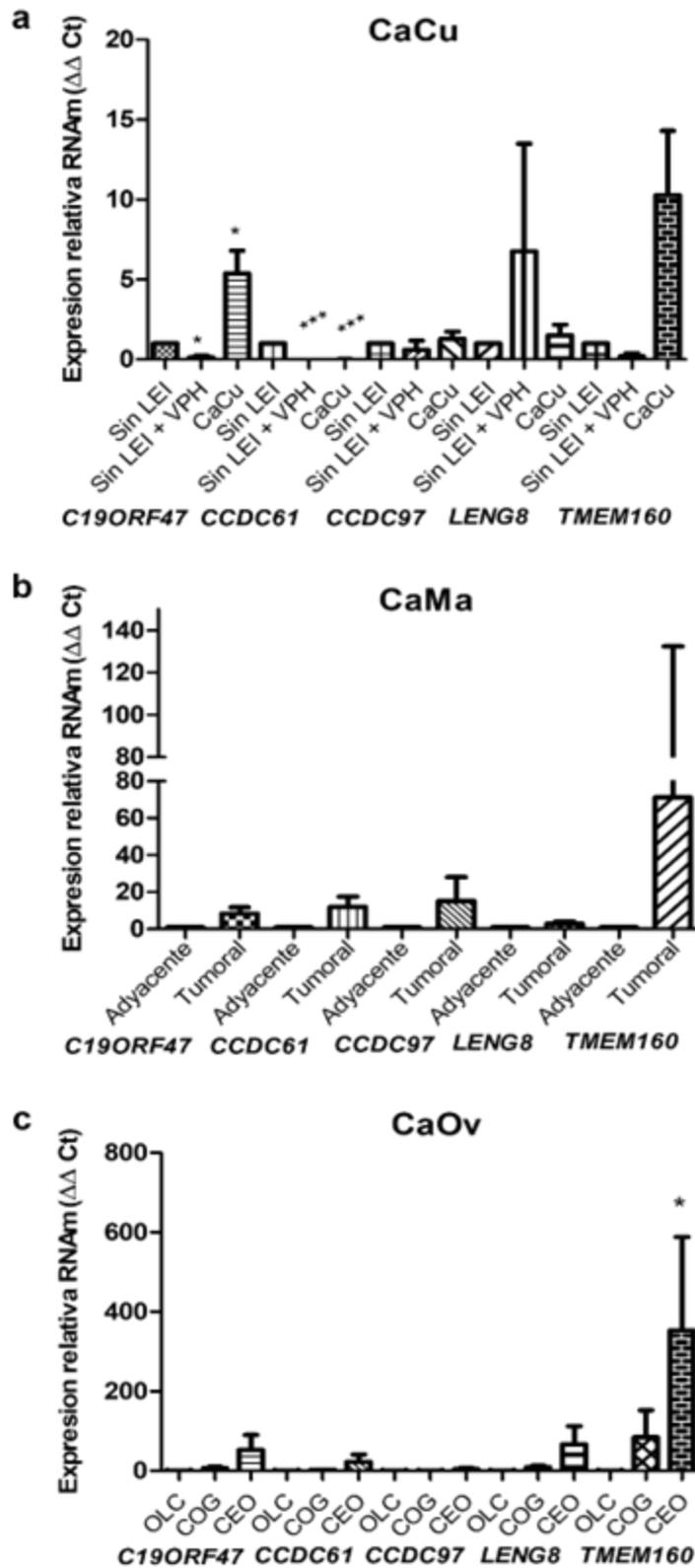


Figura 6. Comportamiento general de la expresión de los genes *C19ORF47*, *CCDC61*, *CCDC97*, *LENG8* y *TMEM160* mediante $2^{-\Delta\Delta Ct}$ en pacientes con a) CaCu b) CaMa y c) CaOv. Sin LEI: Sin lesión escamosa intraepitelial, Sin LEI + VPH: sin lesión escamosa intraepitelial positiva a infección por el

virus del papiloma humano, CaCu: cáncer cervicouterino, CaMa: cáncer de mama, TN: tejido normal, TT: tejido tumoral, CaOv: cáncer de ovario, OLC: ovario libre de cáncer, COG: teratoma benigno, CEO: tejido tumoral de ovario.

Expresión de C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160 en cáncer cervicouterino, cáncer de mama y cáncer de ovario a nivel de proteína

El nivel proteico de C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160, se analizó por medio de inmunohistoquímica, en pacientes con CaCu (carcinoma de células escamosas), CaMa (carcinoma ductal infiltrante y sólo una de carcinoma lobulillar) y CaOv (Adenocarcinoma seroso papilar y sólo una de cistadenocarcinoma mucinoso) y los datos de expresión fueron comparadas entre los tipos de cáncer en función del número de células positivas al DAB. Estos resultados se compararon con los datos de expresión proteica con la base de datos The Human Protein Atlas <https://www.proteinatlas.org/>.

Los datos de expresión proteica analizadas en la base The Human Protein Atlas muestran que las proteínas CCDC61 y CCDC97 fueron reportadas como negativas para la expresión de las proteínas tanto en tejido normal como en cáncer, los resultados coinciden con el análisis en los 4 tejidos de CaMa y CaOv de este proyecto, debido a que la señal no fue detectada (Figura 7 y 8).

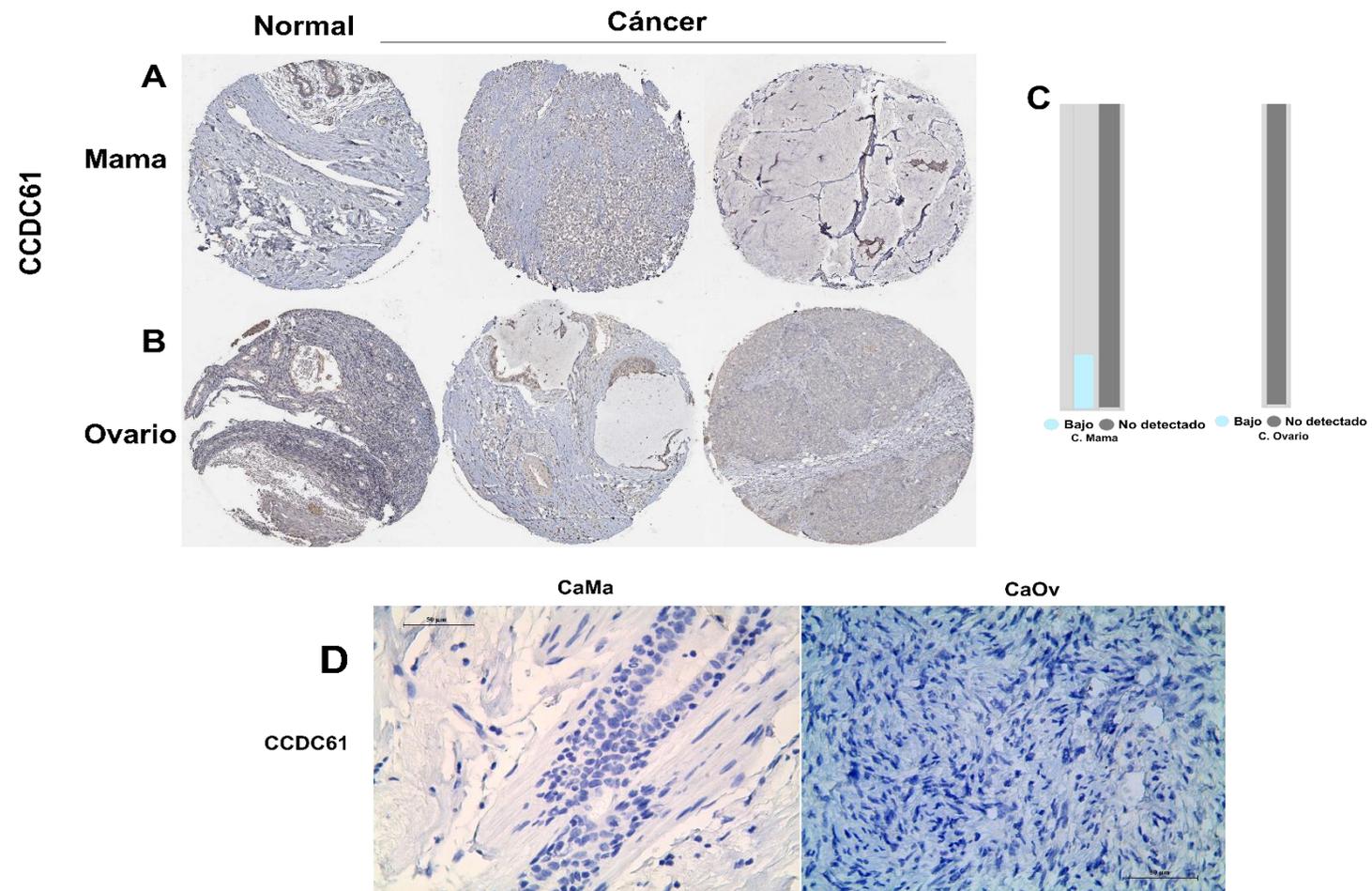


Figura 7. Datos de expresión de CCDC61, base de datos The Human Protein Atlas. A) Cáncer cervicouterino, B) Cáncer de mama, C) Cáncer de ovario., las imágenes de muestran en objetivo 10x, D) Datos semicuantitativos de la fracción de células teñidas (<25%, 25-75% o >75%).

