



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
DOCTORADO EN CIENCIAS BIOMÉDICAS  
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

PERFILES DE EXPRESIÓN DE RNA LARGOS NO CODIFICANTES (lncRNAs) Y RNA MENSAJERO EN CARCINOMA DE MAMA TRIPLE NEGATIVOS.

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
DOCTOR EN CIENCIAS

PRESENTA:  
M. EN C. FREDY OMAR BELTRÁN ANAYA

DIRECTOR DE TESIS  
DR. ALFREDO HIDALGO MIRANDA  
FACULTAD DE MEDICINA, UNAM

COMITÉ TUTOR  
DR. SAMUEL CANIZALES QUINTEROS  
FACULTAD DE MEDICINA, INSTITUTO DE QUÍMICA, UNAM

DR. ALEJANDRO MANUEL GARCÍA CARRANCA  
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

CDMX, OCTUBRE DE 2019



Universidad Nacional  
Autónoma de México



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

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

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

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



PDCB/FM/45/2019.

**C.P. AGUSTIN MERCADO**  
**DIRECTOR DE CERTIFICACION Y CONTROL DOCUMENTAL, UNAM**  
Presente.

Nos permitimos informarle que con base al Artículo 31 del RGEF el Comité Académico de **DOCTORADO EN CIENCIAS BIOMÉDICAS**, en su reunión 450 del 22 de mayo del 2019, designó el siguiente jurado para examen de grado de **DOCTOR EN CIENCIAS** de **FREDY OMAR BELTRÁN ANAYA**, con número de cuenta 514022365 con la tesis titulada "**PERFILES DE EXPRESIÓN DE RNA LARGOS NO CODIFICANTES (LNCRNAS) Y RNA MENSAJERO EN CARCINOMA DE MAMA TRIPLE NEGATIVOS**", dirigida por el Dr. Alfredo Hidalgo Miranda.

PRESIDENTE: DRA. ELIZABETH LANGLEY McCARRON  
SECRETARIO: DR. ALFREDO HIDALGO MIRANDA  
VOCAL: DRA. CATALINA TREJO BECERRIL  
VOCAL: DR. EMILIO ROJAS DEL CASTILLO  
VOCAL: DR. VÍCTOR JULIAN VALDES RODRÍGUEZ

Atentamente

*"Por mi raza hablará el espíritu"*

Ciudad Universitaria, Cd. de Mx., a 13 de septiembre de 2019.

DRA. AUREA OROZCO RIVAS  
COORDINADORA

C.c.p. Dr. Jorge Ramón Morales Pineda, Coordinador de la Unidad de Administración Escolar del Posgrado, UNAM  
Tutor (a) Dr. (a) Alfredo Hidalgo Miranda  
Coordinación de PDCB  
AOR/RGG/evb

---

---

## Agradecimientos

A la Universidad Nacional Autónoma de México por dejarme ser parte de su comunidad y brindarme las herramientas académicas para desarrollarme y crecer como persona. Al Instituto Nacional de Medicina Genómica, por darme la oportunidad de ser parte de su comunidad, por el aprendizaje personal y académico adquirido durante toda mi estancia. A las Unidades de Servicio del INMEGEN.

Al FUCAM y al instituto nacional del tumor de Italia por mantener abiertas sus puertas para colaborar en la recolección de muestras para realizar el presente trabajo. Sobre todo a las pacientes que luchan día a día contra el cáncer de mama.

A un gran amigo y tutor, el Dr. Alfredo Hidalgo Miranda, por darme la oportunidad de formar parte de su laboratorio, por la confianza depositada en mí, por sus enseñanzas, sus consejos en la vida académica y personal, por la calidez humana. Pero sobre todo, por el gran ambiente de trabajo que se tiene día a día en su laboratorio. Un excelente laboratorio. A una gran amiga, la Dra. Silvia Jiménez, por sus enseñanzas, comentarios y colaboraciones. A un gran amigo, el Ing. Luis A. Alfaro Ruíz, por sus consejos, y el apoyo en todos los sentidos durante mi estancia en el laboratorio.

A mis compañeros y amigos del Laboratorio de Genómica del Cáncer, con los que conviví todo el tiempo durante mi estancia en el INMEGEN; Carlos, Rosa, Sandra, Beto, Magda, Diego Rocio, Marigaby, Mireya, Vaitiare, Eduardo, Marquito, Marco Fonseca, Didier, Diana, Jazmin. Siempre poniendo alegría al laboratorio con todos sus comentarios, ocurrencias y anécdotas.

A los miembros del jurado, Dra. Langley, Dra. Trejo, Dr. Rojas, Dr. Valdes; por sus valiosas sugerencias que enriquecieron el presente trabajo.

A mi familia, por brindarme todo su apoyo incondicional para alcanzar, continuar y concluir este objetivo y etapa en mi vida; Mi mamá Dionicia, mi hermano Daniel, mis ti@s, abuel@s. Muchísimas gracias. A mi madre, por todo su trabajo, sus consejos, cariño, paciencia y comprensión. Por ser ese ejemplo a seguir adelante. A Felipe Anaya, por darme y reforzar la confianza en mi para poder llegar a este punto de mi carrera. Mil gracias.

A Oscar del Moral, por todo su apoyo, consejos y ayuda durante esta etapa de mi vida. Muchas gracias.

A Xavier, por sus consejos y ayuda en los momentos que más lo necesitaba.

---

---

---

---

**Agradecimientos al Consejo Nacional de Ciencia y Tecnología por otorgarme la *beca de doctorado* durante el periodo febrero 2014 - enero 2018 (CVU/Becario: 416163/260518) y al programa de estímulo de *ayudante de investigador* (2019).**

**Parte de este trabajo fue financiado con fondos de los proyectos de ciencia básica (proyecto: 258936) y fronteras de la ciencia (proyecto: 1285) otorgados por el Consejo Nacional de Ciencia y Tecnología.**

---

---

---



---

<b>Índice</b>	
<b>Resumen</b> .....	<b>1</b>
<b>Abstract</b> .....	<b>2</b>
<b>Introducción</b> .....	<b>3</b>
Glándula mamaria .....	3
Biología del desarrollo de la glándula mamaria .....	3
Estadísticas y factores de riesgo del cáncer de mama .....	5
Diagnóstico del cáncer de mama y su clasificación molecular .....	7
Cáncer de mama triple negativo (CMTN) .....	11
Clasificación del cáncer de mama triple negativo y sus implicaciones clínicas....	13
Clasificadores adicionales del cáncer de mama triple negativo a lo largo del tiempo .....	17
Los llamados RNAs de la materia oscura .....	20
Clasificación y función de lncRNAs .....	22
Papel de los lncRNAs en el desarrollo de la glándula mamaria .....	25
lncRNAs y cáncer de mama .....	25
Pregunta experimental .....	27
<b>Justificación</b> .....	<b>27</b>
Hipótesis .....	27
Objetivos .....	28
Objetivo general .....	28
Objetivos específicos .....	28
<b>Materiales y métodos</b> .....	<b>29</b>
Obtención de muestras de cáncer de mama triple negativo .....	29
Extracción de RNA .....	29
Análisis de expresión de mRNA y lncRNAs mediante microarreglos de expresión .....	30
Clasificación de subtipos moleculares de cáncer de mama triple negativo basado en mRNA .....	30
Análisis de ontologías de los subtipos moleculares triple negativos .....	31
Análisis de expresión diferencial de lncRNAs entre subtipos de cáncer de mama triple negativo .....	31
Validación de la expresión de lncRNAs en muestras independientes .....	31
Análisis de enriquecimiento de vías de lncRNAs a través de asociación por culpabilidad .....	31
Predicción de inmunofenoscóreo entre subtipos triple negativos .....	32
Cultivo de líneas celulares .....	32
Síntesis de cDNA por RT-PCR .....	32
Análisis de expresión de mRNA y lncRNA por reacción en cadena de la polimerasa cuantitativa (RT-qPCR) .....	33
Generación short hairpin RNAs (shRNAs) .....	33
Modulación de la expresión del lncKLHDC7B por short hairpin RNAs (shRNAs) .....	33
Ensayos de migración e invasión celular .....	34
Ensayo de apoptosis .....	35
Análisis estadísticos .....	35
<b>Resultados</b> .....	<b>36</b>

---



---

Discusión .....	61
Conclusiones .....	67
Perspectivas .....	68
Referencias .....	69
Tablas .....	78
Anexos .....	80
Artículos publicados.....	84

### Índice de figuras

Figura 1. Glándula mamaria y cáncer. ....	4
Figura 2. Estadísticas de cáncer de mama. ....	5
Figura 3. Hipótesis sobre el desarrollo del cáncer de mama.....	8
Figura 4. Clasificación molecular del CaMa y pronóstico clínico. ....	9
Figura 5. Representación de la evaluación inmunohistoquímica de receptores hormonales y HER2 en cáncer de mama.....	12
Figura 6. Clasificación molecular del cáncer de mama triple negativo. ....	16
Figura 7. Estado del arte de la clasificación molecular del cáncer triple negativo. ....	20
Figura 8. Características de los lncRNAs. ....	21
Figura 9. Clasificación de lncRNAs y su participación en salud y enfermedad. ....	23
Figura 10. Análisis de calidad y normalización de microarreglos.....	37
Figura 11. Frecuencia de subtipos de cáncer de mama triple negativo.....	38
Figura 12. Análisis ontológico y rutas biológicas en CMTN. ....	40
Figura 13. Perfiles de lncRNAs expresados diferencialmente entre subtipos CMTN. ....	42
Figura 14 . Validación in silico de la expresión de lncRNAs en muestras independientes.....	44
Figura 15. Panorama general de las vías biológicas mediante el análisis de culpabilidad por asociación a través de los subtipos de TNBC mediante la coexpresión de lncRNAs-mRNAs.....	45
Figura 16. Análisis de correlación entre lncRNA-mRNA co-expresados positivamente del subtipo inmnomodulador. ....	47
Figura 17. Evaluación de la expresión de lncKLHDC7B y KLHDC7B en los subtipos CMTN. ....	49
Figura 18. Expresión de lncKLHDC7B y KLHDC7B en tejido normal y tumoral, modelos in vitro y el impacto de su silenciamiento en la línea HCC1187. ....	52
Figura 19. Análisis de asociación por culpabilidad. ....	53

---

---

<b>Figura 20. Enriquecimiento y análisis funcional por el silenciamiento de IncKLHDC7B en la línea celular HCC1187. ....</b>	<b>55</b>
<b>Figura 21. Análisis funcional del IncKLHDC7B en la línea celular MCF-7. ....</b>	<b>57</b>
<b>Figura 22. Análisis funcional por el silenciamiento del IncKLHDC7B en la línea celular BT-20. ....</b>	<b>58</b>
<b>Figura 23. Implicaciones clínicas de la sub-expresión de IncKLHDC7B y KLHDC7B en CMTN. ....</b>	<b>60</b>
<b>Figura 24. Análisis de sobrevida global en cáncer de mama. ....</b>	<b>60</b>

## **Índice de tablas**

<b>Tabla 1. Clasificación molecular del cáncer de mama y marcadores subrogados. ....</b>	<b>10</b>
<b>Tabla 2 . Características clínico-patológicas de la población. ....</b>	<b>78</b>
<b>Tabla 3. Lista de transcritos de lncRNA y mRNA co-expresados positivamente en el subtipo inmunomodulador. ....</b>	<b>79</b>



---

---

## Resumen

El cáncer de mama triple negativo (CMTN) representa un fenotipo tumoral agresivo con mal pronóstico en comparación con los tumores con receptores de estrógenos (RE), progesterona (RP) y HER2 positivos. Estudios en la última década han demostrado que el CMTN es un enfermedad heterogénea a nivel clínico, celular y molecular. Actualmente, el CMTN se puede dividir en subtipos específicos con base a un análisis de expresión génica. Adicionalmente, las evidencias acumuladas demuestran que los ARNs largos no codificantes (lncRNAs) están involucrados en la regulación de la expresión génica y la biología del cáncer, lo que contribuye a las funciones esenciales de las células cancerosas. En este estudio, se analizó el perfil de expresión de los lncRNAs en distintos subtipos de CMTN de 156 muestras y, además se caracterizó el papel funcional del *LncKLHDC7B* (ENSG00000226738). Se encontró que un total de 710 lncRNAs se expresan diferencialmente entre subtipos de CMTN; y un subconjunto de estos lncRNAs alterados, se validaron de forma independiente. De manera interesante, se observó que el *LncKLHDC7B* actúa como un modulador transcripcional de su gen codificador vecino *KLHDC7B* en el subtipo inmunomodulador.

Además, la subexpresión de *LncKLHDC7B* aumentó la migración y la invasión, y promovió la resistencia a la muerte celular. Nuestros hallazgos confirmaron la contribución de *LncKLHDC7B* a la inducción de la apoptosis y la inhibición de la migración e invasión celular, lo que sugiere que los tumores de CMTN con enriquecimiento de *LncKLHDC7B* pueden mostrar una actividad reguladora distinta, o que esto puede ser un proceso generalizado en el cáncer de mama. Además, el análisis *in silico* confirmó por primera vez que la baja expresión de *KLHDC7B* y *LncKLHDC7B* se asocia con un mal pronóstico en pacientes con cáncer de mama.

---

---

## Abstract

Triple-negative breast cancer (TNBC) represents an aggressive tumor phenotype with poor prognosis compared to tumors positive for estrogen receptor (ER), progesterone (PR) and HER2. Studies in the last decade on TNBC have shown that this pathology is a heterogeneous disease at the clinical, cellular and molecular levels. Currently, TNBC can be divided into specific subtypes based on gene expression analysis. Accumulated evidence demonstrates that long non-coding RNAs (lncRNA) are involved in regulation of gene expression and cancer biology, contributing to essential cancer cell functions. In this study, we analyzed the expression profile of lncRNA in TNBC subtypes from 156 TNBC samples, and then characterized the functional role of LncKLHDC7B (ENSG00000226738). A total of 710 lncRNA were found to be differentially expressed between TNBC subtypes, and a subset of these altered lncRNA were independently validated. We discovered that LncKLHDC7B acts as a transcriptional modulator of its neighboring coding gene KLHDC7B in the immunomodulatory subtype.

Furthermore, LncKLHDC7B knockdown enhanced migration and invasion, and promoted resistance to cell death. Our findings confirmed the contribution of LncKLHDC7B to induction of apoptosis and inhibition of cell migration and invasion, suggesting that TNBC tumors with enrichment of LncKLHDC7B may exhibit distinct regulatory activity, or that this may be a generalized process in breast cancer. Additionally, *in silico* analysis confirmed, for the first time, that low expression of KLHDC7B and LncKLHDC7B is associated with poor prognosis in patients with breast cancer.

---

---

## **Introducción**

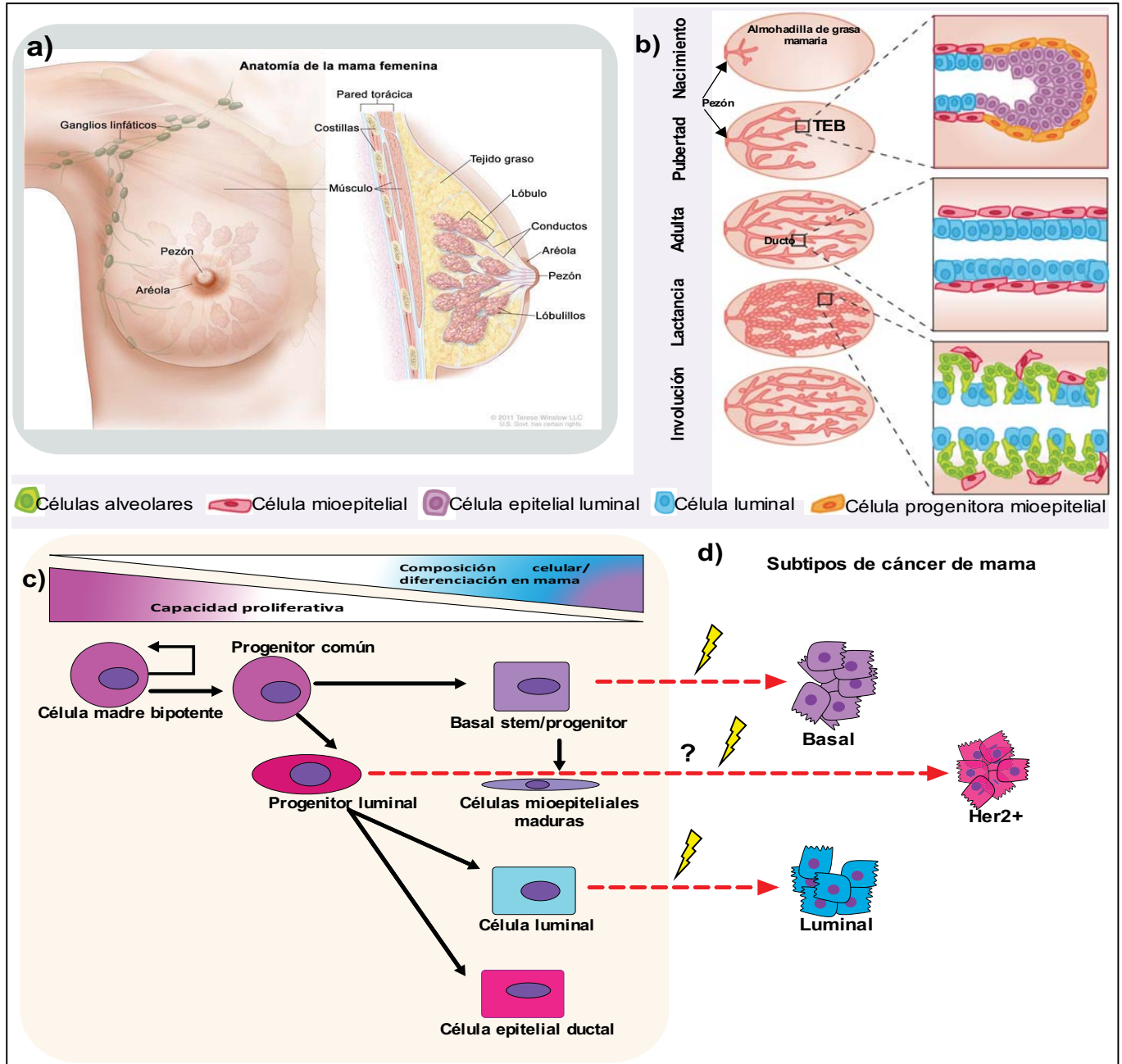
### ***Glándula mamaria***

La glándula mamaria es un órgano único de mamíferos, cuya función es sintetizar, secretar y entregar la leche al recién nacido. En edad adulta, cada seno o mama contiene aproximadamente de 15-20 lóbulos mamarios y cada uno de ellos se ramifica en otros más pequeños llamados unidades ductal lobulares terminales (UDLT). Los lóbulos están revestidos por epitelio glandular, responsables de la producción de leche. Cada lóbulo está rodeado por tejido de sostén y grasa. Otras estructuras localizadas junto a la glándula mamaria son los ganglios linfáticos, estructuras de importancia inmunológica para defender al cuerpo contra microorganismos patógenos (Figura 1a).

### ***Biología del desarrollo de la glándula mamaria***

El desarrollo de la glándula mamaria inicia aproximadamente durante la sexta semana de embriogénesis. Sin embargo, su crecimiento y diferenciación ocurre después del nacimiento. En el nacimiento, las almohadillas de grasa mamaria contienen una estructura ductal epitelial simple (figura 1b). Durante la pubertad, los ovarios y hormonas pituitarias, incluyendo estrógenos, progesterona y hormonas de crecimiento, estimulan la proliferación de las células epiteliales mamarias para generar las estructuras conocidas como Brote de extremo terminal o TEB, por sus siglas en inglés. Los TEBs llevarán a la extensión y bifurcación de los conductos hacia los extremos de las almohadillas adiposas, dando como resultado una estructura ramificada ductal que se convertirá en la glándula mamaria adulta (figura 1b). La glándula mamaria en su etapa adulta será estimulada por los cambios hormonales; principalmente por estrógenos y progesterona. Estructuralmente, los ductos epiteliales estarán compuestos por una capa interna de células epiteliales de tipo luminal y una capa externa compuesta por células epiteliales basales las cuales darán origen a las células mioepiteliales. En resumen, en la etapa adulta, la glándula mamaria está compuesta de

diferentes tipos celulares, incluyendo: células epiteliales, adipocitos, fibroblastos, células del sistema inmunológico, células vasculares, entre otras (figura 1c) [1].



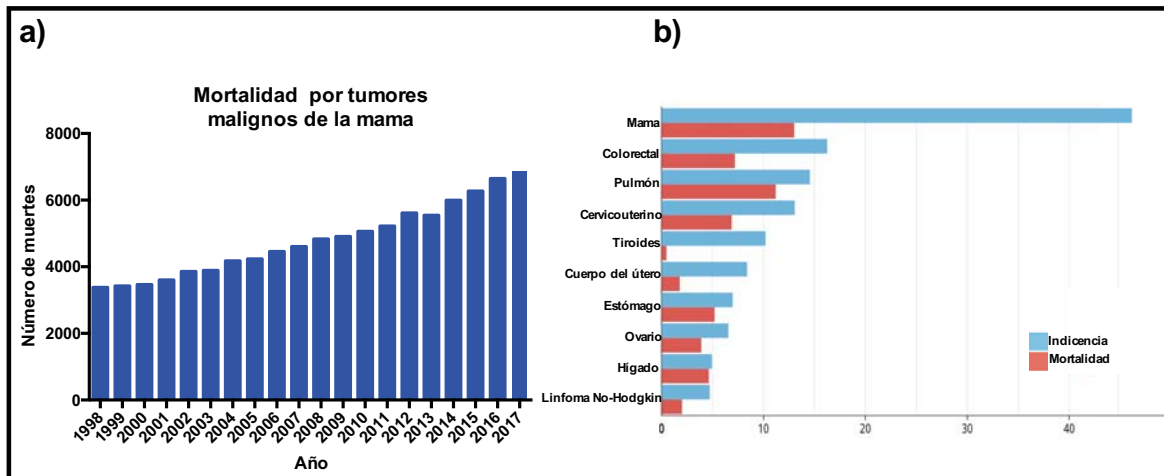
**Figura 1. Glándula mamaria y cáncer.** a) Anatomía típica normal de la mama. Tomado de: <https://www.cancer.gov/images/cdr/live/CDR710874-750.jpg>. b) Desarrollo de la glándula mamaria durante la etapa de crecimiento de la mujer. Brote de extremo terminal; TEB. Modificado de J. Inmal y col.[1]. c) Jerarquización de los componentes celulares de la glándula mamaria. d) Subtipos de cáncer de mama según el origen del tipo celular.

---

---

## Estadísticas y factores de riesgo del cáncer de mama

El cáncer de mama (CaMa) es una patología que afecta principalmente a mujeres alrededor del mundo. En México, de acuerdo al Instituto Nacional de Estadística y Geografía (INEGI), se reportó un aumento en el número de decesos por cáncer de mama a lo largo de la última década, alcanzando 6,650 muertes solo en el 2016 (figura 2a). Por otra parte, la agencia internacional para la Investigación en cáncer (IARC, por sus siglas en inglés) reportó que en México el cáncer de mama sigue siendo la principal causa de muerte por neoplasias, presentando elevadas tasas de incidencia y mortalidad (50 casos y 13 muertes por cada 100 000 mujeres, respectivamente) solo en el 2018 (figura 2b, <http://gco.iarc.fr/today/home>).



**Figura 2. Estadísticas de cáncer de mama. a)** Número de casos de cáncer de mama en México. Fuente: INEGI. Estadística de mortalidad. <http://www.beta.inegi.org.mx/proyectos/registros/vitales/mortalidad/default.html>. **b)** Mortalidad e incidencia en mujeres con cáncer de mama en México a nivel mundial (tasa estandarizada por edad). Fuente: GLOBOCAN. <http://gco.iarc.fr/today/home>. Acceso: 03 diciembre 2018.

El CaMa al igual que otros tipos de cáncer se considera un enfermedad multifactorial; donde su inicio está relacionado con factores o eventos que incrementan el riesgo para el desarrollo de cáncer. Entre estos factores se encuentran [2-6]:

- Edad, la incidencia de cáncer de mama está altamente relacionada con el incremento de la edad de la paciente.

- 
- 
- Etnicidad, otro factor importante para el cáncer de mama es el componente étnico o racial. Se ha observado que las mujeres con un componente étnico africano o hispano comparadas con caucásicas tienen mayores tasa de mortalidad a 5 años. Sin embargo, la incidencia es mayor en caucásicas y estas diferencias de riesgo y mortalidad no son concluyentes.
  - Menarca, el inicio temprano de la menarca podría incrementar el riesgo de cáncer de mama, donde la ventana de exposición hormonal prolongada (estrógenos) pudieran contribuir al desarrollo de esta enfermedad. Varios estudios demuestran que el riesgo para el desarrollo de cáncer de mama en mujeres que iniciaron su menstruación antes de los 11 años comparadas con las que iniciaron a los 13, se elevó en un 20%.
  - Menopausia tardía, si bien este factor es controversial para el desarrollo de cáncer de mama, describe que el alargamiento en el número total de años reproductivos juega un papel relevante que contribuye al desarrollo de esta enfermedad.
  - Alcohol, el consumo excesivo de alcohol se correlaciona con un mayor riesgo de cáncer de mama. Las mujeres que consumen una bebida alcohólica al día tienen un riesgo ligeramente mayor en comparación con los que no beben.
  - Tabaco, aunque la relación entre el hábito de fumar y el riesgo de CaMa sigue siendo controversial, se ha propuesto que los componentes mutágenos del tabaco podrían relacionarse con el desarrollo de CaMa.
  - Obesidad, la acumulación de grasa en el abdomen, provoca un estado inflamatorio que conlleva a la secreción de adipocitocinas proinflamatorias, y el incremento de síntesis de estrógenos por los adipocitos.
  - Genéticos, aproximadamente del 10 al 15% de los tumores de mama poseen un componente hereditario donde los genes afectados como BRCA1 y BRCA2, contribuyen fehacientemente con el desarrollo de CaMa. Estos genes reparadores del daño al DNA no son los únicos genes relacionados con el desarrollo de CaMa. Ya que existen otros genes que

---

---

pueden participar en diversos procesos biológicos relevantes en la transformación celular los cuales también pueden presentar mutaciones que elevan el riesgo de padecer la enfermedad (PTEN, TP53, CDH1, RAD50, CHEK2, ATM, entre otros).

### ***Diagnóstico del cáncer de mama y su clasificación molecular***

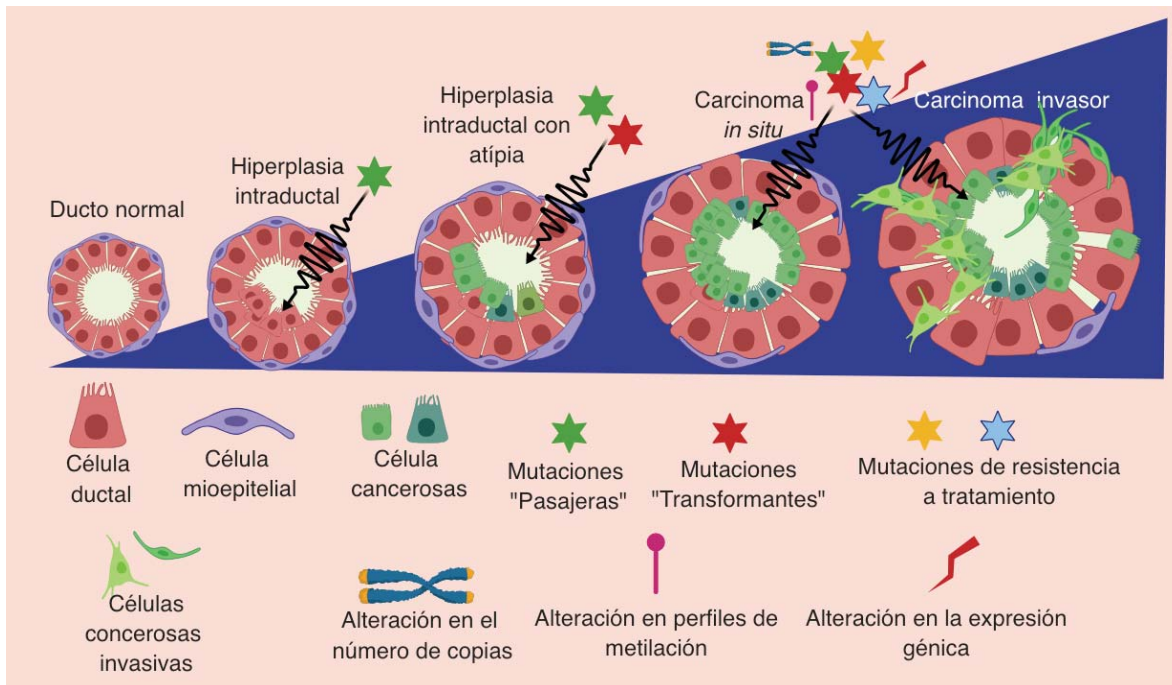
Para el diagnóstico temprano del cáncer de mama, el uso de herramientas de tamizaje ha permitido la detección temprana de este tipo de tumores. Entre las herramientas de tamizaje tenemos: el uso de pruebas de imagen como la mastografía, ultrasonido, resonancia magnética y tomografía. Estas herramientas complementan al diagnóstico de un tumor mamario. Sin embargo, el estándar para el diagnóstico de cáncer de mama se basa en criterios histológicos y características citológicas de los tumores [5]. La evaluación del grado tumoral, caracteriza los estadios de diferenciación, pleomorfismo nuclear y capacidad proliferativa del tumor. La etapa tumoral define el comportamiento de un tumor según el tamaño y la ubicación, así como la afectación a los ganglios linfáticos y la presencia / ausencia de metástasis [5, 7].