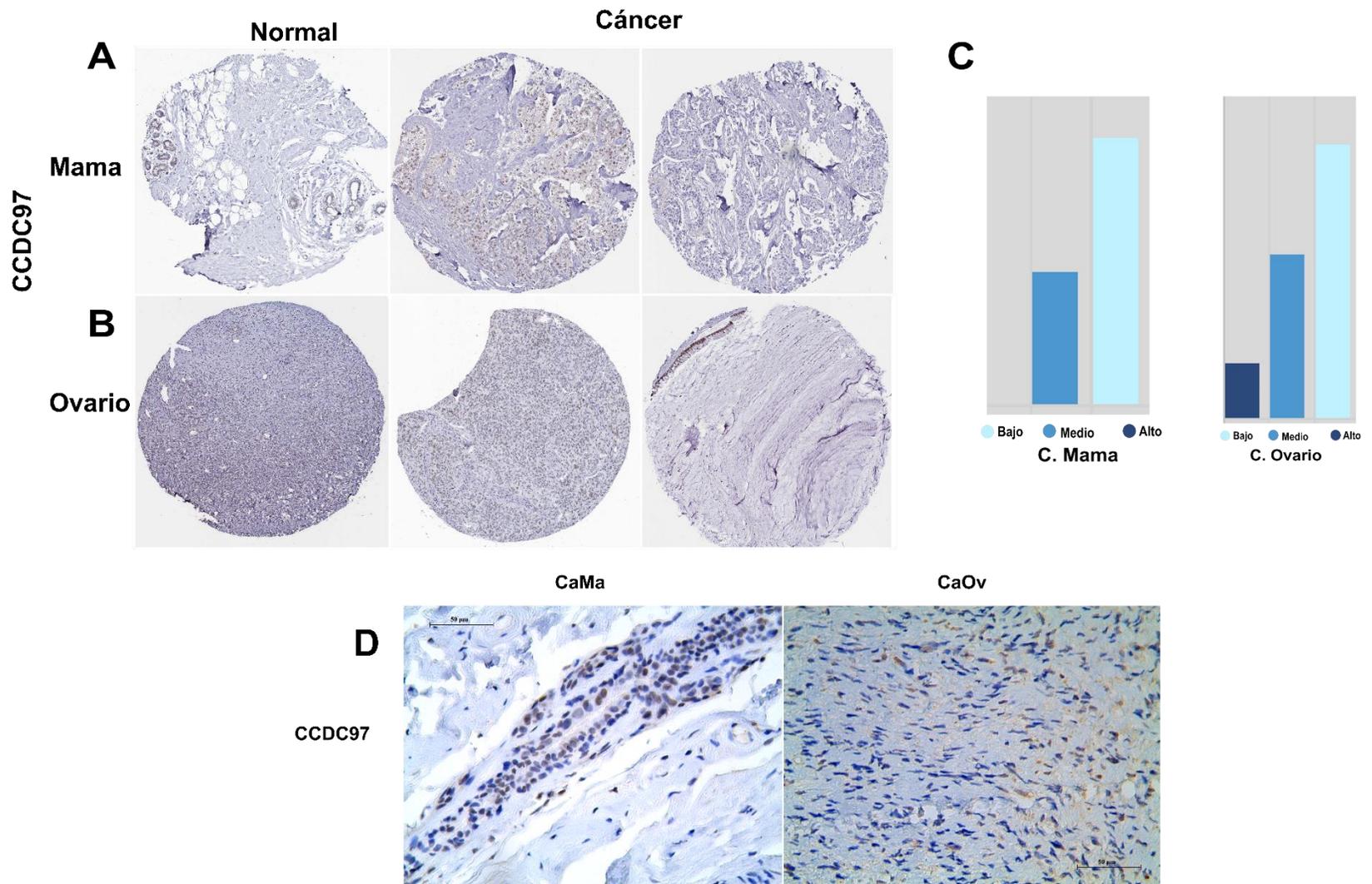


Figura 8. Datos de expresión de CCDC61, base de datos The Human Protein Atlas. A) Cáncer cervicouterino, B) Cáncer de mama, C) Cáncer de ovario., las imágenes de muestran en objetivo 10x, D) Datos semicuantitativos de la fracción de células teñidas (<25%, 25-75% o >75%).

Los datos de expresión proteica para C19ORF47 representados en la base de datos muestran que en CaCu, se observa tinción positiva de débil-fuerte en el >25-75% de las células, con distribución subcelular en citoplasma y núcleo en pacientes con carcinoma de células escamosas y adenocarcinoma. En CaMa, se observa tinción de negativa a moderada de <25-75% de las células, con distribución subcelular en citoplasma y núcleo tanto en carcinoma ductal como en lobulillar. Para CaOv, la tinción fue de negativa a moderada en el <25-75% de las células, con distribución subcelular en núcleo y citoplasma en adenocarcinomas de tipo seroso papilar y mucinoso (Figura 9). Los resultados obtenidos en este trabajo muestran que en CaCu la expresión es de negativa a fuerte, con localización en citoplasma, en CaMa la tinción se observa de débil a fuerte, con localización en citoplasma, y en CaOv de débil a moderada con localización subcelular en citoplasma (Figura 10).

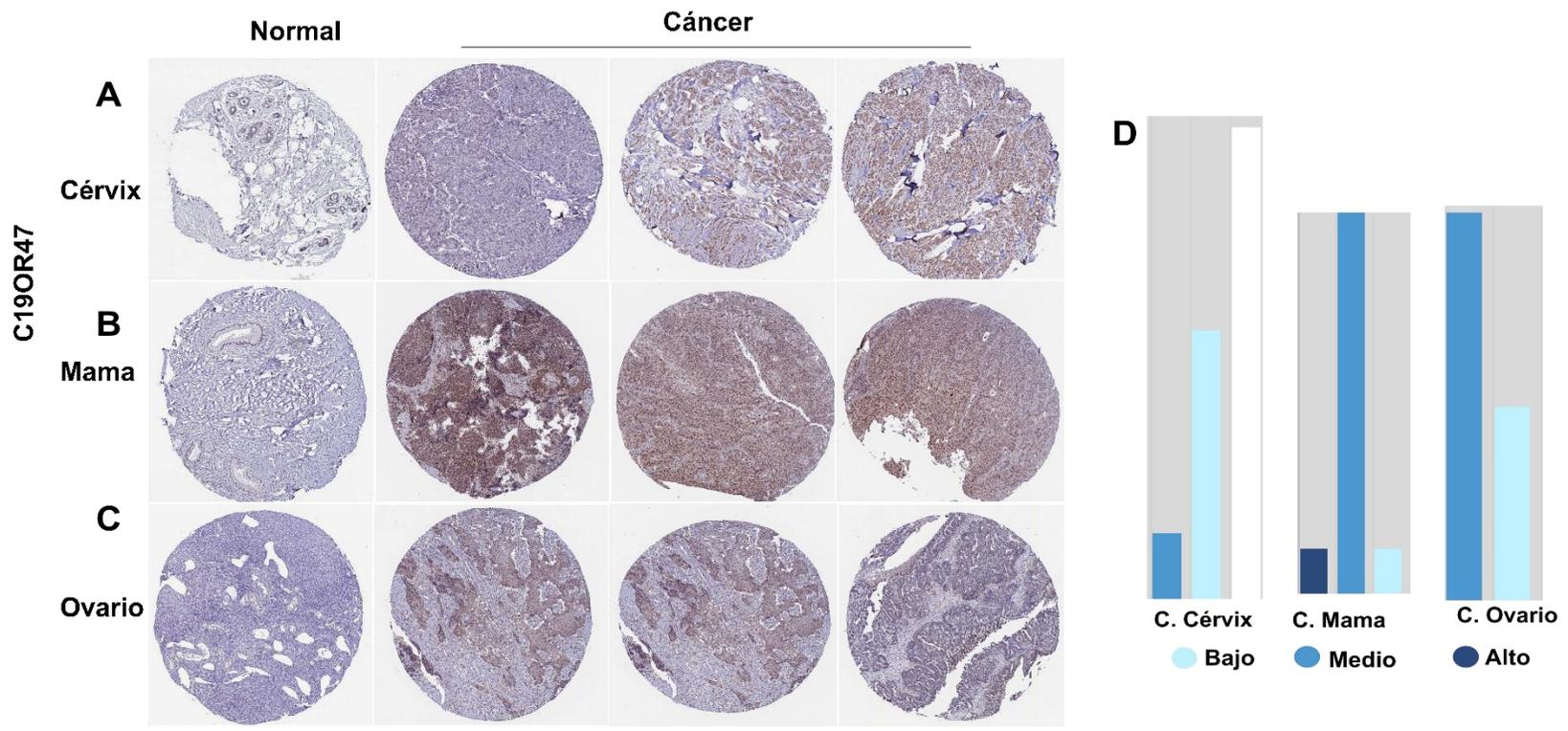


Figura 9. Datos de expresión de C19ORF47, base de datos The Human Protein Atlas. A) Cáncer cervicouterino, B) Cáncer de mama, C) Cáncer de ovario., las imágenes de muestran en objetivo 10x, D) Datos semicuantitativos de la fracción de células teñidas (<25%, 25-75% o >75%).

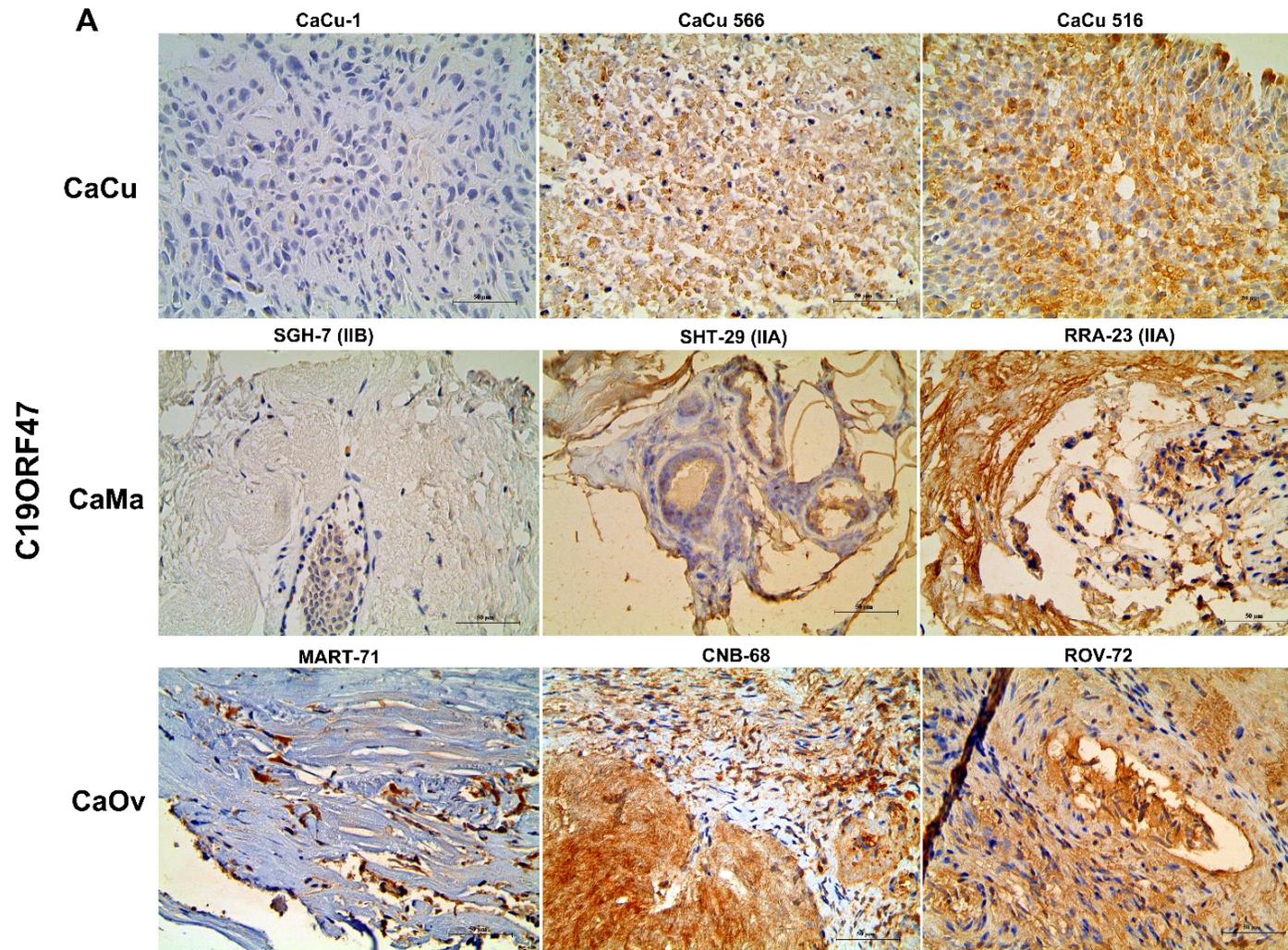


Figura 10. Expresión de C19ORF47 en biopsias de cáncer cervicouterino, mama y ovario. En CaCu la expresión de C19ORF47 (color marrón) se observó principalmente en citoplasma, en CaMa la expresión se observó en el citoplasma de la mayoría de las células estromales y en células mioepiteliales de los ductos mamarios, en CaOv la expresión se observó en el citoplasma de la mayoría de las células.

Los datos de expresión proteica para LENG8 obtenidos en la base The Human Protein Atlas muestran que en tejido normal de cérvix, mama y ovario es negativa, en CaCu, se observa tinción de negativa a moderada en el <25-75% de las células, con distribución subcelular en citoplasma, en pacientes con carcinoma de células escamosas y adenocarcinoma. Para el caso de CaMa, se observa tinción positiva de débil-fuerte en 25>75% de las células, con distribución subcelular en citoplasma tanto en carcinoma ductal como en lobulillar. Para CaOv, la tinción positiva fue de débil-moderada en más del 75% de las células, con distribución subcelular en membrana y citoplasma en adenocarcinomas de tipo seroso papilar y mucinoso (Figura 11). Los resultados obtenidos en este proyecto mostraron que la intensidad de tinción en CaCu es de baja a moderada, en CaMa es de moderada a fuerte y en CaOv es de baja a moderada, con localización subcelular en citoplasma (Figura 12).