En la actualidad se sabe que para el desarrollo del carcinoma de mama invasor, las aberraciones genéticas, como las mutaciones, son un factor importante para inicializar los procesos de carcinogénesis en la mama (figura 3). La proliferación excesiva e hiperplasia de epitelios (ductales o lobulillares) son uno de los primeros eventos que presenta la glándula mamaria. Por lo tanto, la adquisición de aberraciones genéticas a lo largo de la vida de la paciente, podría llevar al desarrollo de un carcinoma invasor .

Actualmente a nivel molecular han sido aceptadas nuevas teorías para el desarrollo del cáncer de mama; 1) alteraciones genéticas como mutaciones somáticas, eliminaciones e inserciones de secuencias de DNA, 2) cambios epigenéticos, 3) alteración en el número de copias, 4) fusión de transcritos y 5) alteración en la expresión de RNAs, los cuales se verán reflejados en su mayoría



en los cambios de expresión de proteínas y en la presencia de determinadas estirpes celulares cancerosas (figura 1c, figura 3) [5, 8].



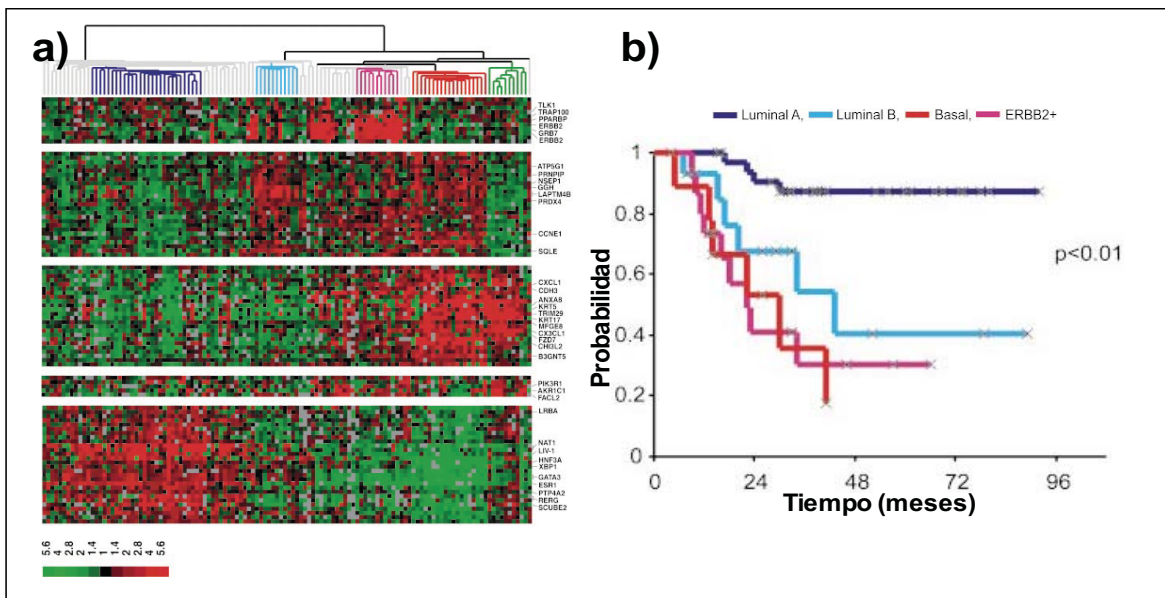
**Figura 3. Hipótesis sobre el desarrollo del cáncer de mama.** Para el desarrollo del cáncer de mama se postula que el proceso de hiperplasia inicia cuando el desarrollo y crecimiento normal de las células se alteran. Procesos como mutaciones en genes relevantes en el control del crecimiento celular así como los responsables de la reparación del DNA son afectados llevando así a la acumulación de mutaciones. Las células que crecen de manera descontrolada, se apilan una sobre la otra y comienzan a adquirir una apariencia anómala. Las células anómalas continúan cambiando de apariencia, multiplicándose y se transforman en un cáncer *in situ* en el que las células cancerosas quedan confinadas en la zona en donde empezaron a crecer. Si el cáncer no es tratado, las células cancerosas adquirirán características que le permitan invadir otros tejidos (metástasis). Asimismo, el surgimiento de mutaciones asociadas a la resistencia a un tratamiento llevará a la recurrencia del tumor [5, 8]. Figura creada con BioRender.com.

En la práctica clínica, la evaluación de marcadores inmunohistoquímicos como el receptor de estrógenos (RE), el de progesterona (RP) así como la sobreexpresión del receptor del factor de crecimiento epidérmico humano tipo 2 (HER2), así como las citoqueratinas (ck5/6, ck14 y ck17), son utilizados en la clasificación, pronóstico y en la toma de decisión en el tratamiento de las pacientes con cáncer de mama [5, 7, 9, 10].

Sin embargo, con el desarrollo de las tecnologías genómicas y transcriptómicas, se ha logrado clasificar molecularmente al cáncer de mama en diferentes subtipos. A lo largo del tiempo, su clasificación se fue refinando hasta obtener hoy en día



cuatro tipos específicos de acuerdo a los perfiles de expresión de RNA mensajero codificante de proteínas (mRNA, por sus siglas en inglés). El algoritmo o clasificador PAM50 es una herramienta que se fundamenta en la medición de 50 genes para clasificar muestras de cáncer de mama entre los cuatro subtipos; luminal A (LumA), luminal B (LumB), el subtipo con amplificación o sobreexpresión del gen *erb2* (HER2) y el subtipo basal (figura 4a). Cada uno de estos subtipos tiene un estadio clínico y pronóstico diferente entre ellos, reportándose que los tumores con fenotipo HER2 y basal presentan un peor pronóstico con menor supervivencia y son más agresivos (figura 4b). En cuanto a su frecuencia, los subtipos luminales son los más detectados, y en menor frecuencia los subtipo HER2 y basal [5, 9, 11, 12].



**Figura 4. Clasificación molecular del CaMa y pronóstico clínico.** a) Identificación de subtipos moleculares del cáncer de mama basados en la expresión diferencial de genes. El mapa de calor muestra los genes sobre y sub-expresados (rojo y verde, respectivamente) entre las muestras de cáncer de mama, el dendrograma muestra los subtipos intrínsecos del CaMa, Luminal A (morado), Luminal B (azul), Basal (rojo) y HER2 (fiusha). b) La curva de supervivencia muestra como los subtipos basal y HER2 son los que presentan menor supervivencia global, considerando a estos subtipo los más agresivos. Imágenes modificadas de [9, 12].

En la búsqueda de una categorización que homologue o subrogue la clasificación de los tumores de mama mediante pruebas moleculares como PAM50 e inmunohistoquímica (tabla 1), a la fecha se reconoce que los tumores luminales expresan los receptores hormonales RE y RP, con presencia o ausencia de HER2; mientras que el subtipo HER2 presenta la amplificación o enriquecimiento del gen

*erbb2* pero no expresa receptores hormonales. Mientras que la mayoría de los tumores basales, no expresan el RE, RP y HER2, por lo tanto al no expresar dichos receptores, los tumores serán conocidos como tumores de mama triple negativo (CMTN) [10, 13].

Tabla 1. Clasificación molecular del cáncer de mama y marcadores subrogados.

Subtipos moleculares	Aproximación IHQ	Marcadores adicionales
Luminal A	RE+ RP>20% Ki67<20% HER2 –	
Luminal B	RE+ RP<20% Ki67<20% HER2 + o –	
Her2	HER2 + RE y RP –	CK5/6 +/- EGFR +/-
Basal	RE- RP- HER2- EGFR+ CK5-6 CK14 y 17 Expresión Ki67 alta	CK5/6 + EGFR +
Triple negativo	RE – RP – HER2 –	

Imunohistoquímica; IHQ, receptor de estrógenos; RE, receptor de progesterona; RP, receptor epidérmico humano 2; HER2, citoqueratinas;ck, receptor del factor de crecimiento epidermal; EGFR.

Sin embargo, el uso del término basal o triple negativo no debe ser utilizado como sinónimo para definir ambos subtipos de cáncer de mama. Si bien la ausencia de los tres marcadores inmunohistoquímicos en el subtipo basal define a un tumor triple negativo, al momento de evaluar los perfiles de expresión de los tumores de CMTN no todos los tumores serán de un fenotipo basal, encontrando así una discordancia entre la clasificación por la expresión inmunohistoquímica de los receptores hormonales y HER2, y la expresión global por mRNA [13, 14].

Estudios demuestran que al analizar la expresión global de genes en los tumores triple negativos clasificados por inmunohistoquímica, el 78.6% de los tumores se

---

---

clasificaron como basales y el 21.4% fue clasificado entre el resto de los subtipos (Luminal A y B, HER2). Por otra parte, la expresión inmunohistoquímica de los tumores con fenotipo basal (clasificados por PAM50) demostró que el 68.5% de los tumores basales eran negativos a los tres marcadores inmunohistoquímicos (RE, RP y HER2) y el 31.5% presentaron positividad a alguno de los marcadores hormonales o HER2 [13]. Por lo tanto, la mayoría de los tumores triple negativos serán de un fenotipo basal, pero no todos los tumores basales serán triple negativos.

### ***Cáncer de mama triple negativo (CMTN)***

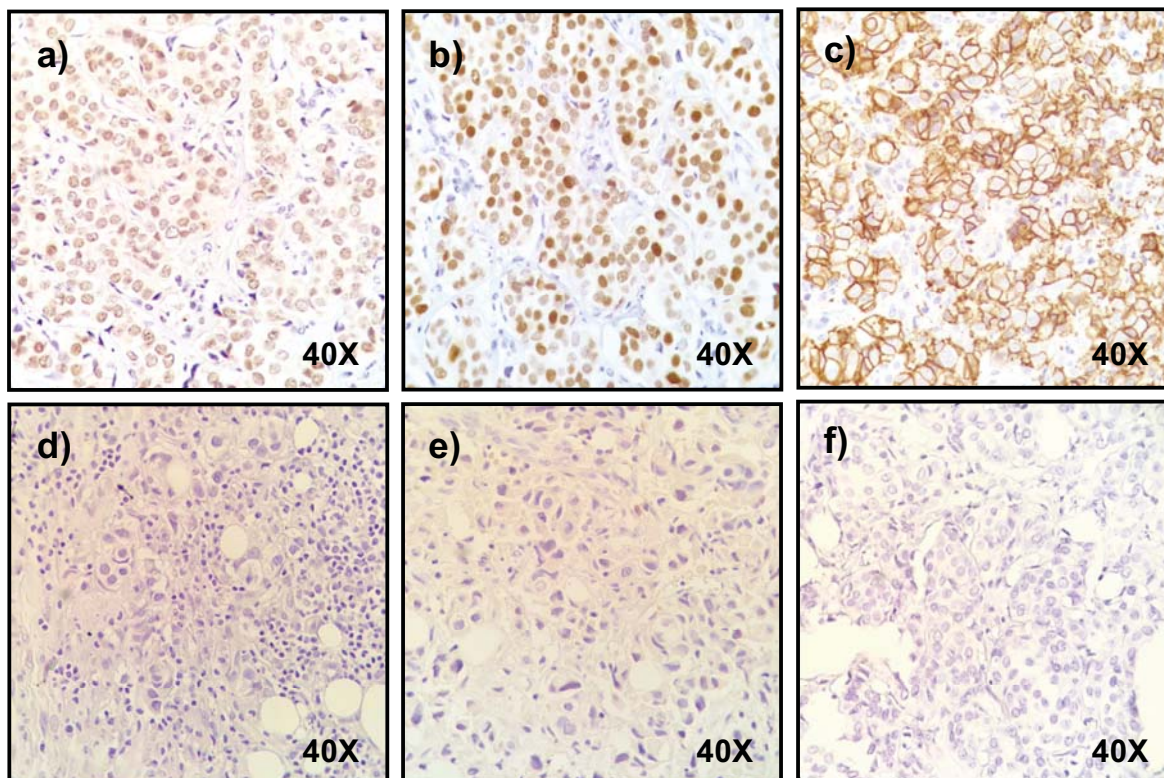
El cáncer de mama triple negativo está definido clínicamente por la negatividad a los receptores de estrógenos, progesterona y la falta de expresión a HER2 (figura 5). Se presenta principalmente en mujeres jóvenes premenopáusicas, caracterizándose por ser de un peor pronóstico clínico y altamente agresivo, tiene las menores tasas de supervivencia global a 5 años y una mayor probabilidad de metástasis en comparación de otros subtipos tumorales. Afecta mayoritariamente a mujeres de etnias afroamericanas o hispanas; está más asociado a un grado tumoral alto, y carece de una terapia dirigida [15-20]. Además, se ha observado altas tasas de CMTN en mujeres con una mayor probabilidad de mutaciones en el gen BRCA1 [21, 22].

El CMTN es una enfermedad morfológica, molecular y clínicamente heterogénea, que presenta una alta inestabilidad genómica, un alto número de alteraciones somáticas en el número de copias de DNA, entre un 50-75% de la pacientes presentan mutaciones en el gen BRCA1 y en un 16-23% en BRCA2 [17, 23-25]. A nivel histopatológico, con frecuencia los tumores triple negativo son carcinomas de tipo ductal invasor (80%) [26] con altos índices mitóticos y con zonas centrales necróticas e infiltrados linfocitarios [27]. En términos de eventos de metástasis, los tumores de CMTN tienden a invadir preferencialmente a nivel visceral (68.7%), pulmón (50.4%), hueso (38.9%), hígado (35.1%) y cerebro (25.2%) [28-32]. La

---

---

recurrencia temprana suele suceder mayormente en CMTN en comparación de otros subtipos [27].



**Figura 5. Representación de la evaluación inmunohistoquímica de receptores hormonales y HER2 en cáncer de mama. a) y b) se muestra la positividad nuclear de la expresión del receptor de estrógenos y progesterona respectivamente. d) y e) representan la ausencia de la expresión de estrógenos y progesterona en las células tumorales en un contexto negativo citoplasmático y estromal, respectivamente. c) positividad y f) negatividad a la expresión de HER2 a nivel membrana en tejido tumoral. Imagen cortesía de la Dra. Lizbett Hidalgo Pérez.**

En cuanto a tratamiento, el uso de agentes quimioterapéuticos sigue siendo el estándar para las pacientes con CMTN. Paradójicamente, aunque las pacientes con CMTN tienen un pobre pronóstico, se ha demostrado que el tratamiento quimioterapéutico ofrece mayor respuesta en comparación de otros subtipos de cáncer de mama no triple negativo [33-36]. Ante un tumor no operable por el tamaño que presenta, está indicado el uso de un tratamiento neoadyuvante con el objetivo de reducir el tamaño del tumor antes del primer tratamiento, que generalmente es la cirugía. El tratamiento basado en antraciclinas (como doxorrubicina) y taxanos (pacli- o docetaxel) ha demostrado ser un estándar para la atención de las pacientes con CMTN[27].

---

---

La respuesta patológica completa (RpC o pCR) es definida por la FDA como la ausencia del cáncer invasivo residual en la evaluación hematoxilina y eosina de la muestra de mama resecada completa y de todos los ganglios linfáticos regionales muestreados después de completar la terapia sistémica neoadyuvante. En CMTN la tasa de RpC ha sido reportada entre un 30-40% con un régimen de tratamiento basado en antraciclinas y taxanos. Sin embargo, se busca aumentar estas tasas de respuesta patológica completa, con el uso de otros agentes quimioterapéuticos (capecitabina, 5-fluorouracilo, vinorelbina, bevacizumab, gemcitabina y agentes de platino) [27].

A nivel mundial, el CMTN representa del 12-20% de todos los cánceres que afectan a la glándula mamaria [15, 20, 37]. Mientras que en países latinoamericanos como Perú, Venezuela, Brasil y Costa Rica se reportó el 21%[38], 24%[39], 19.5%[40] y 22.2%[41] de casos de tumores triple negativos, respectivamente. En México los datos reportados sobre el CMTN son controversiales, presentando frecuencias de 14 -23% de los tumores de cáncer de mama [42-45].

### ***Clasificación del cáncer de mama triple negativo y sus implicaciones clínicas***

Como se ha descrito, el CMTN es una patología compleja tanto a nivel de tratamiento, clínica y biológicamente. Aunado a esto, a través del uso de herramientas de alto rendimiento, como es el caso de las tecnologías genómicas, se ha observado que el CMTN no es una entidad clínica única en su clasificación. En 2011 investigadores de la universidad de Vanderbilt analizaron 386 tumores triple negativos basados en perfiles de expresión de mRNA y por primera vez se identificaron varios subtipos moleculares basado en algoritmos como *k-means* (algoritmo de clasificación no supervisada) y *consensus clustering* (algoritmo para la agrupación por consenso) [46]. Adicional a esta primera identificación, los investigadores ponen a disposición la herramienta para clasificar a los subtipos de CMTN y validaron en una cohorte independiente de 201 muestras [47]. De acuerdo a los resultados de Lehmann y colaboradores, el CMTN se puede



---

---

clasificar en seis subtipos moleculares basados en perfiles de expresión y a sus ontologías o rutas canónicas resultantes de la expresión diferencial de mRNA. Cada uno de estos subtipos se caracteriza por rutas biológicas enriquecidas y sus posibles moléculas blanco como se muestra a continuación (figura 6):

- **Basal 1 (BL1):** un subtipo enriquecido en procesos como la fase G2, G1 y S del ciclo celular, reparación de DNA, vías BRCA/ATR, división celular.
- **Basal 2 (BL2):** este subtipo posee genes que enriquecen procesos como señalización de metabolismo y factores de crecimiento, señalización de las vías *EGF*, *NGF*, *MET*, *WNT*-beta catenina, *IGFR1*, vías como glicólisis y gluconeogénesis.
- **Mesenquimal (M):** está altamente enriquecido en genes codificantes de proteínas relacionados con vías como motilidad celular, interacción del receptor de matriz extracelular (ECM), vías de proliferación celular *WNT*, *TGF*-beta, enriquecimiento del proceso de transición epitelio mesénquima (EMT).
- **Parecidos a troncal mesenquimal (MSL):** este subtipo presenta enriquecimiento de rutas similares al subtipo M incluyendo la motilidad celular por la vía Rho, diferenciación celular, vías de crecimiento (*AKL/TGF*-beta/*WNT*). Sin embargo, lo que lo diferencia del subtipo M es la expresión de genes que enriquecen los procesos como vías de señalización de factores de crecimiento que incluyen el metabolismo de inositol-fosfato, *EGFR*, *PDGF*, señalización mediada por calcio, receptor acoplado a proteínas G, señalización *ERK1/2*, transportadores ABC y señalización de adipocitocinas. Este subtipo se distingue del M por presentar bajos niveles de expresión de genes relacionados con la proliferación, además del enriquecimiento de genes relacionados con células troncales, y baja expresión de claudinas.
- **Inmunomodulador (IM):** presenta enriquecimiento en rutas relacionados con respuesta inmunológica, estos procesos incluyen señalización de citosinas, procesamiento y presentación de antígenos, señalización de la vía *JAK/STAT*, *TNF* y *NFkB*. El análisis por microdissección en tumores IM

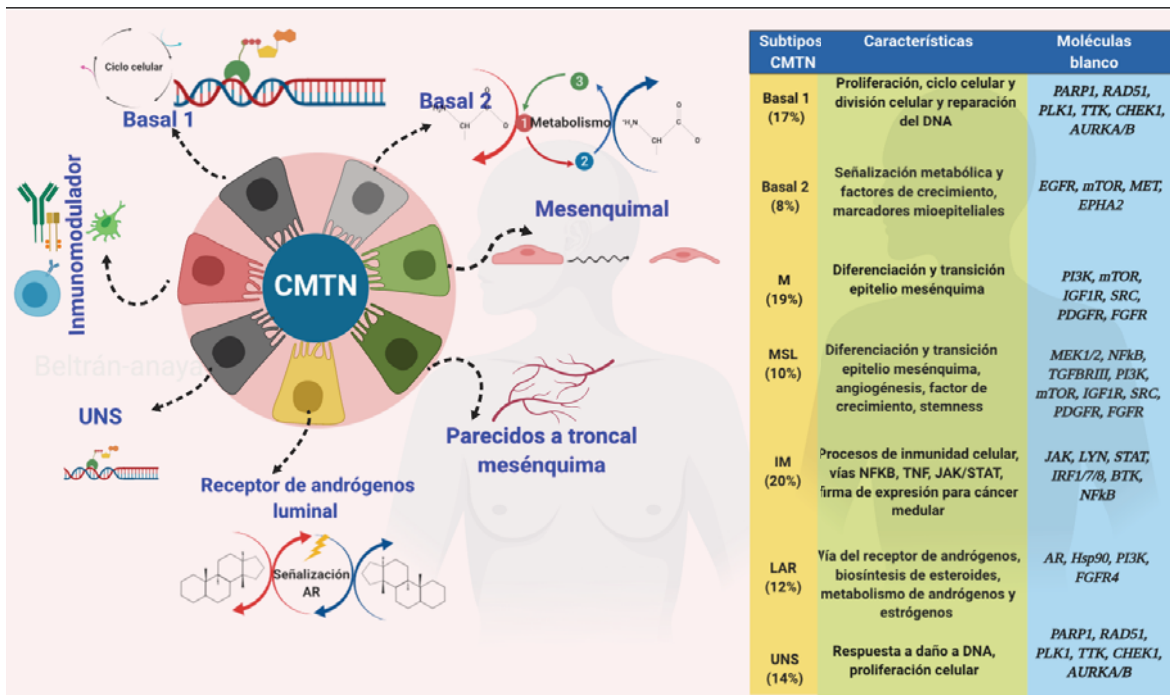
---

---

mostraron genes de señalización inmune que se traslapan con una firma de genes relacionados con cáncer de mama medular, un tipo de tumor histológicamente diferente al CMTN, el cual está asociado a un pronóstico favorable.

- **Receptor de andrógenos luminal (LAR):** un subtipo molecular negativo para el receptor de estrógenos, pero que tiene un enriquecimiento en vías de regulación hormonal incluyendo síntesis de esteroides y metabolismo de andrógenos y estrógenos. Este subtipo sobreexpresa abundantes transcritos río debajo de la vía del receptor de andrógenos así como co-activadores. Este subtipo correlaciona fuertemente con otro subtipo de cáncer de mama conocido como apócrino.
- **Inestables o indeterminados (UNS):** aquellas muestras que mediante su perfil de expresión no logren definir o clasificar en algunos de los subtipos anteriormente mencionados se les conocerá como inestables o no clasificables. Sin embargo, se observó que los tumores UNS y basal 1 correlacionan fuertemente con el subtipo basal determinado por PAM50.

Una vez descrita la clasificación del CMTN se inició la búsqueda de moléculas blanco o posibles tratamientos para este tipo de cáncer. Un primer abordaje fue realizado por el grupo de Lehmann y colaboradores, donde mediante la clasificación molecular del CMTN aplicado a un panel de líneas celulares triple negativo evidenciaron que los agentes antimitóticos como las sales de platino e inhibidores de PARP fueron capaces de reducir la población en un 50% en líneas celulares con un fenotipo basal. Por otra parte, inhibidores de la vía mTOR o SRC/ABL fueron eficaces en fenotipos basales, pero principalmente en el fenotipo mesenquimal. Mientras que en el fenotipo LAR, el uso de bicalutamida (anti-androgénico) mostró mayor efecto en la inhibición de las líneas celulares LAR positivas [46].



**Figura 6. Clasificación molecular del cáncer de mama triple negativo.** Primera clasificación molecular del cáncer de mama triple negativo, basado en perfiles de expresión de mRNA de acuerdo a los reportado por Lehmann y colaboradores [46]. Los subtipos fueron obtenidos por el enriquecimiento de vías y ontologías, las cuales se describen brevemente en la tabla, adicionalmente se observan las posibles moléculas blanco para el tratamiento en CMTN. Mesenquimal; M, parecidos a troncal mesenquimal; MSL, inmunomodulador; IM, inestables; UNS. Figura creada con BioRender.com.

Desde otro punto de vista, Masuda y colaboradores analizaron la implicación clínica de conocer los subtipos de cáncer de mama triple negativo en una cohorte de pacientes bajo tratamiento de quimioterapia neoadyuvante basada en taxanos y antraciclinas. Observaron que las pacientes con un subtipo BL1 presentaron mayor tasa de RpC (52%) que el resto de los subtipos triple negativos (BL2=0%, MSL=23%, LAR10%, M=31%, IM=30%, UNS=33%) [48].

Posteriormente, Lehman y colaboradores analizaron la RpC de cinco ensayos clínicos cuyo esquema de tratamiento se basa en diferentes agentes quimioterapéuticos. En pacientes con tratamiento secuencial basado en antraciclinas y taxanos, las tasas de respuesta más altas fueron observadas en los subtipos BL1y M (>30%). En cambio, los subtipo BL2, M y LAR presentaron tasas de RcP >30% en pacientes tratadas con paclitaxel seguido de una combinación de 5-fluorouracilo, epirubicina y ciclofosfamida. En las pacientes con esquemas de tratamiento basados en antraciclinas y ciclofosfamida seguidas de taxano o



---

---

ixabepilona, los subtipos que mostraron mayor tasa de respuesta patológica completa fueron los subtipos BL1 y 2, así como LAR [49].

De acuerdo a los resultados obtenidos por Santoja y colaboradores, los subtipos BL1y M presentaron mayores tasas de RpC superiores al 70% con tratamiento basado en antraciclinas, taxanos y carboplatino. Mientras que los subtipos basal 2 e IM mostraron una tasa superior al 50% bajo un esquema con antraciclina y taxanos, las pacientes con fenotipo LAR fueron las que obtuvieron una menor RpC (<20%) tanto bajo esquemas de antraciclinas como de taxanos [50].

El estudio reportado por Echavarría y colaboradores demostró que las pacientes con mayor RpC (65.6%) presentaban un fenotipo basal 1 y cuyo tratamiento se basó en taxanos y carboplatino. En sentido opuesto, los subtipos LAR presentaron menor tasa de RpC (21.4%) con un mismo esquema de tratamiento [51].

Como se aprecia en los estudios anteriormente mencionados, existe una alta heterogeneidad en cuanto a la respuesta a tratamiento bajo esquemas con agentes quimioterapéuticos similares entre subtipos de cáncer de triple negativo. Por lo que nuevos esfuerzos se están llevando a cabo para poder redefinir nuevos subtipos de cáncer de mama triple negativo (figura 7).

### ***Clasificadores adicionales del cáncer de mama triple negativo a lo largo del tiempo***

En la búsqueda de nuevos tratamientos específicos basados en los subtipos moleculares del CMTN, nuevos esfuerzos buscan reducir el número de genes y la utilización de diversos niveles de información como DNA, proteínas y estados de metilación (figura 7). Desde sus inicios en el 2011, se definieron seis subtipos principales basados en perfiles de expresión. En 2015, tres diferentes métodos fueron propuestos para estratificar al CMTN, uno de ellos fue dirigido por la universidad de Baylor por Burstein y colaboradores, donde realizaron diferentes subclasificadores basados, no solamente en la expresión de mRNA como lo han hecho el algoritmo de Vanderbilt (TNBCtype<sup>TM</sup>), sino además incluyeron el análisis de los perfiles de alteraciones somáticas en el número de copias de DNA [52]

---

---

(figura 7). En 2015, el grupo de Jezequel y colaboradores [53] proponen la clasificación de tres subtipos basados en expresión génica los cuales incluyen Basal de baja respuesta inmune (LIR), Basal de alta respuesta inmune (HIR), Receptor de andrógenos luminal (LAR) (figura 7). Otra clasificación, usando otro nivel de información, es la propuesta por Stirzaker y colaboradores [54] donde proponen tres subtipos de triple negativos basados en perfiles de metilación. Uno de ellos es el subtipo con baja metilación, otro de metilación media y un tercer grupo de alta metilación.

Para el 2016, Ring y colaboradores [55] reducen la lista de genes que inicialmente usaba el grupo de Lehmann para clasificar a los triple negativos, usando un total de 101 genes de los 2188 que originalmente se evaluaban [46, 47]. Adicionalmente, se elimina el subtipo IM. Con el nuevo refinamiento o minimización de la firma de expresión, se logró clasificar al CMTN en subtipo Basal 1, Basal 2, M, MSL, LAR. Como se observa, la eliminación del subtipo IM para la clasificación del CMTN está basada en que este fenotipo podría ser el reflejo de los linfocitos infiltrantes del tumor (LITs) [55].