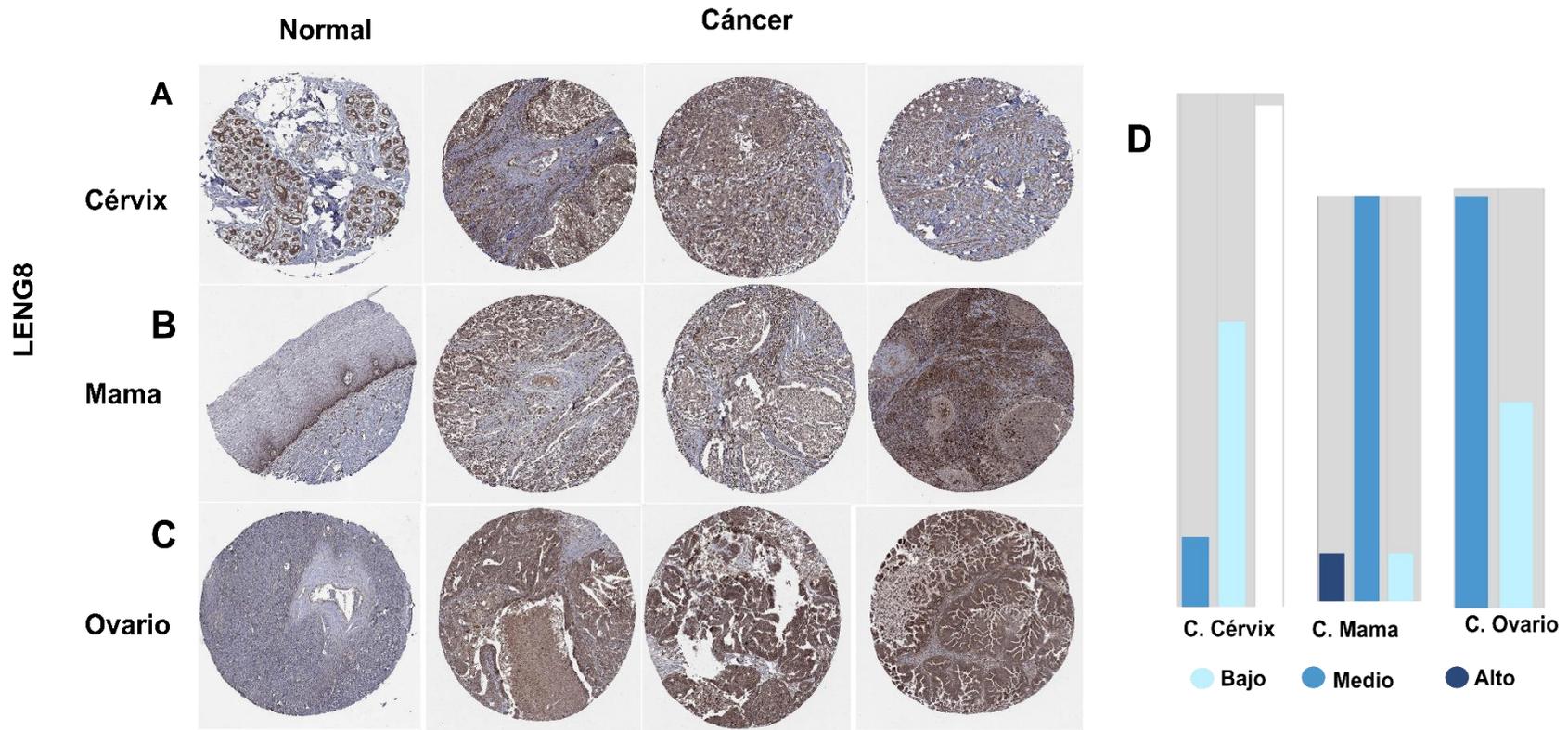


Figura 11. Datos de expresión de LENG8, base de datos The Human Protein Atlas. A) Cáncer cervicouterino, B) Cáncer de mama, C) Cáncer de ovario., las imágenes de muestran en objetivo 10x, D) Datos semicuantitativos de la fracción de células teñidas (<25%, 25-75% o >75%).

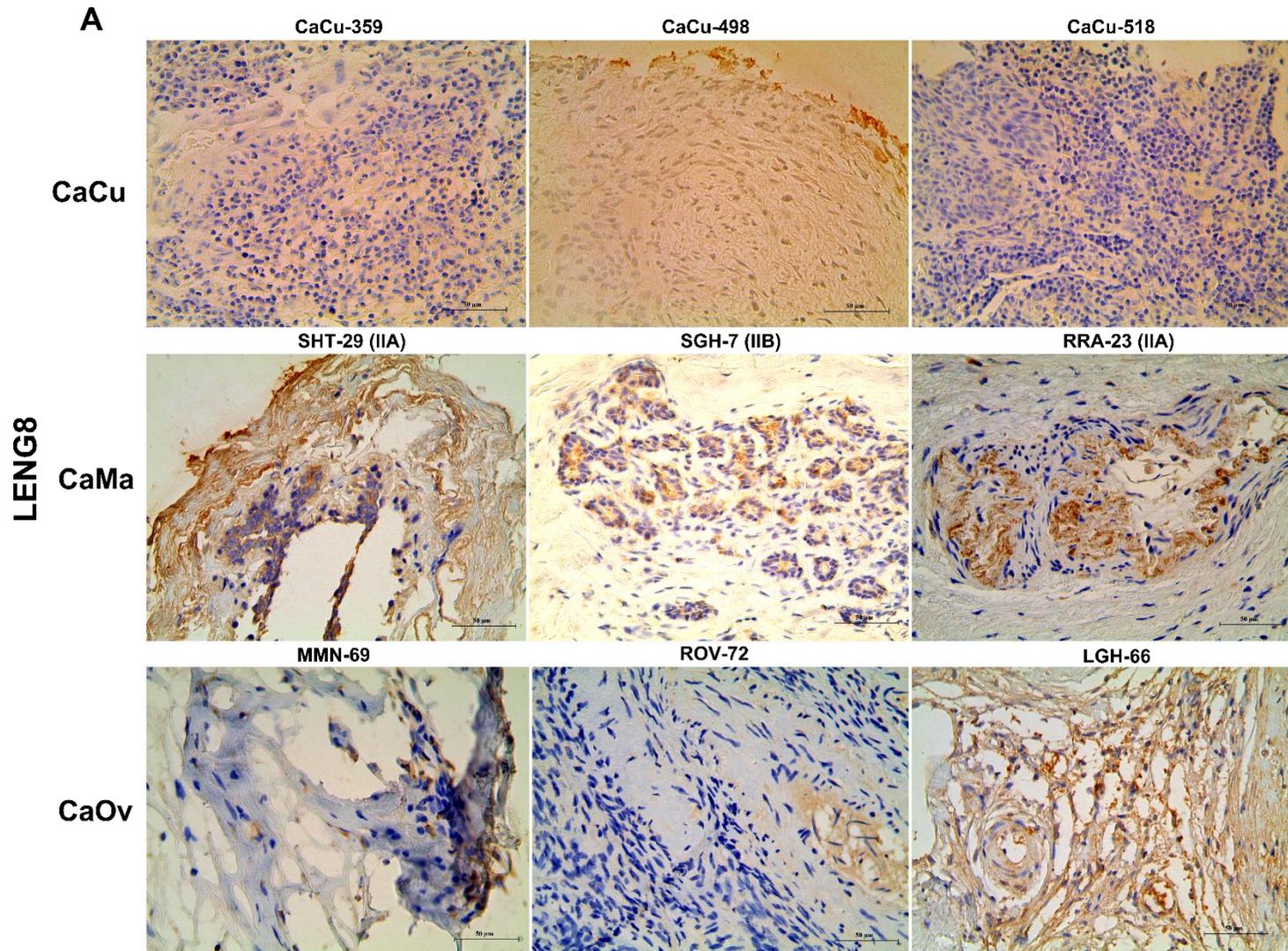


Figura 12. Expresión de LENG8 en biopsias de cáncer cervicouterino, mama y ovario. En CaCu la expresión de LENG8 (color marrón) se observó principalmente en citoplasma, en CaMa la expresión se observó en el citoplasma de las células mioepiteliales de los ductos mamarios, en CaOv la expresión se observó en el citoplasma de las células.

Los datos de expresión proteica para TMEM160 representados en la base de datos muestran que para el caso de CaCu, se observa tinción positiva de débil-alta en el 25-75% de las células, con distribución subcelular en membrana y citoplasma en pacientes con carcinoma de células escamosas. Para el caso de CaMa, se observa tinción positiva de moderada-alta en 25-75% de las células, con distribución subcelular en membrana y citoplasma tanto en carcinoma ductal como en lobulillar. Para CaOv, la tinción positiva fue de moderada-fuerte en más del 75% de las células, con distribución subcelular en membrana y citoplasma en adenocarcinomas de tipo seroso papilar y mucinoso (Figura 13). Los resultados obtenidos en el grupo de trabajo muestran que la expresión en CaCu es de negativa-a moderada, con localización subcelular en membrana y citoplasma, en CaMa de moderada a fuerte y en CaOv con intensidad baja-fuerte con localización en membrana y citoplasma (Figura 14).

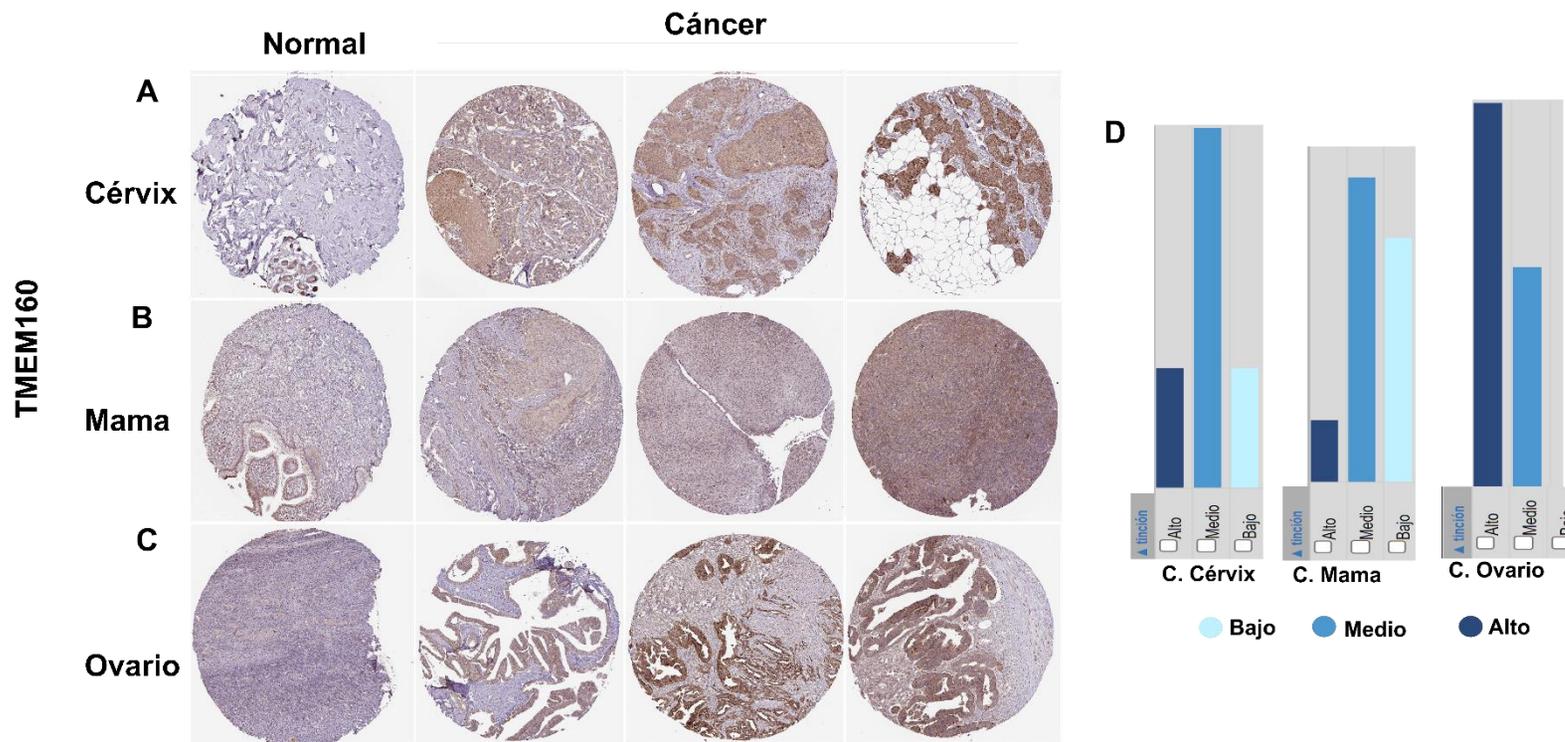


Figura 13. Datos de expresión de TMEM160, base de datos The Human Protein Atlas. A) Cáncer cervicouterino, B) Cáncer de mama, C) Cáncer de ovario., las imágenes de muestran en objetivo 10x, D) Datos semicuantitativos de la fracción de células teñidas (<25%, 25-75% o >75%).

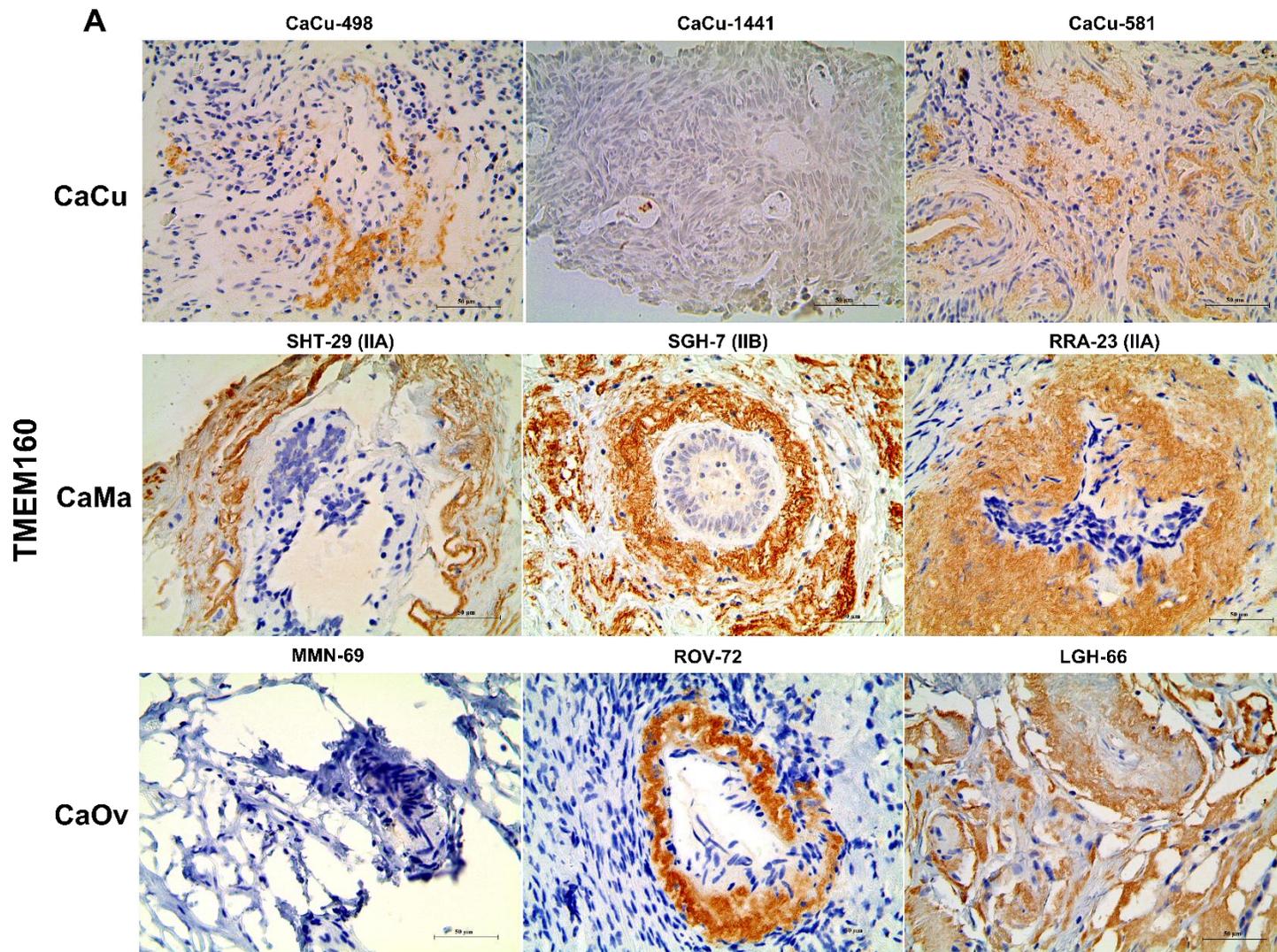


Figura 14. Expresión de TMEM160 en biopsias de cáncer cervicouterino, mama y ovario. En CaCu la expresión de TMEM160 (color marrón) se observó principalmente en membrana, en CaMa la expresión se observó principalmente en la membrana de los ductos mamarios, en CaOv la expresión se observó en membrana.

Los datos de intensidad en el total de las muestras al ser comparadas entre los diferentes tipos de cáncer, se observa que para C19ORF47, la expresión es similar entre los tres tipos de cáncer, mientras que en LENG8 se observa una mayor expresión en cáncer de mama ($p < 0.05$) y ovario ($p \leq 0.0092$) con respecto a CaCu. Para TMEM160 se observó una diferencia mayor de expresión en mama ($p < 0.05$) en comparación con CaCu (Figura 15).