Sin embargo, Lehmann y colaboradores previamente demostraron que al realizar micro-disección de tumores y clasificar nuevamente estas muestras, una proporción de los tumores IM se mantuvo, lo que indicó que probablemente algunas características inmunomoduladoras es específica del tumor y no un reflejo del infiltrado linfocitario [46]. De manera interesante, poco después de la clasificación de Ring, Lehmann y colaboradores determinaron, basados en la expresión de 101 genes, que el subtipo MSL debería salir de la clasificación del CMTN por presentar alta correlación con genes del estroma, lo que podría indicar que los tumores MSL contienen mayor abundancia de tejido mesenquimal asociado a tumor [49, 55].

Por otra parte, Liu y colaboradores [56] identifican varios subtipos, reduciendo de igual manera cuatro grupos que recapitulan la información de los seis subtipos descritos por Lehmann. Otro abordaje novedoso para la clasificación del CMTN fue realizado por Masuda y colaboradores [57], mediante análisis de arreglos de

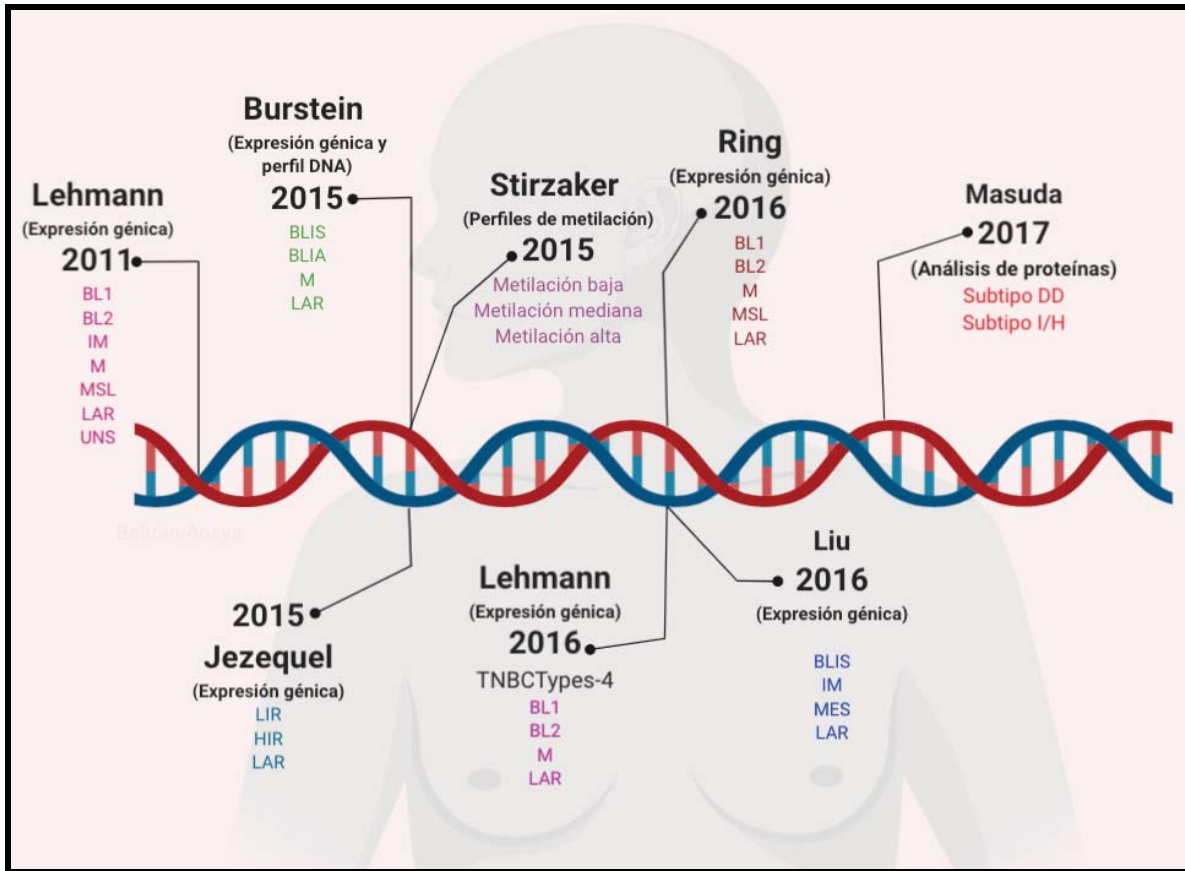
---

---

proteínas, identificaron dos subtipos principales de triple negativos; el enriquecido en daño a DNA (DD) y el enriquecido en vías de inmunidad y hormonas (IH).

Durante un periodo no mayor a diez años, se han generado diferentes algoritmos para la clasificación de los tumores de mama triple negativos (figura 7), observándose una diversidad de clasificadores. Sin embargo, muchos de estos clasificadores convergen o recapitulan la clasificación que originalmente propuso Lehmann en el 2011. Actualmente, el estándar para la clasificación de tumores de CMTN sigue siendo el algoritmo implementado por la universidad de Vanderbilt, que identifica a 4 subtipos moleculares (Basal 1, Basal 2, M y LAR) y adicionalmente se genera un clasificador inmunomodulador complementario que podría ayudar a la predicción de respuesta inmuno-oncológica [49, 58].

Como se observó, los algoritmos para clasificar a los tumores triple negativos se basan muchos de ellos en los perfiles de expresión de mRNA los cuales representan una parte del transcriptoma global. Sin embargo, existe otra variedad de RNAs sin capacidad codificante de proteínas, entre estos se encuentran dos grandes grupos, los RNAs pequeños y los RNAs largos no codificantes.

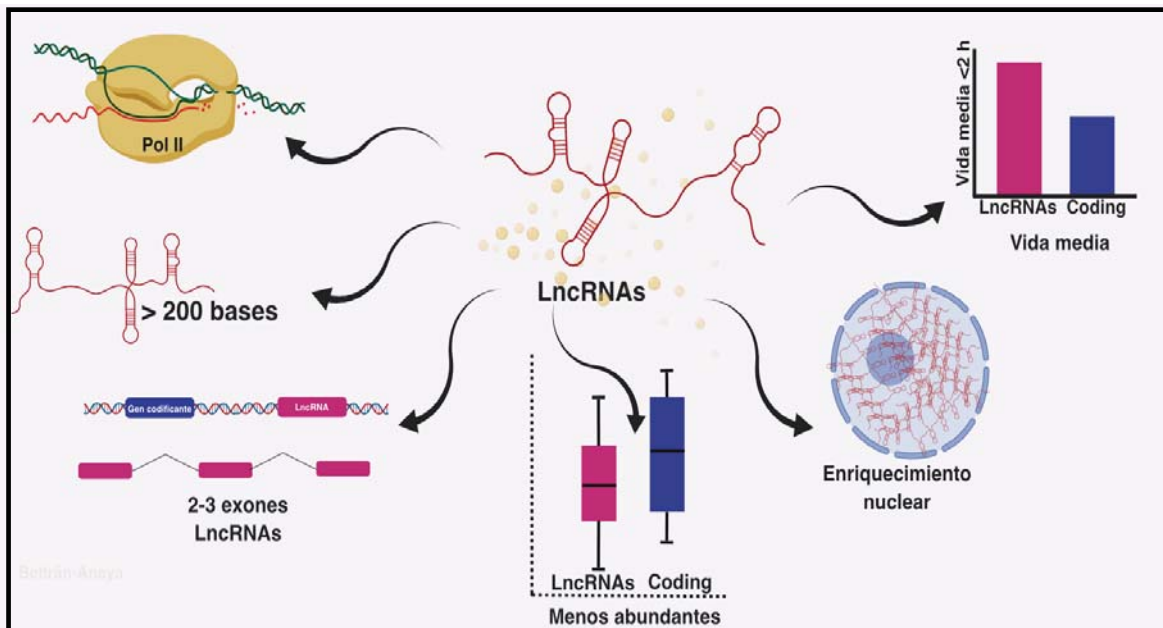


**Figura 7. Estado del arte de la clasificación molecular del cáncer triple negativo.** La imagen muestra como se fueron desarrollando diferentes tipos de clasificación para los tumores triple negativos. Desde sus inicios en el 2011, se definieron seis subtipos principales basados en perfiles de expresión. En el 2015 dos diferentes trabajos proponen nuevas clasificaciones basados en la expresión de mRNA (Jezequel 2015 [53]) mientras que otro estudio además de evaluar la expresión de mRNA, evaluó el número de copias de DNA (Burstein 2015 [52]). Por otra parte el abordaje realizado con perfiles de metilación fue sugerido por Stirzaker y colaboradores. El grupo de metilación baja o hipometilado se asoció a un mejor pronóstico en comparación del grupo hipermetilado, mientras que el grupo de metilación intermedia se relacionó con peor pronóstico. Para el 2016, Ring y colaboradores reducen la lista de genes que inicialmente usaba el grupo de Lehmann para clasificar a los triple negativos, usando un total de 101 genes de los 2188 que originalmente se evaluaban. Adicionalmente, se elimina el subtipo IM. Posteriormente, Lehmann en 2016 refina la clasificación de los tumores y eliminan al subtipo MSL, quedándose con la identificación de 4 subtipos. Para el 2016 Liu y colaboradores crean el clasificador basado en mRNA e identifican 4 subtipos, recapitulando así la información de los seis subtipos descritos por Lehmann. En 2017 Masuda y colaboradores hacen la clasificación del CMTN basado en análisis de proteínas, se identifican dos subtipos de triples negativos, el subtipo DD o de "Daño a DNA" y al subtipos enriquecido en la vía "hormonal y de inmunidad" (I/H). BL1: basal 1, BL2: basal 2, IM: inmunomodulador, M: mesenquimal, MSL: parecidos a trocal mesénquima, LAR: receptor de andrógenos luminal, UNS: inestables o no caracterizados, LIR: basal de baja respuesta inmune, HIR: basal de alta respuesta inmune, BLIS: inmunosuprimidos de tipo basal, BLIA: inmunoactivados de tipo basal, MES: parecidas a mesenquimal. Figura creada con BioRender.com.

## Los llamados RNAs de la materia oscura

Por años a nivel de genoma se ha hablado de la existencia de DNA "basura", dado a que no se producían proteínas de dichas regiones intergénicas. Conforme ha avanzado la tecnología, se observó que dichas regiones se transcriben sin obtener

una proteína, llamando a estos transcritos como “ruido transcripcional”, el cual posteriormente se identificó como RNA no codificante [59]. Adicionalmente, los primeros estudios del genoma humano, determinaron que menos del 5% del DNA era codificante [60]; actualmente se sabe que del total de RNA transcrito, el 66% pertenece a RNA no codificante[61]. Este tipo de RNA se divide en dos grandes grupos: los RNAs pequeños como los microRNAs (~22 nucleótidos) y RNAs largos no codificantes (lncRNAs, por sus siglas en inglés), estos últimos se caracterizan por su longitud mayor a 200 nucleótidos hasta ~100 kb. Son procesados por la polimerasa tipo II, tienen localización preferencialmente nuclear, en promedio presentan de 2 a 3 exones, pueden tener un origen intergénico o intragénico, tienen una expresión menor a los RNA mensajeros y sobre todo, no codifican proteínas [62] (figura 8). Además, se ha considerado que estas moléculas podrían ser RNAs intermediarios durante el proceso de evolución de nuevas proteínas [63].



**Figura 8. Características de los lncRNAs.** En la figura se muestra las características generales de los lncRNAs, como su procesamiento, tamaño, abundancia, compartimentalización celular y vida media. Figura creada con BioRender.com.

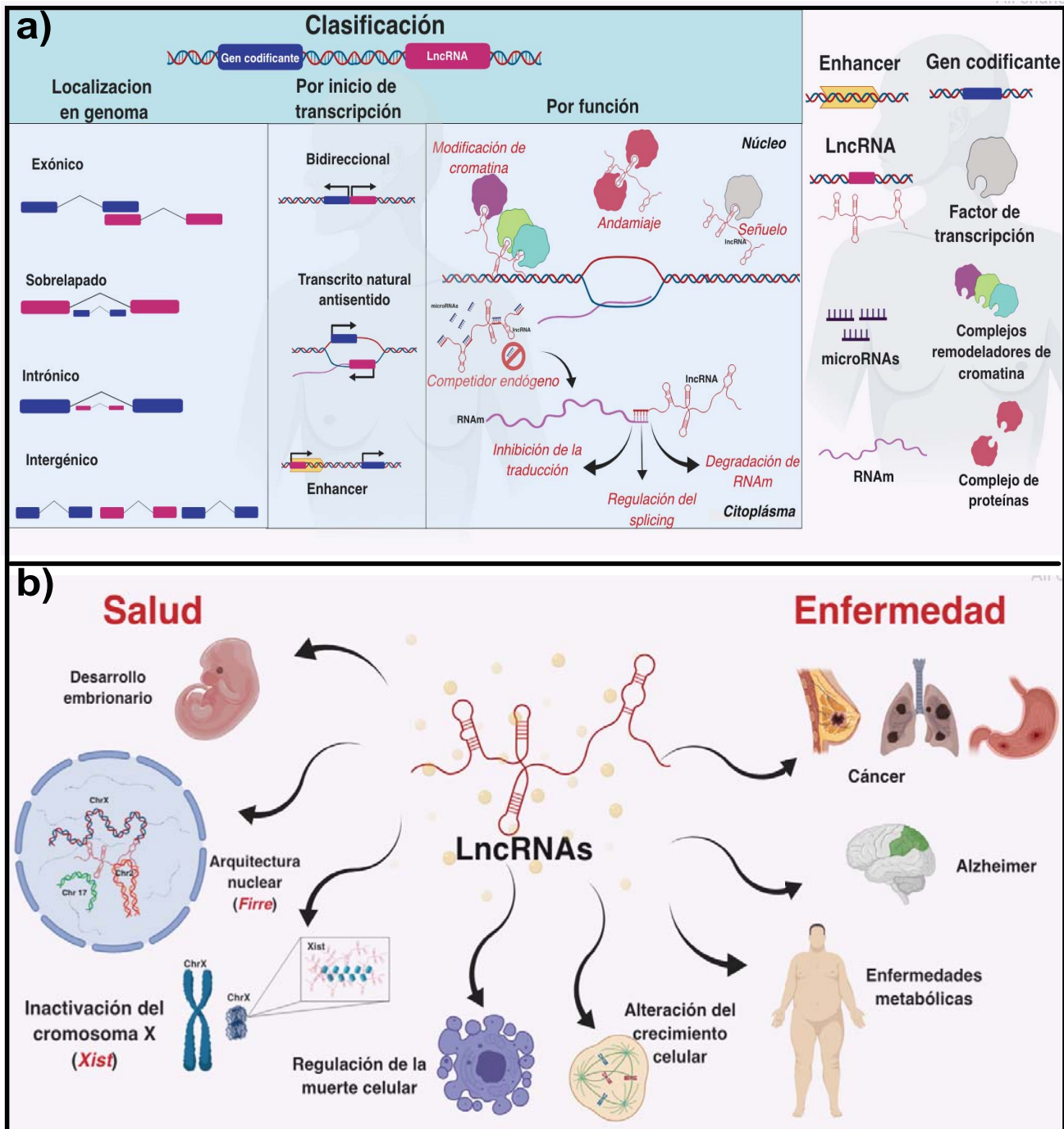
---

---

## ***Clasificación y función de lncRNAs***

En los últimos años, mediante análisis funcionales se ha demostrado que los lncRNAs desempeñan papeles importantes en los procesos como regulación de la transcripción en *cis* o *trans*, organización de dominios nucleares, así como la regulación de proteínas y mRNA [64]. La clasificación de lncRNA debe considerar varios aspectos para asignarlos en un grupo determinado; es decir, los lncRNA pueden clasificarse con base a su tamaño, a su asociación con genes codificantes de proteínas o mediante elementos funcionales del DNA como promotores, secuencias potenciadoras (*enhancer*) o regiones teloméricas [65]. De acuerdo a su localización en el genoma, los lncRNAs se clasifican en dos grandes grupos: intergénicos y génicos. Este último grupo puede dividirse en intrónicos si el lncRNA está albergado en el intrón del gen codificante de proteínas; sobrelapados, si el gen codificante de proteínas está en el intrón del lncRNA, o exónico, si un exón del lncRNA sobrelapa con el exón del gen codificante (a) [62].





**Figura 9. Clasificación de lncRNAs y su participación en salud y enfermedad. a)** En la figura, se muestra las diversas clasificaciones del lncRNA de acuerdo a su localización en el genoma, por su inicio de transcripción y por sus funciones moleculares. En cuanto a la función, un lncRNA puede actuar ya sea remodelando cromatina y estar enriquecido en núcleo, pero también podría estar enriquecido en citoplasma, probablemente ejerciendo una función esponja o como competitivo endógeno hacia microRNA que deberían ejercer su función hacia sus mRNA blanco. Los lncRNAs enriquecidos en núcleo actúan ya sea manteniendo la organización nuclear, como moléculas de andamiaje, ayudando a formar complejos de proteína-proteína, así como remodelando la cromatina para activar o reprimir genes. **b)** Representación de las funciones descritas por los lncRNA en procesos de salud y enfermedad, tanto por ausencia o presencia en su expresión. Evidencias sugieren a los lncRNAs pueden regular el desarrollo embrionario, mantenimiento de la arquitectura nuclear por el lncRNA Firre, la inactivación del cromosoma X en mujeres llevada a cabo por el lncRNA Xist y la regulación de procesos celulares como muerte celular. Sin embargo, la alteración en la expresión de estas moléculas favorece la alteración de procesos celulares (como el descontrol del

---

---

crecimiento celular) o a la alteración funcional de forma sistémica, por ejemplo enfermedades metabólicas, neuropatologías y cáncer. Figura creada con BioRender.com.

Sin embargo, la mayoría de los lncRNAs se clasifican de acuerdo a su función biológica; entre las cuales se encuentran las relacionadas con las modificaciones de histonas, ya sea para activar o reprimir la expresión de genes blanco como es el caso prototipo de HOTAIR [66, 67] y XIST [68] que puede actuar a nivel *cis* o *trans*. La función señuelo, es una forma más de cómo actúan los lncRNAs como el caso de GAS5, donde se une al receptor de glucocorticoides (RG), lo que lleva a que el RG no pueda ejercer su efecto transcripcional y por lo tanto, no se expresen genes respondedores a RG [69]. La función de andamiaje, es una característica que presentan algunos lncRNAs como HOTAIR, GCInc1 y LINC00511, los cuales pueden reclutar modificadores de cromatina como EZH2, WDR5, KAT2A [67, 70-72].

A la fecha, son pocos los lncRNAs reportados con capacidad de regulación del mRNA; ya sea a nivel de procesos de traducción, como el lincRNA-p21, el cual se asocia con el mRNA de JUNB y CTNNB1 bajando su traducción [73]. Por otra parte, la interacción entre lncRNA-mRNA puede llevar a la degradación del mRNA blanco, tal es el caso de Staufen1 [74]. En cuanto a la regulación del corte y empalme (splicing, por sus siglas en inglés) de mRNA se encuentran como principales ejemplos al lncRNA NEAT1, MALAT, FGFR2-AS [75-77]. Una función actualmente evaluada en varios lncRNAs es la función esponja o competidor endógeno; este proceso post-transcripcional tiene como objetivo impedir que los microRNAs degraden mRNA al presentar “*sitios semilla*” similares al RNA blanco [78, 79]. Entre algunos ejemplos podemos mencionar a H19 [80], GAS5 [81], PTEN [82], linc-ROR [83] entre otros, los cuales repercuten en los “*hallmarks*” del cáncer [84].

Durante el desarrollo del cáncer, los lncRNA se presentan como nuevos protagonistas involucrados en la alteración de la función normal de una célula (figura 9b); sin embargo existen lncRNAs que tienen la función del control y desarrollo normal de un organismo, como por ejemplo, durante el desarrollo de la glándula mamaria.



---

---

## ***Papel de los lncRNAs en el desarrollo de la glándula mamaria***

Como se mencionó previamente, la glándula mamaria está compuesta por una gran diversidad de componentes celulares, los cuales deben de presentar una homeostasis durante su desarrollo. El papel de los lncRNAs ha sido escasamente explorado durante estos procesos de vital importancia; por lo tanto, son pocos los lncRNAs descritos a la fecha directamente involucrados en el desarrollo de la glándula mamaria. Entre estos tenemos a H19 y SRA1 (RNA activador del receptor de esteroides 1). En modelos animales, se demostró que la inducción de la expresión de H19 por estrógenos, se enriquece en células epiteliales de los brotes de extremo terminal (TEB, por sus siglas en inglés) de la glándula mamaria. De manera inicial, la función descrita de H19 fue en el control del crecimiento durante el desarrollo embrionógico, por lo que podría tener una función equivalente en el desarrollo de la glándula mamaria [85]. SRA1 es un lncRNA de escasa expresión en la glándula mamaria *virgen* [86], modula la actividad de los receptores hormonales de tipo nuclear y aumenta las ramificaciones del lado ductal e induce la diferenciación alveolar precoz cuando se sobreexpresa en células mamarias epiteliales en modelos de ratón [87].

El mPINC es un lncRNA descrito originalmente en ratones, localizado abundantemente en progenitores alveolares y escasamente en progenitores luminales. Su subexpresión llevó al incremento de la proliferación en células del epitelio de la glándula mamaria derivadas de ratones a mitad del embarazo [88]. Zfas1 es un lncRNA localizado en células ductales así como alveolares de la glándula mamaria; al igual que mPINC, disminuyen durante la diferenciación lactogénica de la glándula mamaria [89].

## ***LncRNAs y cáncer de mama***

Recientemente los lncRNAs han tomado relevancia debido a que se ha comprobado que la alteración en su expresión juega un papel fundamental en el desarrollo de enfermedades humanas, incluyendo el cáncer [90, 91].

---

---

Uno de los estudios más relevantes sobre las alteraciones de los perfiles de expresión de los lncRNAs en cánceres humanos, ha sido reportado por el consorcio del Atlas del Genoma del Cáncer (TCGA, por sus siglas en inglés) en el cual se analizaron 13 tipos de cáncer diferentes. Esto permitió corroborar que los lncRNAs presentan un perfil de expresión tejido específico, el cual se ve modificado significativamente en los tumores, debido a alteraciones a nivel transcripcional, genómico y epigenómico [92].

En cáncer de mama, los perfiles de expresión de los lncRNAs entre los subtipos intrínsecos (luminal A, luminal B, Basal y HER2) han sido explorado por el consorcio del TCGA, demostrando que al menos dos terceras partes de los lncRNAs presentan marcas de cromatina H3K27ac, una marca asociada a regiones potenciadoras de expresión de genes (“enhancer”). Adicionalmente, se observó que los lncRNAs se correlacionaban de manera positiva con la expresión de genes codificantes de proteínas adyacentes [93].

En cambio, Chen y colaboradores analizaron un número limitado de tumores triple negativos y tejido normal, y demostraron que una gran parte de los lncRNAs alterados se encuentran sub-expresados en el tumor en relación al tejido no canceroso [94]. Como se ha expuesto, diversos estudios han mostrado la existencia de alteraciones en los perfiles de expresión, tanto de RNAs codificantes y no codificantes en cáncer de mama [95, 96].

Se han desarrollado diversas firmas moleculares que evalúan la expresión de diferentes genes con la finalidad de subclasificar a los tumores en grupos de riesgo específico, lo que ha permitido generar herramientas de utilidad clínica. Sin embargo, el papel de los lncRNAs en cáncer de mama triple negativo ha sido escasamente explorado. Una forma de iniciar, es la identificación de los perfiles de expresión y caracterización funcional de los lncRNAs con relevancia biológica en el desarrollo del tumor.

---

---

## ***Pregunta experimental***

En los tumores cáncer de mama triples negativos, ¿Existirán patrones diferenciales de expresión de lncRNAs entre los subtipos de CMTN definidos por mRNAs? ¿La caracterización funcional de un lncRNA aportará información sobre procesos celulares relacionados con cáncer?

## **Justificación**

El cáncer de mama triple negativo es una enfermedad clínica y biológicamente heterogénea que afecta principalmente a mujeres jóvenes y premenopáusicas. Tiene un tratamiento limitado, es más agresivo y con menor supervivencia global. Aunque existen estudios donde se ha clasificado a los tumores triple negativos en subtipos específicos, estos se basan solo en el análisis de expresión de genes codificantes. Actualmente se sabe que el 66% del transcriptoma corresponde a RNAs no codificantes, entre ellos los lncRNAs, los cuales han sido involucrados en la alteración de la expresión global de genes, inestabilidad genómica y asociados con menor supervivencia en cáncer. Sin embargo, muchos de los lncRNAs no han sido completamente caracterizados, no solo a nivel de expresión, si no también a nivel funcional. Por lo tanto, dilucidar los perfiles de expresión de lncRNAs asociados a los subtipos de cáncer de mama triple negativo, así como su efecto funcional, permitirá conocer más acerca de los procesos celulares relacionados con la biología del tumor y su correlación con parámetros clínicos como supervivencia.

## ***Hipótesis***

Si existe variación de los perfiles de expresión de los lncRNAs en subtipos de cáncer de mama triple negativo, entonces será posible identificar la heterogeneidad de la expresión de lncRNAs entre subtipos moleculares de los tumores triple negativo, los cuales se asociarán a rutas biológicas particulares. Además, la alteración en la expresión de un lncRNA estará involucrada en la modulación de procesos celulares implicados en la biología del tumor.

---

---

## **Objetivos**

### **Objetivo general**

Determinar los perfiles de expresión de RNAs largos no codificantes y RNAs mensajeros en carcinomas de mama triple negativos.

### **Objetivos específicos**

1. Analizar los perfiles de expresión de mRNA para clasificar a los subtipos moleculares de cáncer de mama triple negativo.
2. Identificar los perfiles de expresión de lncRNAs entre los subtipos de cáncer de mama triple negativo.
3. Validación de la expresión de lncRNAs *in silico* en muestras independientes de cáncer de mama triple negativo derivado de bases de datos públicas.
4. Determinar la correlación de mRNA y lncRNAs para el enriquecimiento de rutas biológicas mediante análisis de culpabilidad en cáncer de mama triple negativo.
5. Identificar lncRNAs y mRNAs correlacionados positivamente en su expresión para su validación funcional.
6. Caracterización funcional de un lncRNA en modelos *in vitro*.

---

---

## **Materiales y métodos**

### **Obtención de muestras de cáncer de mama triple negativo**

Un total de 160 tumores de cáncer de mama derivados de bloques de parafina fueron colectados del Instituto de Enfermedades de la Mama FUCAM de México (N=67) y del Istituto Nazionale dei Tumori de Italia (N=93) a través de una colecta retrospectiva del 2007 al 2015. Las muestras fueron definidas como triple negativas usando criterios internacionales; menos del 1% de positividad para los receptores de estrógenos (RE), progesterona (RP), y del receptor epidérmico humano 2 (HER2). La detección inmunohistoquímica de los tres marcadores fue realizada a través de la valoración de dos patólogos usando los siguientes anticuerpos: RE (clona ID5, Dako, Dinamarca) RP (clona PgR636, Dako, Dinamarca), HER2 (K5204, Dako, Dinamarca). Adicionalmente, con la finalidad de eliminar falsos negativos, se analizó la expresión de citoqueratinas (clona D5/6, Dako, Dinamarca) y Ki67 (16B4, Dako, Dinamarca). Todos los procedimientos fueron realizados de acuerdo con los criterios de la declaración de Helsinki, revisados y aprobados por el comité de ética del Instituto Nacional de Medicina Genómica (número CEI2016/13). Un patólogo realizó la identificación de zonas del tumor con mayor cantidad de tejido canceroso y posteriormente las zonas fueron colectadas mediante una aguja de 2 milímetros de diámetro para realizar la extracción de RNA.

### **Extracción de RNA**

El RNA total de los bloques de parafina fue extraído mediante métodos de columna usando el kit AllPrep® FFPE (Quiagen, CA) de acuerdo a las instrucciones del fabricante. El RNA total de las líneas celulares fue extraído usando el método por Trizol (Invitrogen, CA, USA) de acuerdo a las instrucciones del fabricante. La concentración de RNA se determinó mediante espectrometría con el sistema Nanodrop y fue almacenado a -80°C hasta su procesamiento.

---

---

## **Análisis de expresión de mRNA y lncRNAs mediante microarreglos de expresión**

El perfil transcripcional de RNAs mensajeros y RNA largos no codificantes fueron medidos usando los microarreglos Human Transcriptome Array V2.0 de Affymetrix, de acuerdo a las instrucciones del fabricante. Brevemente, ~200 ng de RNA total fueron convertidos a DNA complementario (cDNA) y marcado con el kit SentationPlus FFPE amplification and WT Labeling de Affymetrix (Affymetrix, Santa Clara CA, USA). El cDNA fue hibridado sobre el microarreglo, posteriormente lavado, teñido y escaneado usando el equipo Genechip Scanner 300 7G (Affymetrix, Santa Clara Ca, USA). Los arreglos fueron leídos en la unidad de microarreglos del Instituto Nacional de Medicina Genómica.