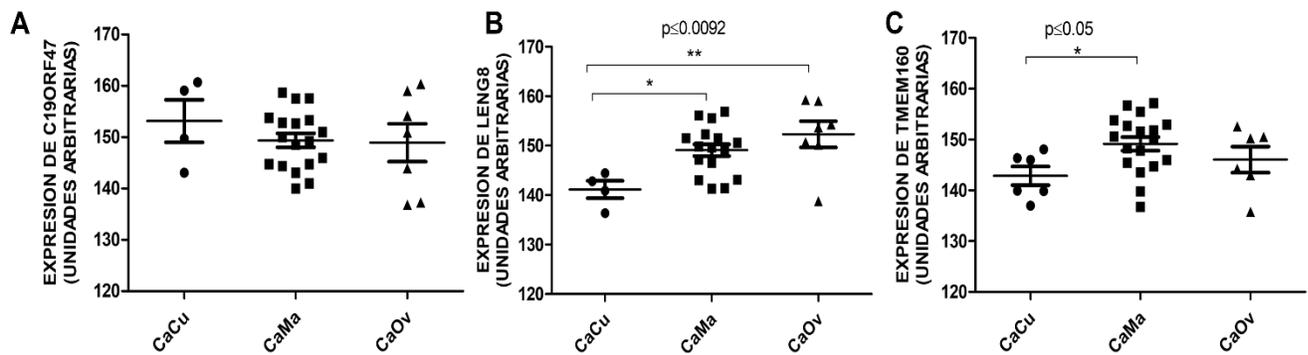


Figura 15. Análisis de la expresión de C19ORF47, LENG8 y TMEM160 en cáncer cervicouterino, mama y ovario. Los gráficos de dispersión muestran los valores promedio de intensidad de cada muestra analizada, la comparación se realizó entre los tipos de cáncer. Los valores significativos son considerados ($p < 0.05$).

Discusión

El proyecto piloto C-HPP lanzado en 2018 tiene como objetivo caracterizar proteínas PE1 funcionalmente no caracterizadas pero identificadas (proteínas uPE1) que corresponden aproximadamente al 10% del “dark side”. Las proteínas C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160 localizadas en el cromosoma 19, fueron identificadas por medio de espectrometría de masa, en líneas celulares de cáncer cervical.

La caracterización de las proteínas sin función asignada es muy importante en la investigación biológica, así como en el diagnóstico de enfermedades, en la búsqueda de nuevos blancos terapéuticos, en el diseño de fármacos y en la generación de conocimiento sobre los procesos biológicos involucrados en la progresión de enfermedades como el cáncer. En este sentido, uno de los primeros enfoques ha sido el análisis de la expresión de proteínas sin función asignada, en cánceres con elevado índice de incidencia y mortalidad como el cérvicouterino, mama y ovario, como una estrategia para el esclarecimiento de su función biológica y localización subcelular. Sin embargo, la falta de información, sobre las características y funciones en las que las proteínas C19ORF47, CCDC61, CCDC97, LENG8 y TMEM160 podrían estar implicadas aún se desconoce, por lo que los resultados obtenidos en este proyecto resultan limitados en el sentido de la escasez de datos comparativos. Una de las estrategias utilizadas fue el uso de bases de datos masivos como GEPIA y The Human Protein Atlas, como herramientas de referencia, sin embargo, los datos de expresión reportados cuentan con limitada información clínica y la comparación se realiza contra tejido paratumoral (adyacente) sin una clasificación histopatológica/molecular clara, proveniente de datos proteómicos, y transcriptómicos, considerando un número limitado de tejidos (proteínas) o con el manejo de los tipos histológicos más frecuentes. En este trabajo los histotipos analizados fueron para CaCu (carcinoma de células escamosas), CaOv (adenocarcinoma seroso papilar) y CaMa (carcinoma ductal), las comparaciones, aunque limitadas en término de número de muestras fueron entre tejido normal y adyacente proveniente de diferentes pacientes, sin embargo, la

comparación entre tejidos de mama fue entre tejido tumoral y adyacente del mismo paciente en el análisis de la expresión génica.

Los datos de expresión proteica en la base de datos The Human Protein Atlas muestran la inmunorreactividad presente en los diferentes tipos de células presentes en los tejidos normales o cancerosos. Los parámetros básicos de anotación incluyen el análisis de la intensidad de la tinción (negativa, débil, moderada o fuerte), el porcentaje de células teñidas (<25%, 25-75% o >75%) y la localización subcelular (nuclear y/o citoplasmática/membranosa). Estos mismos criterios fueron considerados en estos resultados. El análisis comparativo de los resultados obtenidos y los datos mostrados en las bases de datos, muestran que, para las proteínas CCDC61 y CCDC97, los niveles de expresión reportados a nivel de mensajero y proteína coinciden con los resultados obtenidos en este trabajo, para los tres tipos de cánceres CaCu, CaMa y CaOv, que se analizaron en este estudio. Aun cuando existe evidencia de la participación de estas proteínas o miembros de sus familias en funciones biológicas como la división celular, el eflujo de fármacos, y la regulación de la expresión génica (Frezza y Montclare., 2015; McFarlane *et al.*, 2018), además de asociarse con el desarrollo y la progresión de varios tumores generando metástasis en cánceres como el de tiroides (Yin *et al.*, 2016), en carcinoma de células escamosas de esófago (Hu *et al.*, 2018), y en cáncer gástrico (Wang *et al.*, 2020). Estas proteínas no se observaron que puedan estar asociadas a estos tipos de cáncer.

Sin embargo, para el análisis de la expresión de C19ORF47, los resultados observados en GEPIA y The Human Protein Atlas, son similares a los encontrados en este estudio, donde se observa que los niveles de expresión tanto de mensajero como de proteína aumentan en los tres tipos de cáncer en comparación con tejido normal, su localización subcelular es principalmente en citoplasma, mientras que en CaCu, el nivel de expresión parece aumentar en etapas clínicas más avanzadas de la enfermedad. The Human Protein Atlas, aunque en esta base de datos reportan la localización subcelular en citoplasma y núcleo, sólo se observó la expresión en

citoplasma. No hay información reportada hasta el momento sobre su participación en procesos biológicos.

LENG8 conocido como el miembro 8 del grupo de receptores de leucocitos humanos que se albergan en la región cromosómica 19q13.4, hasta ahora su papel biológico sigue siendo desconocido, de acuerdo a lo analizado en las bases de datos GEPIA y The Human Protein Atlas, a nivel de mensajero se encuentran disminuidos, sin embargo, el nivel proteico de expresión es alto en los tres tipos de cáncer, CaCu, CaMa y CaOv, estos resultados fueron similares a los observados en este estudio, a nivel de proteína, y se localiza principalmente en el citoplasma de las células.

El análisis de TMEM160 en las bases de datos concuerda con los resultados obtenidos, en este estudio, aunque particularmente se desconoce información sobre la localización y función de TMEM160, de acuerdo con la familia de proteínas transmembrana (TMEM) a la que pertenece, se ha reportado que es una familia de proteínas que atraviesan la bicapa lipídica y componen las membranas biológicas del retículo endoplásmico, los lisosomas, las mitocondrias y el aparato de Golgi (Beasley et al., 2021; Koteluk et al., 2021). Aunque se ha sugerido la participación de estas proteínas en procesos que conducen a la diseminación de células cancerosas, como la migración y la remodelación de la matriz (Schmit *et al.*, 2020), también se ha reportado que varias proteínas TMEM están involucradas en el crecimiento tumoral a través de la activación de TGF- β para facilitar la progresión maligna y la progresión de EMT (Sun *et al.*, 2022).

En este trabajo describen los niveles de expresión (mensajeros y proteína) de CCDC61, CCDC97, C19ORF47, LENG8 y TMEM160, en biopsias de pacientes y líneas celulares, de cáncer cervical, mama y ovario, además se comparan los datos derivados del análisis transcriptómico y proteómico reportado en bases de datos internacionales. Estos resultados muestran diferencias de expresión entre los tejidos analizados, sin embargo, no se consideran datos clínicos como: tipo histológico, la etapa clínica, así como algunos marcadores moleculares, debido a que el primer enfoque es la detección de la sobreexpresión de estas proteínas en

cáncer para posteriormente, analizar el papel funcional de estas proteínas en modelos celulares donde su silenciamiento permitirá, evaluar su papel en el desarrollo y progresión del cáncer.