Los valores de intensidad de señal fueron corregidos por el algoritmo robust multiarray analysis (RMA, por sus siglas en inglés) y sumariados de forma conjunta mediante el programa expression console de Affymetrix. Los efectos de lote (variación no relacionada con la biología) fueron ajustados mediante abordaje Bayesiano no paramétrico mediante el algoritmo ComBat, como se ha reportado previamente [97]. Los datos de este estudio están públicamente disponibles en la plataforma GEO: GSE86948).

## **Clasificación de subtipos moleculares de cáncer de mama triple negativo basado en mRNA**

El algoritmo TNBCtype (<http://cbc.mc.vanderbilt.edu/tnbc/>) fue usado para la clasificación de los subtipos de cáncer de mama triple negativo (BL1, BL2, IM, M, MSL, LAR, UNS) basados en la expresión de mRNA [47]. Cuatro de 160 muestras inicialmente incluidas para el estudio fueron excluidas para el análisis de perfiles de expresión de lncRNAs de acuerdo al filtro bimodal del algoritmo TNBCtype que evalúa la expresión de mRNA de los genes de receptores de estrógeno y progesterona, y del factor epidérmico humano 2 [47].

---

---

### **Análisis de ontologías de los subtipos moleculares triple negativos**

Los genes diferencialmente expresados (mRNAs) de cada uno de los subtipos TN previamente clasificados fue usado para el análisis ontológico y rutas biológicas usando el programa DAVID [98]. Genes sobre y subexpresados fueron analizados para corroborar las rutas biológicas enriquecidas en cada uno de los subtipos TN, como se ha descrito previamente [46].

### **Análisis de expresión diferencial de lncRNAs entre subtipos de cáncer de mama triple negativo**

Los perfiles de expresión de lncRNAs fueron analizados usando el software Transcriptome Analysis Console (TAC) V.3.0 de Affymetrix. Para la obtención de cada perfil de lncRNAs se realizó lo siguiente: Se conjuntaron todas las muestras pertenecientes a un subtipo específico y se comparó contra el resto de las muestras no pertenecientes al subtipo a interrogar. Una tasa de cambio  $\geq 1.5$  o  $\leq -1.5$ , valor de  $p < 0.05$  de ANOVA y FDR (false discovery rate)  $< 0.05$  fueron considerados significativos excepto para el análisis del subtipo BL2 donde el FDR a considerar fue de  $< 0.5$ . Paquetes del programa R fueron usados para la creación de los mapas de calor (heatmap2), boxplot (ggplot2), corrección de lote (SVA) y análisis de correlación.

### **Validación de la expresión de lncRNAs en muestras independientes**

162 muestras de una cohorte independiente de tumores TN (GEO:GSE76250) [56] fueron usadas para la validación *in silico* de la expresión de lncRNAs detectados diferencialmente en nuestro estudio. Los valores de intensidad de señal fueron corregidos por el algoritmo RMA y sumados de forma conjunta mediante el programa expression console de Affymetrix.

### **Análisis de enriquecimiento de vías de lncRNAs a través de asociación por culpabilidad**

El análisis de enriquecimiento de vías fue realizado por el análisis de asociación por culpabilidad (guilt-by-association analysis) [99]. Brevemente, se construyó una matriz de correlación con los datos de expresión normalizada de los genes

---

---

diferencialmente expresados (mRNA y lncRNA) para cada uno de los subtipos TN. La correlación se obtuvo mediante correlación de Spearman y valor de p para cada lncRNA y mRNA.

### **Predicción de inmunofenoscóreo entre subtipos triple negativos**

La predicción del inmunofenoscóreo (inmunogenicidad del tumor) fue evaluado usando el algoritmo *Immunophenoscore* disponible en GitHub (<https://github.com/mui-icbi/Immunophenogram>). Brevemente, para cada subtipo TN se enriqueció el algoritmo introduciendo los perfiles de expresión de mRNA específicos derivados del análisis de expresión diferencial anteriormente mencionados. Posteriormente, el algoritmo determinó un puntaje respecto al enriquecimiento en la expresión de cuatro diferentes firmas de expresión inmunológicas (células efectoras, células inmunosupresoras, moléculas del complejo MHC e inmunomoduladores).

### **Cultivo de líneas celulares**

Las líneas celulares BT-20, MDA-MB-468, MCF10A, MDA-MB-231, Hs578-T, MCF-7 y HCC1187 fueron cultivadas bajo condiciones estándar (medio DMEM o RPMI-1640, ATCC, Manassas, VA, USA) de acuerdo a cada requerimiento de la línea celular, suplementado con 10% de suero fetal bovino (SFB, ATCC, Manassas, VA, USA) e incubadas a 37° C con 5% de CO<sub>2</sub>.

### **Síntesis de cDNA por RT-PCR**

El cDNA fue sintetizado usando SuperScript III RT-PCT (Invitrogen, CA, USA) siguiendo las recomendaciones del fabricante. Brevemente, 200 ng de RNA total fueron usados para la síntesis de cDNA en un volumen final de 20µL, la cual contenía 1µL de oligo dT [50µM], 1 µL de hexámeros aleatorios [50ng], 1 µL de dNTPs [10mM], 7µL de agua libre de nucleasas, 2 µL de RT amortiguador [10x], 4 µL de MgCl<sub>2</sub> [25mM], 2 µL de DTT [1M], 1 µL de inhibidor de RNAsa [40U/µL] y 1 µL de enzima SuperScript III [200U/µL]. La mezcla de reacción de PCR fue corrida bajo las siguientes condiciones: 25° C por 10 minutos, 50° C por 50 minutos, 85° C por 5 minutos. Agregar 1 µL de RNAsa H e incubar por 20 minutos a 37°C.



---

---

## **Análisis de expresión de mRNA y lncRNA por reacción en cadena de la polimerasa cuantitativa (RT-qPCR)**

La evaluación de la expresión de mRNA y lncRNA fue realizada en un volumen total de 10µL. La mezcla contenía 1 µL de cDNA, 5 µL de 2X TaqMan Universal Master Mix (Applied Biosystems), 0.5 µL de sonda TaqMan y 3.5 µL de agua libre de nucleasas. GAPDH (Hs99999905\_m1) y U87 (Hs03298717\_s1) fueron usados como controles endógenos para genes codificantes y no codificantes, respectivamente. La tasa de cambio entre las condiciones de tratamiento y condiciones control para cada gen fue calculada por el método de  $2^{-\Delta\Delta Ct}$ , donde  $\Delta\Delta Ct = \Delta Ct \text{ tratamiento (shRNAs)} - \Delta Ct \text{ condición control}$ .  $\Delta Ct = Ct \text{ gen de interés} - Ct \text{ gen endógeno}$ .

## **Generación short hairpin RNAs (shRNAs)**

Los sistemas de silenciamiento shRNAs fueron generados usando BLOCK-iT™ U6 RNAi Entry Vector (Invitrogen, Carlsbad, CA, USA) siguiendo las instrucciones del fabricante (anexo 1). Brevemente, un par de oligonucleótidos complementarios fueron diseñados. La generación de los oligonucleótidos de doble cadena (ds-oligo) fue necesaria para su subsecuente ligación en el vector pENTR™ /U6 el cual fue usado para transformar células competentes de *E. coli* One Shot TOP10. Para garantizar la correcta inserción o posibles cambios de nucleótido en la secuencia del shRNA insertado, se realizó la secuenciación Sanger de los plásmidos shRNAs generados.

## **Modulación de la expresión del lncKLHDC7B por short hairpin RNAs (shRNAs)**

El silenciamiento del lncKLHDC7B fue realizado en tres líneas celulares: HCC1187 (triple negativo), BT-20 (triple negativo) y MCF-7 (luminal A). Dos shRNAs fueron generados para el silenciamiento del lncKLHDC7B (shRNA-1, shRNA-2) y adicionalmente un shRNA control dirigido contra GFP como control negativo (NC, anexo 1). Los shRNAs fueron transfectados con Xfect Transfection Reagent (Clontech, California, USA) como se describe a continuación:  $6.5 \times 10^5$  células de HCC1187,  $5 \times 10^5$  de BT-20 y  $5 \times 10^5$  de MCF-7 fueron sembradas en

---

---

placas de 6 pozos para obtener una confluencia del ~60%. Posteriormente, 3  $\mu$ g de plásmido fueron transfectados por 24 horas. El complejo de nanopartículas fue removido y reemplazado por medio nuevo RPMI enriquecido con 10% de suero fetal bovino (SFB). Después de 48 horas post-transfección, las células fueron despegadas y lavados con PBS 1X y el RNA fue extraído usando Trizol, siguiendo las recomendaciones del fabricante y almacenado a -80 °C hasta su procesamiento. La expresión del lncKLHDC7B fue evaluada por PCR cuantitativa en tiempo real como se describió anteriormente. Un 70% de disminución de la expresión del lncKLHDC7B fue obtenido después de 48 horas post-transfección para la línea celular HCC1187, 58% para MCF-7 y un 30% para BT-20. El panorama transcripcional por el efecto del silenciamiento del lncKLHDC7B en HCC1187 fue evaluado mediante el análisis de expresión de mRNA y lncRNAs usando la plataforma de microarreglos de Affymetrix HTA 2.0 (Affymetrix, Santa Clara, USA; datos públicamente disponibles en GEO: GSE114468) como se describió anteriormente. Genes con una tasa de cambio de  $> 2$  y  $< 2$ , p-valor  $< 0.05$  fueron considerados como significativos y seleccionados para el análisis de vías biológicas usando el software Ingenuity Pathway Analysis (IPA) software y GeneTrail2 [100]

### **Ensayos de migración e invasión celular**

Los ensayos de migración e invasión celular fueron realizados usando cámaras Transwell con poros de 8  $\mu$ m de diámetro (Corning, NY, USA). Células transfectadas (48 h post-transfección) fueron despegadas y resuspendidas en RPMI sin SFB a una concentración de  $5 \times 10^4$  células en 100  $\mu$ L y sembradas en el interior del inserto transwell. Posteriormente 600  $\mu$ L de medio RPMI con 10% de SFB fueron colocados en el pozo de la placa. Para el ensayo de invasión,  $1 \times 10^5$  células fueron colocadas en los insertos transwell previamente recubiertos con 1.6 mg de Matrigel (Corning, NY, USA). Las células fueron incubadas para el ensayo de migración (HCC1187 y MCF-7 por 24 horas y BT-20 por 6 horas) e invasión (HCC1187 por 36 horas, BT-20 y MCF-7 por 48 h) a 37 °C con 5% de CO<sub>2</sub>. Cada experimento fue realizado por triplicado. Al final de cada experimento, las células que migraron al reverso de la membrana transwell fueron fijadas con

---

---

paraformaldehído al 3.7%, y teñidas con cristal violeta. Las células fueron cuantificadas usando el software ImageJ [101].

### **Ensayo de apoptosis**

La apoptosis fue inducida con ~50  $\mu$ M de QBS (Quillaja Bark Saponin: 2-Amino-N-quinolin-8-yl-benzenesulfonamide; A3105, SIGMA-ALDRICH) por 24 h antes de finalizar las 48 h post-transfección. Inmediatamente, las células fueron despegadas con tripsina y lavadas con PBS al 1X. Las células fueron re-suspendidas en la solución amortiguadora de unión y teñidas con anexina V y Ioduro de propidio (IP) (FITC Annexin V/ Dead cell apoptosis kit, Invitrogen, Carlsbad, CA, USA) por 15 minutos en oscuridad a temperatura ambiente. Las células teñidas fueron medidas mediante ensayo de citometría. Las células fueron categorizadas como apoptosis temprana (tinción positiva a fosfatidil-serina) y tardía (tinción positiva a fosfatidil-serina + tinción DNA) indicadas en el recuadro rojo del análisis por citometría.

### **Análisis estadísticos**

Las significancias estadísticas fueron analizadas usando los software GraphPad Prism (versión 6, San Diego, USA) y STATA (Versión 12, Texas, USA). Para las pruebas Kruskal-Wallis, ANOVA y Chi-cuadrado fueron realizadas para todas las comparaciones que involucraban categorías. La correlación entre variables fue determinada mediante el coeficiente de correlación de Spearman. Análisis de Kaplan-Meier y supervivencia entre grupos fue realizado usando la prueba de long-rank. Los valores de  $p < 0.05$  fueron considerados como significativos (\*  $<0.05$ , \*\* $<0.01$ , \*\*\* $<0.001$ ).

---

---

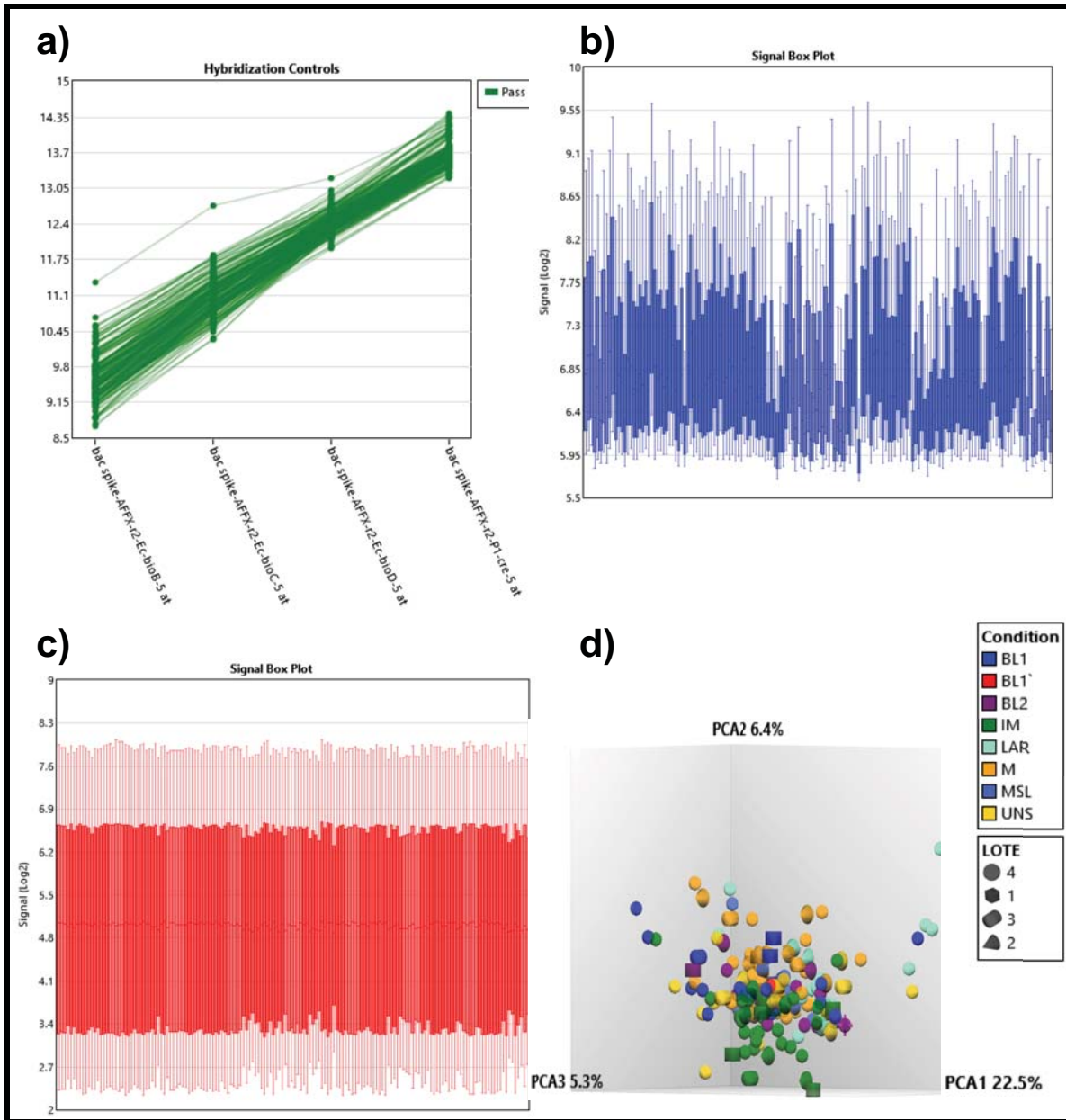
## Resultados

### Análisis de microarreglos

Ciento cincuenta y seis muestras de RNA extraídas de cáncer de mama triple negativo fueron hibridadas en microarreglos “Human Transcriptome Array” versión 2.0 de Affymetrix, los cuales evaluaron la expresión de mRNAs y lncRNAs. En la figura 10a, se muestra el análisis del control de calidad de la hibridación cada uno de los microarreglos analizados. Como se observa, la hibridación de cada muestra presentó una fluorescencia que fue incrementando de manera proporcional a la concentración de los siguientes controles de hibridación como lo especifica el manual de procedimiento: *AFFX-r2-Ec-BioB* > *AFFX-r2-Ec-BioC* > *AFFX-r2-Ec-BioD* > *AFFX-r2-P1-Cre*. Por lo tanto, los resultados derivados de este trabajo serán producto de la variación biológica de las muestras, y no por variaciones técnicas o metodológicas de acuerdo a los resultados de los controles.

Los datos crudos (.CEL) de fluorescencia (figura 10b), fueron normalizados (figura 10c), a través de una corrección de fondo para su posterior análisis de expresión diferencial de mRNAs y lncRNAs entre las diferentes condiciones biológicas (subtipos de cáncer de mama triple negativo).

Posteriormente, se realizó un análisis de componentes principales (PCA, por sus siglas en inglés) para reducir la multidimensionalidad de los datos de microarreglos, pero al mismo tiempo conservando la mayor cantidad de información posible. Mediante una simplificación o transformación de los datos de fluorescencia, el PCA mostró una baja variabilidad de las muestras analizadas entre cada componente evaluado (figura 10d). La figura muestra escasa dispersión entre las muestras analizadas, esta baja variabilidad podría deberse a que todas las muestras pertenecen a cáncer, y tienen poca dispersión. Sin embargo, podemos observar en la figura 10d, una discreta agrupación de algunos subtipos de moleculares, como los inmunomoduladores (verde) y mesenquimales (amarillo).



**Figura 10. Análisis de calidad y normalización de microarreglos.** a) La gráfica muestra la fluorescencia (eje Y) de los controles de hibridación (eje x) para cada microarreglo analizado (n=156). La línea verde representa cada una de las muestras analizadas, los puntos verdes muestran los niveles de fluorescencia de cada control de forma creciente (AFFX-r2-Ec-BioB, AFFX-r2-Ec-BioC, AFFX-r2-Ec-BioD, AFFX-r2-P1-Cre). b) Diagrama de cajas y bigotes. Archivos .CEL, datos no normalizados, el eje Y representa los niveles de fluorescencia por cada microarreglo (eje X) analizado (n=156). c) Diagrama de cajas y bigotes. Intensidades de fluorescencia normalizados por el método robusto de promedio de matriz múltiple (RMA, por sus siglas en inglés) para el análisis de expresión diferencial. Eje Y muestra la intensidad de fluorescencia, eje X representa cada microarreglo analizado. d) Análisis de componentes principales (PCA, por sus siglas en inglés) de las muestras de cáncer de mama triple negativo analizadas. Cada color representa a un subtipo de CMTN y cada forma (círculo, cuadrado, triángulo, cilindro) el número de lote de procesamiento de cada muestra. El PCA1 describe la variabilidad del 22.5%, el PCA2 el 6.4% y el PCA3 el 5.3%. Basal 1, BL-1; Basal 2, BL-2; Inmunomodulador, IM; Mesenquimal, M; Parecidos a troncal mesénquima, MSL; Receptor de andrógenos luminal, LAR; Inestables o indeterminados, UNS.

## Características de la población y distribución de subtipos triple negativos

Un total de 156 muestras derivadas de pacientes con cáncer de mama triple negativo fueron incluidas en este estudio. La media de edad de las participantes fue de 53 años (rango de 26 a 88 años), siendo el carcinoma ductal infiltrante el tumor de mama más frecuente. La información clínico-patológica de las pacientes con tumores triples negativos incluidos en este estudio se muestran en la tabla II. Los subtipos de CMTN fueron identificados usando el algoritmo TNBCtype, el cual asignó a cada una de las muestras en uno de los 6 subtipos de cáncer de mama triple negativo. Analizando nuestra cohorte de CMTN, identificamos que la mayor parte de los tumores se clasificaron como IM (22.4%, n= 35/156) y M (22.4%, n= 35/156), seguido por los BL1 (14.7%, n= 12/156) y subtipo LAR (12.2%, n= 19/156), BL2 con un 7.7% (n= 12/156), el subtipo MSL (7%, n= 11/156) pero un 13.5% (n= 21/156) fueron categorizadas como no clasificables o UNS (figura 11a). La distribución de las frecuencias de los subtipos de CMTN de nuestra cohorte fue similar a las previamente reportadas por otros estudios (figura 11b) [8, 46, 48, 50, 56].

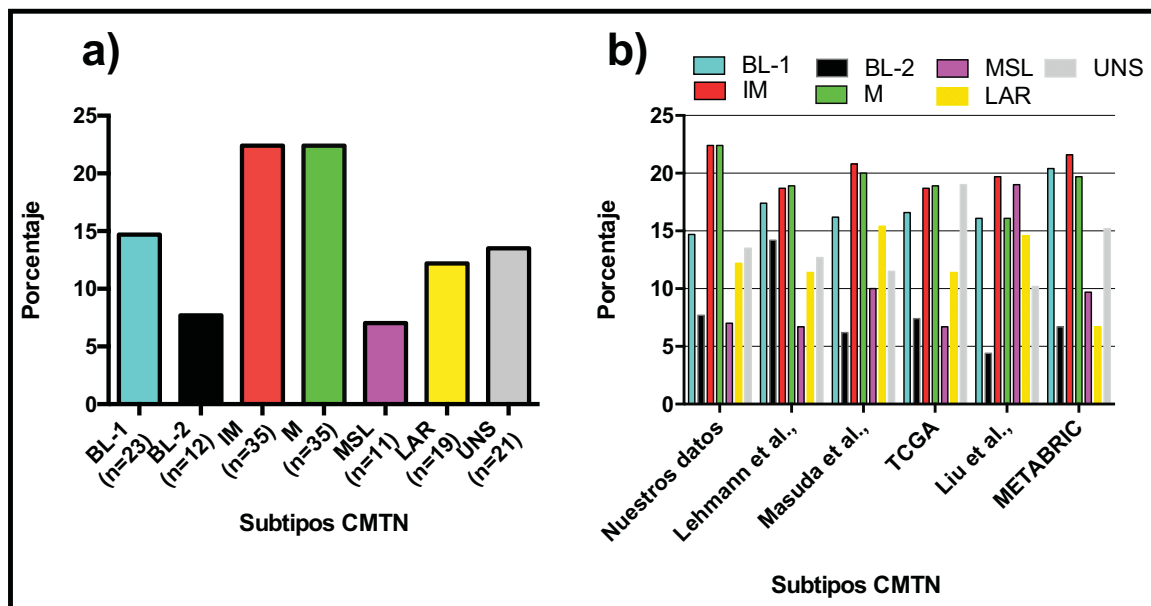


Figura 11. Frecuencia de subtipos de cáncer de mama triple negativo. a) Frecuencia de subtipos de CMTN analizados en este trabajo. b) Comparación de frecuencias de subtipos de CMTN entre diferentes estudios. Basal 1, BL-1; Basal 2, BL-2; Inmunomodulador, IM; Mesenquimal, M; Parecidos a troncal mesenquimal, MSL; Receptor de andrógenos luminal, LAR; Inestables o indeterminados, UNS.

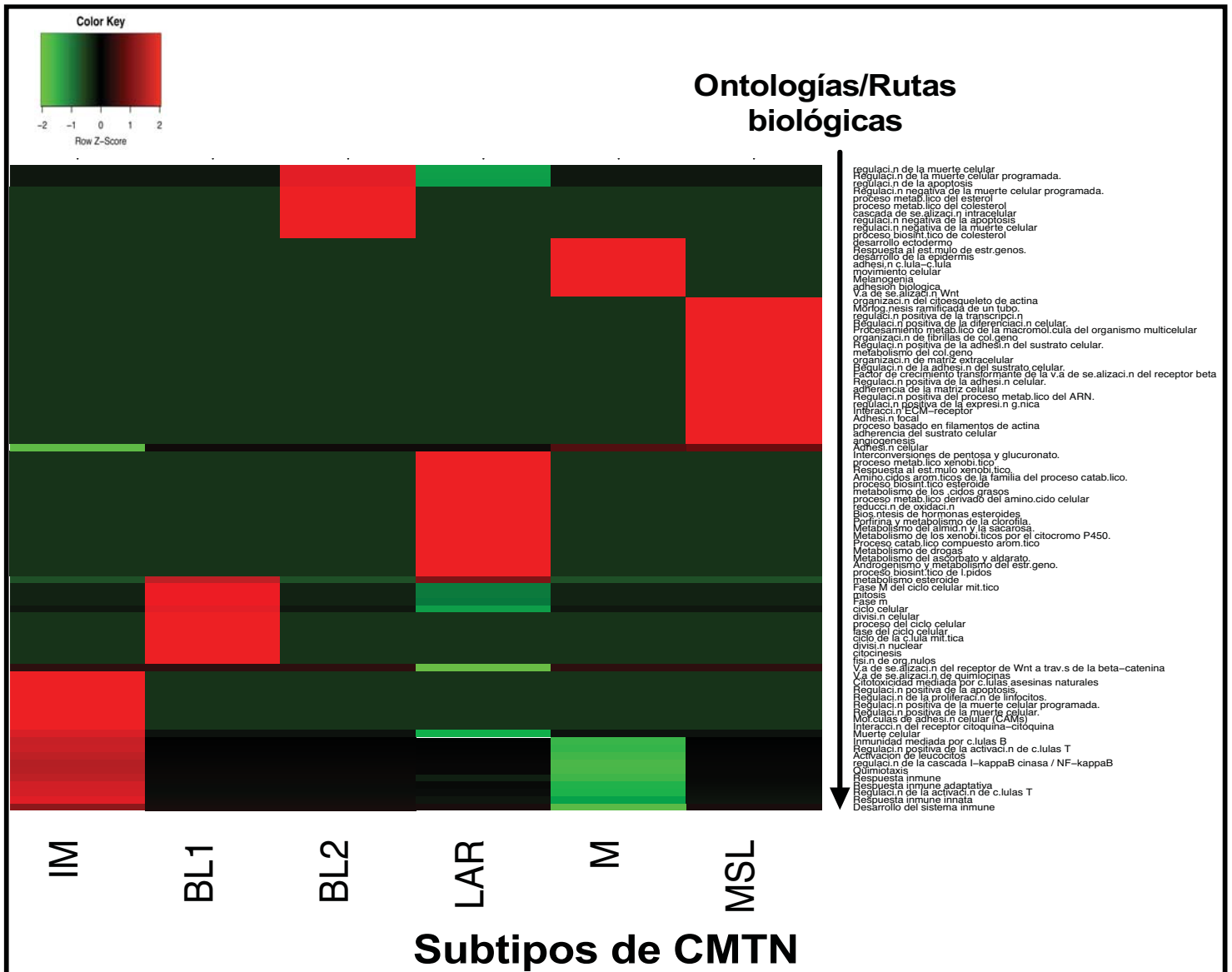
---

---

### **Análisis de ontología génica y enriquecimiento de rutas biológicas**

Un análisis de ontologías y rutas biológicas fue realizado mediante el análisis de los genes diferencialmente expresados de cada uno de los subtipos CMTN. Como se muestra en la figura 12, el subtipo IM mostró un enriquecimiento en los procesos inmunológicos como respuesta inmune, activación de leucocitos, desarrollo del sistema inmunitario, regulación de células T, respuesta inmune innata, quimiotaxis, regulación de la cascada I-kappaB quinasa / NF-kappaB. Sin embargo, también se observaron procesos como la regulación de la muerte celular programada, la muerte celular y la apoptosis. El subtipo BL1 mostró un enriquecimiento en los procesos celulares como el ciclo celular, la división celular y ciclo mitótico, pero el BL2 presentó un mayor enriquecimiento en la regulación de la apoptosis y la muerte celular programada. De manera interesante, las rutas del metabolismo del colesterol y el proceso de biosíntesis fueron identificados en este subtipo. Los subtipos M y MSL mostraron un enriquecimiento en procesos descritos anteriormente, como la adhesión célula-célula, el movimiento celular, la vía de señalización Wnt, y adhesión de la matriz celular, adhesión focal, angiogénesis, interacción ECM-receptor, organización del citoesqueleto de actina, vía de señalización del receptor TGFB, respectivamente. Además, diferentes procesos como el metabolismo de los ácidos grasos, biosíntesis y metabolismo de esteroides, así como el metabolismo de andrógenos y estrógenos, biosíntesis de lípidos e interconversiones pentosa-glucuronato fueron los procesos con mayor relevancia en el subtipo LAR.





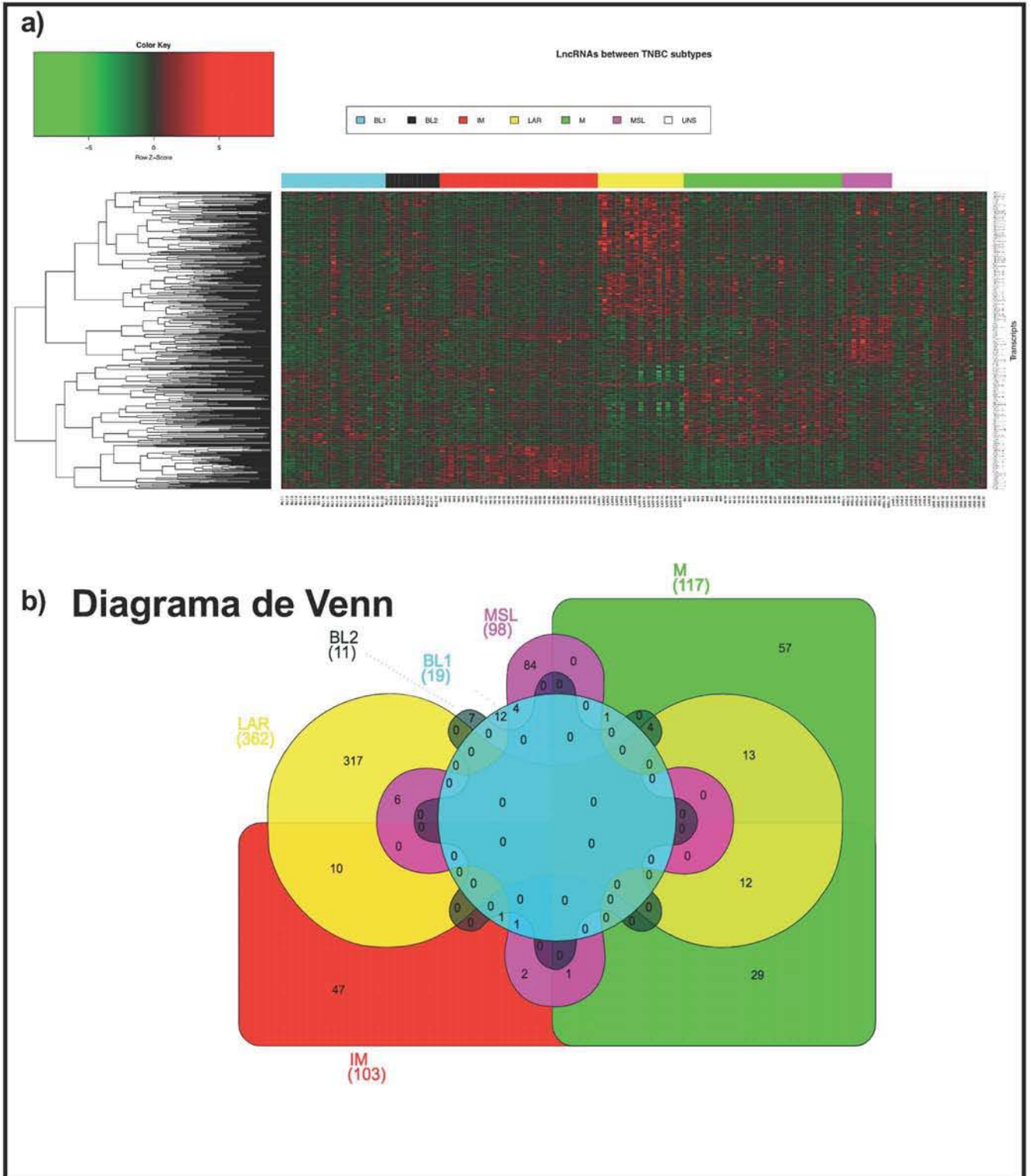
**Figura 12. Análisis ontológico y rutas biológicas en CMTN.** El mapa de calor representa los procesos celulares enriquecidos (derecha) de manera positiva (en rojo) y negativa (en verde) derivado de los genes diferencialmente expresados entre los subtipos de cáncer de mama triple negativo. Valores de z-score de cada proceso biológico enriquecido de manera significativa ( $p < 0.05$ ) fue graficado para cada uno de los subtipos de cáncer de mama triple negativo. <https://febs.onlinelibrary.wiley.com/doi/10.1002/1878-0261.12446>. Basal 1, BL-1; Basal 2, BL-2; Inmunomodulador, IM; Mesenquimal, M; Parecidos a troncal mesénquima, MSL; Receptor de andrógenos luminal, LAR.

---

---

## Perfiles de expresión de RNA largos no codificantes en cáncer de mama triple negativo

Una vez clasificadas las muestras entre los subtipos intrínsecos de CMTN, analizamos los perfiles de expresión de RNA largos no codificantes. Identificamos un total de 710 lncRNAs diferencialmente expresados entre los subtipos de cáncer de mama (figura 13a, material suplementario, <https://febs.onlinelibrary.wiley.com/doi/10.1002/1878-0261.12446>), 84 de los mismos se expresan en al menos dos subtipos de CMTN, mientras que 524 se distribuyeron de manera particular entre los subtipos de cáncer de mama triple negativo, como se muestra en la figura 13b. De manera interesante, el subtipo LAR presentó un mayor número de lncRNAs expresados (50.9%, 362/710) entre los cuales encontramos a RP11-206M11.7, LINC00993 y PCAT18, por mencionar algunos. El subtipo mesenquimal presentó 117 lncRNAs (16.5%) alterados, por ejemplo RP11-84E24.2, LINC01152, H19, ANO1-AS1, CASC15, CTA-384D8.31, ANKRD44-IT1, PSMB8-AS1, RP11-142M10.2, ANKRD44-IT1, ELMO1-AS1, IFNG-AS1, CFLAR-AS1 fueron los lncRNAs mayormente sobre expresados en el subtipo inmunomodulador. Por otra parte, el subtipo MSL mostró noventa y ocho lncRNA alterados como DNM3OS, ADAMTSL4-AS1, LATS2-AS1, LINC00312 and MEG8. Los subtipos BL1 y BL2 fueron los subtipos que presentaron 19 y 11 lncRNAs expresados, respectivamente (<https://febs.onlinelibrary.wiley.com/doi/10.1002/1878-0261.12446>).



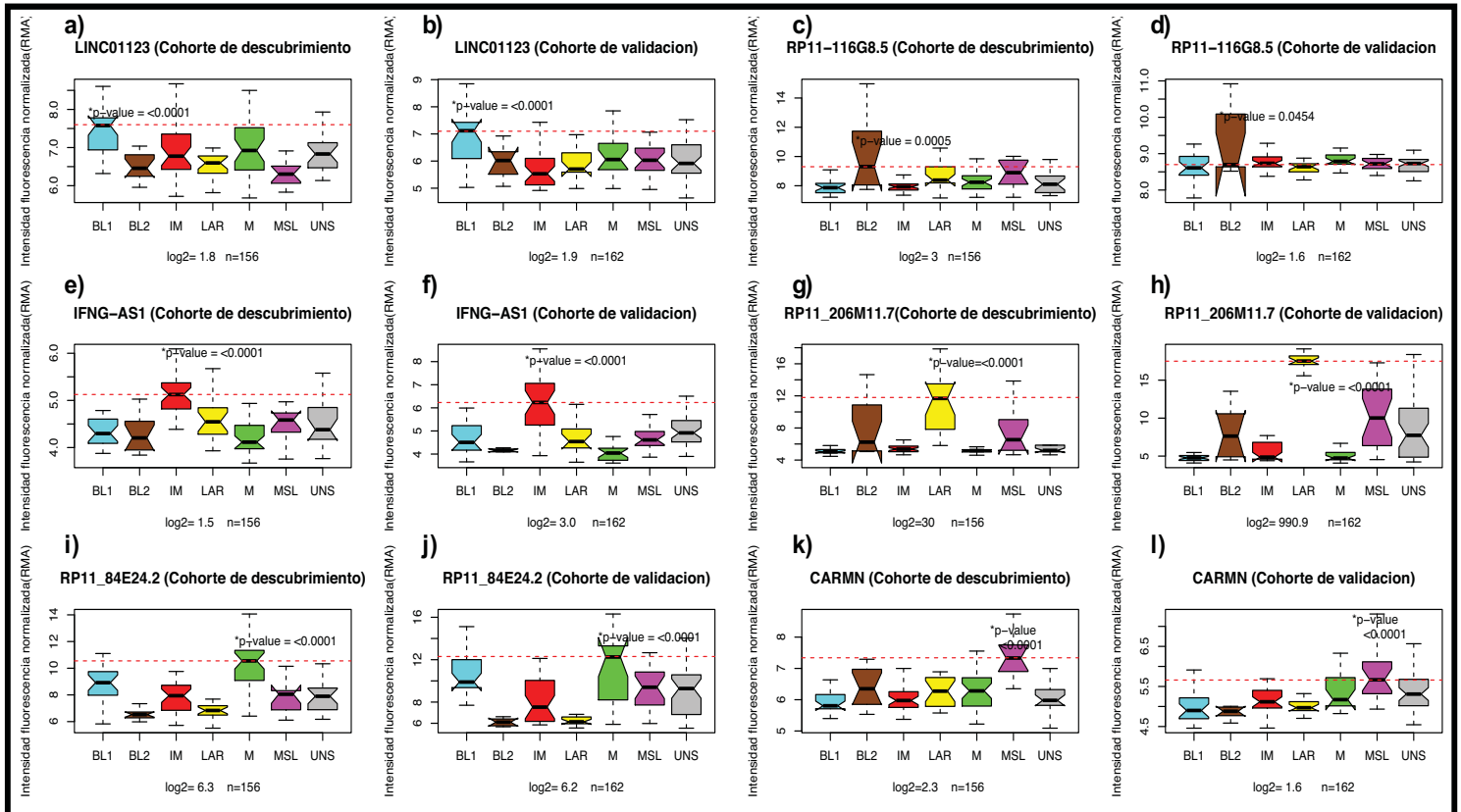
**Figura 13. Perfiles de lncRNAs expresados diferencialmente entre subtipos CMTN. a)** El mapa de calor representa en rojo a los 710 lncRNAs sobreexpresados y/o en verde subexpresados. En la parte superior del mapa se representan cada uno de los subtipos de cáncer de mama triple negativo y su código de color correspondiente. A nivel de columna se representa el perfil de expresión por muestra (n=156). **b)** Diagrama de Venn muestra el número de lncRNAs compartidos entre los subtipos CMTN y aquellos exclusivos para cada subtipo.

---

---

## Validación de la expresión *in silico* de lncRNAs en muestras independientes de cáncer de mama triple negativos

Para corroborar la expresión de lncRNAs identificados en nuestra cohorte de estudio, validamos la expresión de lncRNAs en una cohorte independiente de tumores de mama triple negativos, obtenida de la base de datos públicos GEO (figura 14). Observamos que ~34% de lncRNAs (resultados no mostrados) identificados en nuestra cohorte fueron validados en la cohorte independiente, por ejemplo RP11-206M11.7 en el subtipo LAR (tasa de cambio de 30 veces más que otro subtipo ( $p < 0.0001$ ) en nuestra cohorte, 991 veces más expresado ( $p < 0.0001$ ) en el subtipo LAR en la corte de validación). INFG-AS1 es un lncRNA relacionado como un regulador de la respuesta inflamatoria mediada por interferón gama (IFNG) el cual fue encontrado en el subtipo inmunomodulador con una tasa de cambio de 3 veces más en la cohorte de validación ( $p < 0.0001$ ) mientras que en este trabajo se detectó 1.5 veces más expresado ( $p < 0.0001$ ). Por otra parte, en los subtipos M y MSL el lncRNA RP11-84E24.2 y CARMN fueron validados y corroborado su incremento de expresión en cada subtipo, respectivamente (figura 14). En el subtipo BL1 fue encontrado el enriquecimiento del LINC01123 con 1.8 veces más expresado en este estudio, mientras que en la cohorte de validación fue encontrado incrementado 1.9 veces más de manera significativa ( $p < 0.0001$ ). Sin embargo, el lncRNA RP11-116G8.5 expresado en BL2 mostró una tasa de expresión de 3 veces más en nuestro estudio con una significancia marginal ( $p = 0.0265$ ), mientras que en la cohorte de validación fue de 1.6 veces más su cambio ( $p < 0.0001$ ).

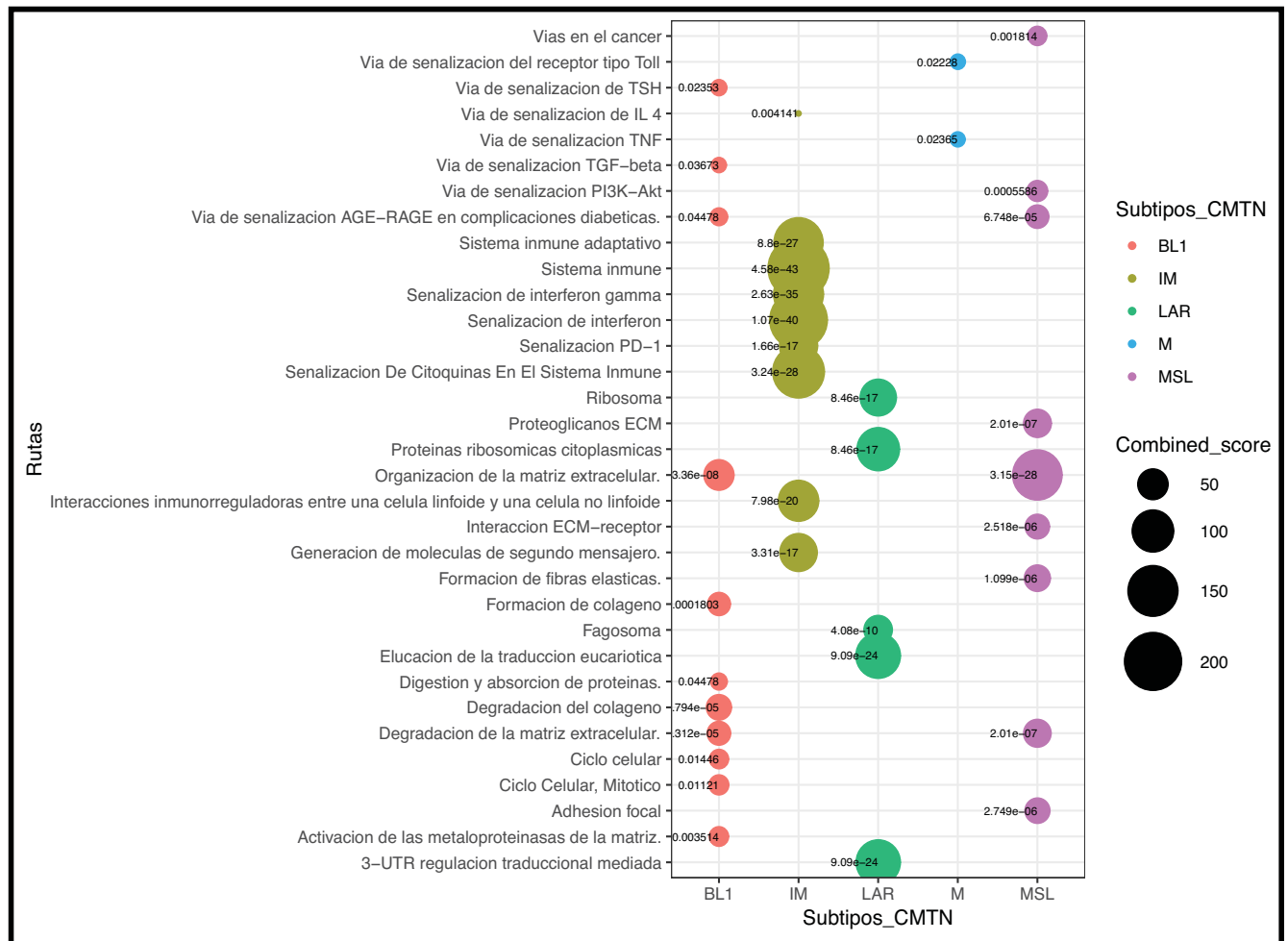


**Figura 14 . Validación *in silico* de la expresión de lncRNAs en muestras independientes.** En la figura se muestra la medias de fluorescencia normalizadas para RMA de algunos lncRNA seleccionados para cada uno de los subtipos de CMTN. Los incisos **a), c), e), g), i) y k)** representan los datos de nuestra cohorte de descubrimiento (n=156) y **b), d), f), h), j) y l)** representan la expresión validada en muestras independientes disponible públicamente en la base de datos GEO (n=156).

## **Análisis de enriquecimiento de rutas biológicas basado en la expresión de lncRNAs en mama triple negativos**

Para comprender mejor la importancia biológica de los lncRNAs alterados, realizamos un análisis de asociación por culpabilidad para investigar su relación con los mRNA desregulados y su posible impacto en diferentes rutas biológicas. Este análisis reveló una asociación significativa entre algunos de los lncRNAs y mRNA expresados diferencialmente, así como el enriquecimiento en vías clave relacionadas con el cáncer de mama de genes codificantes asociados para cada subtipo de TNBC (figura 15). Los lncRNAs alterados en BL1 se correlacionaron más a procesos de organización de la matriz extracelular, al ciclo celular y a la señalización del factor de crecimiento transformante beta (TGF- $\beta$ ). Los lncRNA

derivados del subtipo IM están relacionados con procesos como el procesamiento de antígenos, señalización de interferón y al sistema inmunológico. Por otro lado, los lncRNAs del subtipo LAR están potencialmente involucrados en la regulación de la traducción. Los lncRNAs alterados en los tumores M se correlacionaron con el receptor tipo toll y la señalización del TNF. Finalmente, la red de correlación de coexpresión de lncRNAs-mRNAs alterados en el subtipo MSL se enriqueció en procesos de adhesión focal, interacción con componentes de matriz extracelular y la vía de señalización PI3K-Akt (figura 15).



**Figura 15. Panorama general de las vías biológicas mediante el análisis de culpabilidad por asociación a través de los subtipos de TNBC mediante la coexpresión de lncRNAs-mRNAs.** Análisis de asociación por culpabilidad muestra las vías enriquecidas significativas de la co-expresión resultante de los mRNA-lncRNAs. El eje Y representa las vías y el eje X representa los subtipos de CMTN. El gráfico de burbujas muestra el enriquecimiento de las rutas biológicas mediante los datos de expresión diferencial de lncRNAs en cada subtipo de CMTN. El tamaño y el color de la burbuja representan la puntuación de cada vía y los subtipos de CMTN, respectivamente.

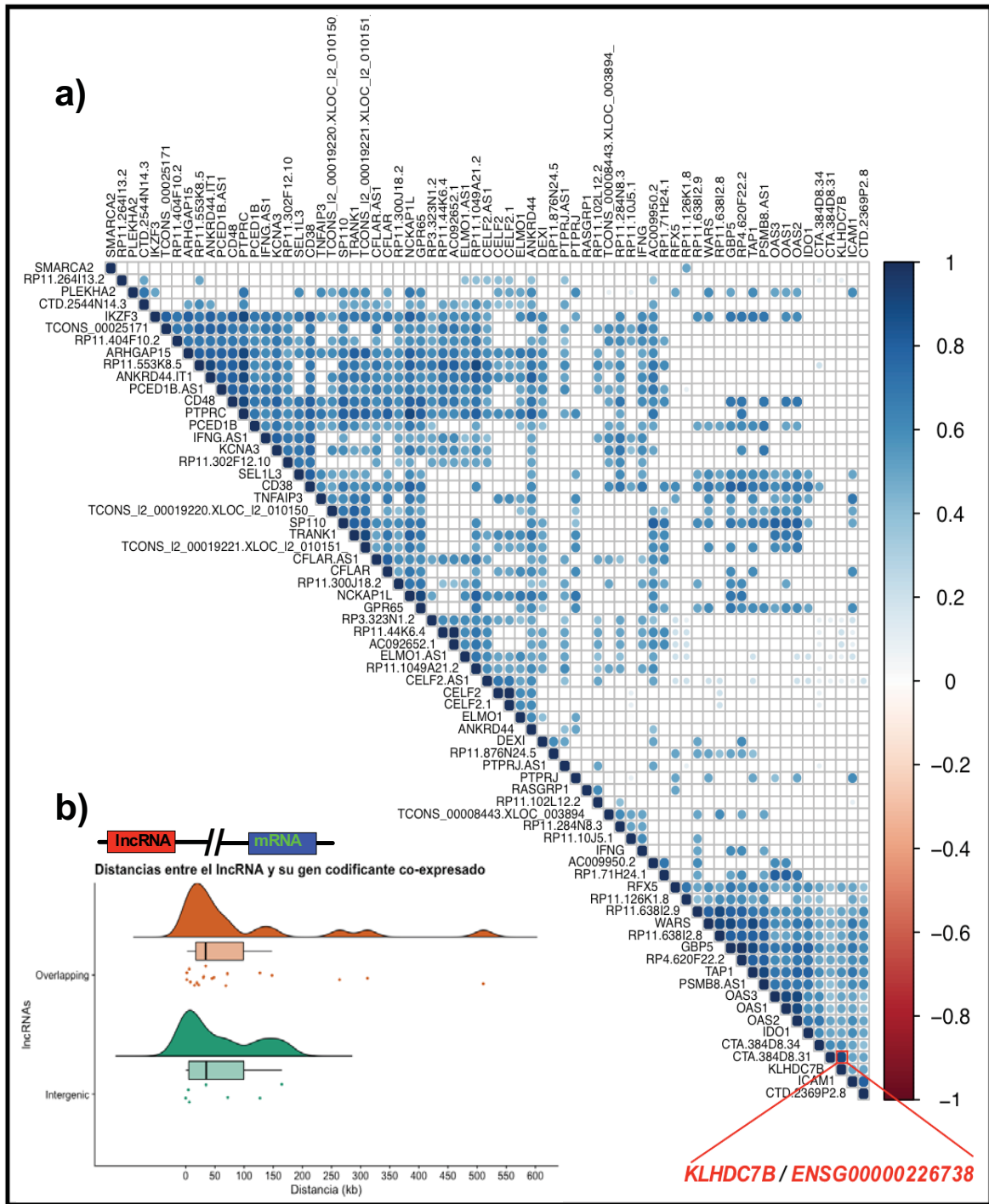
---

---

## **Análisis de correlación de lncRNAs-mRNA en el subtipo inmunomodulador de cáncer de mama triple negativo**

Se ha descrito que entre las diversas funciones que poseen los lncRNAs, se encuentra la capacidad de regular la expresión génica. Apoyados en estas evidencias, se realizó un análisis basado en correlaciones de la expresión positiva de lncRNAs-mRNAs para priorizar la selección de probables lncRNAs candidatos con la capacidad de regular a genes codificantes (intergénicos o intragénicos) en uno de los subtipos más frecuentemente detectados en este estudio, el inmunomodulador. Identificamos 37 pares de transcritos de lncRNA-mRNA co-expresados positivamente (figura 16). Se identificó que los transcritos con mayor correlación fueron TRANK1 y TCONS\_I2\_00019221 ( $\rho=0.9$ ), mientras que los menos correlacionados fueron RP11-44K6.4 e IDO1 ( $\rho=0.4$ ) como se muestra en la tabla III.





**Figura 16. Análisis de correlación entre lncRNA-mRNA co-expresados positivamente del subtipo inmmomodulador. a)** El mapa de calor representa en círculos el grado de correlación computado entre todos lncRNAs y mRNAs expresados positivamente. Las correlaciones con significativas  $<0.0001$  son presentadas en esta figura. La intensidad del color se basa en el grado de correlación entre los transcritos ( $n=37$  pares). **b)** Gráfica representativa de las distancias genómicas entre el lncRNA y su gen codificante (mRNA) co-expresado positivamente. El eje x se representan las distancias en kilobase (kb), el eje y representa los tipos de lncRNAs identificados. Los puntos representan la distancia del par de los transcritos de lncRNA-mRNA, observando una media de 58 kb para los lncRNAs intergénicos, pero una media de 93 kb de distancia entre lncRNA y el gen codificante en los lncRNAs sobrelapados.

---

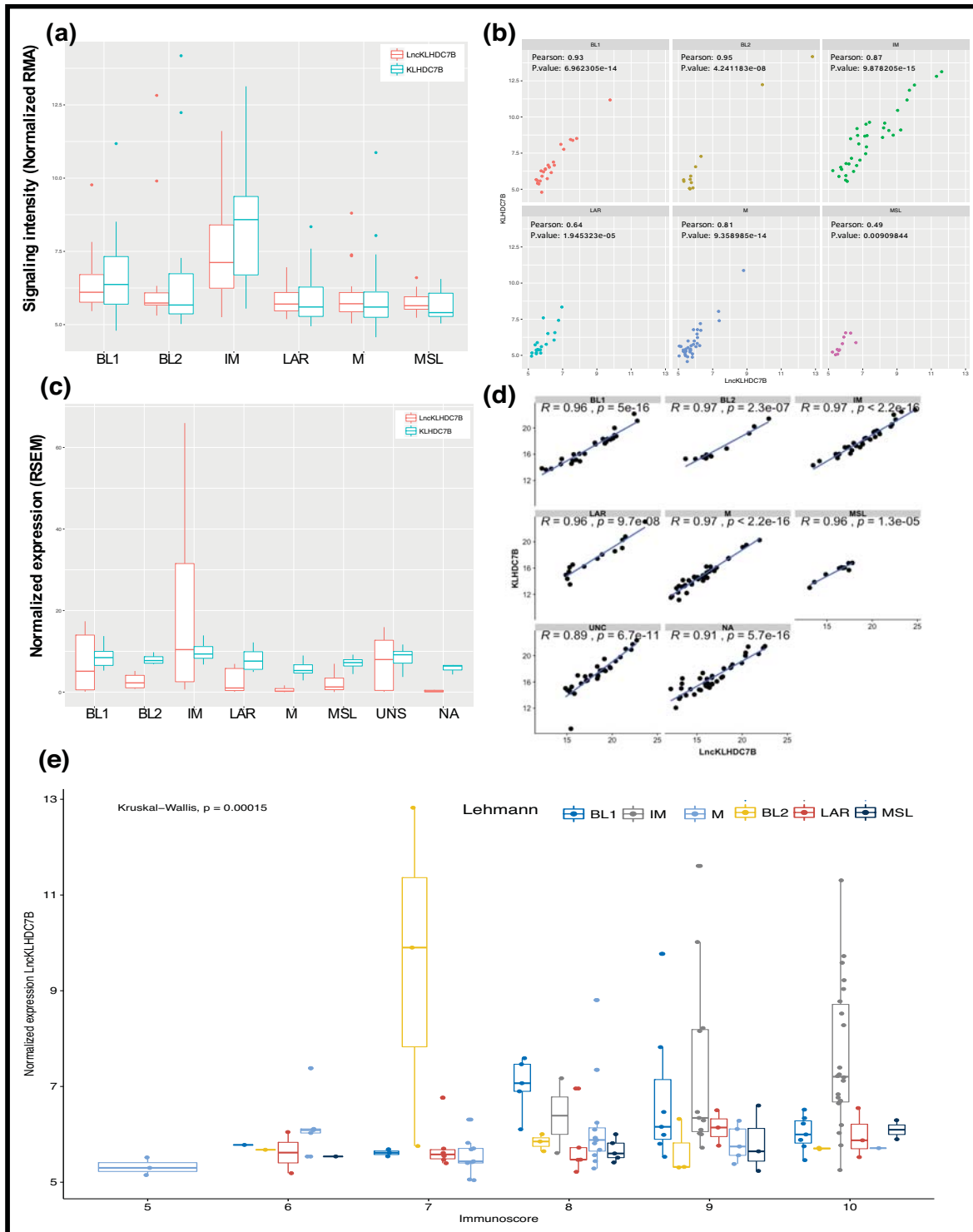
---

**El RNA largo no codificante *ENSG00000226738 (LncKLHDC7B)* y el gen codificante *KLHDC7B* muestran un patrón de expresión correlacionado al fenotipo tumoral inmunomodulador**

El lncRNA *ENSG00000226738* (que a partir de aquí llamaremos *LncKLHDC7B*) y el gen codificante *KLHDC7B* (miembro de la familia de las proteínas con dominios Kelch) fueron uno de los pares de transcritos lncRNA-mRNA mayormente correlacionados con una  $r^2= 0.9$  y  $p < 2.2e^{-16}$  en el fenotipo IM (tabla III). La sobreexpresión del par *LncKLHDC7B/KLHDC7B* así como su correlación en el fenotipo IM fue demostrada en nuestra cohorte de estudio (figura 17a, b) y fue validada en muestras independientes de bases de datos como el TCGA (figura 17c,d). Adicionalmente, evaluamos la expresión de ambos transcritos en muestras de tejido normal y tumoral de bases de datos (TCGA, figura 18a) y en un panel de líneas celulares de cáncer de mama, observando una expresión incrementada en los tumores, así como en varias líneas celulares triple negativo en contraste con la línea celular no tumorigénica MCF10A. Entre las líneas celulares, lo encontramos expresado en HCC1187, la cual presenta un fenotipo inmunomodulador (figura 18b). Para validar la relación del *LncKLHDC7B* con un inmunofenotipo, calculamos el inmuno-puntaje o inmunofenoscore (análisis de inmunogenecidad) basado en la expresión del lncRNA en nuestras muestras analizadas. Nuestros resultados mostraron un patrón de expresión del lncRNA mayormente relacionadas con un alto inmunofenoscore (8 y 9) sobre todo en los tumores con fenotipo inmunomodulador (figura 17e).