Conclusiones

La expresión negativa de CCDC61 y CCDC97 en CaCu, CaMa y CaOv sugiere poca relevancia en estos tipos de cáncer.

La base de datos GEPIA reporta para **LENG8** una subexpresión en CaCu, CaMa y CaOv ($p \leq 0.05$); mientras que para **TMEM160**, una sobreexpresión en CaOv, respecto a CaCu y CaMa ($p < 0.05$).

En la línea celular HeLa (CaCu), se observó la sobreexpresión de **C19ORF47** y **TMEM160** en comparación con la línea no tumoral HaCaT ($p > 0.001$). En las líneas celulares de CaMa se observó la sobreexpresión de **C19ORF47** en MCF7, y **TMEM160** en MCF7 y MDA-MB-231 en comparación con la línea control MCF10A ($p < 0.05$). En la línea celular C33A, **LENG8** está sobreexpresado en comparación con la línea control HaCaT ($p \leq 0.05$).

A nivel de proteína **C19ORF47**, **TMEM160** y **LENG8** se observa una sobreexpresión diferencial en cáncer cérvicouterino, mama y ovario, con localización subcelular en membrana y citoplasma. La expresión de **C19ORF47** es similar entre los tres tipos de cáncer, mientras que en **LENG8** se observa una mayor expresión en CaMa ($p < 0.05$) y CaOv con respecto a CaCu ($p \leq 0.0092$). **TMEM160** está sobreexpresado en CaMa en comparación con CaCu ($p < 0.05$).

Perspectivas

- Seleccionar una línea celular de CaOv y CaMa que pueda ser un buen modelo experimental, para evaluar el efecto de la sobreexpresión o silenciamiento de las proteínas y su relación con cáncer.
- Una vez seleccionado el modelo experimental, evaluar el papel funcional de estas proteínas en la progresión tumoral.
- Explorar en otros tipos de cáncer el nivel de expresión de CCDC61 y CCDC97, para seleccionar un buen modelo experimental, que permita evaluar el papel funcional de ambas proteínas.

Anexos

Tabla 2. Características clínicas de las pacientes con cáncer de mama

Folio	Edad	Tipo histológico	Etapla clínica	Diagnóstico molecular
MBR CT/ST	40 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 7	IIIA	HER2
MGH CT/ST	51 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 7	IIIA	HER2
GFE CT/ST	45 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 6	IIIB	Luminal A
FSC CT/ST	64 años	Carcinoma ductal infiltrante sin patrón de crecimiento, poco diferenciado, con SBR de 8	IIIA	HER2
MRA CT/ST	42 años	Carcinoma ductal infiltrante	IIIA	Triple negativo
JBR CT/ST	34 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 7/9	IIIA	Triple negativo
MRV CT/ST	39 años	Carcinoma lobulillar infiltrante con SBR modificado de 5	IIIA	Triple negativo
FRR CT/ST	68 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 6	IIIB	HER2
MCG CT/ST	48 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 7	IIB	Luminal A
IMB CT/ST	37 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 8	IIIA	Triple negativo
OGC CT/ST	63 años	Carcinoma ductal infiltrante	IIIA	Triple negativo
MMC CT/ST	37 años	Carcinoma ductal infiltrante sin patrón específico con SBR de 7	IIIC	Luminal B

CT: Tejido tumoral, ST: Tejido adyacente (ambas biopsias de la misma paciente).

Tabla 3. Características clínicas de las pacientes con cáncer de ovario

Folio	Edad	Tipo histológico	Etapa clínica
OLC1	40 años		
OLC2	56 años	Ovario libre de cáncer	-
OLC3	45 años		
COG 1	12 años	Teratoma maduro	
COG3	36 años	Teratoma quístico maduro	
COG4	54 años	Teratoma quístico maduro	N/A
COG5	48 años	Teratoma quístico maduro	
COG6	68 años	Teratoma quístico maduro	
CEO1	84 años	Cistadenoma mucinoso	
CEO2	54 años	Carcinoma seroso papilar de alto grado con extensa necrosis limitada al ovario.	IIIC
CEO3	45 años	Cistadenoma mucinoso	
CEO4	63 años	Adenocarcinoma seroso papilar	
CEO5	41 años	Adenocarcinoma mucinoso Cistoadenocarcinoma mucinoso papilar.	IC
CEO6	32 años	Ca. De ovario bilateral + ascitis a tensión (Después de cirugía carcinomatosis)	IIIC
CEO9	50 años	C. Ovario derecho limítrofe	
CEO10	69 años	Cistoadenocarcinoma mucinoso	IA
CEO11	19 años	Cistoadenocarcinoma seroso con infiltrado moderado	III C
CEO12			
CEO13	43 años	Tumoración abdominopélvica en estudio	

OLC: Ovario libre de cáncer, **COG:** Teratoma benigno, **CEO:** Tejido tumoral de cáncer de ovario.

Tabla 4. Características clínicas de las pacientes con cáncer de mama analizadas por inmunohistoquímica.

Folio	Tipo histológico	Etapla clínica	Diagnóstico molecular
KDA-2	Carcinoma ductal infiltrante sin patrón específico	IIB	Luminal A
MCAZ-5	Carcinoma ductal infiltrante sin patrón específico	IIA	Her2 Positivo
LPM-6	Carcinoma ductal infiltrante sin patrón específico	IIA	Triple Negativo
SGH-7	Carcinoma ductal infiltrante sin patrón específico	IIB	Triple Negativo
MIPA-11	Carcinoma ductal infiltrante sin patrón específico	IIB	Her2 Positivo
AM-13	Carcinoma ductal infiltrante sin patrón específico	IA	Luminal A
GAB-14	Carcinoma ductal infiltrante sin patrón específico	IIB	Luminal A
MVM-16	Carcinoma ductal infiltrante sin patrón específico	IIB	Luminal B
FAS-20	Carcinoma ductal infiltrante sin patrón específico	IIIA	Luminal A
RRA-23	Carcinoma ductal infiltrante sin patrón específico	IIA	Luminal B
GCP-26	Carcinoma ductal infiltrante sin patrón específico	IA	
SHT-29	Carcinoma ductal infiltrante sin patrón	IIIA	Triple Negativo
ALB-33	Carcinoma ductal infiltrante sin patrón específico	IIA	Luminal B
CGR-50	Carcinoma ductal infiltrante sin patrón específico	IIA	Luminal A
PLG-51	Carcinoma ductal infiltrante sin patrón específico	IIB	Luminal A
LCC-47	Carcinoma ductal infiltrante sin patrón específico	IIA	Triple Negativo
LGB -49	Carcinoma ductal infiltrante sin patrón específico	IIA	Triple Negativo
LHM-35	Carcinoma lobulillar infiltrante	IIA	Her2 Positivo

CT: Tejido tumoral, **ST:** Tejido adyacente (ambas biopsias de la misma paciente).

Referencias

- Atlas de Proteínas Humanas (no date). Available at: <https://www.proteinatlas.org/> (Accessed: 6 November 2022).
- Babcock, J.J., Li, M. (2014). Deorphanizing the human transmembrane genome: A landscape of uncharacterized membrane proteins. *Acta Pharmacol. Sin.* 35, 11–23.
- Burkhard, P., et al. (2001). Coiled coils: a highly versatile protein folding motif, *Trends in Cell Biology*, 11(2), pp. 82–88.
- C-33 A [c-33a] | ATCC (no date). Available at: <https://www.atcc.org/products/htb-3> (Accessed: 11 January 2022).
- Chen, K.L., et al. (2017). Impact of Diet and Nutrition on Cancer Hallmarks. *J. Cancer Prev. Curr. Res.* 7.
- Chen, M., et al. (2009). CCDC62/ERAP75 functions as a coactivator to enhance estrogen receptor beta-mediated transactivation and target gene expression in prostate cancer cells. *Carcinogenesis* 30, 841–850.
- da Conceição Braga, L. et al. (2022) 'Identification of best housekeeping genes for the normalization of RT-qPCR in human cell lines', *Acta Histochemica*, 124(1), p. 151821.
- Duvvuri, U., et al. (2012). TMEM16A induces MAPK and contributes directly to tumorigenesis and cancer progression. *Cancer Res.* 72, 3270–3281.
- Expresión de TMEM160 en cáncer - Resumen - Atlas de proteínas humanas [WWW Document], n.d. URL <https://www.proteinatlas.org/ENSG00000130748-TMEM160/pathology> (accessed 11.30.20).
- Flamant, L., et al. (2012). TMEM45A is essential for hypoxia-induced chemoresistance in breast and liver cancer cells. *BMC Cancer* 12, 391.
- Frezzo, J.A. and Montclare, J.K. (2015). Exploring the potential of engineered coiled-coil protein microfibers in drug delivery, *Therapeutic Delivery*, 6(6), pp. 643–646.
- Gil, J., et al. (2017). Lysine acetylation stoichiometry and proteomics analyses reveal pathways regulated by sirtuin 1 in human cells. *J. Biol. Chem.* 292, 18129–18144.
- Grimwood, J., et al. (2004). The DNA sequence and biology of human chromosome 19. *Nature* 428, 529–535.
- Guo, L. et al. (2021). CCDC137 Is a Prognostic Biomarker and Correlates With Immunosuppressive Tumor Microenvironment Based on Pan-Cancer Analysis, *Frontiers in Molecular Biosciences*, 8.
- Gupta, G.D., et al. (2015). A dynamic protein interaction landscape of the human centrosome-cilium interface. *Cell* 163, 1484–1499.
- Hanahan, D., Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D., Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Harris, R.A., Raveendran, M., Worley, K.C., Rogers, J. (2020). Unusual sequence characteristics of human chromosome 19 are conserved across 11 nonhuman primates. *BMC Evol. Biol.* 20.
- HeLa | ATCC (no date). Available at: <https://www.atcc.org/products/ccl-2> (Accessed: 11 January 2022).
- Herramienta de análisis Cancer Hallmarks [WWW Document], n.d. URL <https://chat.lionproject.net/hallmarks> (accessed 11.30.20).
- Hong, J., et al. (2013). Genetic susceptibility, birth weight and obesity risk in young Chinese. *Int. J. Obes.* 37, 673–677.
- Hrašovec, S., et al. (2013). TMEM25 is a candidate biomarker methylated and downregulated in colorectal cancer. *Dis. Markers* 34, 93–104.

- Hu, D.-D. et al. (2018). Overexpression of Coiled-Coil Domain-Containing Protein 34 (CCDC34) and its Correlation with Angiogenesis in Esophageal Squamous Cell Carcinoma', *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 24, pp. 698–705.
- HUPO - About [WWW Document], n.d. URL <https://www.hupo.org/about-hupo> (accessed 11.14.20).
- Legrain, P., et al. (2011). The Human Proteome Project: Current State and Future Direction. *Mol. Cell. Proteomics MCP* 10.
- Li, B., et al. (2015). TMEM140 is associated with the prognosis of glioma by promoting cell viability and invasion. *J. Hematol. Oncol.* 8.
- Liu, Z., et al. (2020). Expression profiles of genes associated with inflammatory responses and oxidative stress in lung after heat stroke. *Biosci. Rep.* 40.
- Lodish, H., et al. (2000). Insertion of Membrane Proteins into the ER Membrane. *Mol. Cell Biol.* 4th Ed.
- Marx, S., et al. (2020). Transmembrane (TMEM) protein family members: Poorly characterized even if essential for the metastatic process. *Semin. Cancer Biol.*, Enigmatic tumor properties associated with metastatic spread 60, 96–106.
- MCF 10A | ATCC (no date). Available at: <https://www.atcc.org/products/crl-10317> (Accessed: 11 January 2022).
- MCF7 | ATCC (no date). Available at: <https://www.atcc.org/products/htb-22> (Accessed: 11 January 2022).
- McFarlane, A.A., Orriss, G.L. and Stetefeld, J. (2009). The use of coiled-coil proteins in drug delivery systems', *European Journal of Pharmacology*, 625(1), pp. 101–107.
- MDA-MB-231 | ATCC (no date). Available at: <https://www.atcc.org/products/crm-htb-26> (Accessed: 11 January 2022).
- México - Estadísticas Vitales. Mortalidad, Defunciones Generales 2017 - variable - V158 [WWW Document], n.d. URL <https://www.inegi.org.mx/rnm/index.php/catalog/407/datafile/F5/V158> (accessed 12.1.20).
- Modjtahedi, N., Tokatlidis, K., Dessen, P., Kroemer, G. (2016). Mitochondrial Proteins Containing Coiled-Coil-Helix-Coiled-Coil-Helix (CHCH) Domains in Health and Disease. *Trends Biochem. Sci.* 41, 245–260.
- Ni, J., et al. (2013). Epithelial cell adhesion molecule (EpCAM) is associated with prostate cancer metastasis and chemo/radioresistance via the PI3K/Akt/mTOR signaling pathway. *Int. J. Biochem. Cell Biol.* 45, 2736–2748.
- Nilsson, C.L., et al. (2012). Chromosome 19 Annotations with Disease Speciation: A First Report from the Global Research Consortium [WWW Document].
- Nilsson, C.L., et al. (2013). Chromosome 19 Annotations with Disease Speciation – A First Report from the Global Research Consortium. *J. Proteome Res.* 12, 135–150.
- Omenn, G.S. (2014). THE STRATEGY, ORGANIZATION, AND PROGRESS OF THE HUPO HUMAN PROTEOME PROJECT. *J. Proteomics* 100, 3–7.
- Paik, Y.-K., et al. (2012). Standard Guidelines for the Chromosome-Centric Human Proteome Project. *J. Proteome Res.* 11, 2005–2013.
- Paik, Y.-K., et al. (2018). Launching the C-HPP neXt-CP50 Pilot Project for Functional Characterization of Identified Proteins with No Known Function. *J. Proteome Res.* 17, 4042–4050.
- Park, S.-J., et al. (2012). Epigenetic alteration of CCDC67 and its tumor suppressor function in gastric cancer. *Carcinogenesis* 33, 1494–1501.

- Rigbolt, K.T.G., et al. (2011). System-Wide Temporal Characterization of the Proteome and Phosphoproteome of Human Embryonic Stem Cell Differentiation. *Sci. Signal.* 4, rs3–rs3.
- Schmick, M., Bastiaens, P.I.H. (2014). The Interdependence of Membrane Shape and Cellular Signal Processing. *Cell* 156, 1132–1138.
- Schmit, K. and Michiels, C. (2018). TMEM Proteins in Cancer: A Review', *Frontiers in Pharmacology*, 9, p. 1345.
- Sever, R., Brugge, J.S. (2015). Signal Transduction in Cancer. *Cold Spring Harb. Perspect. Med.* 5.
- SiHa | ATCC (no date). Available at: <https://www.atcc.org/products/htb-35> (Accessed: 11 January 2022).
- Sun, Z. et al. (2022). Transmembrane protein 88 inhibits transforming growth factor- β 1-induced-extracellular matrix accumulation and epithelial-mesenchymal transition program in human pleural mesothelial cells through modulating TGF- β 1/Smad pathway', *Journal of Receptor and Signal Transduction Research*, 42(1), pp. 60–66.
- Sung, H., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* 71, 209–249.
- Tang, Z. et al. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses', *Nucleic Acids Research*, 45(W1), pp. W98–W102.
- Taylor, C.F., et al. (2008). Guidelines for reporting the use of mass spectrometry in proteomics. *Nat. Biotechnol.* 26, 860–861.
- TMEM160 Gene - GeneCards | TM160 Protein | TM160 Antibody [WWW Document], n.d. URL <https://www.genecards.org/cgi-bin/carddisp.pl?gene=TMEM160&keywords=Q9NX00> (accessed 11.29.20).
- Yin, D.T. et al. (2016). Characterization of the novel tumor-suppressor gene CCDC67 in papillary thyroid carcinoma', *Oncotarget*, 7(5), pp. 5830–5841.
- Zhang, X., et al. (2017). TMEM17 depresses invasion and metastasis in lung cancer cells via ERK signaling pathway. *Oncotarget* 8, 70685–70694.
- Zhong, J., et al. (2013). CCDC134 is down-regulated in gastric cancer and its silencing promotes cell migration and invasion of GES-1 and AGS cells via the MAPK pathway. *Mol. Cell. Biochem.* 372, 1–8.