Con base en todos estos resultados, seleccionamos a la línea celular HCC1187 para investigar el papel funcional del *LncKLHDC7B* en el fenotipo inmunomodulador de cáncer de mama. Seleccionamos el par *LncKLHDC7B/KLHDC7B* por las siguientes razones: i) el *LncKLHDC7B* podría ser un marcador tumoral para el subtipo IM, sin embargo su papel funcional no ha sido descrito hasta ahora; ii) el *LncKLHDC7B* presenta una co-expresión y asociación significativa con el gen codificante *KLHDC7B*, lo cual podría indicar una posible actividad regulatoria; iii) nuestro análisis *in silico*, determinó los posibles procesos

biológicos en los que podría estar involucrado el lncRNA como presentación de antígeno, señalización de interferón y adhesión celular (figura 19).



**Figura 17. Evaluación de la expresión de *LncKLHDC7B* y *KLHDC7B* en los subtipos CMTN. a)** Niveles de expresión de *LncKLHDC7B* y *KLHDC7B* en nuestra cohorte y **c)** validación en base de datos de TCGA. **b)** y **d)** Correlación entre el nivel de expresión de *LncKLHDC7B* y *KLHDC7B* a través de los subtipos de CMTN en nuestros datos y del TCGA, respectivamente. **e)** Relación del Immunoscore con la expresión de

---

---

*LncKLHDC7B* en nuestra cohorte. La prueba de Kruskal-Wallis se realizó para determinar la significancia estadística. La etiqueta del eje Y representa la expresión normalizada de *LncKLHDC7B* y la etiqueta del eje X representa el inmunoscore.

### **El silenciamiento del *LncKLHDC7B* lleva a la subexpresión del gen codificante *KLHDC7B* e impacta en la expresión de múltiples genes relacionados con cáncer**

Para evaluar el papel del *LncKLHDC7B* en la línea celular HCC1187 con fenotipo inmunomodulador, realizamos el silenciamiento del lncRNA hasta un 75% de eficiencia mediante sistemas de shRNAs, observando una subexpresión del gen codificante *KLHDC7B* hasta un 50% (figura 18c). Sin embargo, un análisis transcriptómico completo después del silenciamiento de *LncKLHDC7B*, identificó un impacto en la expresión de 1265 transcritos de RNA (figura 17d), incluyendo varios transcritos miembros de la familia de las proteínas con dominios Kelch. Estos datos sugieren que el *LncKLHDC7B* podría regular la expresión de *KLHDC7B* mediante un evento en *cis*. Sin embargo, la desregulación de genes distantes a *LncKLHDC7B*, podría sugerir que el lncRNA pudiera tener una función adicional en *trans* o el silenciamiento del lncRNA indirectamente afecta la expresión de otros genes que no necesariamente implicaría una regulación por efecto *trans*.

Los lncRNAs muestran una diversidad de funciones a través de múltiples mecanismos, entre los cuales se encuentran las interacciones con moléculas como proteínas, DNA y RNA. Siendo las más representativas las interacciones con moléculas de RNA [102]. Debido a que los sitios de unión de las proteínas estabilizadoras del mRNA (RBP, por sus siglas en inglés) pueden ser compartidos con los lncRNAs, la estabilidad o desestabilidad del mRNA puede ser regulada por los RNAs largos no codificantes que posean esta característica[103].

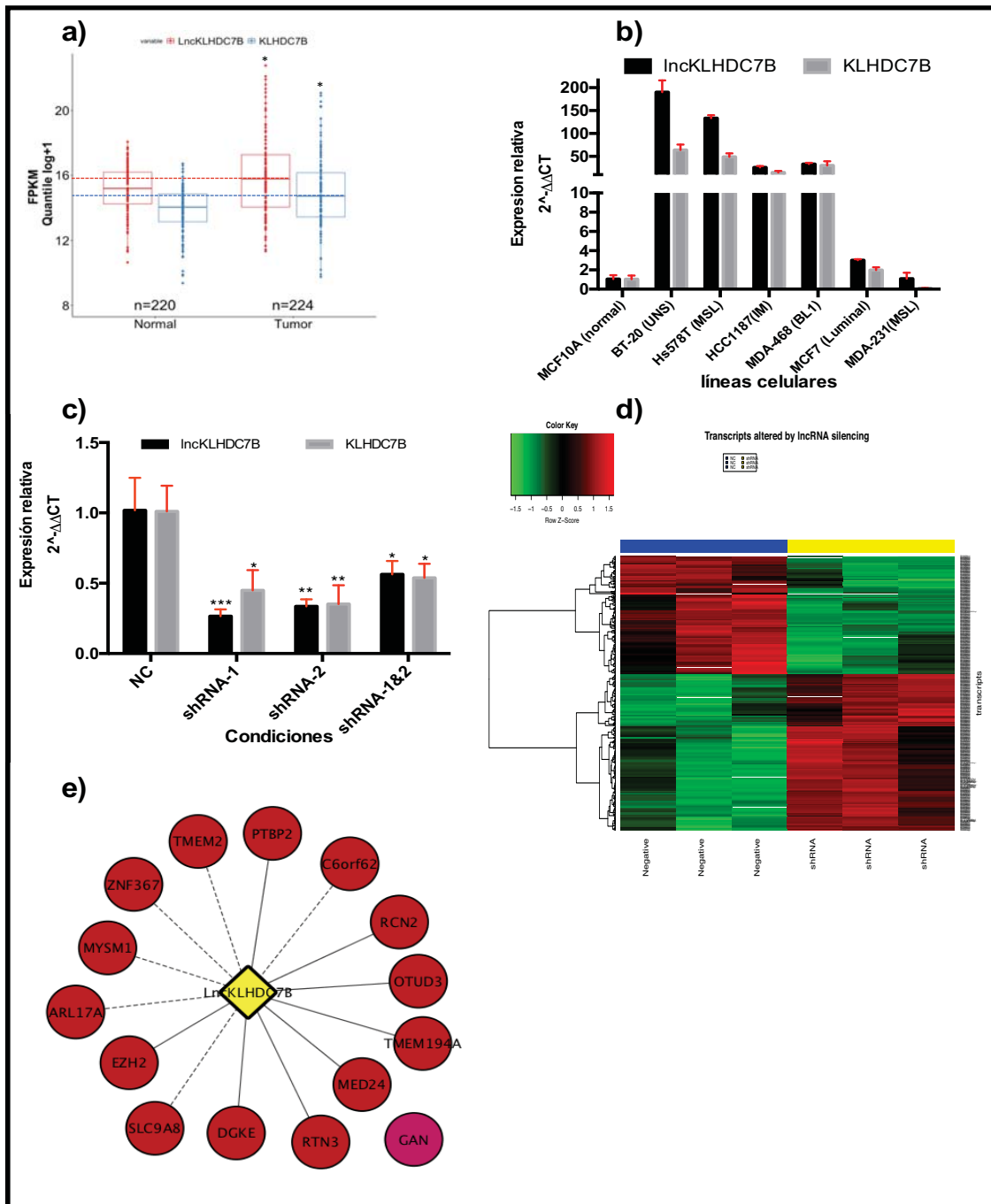
Por lo tanto, una primera aproximación usada para explorar la función del lncRNA y su relación con la alteración en la expresión de algunos transcritos, se aplicó una predicción de interacción mRNA-lncRNA mediante un método basado en parámetros termodinámicos usando la herramienta LncTar [104]. Al analizar los genes subexpresados por el silenciamiento del *LncKLHDC7B*, se observó la

---

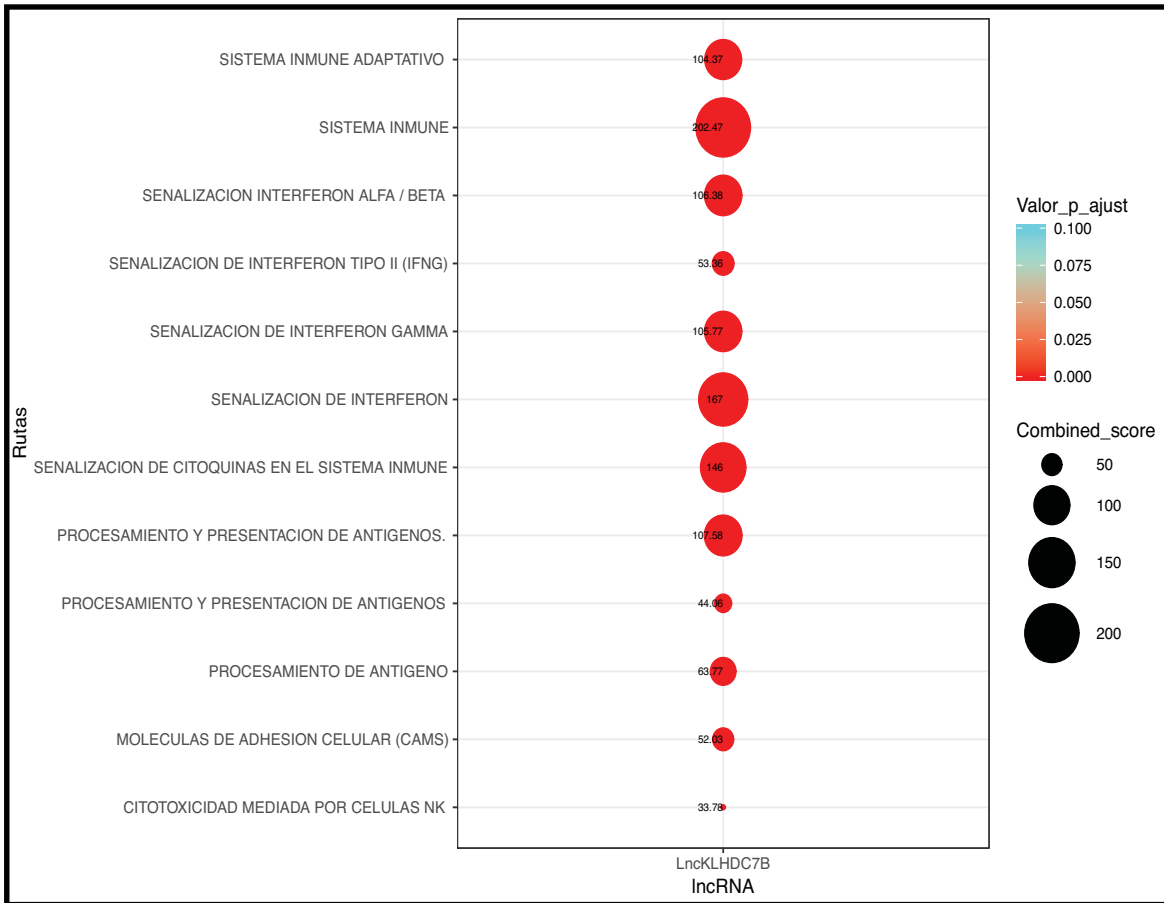
---

interacción predicha para algunos mRNAs, indicando una probabilidad de interacción con mRNA y *LncKLHDC7B* (figura 18e). Entre las interacciones obtenidas se encontró al mRNA de EZH2, un gen que forma parte del complejo de silenciamiento de genes Polycomb [105]. Por lo tanto, la alteración observada en la expresión de varios genes podría ser una consecuencia indirecta por el silenciamiento de *LncKLHDC7B*, y la falta de estabilidad del mRNA de EZH2.

Este enfoque proporcionó información para identificar las interacciones relevantes de lncRNA-mRNA para predecir probables dianas del mRNA, y determinar la posible función biológica del lncRNA. Además, estudios relacionados a estructuras secundarias o motivos estructurales por parte del lncRNA podrían ayudar a dilucidar la funcionalidad del lncRNA basado en la formación de estructuras que reflejen los cambios dinámicos, organizacionales y funcionales [106]. Sin embargo, estudios adicionales tanto computacionales y experimentales son necesarios para validar la hipótesis como regulador post-transcripcional entre el lncRNA y sus moléculas blanco predichas.



**Figura 18. Expresión de *lncKLHDC7B* y *KLHDC7B* en tejido normal y tumoral, modelos in vitro y el impacto de su silenciamiento en la línea HCC1187.** a) Expresión de *lncKLHDC7B* y *KLHDC7B* en tejidos adyacentes normales y tejidos tumorales a partir de datos de TCGA. b) Validación de la expresión del *lncKLHDC7B* y *KLHDC7B* en líneas celulares de cáncer de mama. c) Niveles de expresión del *lncKLHDC7B* en la línea celular HCC1187 (fenotipo inmunomodulador) de cáncer de mama después de silenciar por el sistema de shRNA. La expresión del gen codificación *KLHDC7B* se ve afectada por el silenciamiento *lncKLHDC7B*. Figura b) y c) muestran la media  $\pm$  desviación estándar de al menos tres experimentos independientes. La prueba t de Student se realizó para determinar la significancia, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  de NC vs shRNAs. d) El mapa de calor representa los transcritos alterados por *lncKLHDC7B* silenciado. Los microarreglos se realizaron por triplicado para la condición NC (barra azul) y shRNA-1 (barra amarilla), el rojo representa una expresión elevada y en verde menos expresado. e) Predicción de las interacciones entre los blancos de lncRNA-mRNA subexpresados por el silenciamiento del *lncKLHDC7B*. Las líneas continuas y punteadas una interacción media y alta, respectivamente.



**Figura 19. Análisis de asociación por culpabilidad.** El gráfico de burbujas muestra el enriquecimiento de las rutas biológicas correlacionadas por la expresión de *lncKLHDC7B*. El eje Y representa las vías. El tamaño y el color de la burbuja representan la puntuación de cada vía y significancia, respectivamente.

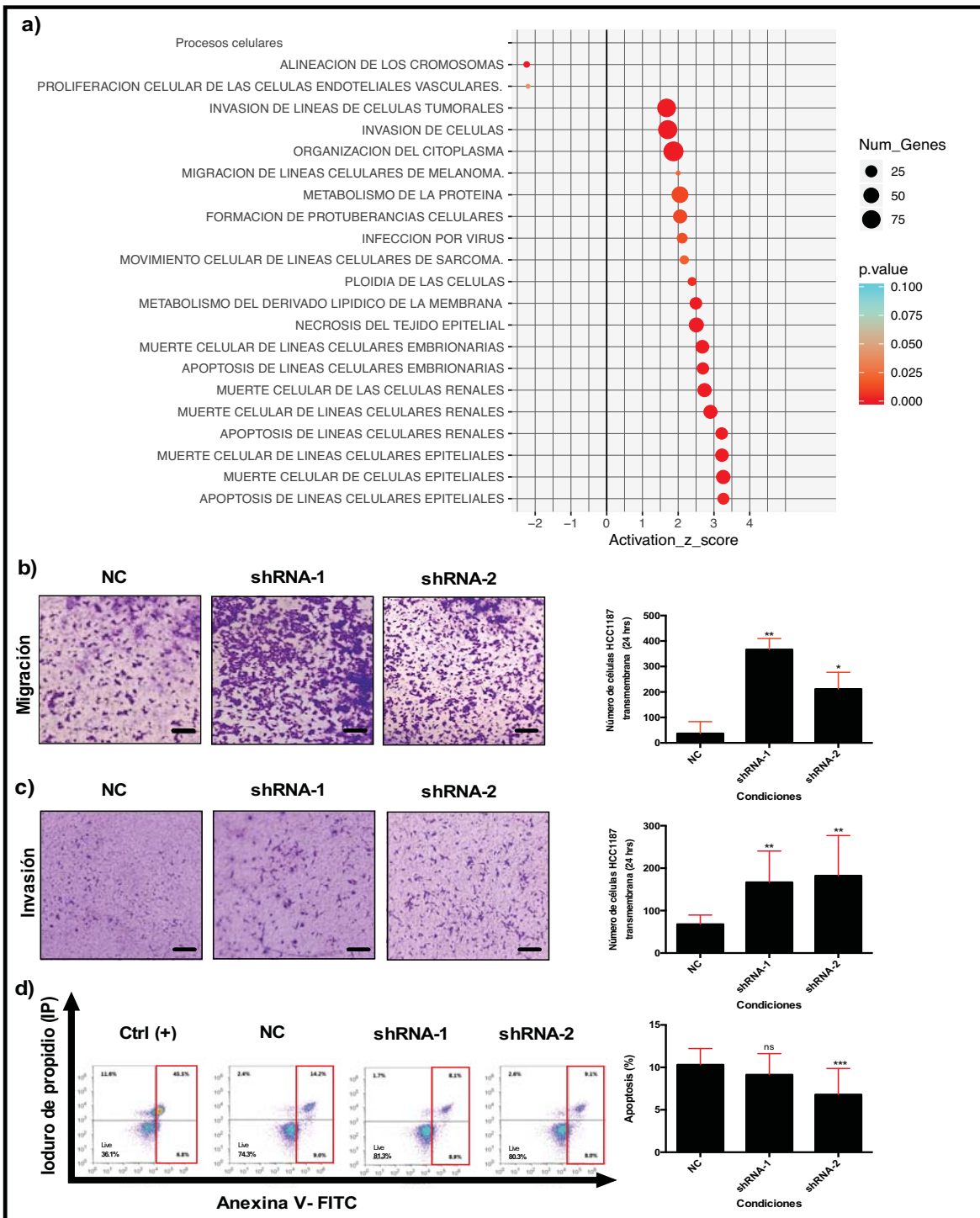


---

---

### ***LncKLHDC7B* modula la expresión de genes asociados a características relevantes del cáncer**

Para evaluar el impacto global en genes por el silenciamiento *LncKLHDC7B* en la línea celular con fenotipo inmunomodulador HCC1187, realizamos un análisis de enriquecimiento de rutas biológicas utilizando los genes alterados, genes codificantes de proteínas significativamente sobreexpresados y subexpresados. De acuerdo a los resultados, procesos como migración, invasión y apoptosis celular fueron representativos por el silenciamiento del *LncKLHDC7B* (figura 20a). Sin embargo, se identificaron otros procesos como metabolismo de proteínas e infecciones virales. Por una parte, el enriquecimiento del proceso de infección viral se debió a la presencia de algunos genes involucrados en la replicación o ciclo vital del virus. Por ejemplo, genes como *AP2A1* y *ARRB1* se encontraron alterados en este trabajo, pero previamente han sido repostados como genes clave en la endocitosis de enterovirus [107]. Adicionalmente, la baja expresión de *KLHDC7B* en conjunto con otros genes, está relacionada con la baja replicación del virus de hepatitis C [108]. Por la motivo, la identificación de rutas biológicas diferentes al cáncer, pueden deberse a la presencia de algunos genes que han sido reportados no necesariamente en los procesos de carcinogénesis.



**Figura 20. Enriquecimiento y análisis funcional por el silenciamiento de *IncKLHDC7B* en la línea celular HCC1187. a)** Enriquecimiento de procesos biológicos por la alteración transcripcional del *IncKLHDC7B* en la línea celular HCC1187. El tamaño del círculo representa el número de genes involucrados en el proceso biológico, el color indica la significación estadística. **b)** Ensayo de migración e **c)** invasión celular al silenciar al lncRNA en HCC1187. Imágenes representativas son mostradas a la izquierda y sus cuantificaciones a la derecha. Barra de escala = 50µm. **d)** Análisis de apoptosis por citometría de flujo (apoptosis temprana y tardía) en HCC1187 transfectadas con shRNA NC, shRNA-1 y shRNA-2. Los resultados son mostrados en medias  $\pm$  DE de al menos tres experimentos independientes. Prueba T de student fue realizada para determinar la significancia estadística \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , NC vs shRNAs.

---

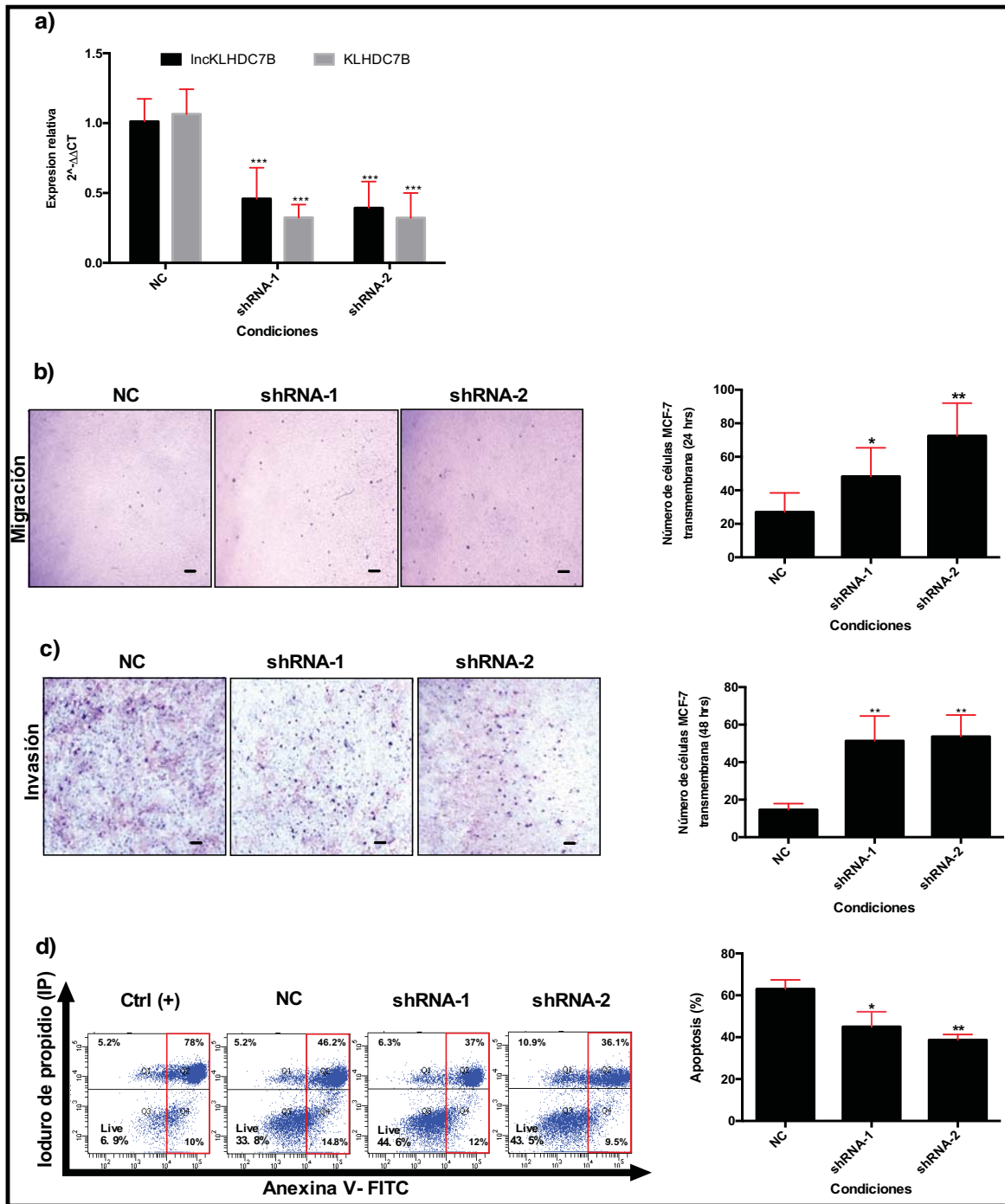
---

## **El silenciamiento del *LncKLHDC7B* incrementa los procesos de migración e invasión celular, pero reduce la muerte celular**

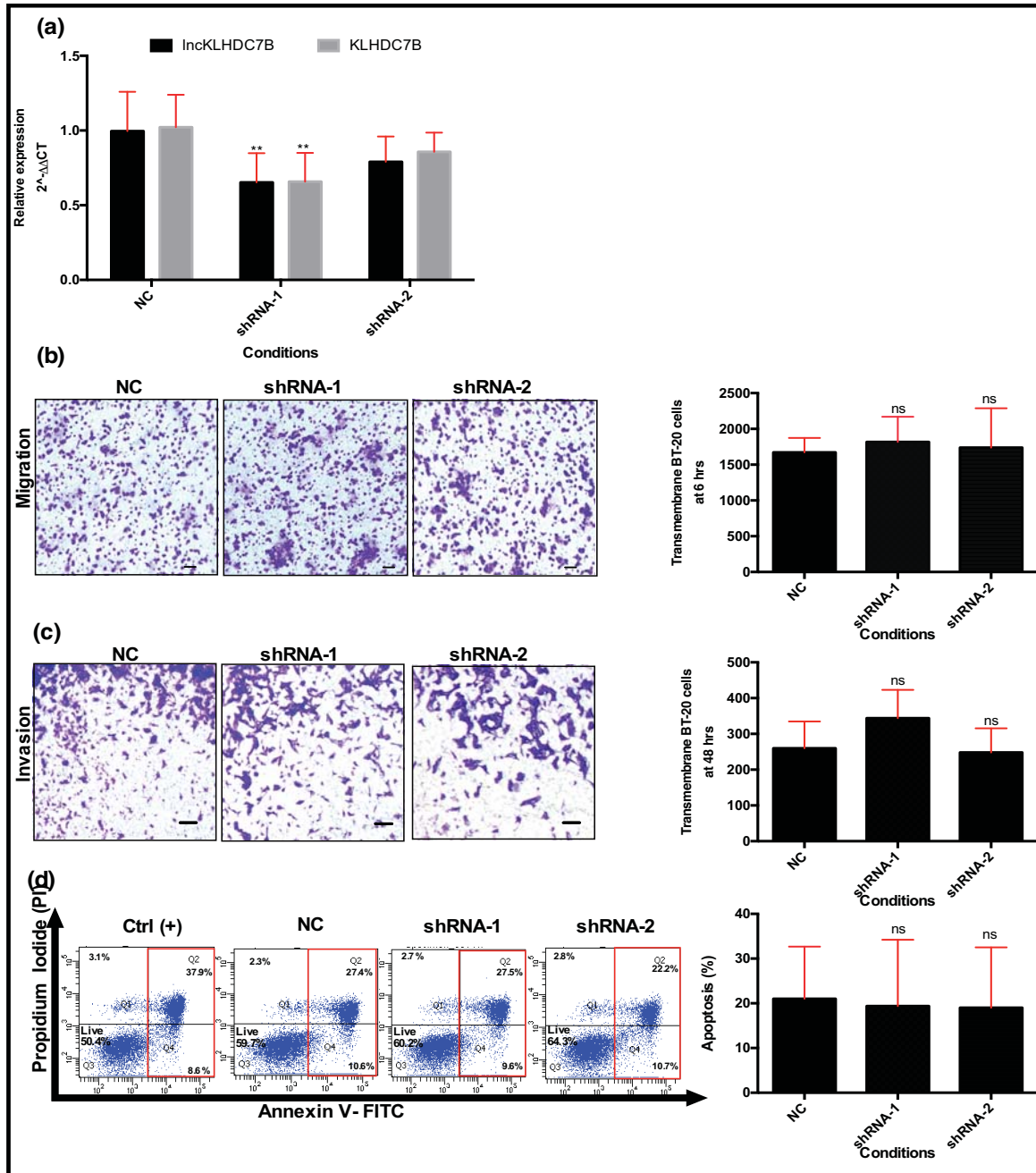
El silenciamiento del *LncKLHDC7B* incrementó significativamente la habilidad migratoria de las células HCC1187 en comparación con el grupo control (figura 20b). Este mecanismo podría deberse a la alteración en la sobreexpresión de genes relacionados con estos procesos celulares como se han reportado previamente. Tal es el caso de *TNC*, *CDH3*, *APP*, *EPS8* y *CAV1* [113-117]. Como se muestra en la figura 19c, la capacidad de invasión de HCC1187 se incrementó al silenciar el lncRNA en comparación al grupo control. Este mecanismo podría estar relacionado con la sobreexpresión de genes que han sido previamente involucrados en procesos de invasión; tales como, *MMP2* [118, 119], *ADAM9* [120] y *AMOT* [121]. Por otra parte, el silenciamiento del lncRNA disminuyó el número de células apoptóticas de HCC1187 en ~23%, al inducir la muerte celular con el agente QBS (figura 20d). Este efecto podría ser explicado, por una parte a la subexpresión de genes relacionados con la inducción de apoptosis como *PMAIP1* [122], *RAD21* [123, 124], *CCAR2* también conocido como *DBC1* [125].

Para determinar si el efecto del silenciamiento del *LncKLHDC7B* y los procesos de migración, invasión y apoptosis son exclusivos en cáncer de mama triple negativo, silenciamos al lncRNA en dos líneas celulares adicionales de cáncer de mama (BT-20 y MCF-7). Los resultados mostraron que el apagado del *LncKLHDC7B* en MCF-7 con fenotipo luminal, afectó la expresión de *KLHDC7B* (figura 21a) como se observó en HCC1187. Además, el silenciamiento promovió la migración e invasión de MCF-7 (figura 21b y c); así como la resistencia a la apoptosis (18-24%) comparado con el grupo control (figura 21d).

Sin embargo, en un análisis adicional en la línea celular BT-20 triple negativo (fenotipo UNS), se obtuvo un silenciamiento del *LncKLHDC7B* no mayor al 35% por uno de los shRNAs (figura 22a), observando una disminución de *KLHDC7B* hasta un 34% (figura 22a). Además, los procesos de migración (figura 22b) e invasión (figura 22c) y resistencia a la apoptosis (figura 22d) fueron ligeramente alteradas de manera no significativa por el shRNA-1, el cual tuvo mayor efecto en el silenciamiento del lncRNA.



**Figura 21. Análisis funcional del *IncKLHDC7B* en la línea celular MCF-7.** a) Niveles de expresión del *IncKLHDC7B* en la línea celular MCF-7 (fenotipo luminal) de cáncer de mama después de silenciar por el sistema de shRNA. La expresión del gen codificación *KLHDC7B* se ve afectada por el silenciamiento *IncKLHDC7B*. b) Ensayo de migración e **c)** invasión celular al silenciar al lncRNA en MCF-7. Imágenes representativas son mostradas en a la izquierda y sus cuantificaciones a la derecha. Barra de escala = 50µm. **d)** Análisis de apoptosis por citometría de flujo (apoptosis temprana y tardía) en MCF-7 transfectadas con shRNA NC, shRNA-1 y shRNA-2. Los resultados son mostrados en medias ± DE de al menos tres experimentos independientes. Prueba T de student fue realizada para determinar la significancia estadística \*p<0.05, \*\*p<0.01, \*\*\*p<0.001, NC vs shRNAs.



**Figura 22. Análisis funcional por el silenciamiento del *IncKLHDC7B* en la línea celular BT-20.** a) La expresión del *IncKLHDC7B* y *KLHDC7B* en células BT-20 transfectadas con shRNA NC o shRNA-1 o -2 hacia el *IncKLHDC7B* fue determinado mediante qRT-PCR. b) El ensayo de migración e c) invasión por el silenciamiento del *IncKLHDC7B* en BT-20. Las imágenes representativas se muestran a la izquierda y su cuantificación a la derecha. Barra de escala = 50 $\mu$ m. d) Análisis de citometría para el proceso de apoptosis (temprana y tardía) en células BT-20 transfectadas con control y shRNAs-1 y -2 después de la tinción con Annexin-V / PI. Todos los datos se muestran como la media  $\pm$  desviación estándar de al menos de tres experimentos independientes. La prueba t de Student se realizó para determinar la significancia estadística, ns =  $p > 0.05$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  de NC vs shRNAs.

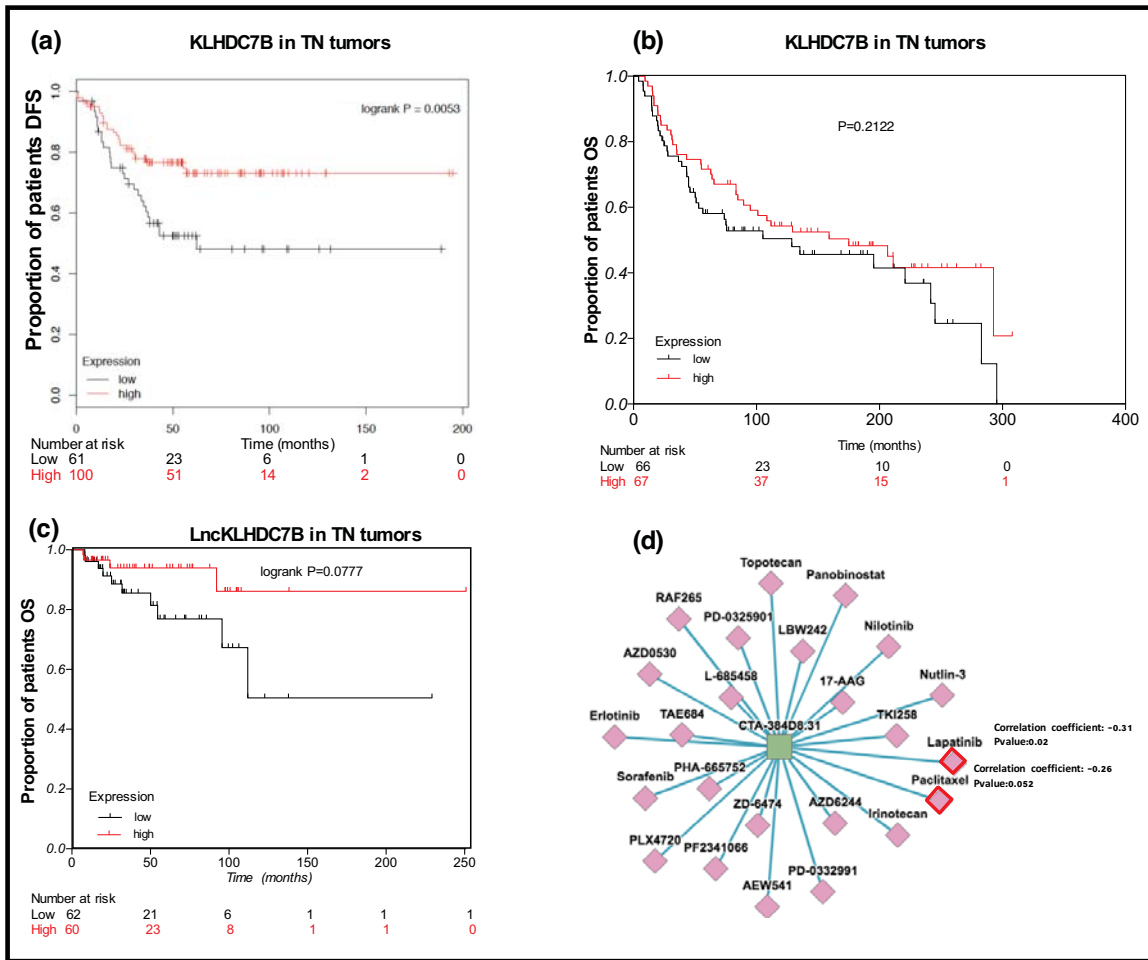
---

---

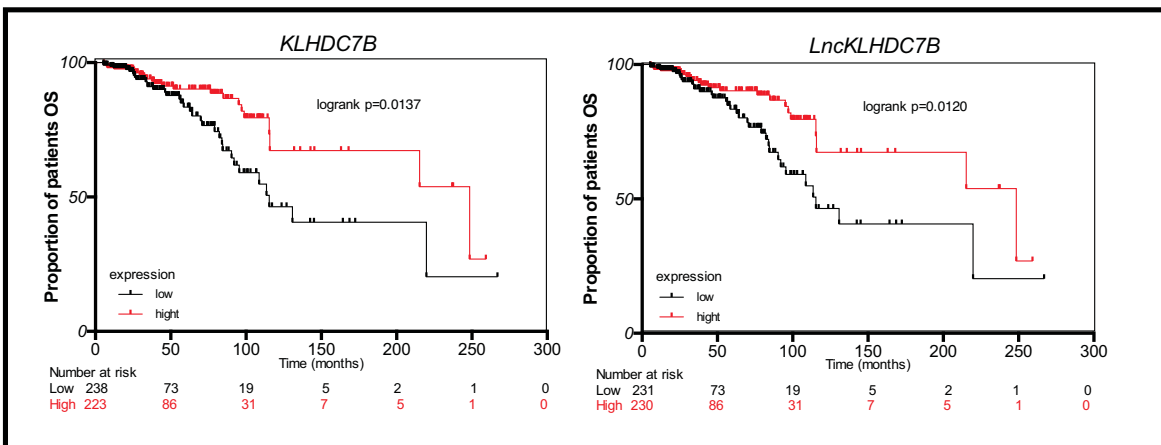
## Expresión del *LncKLHDC7B* y *KLHDC7B*, implicaciones clínicas en cáncer de mama

Para investigar la implicación clínica de la expresión del *LncKLHDC7B* y *KLHDC7B* en cáncer de mama triple negativo, se analizó e identificó los niveles de expresión del lncRNA y mRNA e información clínica de bases de datos públicos [8, 126, 127]. Por una parte, la subexpresión de *KLHDC7B* fue asociada con una alta probabilidad de recurrencia o metástasis (n=161,  $p= 0.0053$ , figura 23a), pero no al riesgo de muerte comparado con las muestras que mostraron mayor expresión del gen codificante (n= 133,  $p= 0.2122$ , figura 23b). Sin embargo, no se detectó una probabilidad significativa de muerte por la baja expresión del lncRNA en pacientes con cáncer de mama triple negativo (figura 23c). Sin embargo, un análisis generalizado en cáncer de mama, mostró que la baja expresión del *LncKLHDC7B* y *KLHDC7B* se asoció significativamente con una menor supervivencia (HR:1.910, IC 95%= 1.147 a 3.233, logrank test  $p= 0.0137$ ; HR: 1.933, IC 95%= 1.161 a 3.273, logrank test  $p= 0.0120$ , respectivamente) comparado con pacientes con mayor expresión de *LncKLHDC7B* o *KLHDC7B* (figura 24).





**Figura 23. Implicaciones clínicas de la sub-expresión de *IncKLHC7B* y *KLHC7B* en CMTN.** a) Curva de análisis de Kaplan-Meier de supervivencia libre de enfermedad (DFS, por sus siglas en inglés) en CMTN de acuerdo a la expresión de *KLHC7B*. b) Análisis de supervivencia global (OS) de acuerdo a la expresión de *KLHC7B*. c) Análisis de sobrevida global Kaplan-Meier de acuerdo a la expresión de *IncKLHC7B* a partir de datos públicos. (d) Predicción in silico de la susceptibilidad a fármacos aprobados por FDA relacionados con la expresión de *IncKLHC7B* en la línea celular de cáncer de mama triple negativo.



**Figura 24. Análisis de sobrevida global en cáncer de mama.** El análisis de Kaplan-Meier se analizó de acuerdo con los niveles de expresión de *KLHC7B* y *IncKLHC7B* a partir de los datos de TCGA BRCA [126, 128]. El valor de p se obtuvo usando la prueba log-rank.



---

---

Estudios previos y así como los nuestros, indican que los lncRNAs se expresan diferencialmente entre los subtipos de cáncer de mama; por lo que los lncRNAs podrían explorarse como una nueva clase de biomarcadores clínicos de respuesta a terapia o para el descubrimiento de dianas farmacológicas [129-131]. Para identificar lncRNAs relacionados a fármacos en cáncer de mama, se exploró la asociación de los niveles de expresión del *LncKLHDC7B* con la sensibilidad hacia fármacos derivado de bases de datos. Para este fin, se utilizó la herramienta IncMAP [128] que calcula el coeficiente de correlación de Spearman entre la expresión de lncRNA y los valores de CI50 de 24 fármacos aprobados para fines oncológicos en líneas celulares de CMTN. Se observó una correlación significativamente negativa entre la expresión de *LncKLHDC7B* y las concentraciones de Lapatinib y Paclitaxel (figura 23d), los cuales han sido evaluados en diferentes ensayos clínicos (<https://clinicaltrials.gov/ct2/home>).

## Discusión

El cáncer de mama triple negativo es una enfermedad heterogénea a nivel celular, molecular y clínico [46-48, 56]. Recientemente, los lncRNAs han emergido como moléculas involucradas en diversos procesos biológicos [132-134], así como potenciales biomarcadores o blancos terapéuticos [117, 129-131, 135]. En el presente trabajo, nos enfocamos en la clasificación de los subtipos moleculares de los tumores de cáncer de mama triple negativo, usando la expresión de mRNAs [46, 47] para posteriormente identificar los perfiles de expresión de lncRNAs de cada subtipos de CMTN. De acuerdo a los controles de calidad derivados del análisis de microarreglos y el análisis componentes principales, se presentó una baja variabilidad de las muestras analizadas entre cada componente evaluado. Esta baja variabilidad podría deberse a que todas las muestras pertenecen a cáncer, y tienen poca dispersión. Sin embargo, se observó una discreta agrupación de algunos subtipos de moleculares, como los inmunomoduladores y mesenquimales.

---

---

En cuanto al subtipo de CMTN y las características clínicas, no detectamos ninguna asociación significativa entre los subtipos de TNBC y el tamaño del tumor, histología o datos de supervivencia, excepto grado de tumor. Sin embargo, observamos que los subtipos mesenquimales e inmunomoduladores fueron los más frecuentes en nuestro estudio, siendo estos resultados similares a los previamente reportados por otros trabajos [46, 48, 56].

Mediante un análisis supervisado, identificamos un perfil de 710 lncRNAs expresados entre los subtipos de CMTN de las 156 muestras analizadas. De manera interesante, el subtipo LAR presentó el mayor número de lncRNAs alterados; entre los cuales podemos destacar a CBR3-AS1 descrito sobreexpresado en cáncer gástrico [136], *LINC01207* el cual se ha demostrado que interactúa con el represor transcripcional EZH2 en adenocarcinoma de pulmón y favorece la proliferación celular [137]. Otro lncRNA fue *PCAT18*, un lncRNA implicado en procesos como proliferación, migración e invasión celular dependiente de la activación de la vía del receptor de andrógenos en cáncer de próstata [138]. Dado este último dato, se podría sugerir que los subtipos de cáncer de mama con fenotipo LAR presentaron mayor número de lncRNAs desregulados probablemente por la activación de la vía de andrógenos, un proceso donde se ha demostrado el papel de este receptor en la regulación de la expresión de lncRNAs en cáncer de próstata [139, 140]. Sin embargo, más estudios son necesarios para demostrar la regulación de los lncRNAs por la vía de andrógenos en cáncer de mama triple negativo.

Respecto a los subtipos de CMTN con el menor número de lncRNA alterados, se encontró a los subtipos BL1 y BL2. De manera interesante, la mayoría de los lncRNAs identificados en estos subtipos no han sido descritos o caracterizados funcionalmente; a excepción del *LINC01133*, detectado en el subtipo BL2 y cuya expresión ha sido demostrada alterada en cáncer de pulmón de células escamosas y relacionada con una menor supervivencia [141].

---

---

Por otra parte, en el subtipo M entre los lncRNAs identificados se observó a *H19*, un lncRNA que promueve la transición epitelio mesénquima (TEM) [142], y la alteración en su expresión ha sido asociada con un pobre pronóstico en diferentes tipos de cáncer [142, 143]. *MEG3* y *MEG8* fueron otros lncRNAs encontrados sobreexpresados en el subtipo MSL y en recientes estudios, muestran que *MEG3* y *MEG8* contribuyen en el proceso de TEM [144] [145], así como en la diferenciación de células troncales mesenquimales en mieloma múltiple.

Mediante el uso de bases de datos, se validó la expresión de un grupo de lncRNAs identificados entre los subtipos de CMTN. De manera interesante, solo el ~34% de los lncRNAs de nuestro estudio fueron validados, en una cohorte de muestras independientes de tejido fresco congelado. Como se ha reportado previamente, un análisis realizado en tejidos frescos congelados contra tejidos parafinados de CMTN, es mayor la falta de concordancia de la expresión de RNAs no codificantes (78%), en comparación del RNA codificante de proteínas (27%) [146]. Esto podría explicar el bajo número de lncRNAs validados en este estudio. Sin embargo, factores como: la metodología para la obtención del RNA [147], las diferencias entre las químicas usadas para la detección del transcriptoma, la captura del RNA (por poli-A) [146], la eliminación del RNA ribosomal [147] o la longitud del transcrito de RNA [146], podrían ser factores clave para tomar en cuenta las diferencias encontradas a nivel de expresión, sobre todo para RNAs no codificantes.

Por otra parte, para analizar de una forma global el papel de los lncRNAs diferencialmente expresados identificados en cada uno de los subtipos de CMTN, realizamos un análisis de asociación por culpabilidad, el cual reveló las vías o rutas biológicas asociadas con los lncRNAs expresados para cada uno de los subtipos de CMTN; excepto para el subtipo BL2, posiblemente por la falta de correlación entre la expresión de mRNAs y lncRNAs. Nuestros hallazgos indican que los perfiles de expresión pueden distinguir los subtipos de CMTN, y que los lncRNAs pueden estar involucrados en importantes vías relacionadas con la

---

---

biología del tumor, contribuyendo a la regulación de los circuitos entre lncRNAs y mRNAs.

Entre los diferentes lncRNAs desregulados en el subtipo IM, se encontró a los lncRNAs *IFNG-AS1* y *LncKLHDC7B* sobreexpresados; *IFNG-AS1* es un lncRNA involucrado en la respuesta inflamatoria y regula positivamente la expresión de su gen codificante adyacente *IFNG* [148]. Corroborando esta información, se identificó que el *IFNG-AS1* y el *IFNG* mostraron una correlación positiva en la expresión de estos dos genes ( $r^2=0.6$ ,  $p=2.2e-16$ ). Por otra parte, otro par de transcritos sobreexpresados positivamente fueron: el gen codificante *KLHDC7B* y *LncKLHDC7B* con una correlación de  $r^2=0.9$  ( $p=2.2e-16$ ), sugiriendo que el lncRNA pudiera regular la expresión del gen codificante *KLHDC7B*.

Dado que hubo una relación entre los lncRNAs y el subtipo inmunomodulador, el cual fue uno de los subtipos tumorales más frecuentes, y con un número de lncRNAs importantes en vías implicadas en el establecimiento del fenotipo inmune y por tener un mejor pronóstico clínico [149], se decidió caracterizar a este subtipo tumoral. Varias evidencias indican la importancia del par de transcritos *LncKLHDC7B/KLHDC7B*, incluyendo la expresión específica o enriquecimiento del *LncKLHDC7B* en el subtipo IM. Los niveles de expresión del *LncKLHDC7B* están significativamente correlacionados con su gen codificante (localizado en el mismo locus) y nuestro análisis de asociación por culpabilidad resalta la correlación entre el lncRNA y las vías clave relacionadas en cáncer.

*KLHDC7B* (proteína 7B que contiene dominios Kelch) es un miembro de la superfamilia Kelch; las cuales son proteínas relacionadas con procesos celulares como el rearrreglo del citoesqueleto y degradación de proteínas [150]. Alteraciones en este tipo de proteínas han sido asociadas con varios tipos de cáncer, incluyendo: cáncer de pulmón, leucemia, cáncer de próstata, de cerebro y la enfermedad de Hodgkin [151]. Hasta donde tenemos conocimiento, el papel de *KLHDC7B* es poco conocido en cáncer de mama. Por una parte se ha detectado la hipermetilación y sobreexpresión de este gen en cáncer de mama, sugiriendo

---

---

un posible papel como marcador epigenético [152] y su papel en la resistencia a muerte celular [153]. Sin embargo, el papel funcional del *LncKLHDC7B* es desconocido.

El *LncKLHDC7B* es un transcrito de 260 bases y cuenta con dos exones. A nivel genómico es un lncRNA de ~1.26 Kb, localizado en el brazo largo q13.33 del cromosoma 22 y comparte el mismo locus que *KLHDC7B* [154, 155]. Diversos estudios han demostrado el papel de los lncRNAs como reguladores de la expresión de genes en *cis* o *trans* [66, 156]. En este estudio, primero corroboramos la correlación de la expresión del *LncKLHDC7B* y *KLHDC7B* en muestras independientes de bases de datos y líneas celulares. Esta sobreexpresión fue corroborada en la línea celular HCC1187 con fenotipo inmunomodulador. Posteriormente, se silenció el *LncKLHDC7B* usando sistemas de shRNAs, lo que resultó en la subexpresión del gen codificante *KLHDC7B*. Así, los resultados sugieren que el *LncKLHDC7B* pudiera regular la transcripción de *KLHDC7B*. Adicional a esta evaluación, se analizó el panorama transcripcional global por el efecto del silenciamiento del *LncKLHDC7B* en HCC1187. Se encontró que la expresión de genes se modificó después del silenciamiento; lo que podría sugerir el papel potencial en *trans* por parte del lncRNA.

Mediante análisis de enriquecimiento de vías se identificó procesos celulares alterados; como muerte celular, necrosis, apoptosis, movimiento celular, migración, invasión y organización del citoplasma por el silenciamiento del lncRNA. De manera interesante, se observó varios genes miembro de la familia Kelch sobre y subexpresados; algunos de ellos, relacionados con procesos de migración e invasión celular, así como en apoptosis [157, 158]. Estudios previos mostraron que el silenciamiento de *KLHL39*, otro miembro de la familia Kelch, incrementa la migración e invasión celular, pero decrementa la muerte celular por anoikis [159]. De acuerdo con nuestros datos, se mostró que el silenciamiento del *LncKLHDC7B* incrementó significativamente la migración e invasión celular, así como la resistencia a la apoptosis de manera significativa en las líneas celulares HCC1187 y MCF-7 (fenotipo luminal). Estas observaciones fueron también

---

---

apoyadas por otro análisis en la línea celular BT-20 (la cual mostró mayor expresión del *LncKLHDC7B* y *KLHDC7B*). De manera interesante, el silenciamiento del ~30% fue observado por uno de los shRNAs en BT-20. Aunque de forma no significativa, se observó un incremento en los procesos de migración e invasión y una tendencia a la resistencia a la apoptosis. En general, estas observaciones sugieren que el efecto de la inhibición del *LncKLHDC7B* y *KLHDC7B* pudiera ser aplicable no solo en cáncer de mama triple negativo, si no en cáncer de mama en general.

Estudios previos sugieren que la sobreexpresión de proteínas de la familia Kelch, como *KLHL2*, incrementan la apoptosis [160] o el silenciamiento de *KLHDC7B* favorece la resistencia a muerte [153]. Aquí mostramos que el silenciamiento del *LncKLHDC7B* llevó a la resistencia a la muerte celular; resultados similares han sido reportados donde el silenciamiento de miembros de la familia Kelch llevan a la resistencia a muerte celular por anoikis y está asociado con menor sobrevida en cáncer colon [159].

Estos resultados apoyan el papel potencial del *LncKLHDC7B* como un gen regulador de la expresión de *KLHDC7B* y su asociación con el rearrreglo del citoesqueleto, migración e invasión celular. Otra evidencias del papel potencial de los genes *KLHDC7B* y lncRNAs, provienen del trabajo realizado por Guo y colaboradores [161], donde describen un nuevo sitio alternativo de *splicing* 5' de *KLHDC7B* en carcinoma cervical. Identificando este nuevo sitio, observamos que corresponde a una región que codifica a un lncRNA (Tcons\_00029745) y se relacionó con diferenciación celular y el tamaño del tumor en carcinoma cervical [161].

Nuestros datos indican que el *LncKLHDC7B* es requerido para inhibir el proceso de migración e invasión celular, contribuyendo en la inducción a la apoptosis, sugiriendo su importante actividad reguladora en tumores de CMTN con enriquecimiento del *LncKLHDC7B* o probablemente podría ser un proceso generalizado en cáncer de mama.

---

---

Para investigar la correlación de la expresión del *LncKLHDC7B* o *KLHDC7B* y la significancia clínica en cáncer de mama triple negativo, el análisis de curvas de supervivencia fue realizado usando datos públicamente disponibles. Por una parte, la baja expresión de *KLHDC7B* de muestras de CMTN se asoció a una menor sobrevida libre de enfermedad. Mientras la baja expresión del *LncKLHDC7B* y *KLHDC7B* en pacientes de CMTN presentaron menor sobrevida global de manera no significativa. Sin embargo, un análisis general en cáncer de mama mostró una mayor sobrevida global de manera significativa en pacientes con alta expresión de *LncKLHDC7B* y *KLHDC7B*.

Hasta donde sabemos este es el primer reporte de la asociación de la alta expresión de *LncKLHDC7B* y *KLHDC7B* con un mejor pronóstico en cáncer de mama. Además, un análisis *in silico* develó las probables implicaciones clínicas de la expresión de *LncKLHDC7B* con drogas oncológicas. Los resultados sugieren que hay una mayor sensibilidad a lapatinib y paclitaxel en células que sobreexpresan al *LncKLHDC7B*; sin embargo, la limitación de este análisis fue la validación en un modelo *in vitro*.

## Conclusiones

En este trabajo se exploró el panorama transcripcional de mRNA y lncRNAs de una enfermedad muy compleja y heterogénea representada por el cáncer de mama triple negativo. Se identificaron varios lncRNAs diferencialmente expresados entre los subtipos de cáncer de mama triple negativo. Al analizar la correlación de la expresión de lncRNAs y mRNA en el subtipo inmunomodulador de CMTN, se identificó al par de transcritos codificante y no codificante de proteína *LncKLHDC7B/KLHDC7B*. El silenciamiento del *LncKLHDC7B* llevó a la subexpresión de *KLHDC7B*, un miembro de la familia Kelch. La baja expresión de ambos transcritos resultó en el aumento de la migración e invasión celular con resistencia a la muerte celular. Estos resultados sugieren que el *LncKLHDC7B* podría regular la agresividad del subtipo IM de cáncer de mama en parte por la regulación de la expresión de su gen huésped *KLHDC7B*. Además, se observó



---

---

una tendencia constante hacia una mayor supervivencia global en las pacientes con mayor expresión de *LncKLHDC7B* y *KLHDC7B*.

Aunque la relevancia clínica de la expresión del *LncKLHDC7B* debe de ser analizada mas a fondo, nuestros resultados reafirman que la subexpresión de estos transcritos están asociados con un pobre pronóstico en pacientes con cáncer de mama triple negativo. Sin embargo, el mecanismo exacto de cómo actúa este lncRNA deberá ser dilucidado para entender mejor su acción en cáncer.

## **Perspectivas**

- Identificar las proteínas asociadas al *LncKLHDC7B* mediante el ensayo de inmunoprecipitación de RNA (RIP) para determinar el mecanismo por el cuál el lncRNA hace su función.
- Realizar ensayos de aislamiento de cromatina por purificación de RNA para determinar la unión del *LncKLHDC7B* en las regiones promotoras de los genes alterados por la modulación en su expresión.
- Continuar con la caracterización de los lncRNAs identificados en cada uno de los subtipos de cáncer de mama triple negativo.
- Evaluar si la activación del receptor de andrógenos modula la expresión de los lncRNAs identificados en el subtipo de cáncer de mama triple negativo con fenotipo LAR.
- Aumentar el número de muestras de subtipos de CMTN menos representados (Basal tipo 2) para enriquecer en análisis de rutas biológicas de los lncRNAs expresados en este subtipo.

---

---

## Referencias

1. Inman, J.L., et al., *Mammary gland development: cell fate specification, stem cells and the microenvironment*. *Development*, 2015. **142**(6): p. 1028-42.
2. Sun, Y.S., et al., *Risk Factors and Preventions of Breast Cancer*. *Int J Biol Sci*, 2017. **13**(11): p. 1387-1397.
3. Winters, S., et al., *Breast Cancer Epidemiology, Prevention, and Screening*. *Prog Mol Biol Transl Sci*, 2017. **151**: p. 1-32.
4. Subramani, R. and R. Lakshmanaswamy, *Pregnancy and Breast Cancer*. *Prog Mol Biol Transl Sci*, 2017. **151**: p. 81-111.
5. Moamen Bydoun, P.M., Graham Dellaire,, *Chapter 13 - Breast Cancer Genomics*, in *Cancer Genomics*, J.N.B. Graham Dellaire, Robert J. Arceci,, Editor. 2014, Academic Press,. p. 213-232,.
6. Chen, S.I. and C.C. Hsieh, *Why are women with obesity more likely to develop breast cancer*. *Future Oncol*, 2018. **14**(16): p. 1523-1526.
7. Cancer, I.A.f.R.o., *Breast Cancer Screening* Vol. 15. 2016.
8. Curtis, C., et al., *The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups*. *Nature*, 2012. **486**(7403): p. 346-52.
9. Weigelt, B., et al., *Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer*. *Cancer Res*, 2005. **65**(20): p. 9155-8.
10. mamario, C.M.s.d.y.t.d.c. *Consenso Mexicano sobre diagnóstico y tratamiento del cáncer mamario*. 2017; Available from: [http://consensocancermamario.com/documentos/FOLLETO\\_CONSENSO\\_DE\\_CANCER\\_DE\\_MAMA\\_7aRev2017c.PDF](http://consensocancermamario.com/documentos/FOLLETO_CONSENSO_DE_CANCER_DE_MAMA_7aRev2017c.PDF).
11. Bastien, R.R., et al., *PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers*. *BMC Med Genomics*, 2012. **5**: p. 44.
12. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets*. *Proc Natl Acad Sci U S A*, 2003. **100**(14): p. 8418-23.
13. Prat, A., et al., *Molecular characterization of basal-like and non-basal-like triple-negative breast cancer*. *Oncologist*, 2013. **18**(2): p. 123-33.
14. Badve, S., et al., *Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists*. *Mod Pathol*, 2011. **24**(2): p. 157-67.
15. Bauer, K.R., et al., *Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry*. *Cancer*, 2007. **109**(9): p. 1721-8.
16. Dent, R., et al., *Triple-negative breast cancer: clinical features and patterns of recurrence*. *Clin Cancer Res*, 2007. **13**(15 Pt 1): p. 4429-34.
17. Foulkes, W.D., I.E. Smith, and J.S. Reis-Filho, *Triple-negative breast cancer*. *N Engl J Med*, 2010. **363**(20): p. 1938-48.
18. Rastelli, F., et al., *Triple-negative breast cancer: current state of the art*. *Tumori*, 2010. **96**(6): p. 875-88.

- 
- 
19. Mohar-Betancourt, A., et al., *[Prognostic factors in patients with breast cancer and brain metastasis as the first site of recurrence]*. *Salud Publica Mex*, 2018. **60**(2): p. 141-150.
  20. Li, X., et al., *Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer*. *Breast Cancer Res Treat*, 2017. **161**(2): p. 279-287.
  21. Villarreal-Garza, C., et al., *The prevalence of BRCA1 and BRCA2 mutations among young Mexican women with triple-negative breast cancer*. *Breast Cancer Res Treat*, 2015. **150**(2): p. 389-94.
  22. Pal, T., et al., *A high frequency of BRCA mutations in young black women with breast cancer residing in Florida*. *Cancer*, 2015. **121**(23): p. 4173-80.
  23. Pareja, F., et al., *Triple-negative breast cancer: the importance of molecular and histologic subtyping, and recognition of low-grade variants*. *NPJ Breast Cancer*, 2016. **2**: p. 16036.
  24. Atchley, D.P., et al., *Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer*. *J Clin Oncol*, 2008. **26**(26): p. 4282-8.
  25. Couch, F.J., et al., *Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer*. *J Clin Oncol*, 2015. **33**(4): p. 304-11.
  26. Harano, K., et al., *Rates of immune cell infiltration in patients with triple-negative breast cancer by molecular subtype*. *PLoS One*, 2018. **13**(10): p. e0204513.
  27. Chaudhary, L.N., K.H. Wilkinson, and A. Kong, *Triple-Negative Breast Cancer: Who Should Receive Neoadjuvant Chemotherapy?* *Surg Oncol Clin N Am*, 2018. **27**(1): p. 141-153.
  28. Lin, N.U., et al., *Clinicopathologic features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network*. *Cancer*, 2012. **118**(22): p. 5463-72.
  29. Sihto, H., et al., *Breast cancer biological subtypes and protein expression predict for the preferential distant metastasis sites: a nationwide cohort study*. *Breast Cancer Res*, 2011. **13**(5): p. R87.
  30. Bartmann, C., et al., *Pattern of metastatic spread and subcategories of breast cancer*. *Arch Gynecol Obstet*, 2017. **295**(1): p. 211-223.
  31. Tseng, L.M., et al., *Distant metastasis in triple-negative breast cancer*. *Neoplasma*, 2013. **60**(3): p. 290-4.
  32. Dent, R., et al., *Pattern of metastatic spread in triple-negative breast cancer*. *Breast Cancer Res Treat*, 2009. **115**(2): p. 423-8.
  33. Rouzier, R., et al., *Breast cancer molecular subtypes respond differently to preoperative chemotherapy*. *Clin Cancer Res*, 2005. **11**(16): p. 5678-85.
  34. Carey, L.A., et al., *The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes*. *Clin Cancer Res*, 2007. **13**(8): p. 2329-34.
  35. Liedtke, C., et al., *Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer*. *J Clin Oncol*, 2008. **26**(8): p. 1275-81.
  36. Ye, L.L., *Professor I. Craig Henderson: triple negative breast cancer is most chemo-sensitive of all subtypes of cancer*. *Chin Clin Oncol*, 2016. **5**(6): p. 84.

- 
- 
37. Sharma, P., *Update on the Treatment of Early-Stage Triple-Negative Breast Cancer*. *Curr Treat Options Oncol*, 2018. **19**(5): p. 22.
  38. Vallejos, C.S., et al., *Breast cancer classification according to immunohistochemistry markers: subtypes and association with clinicopathologic variables in a peruvian hospital database*. *Clin Breast Cancer*, 2010. **10**(4): p. 294-300.
  39. Moraima Márquez, S., et al., *Sobrevida en pacientes con cáncer de mama triple negativo*. *Revista de Obstetricia y Ginecología de Venezuela*, 2012. **72**(3): p. 152-160.
  40. Goncalves, H., Jr., et al., *Survival Study of Triple-Negative and Non-Triple-Negative Breast Cancer in a Brazilian Cohort*. *Clin Med Insights Oncol*, 2018. **12**: p. 1179554918790563.
  41. Quirós Alpízar, J.L. and K. Espinoza Morales, *Supervivencia a 5 años de pacientes con cáncer de mama triple negativo*. *Medicina Legal de Costa Rica*, 2017. **34**: p. 59-72.
  42. Lara-Medina, F., et al., *Triple-negative breast cancer in Hispanic patients: high prevalence, poor prognosis, and association with menopausal status, body mass index, and parity*. *Cancer*, 2011. **117**(16): p. 3658-69.
  43. Martinez, M.E., et al., *Reproductive factors, heterogeneity, and breast tumor subtypes in women of mexican descent*. *Cancer Epidemiol Biomarkers Prev*, 2013. **22**(10): p. 1853-61.
  44. Perez-Rodriguez, G., *[Prevalence of breast cancer sub-types by immunohistochemistry in patients in the Regional General Hospital 72, Instituto Mexicano del Seguro Social]*. *Cir Cir*, 2015. **83**(3): p. 193-8.
  45. Ruvalcaba Limon, E., et al., *[Neoadjuvant Chemotherapy (NC) Response in Patients with Breast Cancer According to Immunohistochemical Intrinsic Subtypes (IHC)]*. *Gac Med Mex*, 2014. **150 Suppl 2**: p. 161-70.
  46. Lehmann, B.D., et al., *Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies*. *J Clin Invest*, 2011. **121**(7): p. 2750-67.
  47. Chen, X., et al., *TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer*. *Cancer Inform*, 2012. **11**: p. 147-56.
  48. Masuda, H., et al., *Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes*. *Clin Cancer Res*, 2013. **19**(19): p. 5533-40.
  49. Lehmann, B.D., et al., *Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection*. *PLoS One*, 2016. **11**(6): p. e0157368.
  50. Santonja, A., et al., *Triple negative breast cancer subtypes and pathologic complete response rate to neoadjuvant chemotherapy*. *Oncotarget*, 2018. **9**(41): p. 26406-26416.
  51. Echavarria, I., et al., *Pathological Response in a Triple-Negative Breast Cancer Cohort Treated with Neoadjuvant Carboplatin and Docetaxel According to Lehmann's Refined Classification*. *Clin Cancer Res*, 2018. **24**(8): p. 1845-1852.

- 
- 
52. Burstein, M.D., et al., *Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer*. Clin Cancer Res, 2015. **21**(7): p. 1688-98.
  53. Jezequel, P., et al., *Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response*. Breast Cancer Res, 2015. **17**: p. 43.
  54. Stirzaker, C., et al., *Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value*. Nat Commun, 2015. **6**: p. 5899.
  55. Ring, B.Z., et al., *Generation of an algorithm based on minimal gene sets to clinically subtype triple negative breast cancer patients*. BMC Cancer, 2016. **16**: p. 143.
  56. Liu, Y.R., et al., *Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer*. Breast Cancer Res, 2016. **18**(1): p. 33.
  57. Masuda, H., et al., *Reverse phase protein array identification of triple-negative breast cancer subtypes and comparison with mRNA molecular subtypes*. Oncotarget, 2017. **8**(41): p. 70481-70495.
  58. GENETICS, I. *INSIGHT MOLECULAR LABS*. Available from: <http://www.insightmdx.com/testmenu>.
  59. Kapranov, P., et al., *The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA*. BMC Biol, 2010. **8**: p. 149.
  60. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
  61. Frankish, A., et al., *GENCODE reference annotation for the human and mouse genomes*. Nucleic Acids Res, 2019. **47**(D1): p. D766-D773.
  62. Derrien, T., et al., *The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression*. Genome Res, 2012. **22**(9): p. 1775-89.
  63. Ruiz-Orera, J., et al., *Long non-coding RNAs as a source of new peptides*. Elife, 2014. **3**: p. e03523.
  64. Kopp, F. and J.T. Mendell, *Functional Classification and Experimental Dissection of Long Noncoding RNAs*. Cell, 2018. **172**(3): p. 393-407.
  65. St Laurent, G., C. Wahlestedt, and P. Kapranov, *The Landscape of long noncoding RNA classification*. Trends Genet, 2015. **31**(5): p. 239-51.
  66. Gupta, R.A., et al., *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis*. Nature, 2010. **464**(7291): p. 1071-6.
  67. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs*. Cell, 2007. **129**(7): p. 1311-23.
  68. Lee, J.T., *Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control*. Nat Rev Mol Cell Biol, 2011. **12**(12): p. 815-26.
  69. Kino, T., et al., *Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor*. Sci Signal, 2010. **3**(107): p. ra8.



- 
- 
70. Sun, T.T., et al., *LncRNA GCInc1 Promotes Gastric Carcinogenesis and May Act as a Modular Scaffold of WDR5 and KAT2A Complexes to Specify the Histone Modification Pattern*. *Cancer Discov*, 2016. **6**(7): p. 784-801.
  71. Tsai, M.C., et al., *Long noncoding RNA as modular scaffold of histone modification complexes*. *Science*, 2010. **329**(5992): p. 689-93.
  72. Sun, C.C., et al., *Long Intergenic Noncoding RNA 00511 Acts as an Oncogene in Non-small-cell Lung Cancer by Binding to EZH2 and Suppressing p57*. *Mol Ther Nucleic Acids*, 2016. **5**(11): p. e385.
  73. Yoon, J.H., et al., *LincRNA-p21 suppresses target mRNA translation*. *Mol Cell*, 2012. **47**(4): p. 648-55.
  74. Gong, C. and L.E. Maquat, *lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements*. *Nature*, 2011. **470**(7333): p. 284-8.
  75. West, J.A., et al., *The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites*. *Mol Cell*, 2014. **55**(5): p. 791-802.
  76. Tripathi, V., et al., *The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation*. *Mol Cell*, 2010. **39**(6): p. 925-38.
  77. Gonzalez, I., et al., *A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature*. *Nat Struct Mol Biol*, 2015. **22**(5): p. 370-6.
  78. Militello, G., et al., *Screening and validation of lncRNAs and circRNAs as miRNA sponges*. *Brief Bioinform*, 2017. **18**(5): p. 780-788.
  79. Sen, R., et al., *Competing endogenous RNA: the key to posttranscriptional regulation*. *ScientificWorldJournal*, 2014. **2014**: p. 896206.
  80. Liang, W.C., et al., *The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer*. *Oncotarget*, 2015. **6**(26): p. 22513-25.
  81. Wang, Y. and D. Kong, *LncRNA GAS5 Represses Osteosarcoma Cells Growth and Metastasis via Sponging MiR-203a*. *Cell Physiol Biochem*, 2018. **45**(2): p. 844-855.
  82. Poliseno, L. and P.P. Pandolfi, *PTEN ceRNA networks in human cancer*. *Methods*, 2015. **77-78**: p. 41-50.
  83. Pan, Y., et al., *The Emerging Roles of Long Noncoding RNA ROR (lincRNA-ROR) and its Possible Mechanisms in Human Cancers*. *Cell Physiol Biochem*, 2016. **40**(1-2): p. 219-229.
  84. Abdollahzadeh, R., et al., *Competing endogenous RNA (ceRNA) cross talk and language in ceRNA regulatory networks: A new look at hallmarks of breast cancer*. *J Cell Physiol*, 2018.
  85. Shore, A.N. and J.M. Rosen, *Regulation of mammary epithelial cell homeostasis by lncRNAs*. *Int J Biochem Cell Biol*, 2014. **54**: p. 318-30.
  86. Lanz, R.B., et al., *Steroid receptor RNA activator stimulates proliferation as well as apoptosis in vivo*. *Mol Cell Biol*, 2003. **23**(20): p. 7163-76.
  87. Lanz, R.B., et al., *A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex*. *Cell*, 1999. **97**(1): p. 17-27.
  88. Ginger, M.R., et al., *A noncoding RNA is a potential marker of cell fate during mammary gland development*. *Proc Natl Acad Sci U S A*, 2006. **103**(15): p. 5781-6.

- 
- 
89. Askarian-Amiri, M.E., et al., *SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer*. RNA, 2011. **17**(5): p. 878-91.
  90. Li, C.H. and Y. Chen, *Targeting long non-coding RNAs in cancers: progress and prospects*. Int J Biochem Cell Biol, 2013. **45**(8): p. 1895-910.
  91. Wapinski, O. and H.Y. Chang, *Long noncoding RNAs and human disease*. Trends Cell Biol, 2011. **21**(6): p. 354-61.
  92. Yan, X., et al., *Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers*. Cancer Cell, 2015. **28**(4): p. 529-40.
  93. Su, X., et al., *Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes*. Oncotarget, 2014. **5**(20): p. 9864-76.
  94. Chen, C., et al., *Microarray expression profiling of dysregulated long non-coding RNAs in triple-negative breast cancer*. Cancer Biol Ther, 2015. **16**(6): p. 856-65.
  95. Maruyama, R., et al., *Altered antisense-to-sense transcript ratios in breast cancer*. Proc Natl Acad Sci U S A, 2012. **109**(8): p. 2820-4.
  96. Brunner, A.L., et al., *Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers*. Genome Biol, 2012. **13**(8): p. R75.
  97. Códoba, S.L.R., *Clasificación molecular de tumores de mama triple negativos mediante análisis de perfiles de expresión de microRNAs y marcadores inmunohistoquímicos*. , in *Facultad de Medicina*. 2015, Universidad Nacional Autónoma de México.
  98. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
  99. Guttman, M., et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals*. Nature, 2009. **458**(7235): p. 223-7.
  100. Stockel, D., et al., *Multi-omics enrichment analysis using the GeneTrail2 web service*. Bioinformatics, 2016. **32**(10): p. 1502-8.
  101. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis*. Nat Methods, 2012. **9**(7): p. 671-5.
  102. Li, R., H. Zhu, and Y. Luo, *Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures*. Int J Mol Sci, 2016. **17**(5).
  103. He, R.Z., D.X. Luo, and Y.Y. Mo, *Emerging roles of lncRNAs in the post-transcriptional regulation in cancer*. Genes Dis, 2019. **6**(1): p. 6-15.
  104. Li, J., et al., *LncTar: a tool for predicting the RNA targets of long noncoding RNAs*. Brief Bioinform, 2015. **16**(5): p. 806-12.
  105. Vire, E., et al., *The Polycomb group protein EZH2 directly controls DNA methylation*. Nature, 2006. **439**(7078): p. 871-4.
  106. Ganser, L.R., et al., *The roles of structural dynamics in the cellular functions of RNAs*. Nat Rev Mol Cell Biol, 2019. **20**(8): p. 474-489.
  107. Hussain, K.M., et al., *The essential role of clathrin-mediated endocytosis in the infectious entry of human enterovirus 71*. J Biol Chem, 2011. **286**(1): p. 309-21.



- 
- 
108. Papic, N., et al., *RNA-sequencing analysis of 5' capped RNAs identifies many new differentially expressed genes in acute hepatitis C virus infection*. *Viruses*, 2012. **4**(4): p. 581-612.
  109. Zhu, X., et al., *TMEM2 inhibits hepatitis B virus infection in HepG2 and HepG2.2.15 cells by activating the JAK-STAT signaling pathway*. *Cell Death Dis*, 2016. **7**(6): p. e2239.
  110. Zhao, L., et al., *The Downregulation of MicroRNA hsa-miR-340-5p in IAV-Infected A549 Cells Suppresses Viral Replication by Targeting RIG-I and OAS2*. *Mol Ther Nucleic Acids*, 2019. **14**: p. 509-519.
  111. Li, Y., et al., *Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses*. *Proc Natl Acad Sci U S A*, 2016. **113**(8): p. 2241-6.
  112. Benavente, F.M., et al., *Contribution of IDO to human respiratory syncytial virus infection*. *J Leukoc Biol*, 2019.
  113. Baek, S., et al., *CDH3/P-Cadherin regulates migration of HuCCT1 cholangiocarcinoma cells*. *Anat Cell Biol*, 2010. **43**(2): p. 110-7.
  114. Diaz, J., et al., *Rab5 is required in metastatic cancer cells for Caveolin-1-enhanced Rac1 activation, migration and invasion*. *J Cell Sci*, 2014. **127**(Pt 11): p. 2401-6.
  115. Lim, S., et al., *Amyloid-beta precursor protein promotes cell proliferation and motility of advanced breast cancer*. *BMC Cancer*, 2014. **14**: p. 928.
  116. Chen, C., et al., *Eps8 regulates cellular proliferation and migration of breast cancer*. *Int J Oncol*, 2015. **46**(1): p. 205-14.
  117. Zhang, Z., et al., *Fibroblast-derived tenascin-C promotes Schwann cell migration through beta1-integrin dependent pathway during peripheral nerve regeneration*. *Glia*, 2016. **64**(3): p. 374-85.
  118. Fagan-Solis, K.D., et al., *The RhoA pathway mediates MMP-2 and MMP-9-independent invasive behavior in a triple-negative breast cancer cell line*. *J Cell Biochem*, 2013. **114**(6): p. 1385-94.
  119. Kim, D. and S. Rhee, *Matrix metalloproteinase2 regulates MDAMB231 breast cancer cell invasion induced by active mammalian diaphanous-related formin 1*. *Mol Med Rep*, 2016. **14**(1): p. 277-82.
  120. Micocci, K.C., et al., *ADAM9 silencing inhibits breast tumor cell invasion in vitro*. *Biochimie*, 2013. **95**(7): p. 1371-8.
  121. Zhang, H. and Q. Fan, *MicroRNA-205 inhibits the proliferation and invasion of breast cancer by regulating AMOT expression*. *Oncol Rep*, 2015. **34**(4): p. 2163-70.
  122. Zhao, X., X. Liu, and L. Su, *Parthenolide induces apoptosis via TNFRSF10B and PMAIP1 pathways in human lung cancer cells*. *J Exp Clin Cancer Res*, 2014. **33**: p. 3.
  123. Chen, F., et al., *Caspase proteolysis of the cohesin component RAD21 promotes apoptosis*. *J Biol Chem*, 2002. **277**(19): p. 16775-81.
  124. Pati, D., N. Zhang, and S.E. Plon, *Linking sister chromatid cohesion and apoptosis: role of Rad21*. *Mol Cell Biol*, 2002. **22**(23): p. 8267-77.
  125. Kim, J.E., J. Chen, and Z. Lou, *DBC1 is a negative regulator of SIRT1*. *Nature*, 2008. **451**(7178): p. 583-6.

- 
- 
126. Anaya, J., *OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs*. *PeerJ Computer Science* 2:e67. 2016: PeerJ Computer Science.
  127. Lanczky, A., et al., *miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients*. *Breast Cancer Res Treat*, 2016. **160**(3): p. 439-446.
  128. Li, Y., et al., *LncMAP: Pan-cancer atlas of long noncoding RNA-mediated transcriptional network perturbations*. *Nucleic Acids Res*, 2018. **46**(3): p. 1113-1123.
  129. Prabhakar, B., X.B. Zhong, and T.P. Rasmussen, *Exploiting Long Noncoding RNAs as Pharmacological Targets to Modulate Epigenetic Diseases*. *Yale J Biol Med*, 2017. **90**(1): p. 73-86.
  130. Wu, T. and Y. Du, *LncRNAs: From Basic Research to Medical Application*. *Int J Biol Sci*, 2017. **13**(3): p. 295-307.
  131. Arun, G., S.D. Diermeier, and D.L. Spector, *Therapeutic Targeting of Long Non-Coding RNAs in Cancer*. *Trends Mol Med*, 2018.
  132. Martin, L. and H.Y. Chang, *Uncovering the role of genomic "dark matter" in human disease*. *J Clin Invest*, 2012. **122**(5): p. 1589-95.
  133. Kwok, Z.H. and Y. Tay, *Long noncoding RNAs: links between human health and disease*. *Biochem Soc Trans*, 2017. **45**(3): p. 805-812.
  134. Dong, Y., et al., *Long noncoding RNAs coordinate functions between mitochondria and the nucleus*. *Epigenetics Chromatin*, 2017. **10**(1): p. 41.
  135. Chandra Gupta, S. and Y. Nandan Tripathi, *Potential of long non-coding RNAs in cancer patients: From biomarkers to therapeutic targets*. *Int J Cancer*, 2016.
  136. Gao, J., R. Cao, and H. Mu, *Long non-coding RNA UCA1 may be a novel diagnostic and predictive biomarker in plasma for early gastric cancer*. *Int J Clin Exp Pathol*, 2015. **8**(10): p. 12936-42.
  137. Wang, G., H. Chen, and J. Liu, *The long noncoding RNA LINC01207 promotes proliferation of lung adenocarcinoma*. *Am J Cancer Res*, 2015. **5**(10): p. 3162-73.
  138. Crea, F., et al., *Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer*. *Oncotarget*, 2014. **5**(3): p. 764-74.
  139. Shih, J.W., et al., *Non-Coding RNAs in Castration-Resistant Prostate Cancer: Regulation of Androgen Receptor Signaling and Cancer Metabolism*. *Int J Mol Sci*, 2015. **16**(12): p. 28943-78.
  140. Yang, L., et al., *lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs*. *Nature*, 2013. **500**(7464): p. 598-602.
  141. Zhang, J., N. Zhu, and X. Chen, *A novel long noncoding RNA LINC01133 is upregulated in lung squamous cell cancer and predicts survival*. *Tumour Biol*, 2015. **36**(10): p. 7465-71.
  142. Wang, S.H., et al., *Upregulation of H19 indicates a poor prognosis in gallbladder carcinoma and promotes epithelial-mesenchymal transition*. *Am J Cancer Res*, 2016. **6**(1): p. 15-26.
  143. Chen, J.S., et al., *H19 serves as a diagnostic biomarker and up-regulation of H19 expression contributes to poor prognosis in patients with gastric cancer*. *Neoplasma*, 2016. **63**(2): p. 223-30.

- 
- 
144. Terashima, M., et al., *MEG3 Long Noncoding RNA Contributes to the Epigenetic Regulation of Epithelial-Mesenchymal Transition in Lung Cancer Cell Lines*. J Biol Chem, 2017. **292**(1): p. 82-99.
  145. Terashima, M., et al., *MEG8 long noncoding RNA contributes to epigenetic progression of the epithelial-mesenchymal transition of lung and pancreatic cancer cells*. J Biol Chem, 2018. **293**(47): p. 18016-18030.
  146. Jovanovic, B., et al., *Comparison of triple-negative breast cancer molecular subtyping using RNA from matched fresh-frozen versus formalin-fixed paraffin-embedded tissue*. BMC Cancer, 2017. **17**(1): p. 241.
  147. Guo, Y., et al., *RNA Sequencing of Formalin-Fixed, Paraffin-Embedded Specimens for Gene Expression Quantification and Data Mining*. Int J Genomics, 2016. **2016**: p. 9837310.
  148. Padua, D., et al., *A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation*. Am J Physiol Gastrointest Liver Physiol, 2016. **311**(3): p. G446-57.
  149. Bareche, Y., et al., *Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis*. Ann Oncol, 2018. **29**(4): p. 895-902.
  150. Adams, J., R. Kelso, and L. Cooley, *The kelch repeat superfamily of proteins: propellers of cell function*. Trends Cell Biol, 2000. **10**(1): p. 17-24.
  151. Gupta, V.A. and A.H. Beggs, *Kelch proteins: emerging roles in skeletal muscle development and diseases*. Skelet Muscle, 2014. **4**: p. 11.
  152. Kim, T.W., et al., *Hs.137007 is a novel epigenetic marker hypermethylated and up-regulated in breast cancer*. Int J Oncol, 2010. **36**(5): p. 1105-11.
  153. Jeong, G., et al., *A Kelch domain-containing KLHDC7B and a long non-coding RNA ST8SIA6-AS1 act oppositely on breast cancer cell proliferation via the interferon signaling pathway*. Sci Rep, 2018. **8**(1): p. 12922.
  154. Kent, W.J., et al., *The human genome browser at UCSC*. Genome Res, 2002. **12**(6): p. 996-1006.
  155. Volders, P.J., et al., *An update on LNCipedia: a database for annotated human lncRNA sequences*. Nucleic Acids Res, 2015. **43**(Database issue): p. D174-80.
  156. Engreitz, J.M., et al., *Local regulation of gene expression by lncRNA promoters, transcription and splicing*. Nature, 2016. **539**(7629): p. 452-455.
  157. Lian, Y.F., et al., *Upregulation of KLHDC4 Predicts a Poor Prognosis in Human Nasopharyngeal Carcinoma*. PLoS One, 2016. **11**(3): p. e0152820.
  158. Ohta, Y., et al., *A kelch family protein Nd1-L functions as a metastasis suppressor in cancer cells via Rho family proteins mediated mechanism*. Int J Oncol, 2010. **36**(2): p. 427-34.
  159. Chen, H.Y., et al., *KLHL39 suppresses colon cancer metastasis by blocking KLHL20-mediated PML and DAPK ubiquitination*. Oncogene, 2015. **34**(40): p. 5141-51.
  160. Tseng, L.A. and J.L. Bixby, *Interaction of an intracellular pentraxin with a BTB-Kelch protein is associated with ubiquitylation, aggregation and neuronal apoptosis*. Mol Cell Neurosci, 2011. **47**(4): p. 254-64.
  161. Guo, P., et al., *The landscape of alternative splicing in cervical squamous cell carcinoma*. Onco Targets Ther, 2015. **8**: p. 73-9.

## Tablas

Tabla 2 . Características clínico-patológicas de la población

Características	Número (Total = 156)	BL1 (n=23)	BL2 (n=12)	IM (n=35)	M (n=35)	MSL (n=11)	LAR (n=19)	UNS (n=21)	Valor de p <sup>#</sup>
<b>Edad (años)</b>									
≤ 50	66	12	3	15	15	5	6	10	0.0835 <sup>&amp;</sup>
>50	79	8	9	18	18	6	10	10	
NA	11	3	0	2	2	0	3	1	
<b>Grado de tumor</b>									
≤ II	26	1	1	3	7	1	8	5	0.0148 <sup>\$</sup>
> II	110	19	10	26	24	9	8	14	
NA	20	3	1	6	4	1	3	2	
<b>Tamaño de tumor (cm)</b>									
≤ 2cm	64	6	6	12	15	7	9	9	0.398 <sup>\$</sup>
> 2cm	74	14	5	19	15	3	7	11	
NA	18	3	1	4	5	1	3	1	
<b>Tipo histológico</b>									
CDI	110	16	6	26	27	8	13	14	0.400 <sup>\$</sup>
Otro	30	3	5	6	5	2	3	6	
NA	16	4	1	3	3	1	3	1	
<b>Seguimiento, meses (mediana)</b>									
	55.8	50.5	56.6	55	48	65.8	59.8	62.1	0.5620 <sup>@</sup>
IC	50.7-60.9	34.8-66.2	37.3-75.8	43.9-66.1	37.6-58.4	41.6-90.1	42-77.6	49.2-75.1	
<b>Dead event</b>									
Yes	30	8	4	9	5	1	1	2	0.0663 <sup>\$</sup>
No	112	12	7	23	27	10	15	18	
Unknown	14	3	1	3	3	0	3	1	

BL1: Basal-like1, BL2: Basal-like 2, IM: Inmunomodulador, MSL: Troncales de tipo mesénquima, M: Parecidos a mesénquima, LAR: receptor de andrógenos luminal, UNS: no determinados. CDI: carcinoma ductal infiltrante. IC: intervalo de confianza 95%. NA: información no disponible. <sup>#</sup>: Los datos NA no fueron incluidos para determinar la significancia estadística. <sup>&</sup>: prueba de ANOVA, <sup>\$</sup>: Chi- cuadrado. <sup>@</sup>: prueba Kruskal-Wallis.

Tabla 3. Lista de transcritos de lncRNA y mRNA co-expresados positivamente en el subtipo inmunomodulador.

Genes codificantes	lncRNAs	rho (Spearman)	Valor de p
TRANK1	TCONS_I2_00019221-XLOC_I2_010151_	0.9	< 2.2e-16
GBP5	RP4-620F22.2	0.9	< 2.2e-16
TAP1	PSMB8-AS1	0.9	< 2.2e-16
WARS	RP11-638I2.8	0.9	< 2.2e-16
PTPRC	RP11-553K8.5	0.9	< 2.2e-16
KLHDC7B	ENSG00000226738 (CTA-384D8.31)	0.9	< 2.2e-16
OAS1	RP1-71H24.1	0.8	< 2.2e-16
SP110	AC009950.2	0.8	< 2.2e-16
ARHGAP15	AC092652.1	0.8	< 2.2e-16
WARS	RP11-638I2.9	0.8	< 2.2e-16
ICAM1	CTD-2369P2.8	0.8	< 2.2e-16
NCKAP1L	RP11-1049A21.2	0.8	< 2.2e-16
OAS3	RP1-71H24.1	0.8	< 2.2e-16
IKZF3	TCONS_00025171	0.8	< 2.2e-16
OAS2	RP1-71H24.1	0.8	< 2.2e-16
CD48	RP11-404F10.2	0.8	< 2.2e-16
CELF2	CELF2-AS1	0.8	< 2.2e-16
ANKRD44	ANKRD44-IT1	0.7	< 2.2e-16
KLHDC7B	CTA-384D8.34	0.7	< 2.2e-16
CFLAR	CFLAR-AS1	0.7	< 2.2e-16
TRANK1	TCONS_I2_00019220-XLOC_I2_010150_	0.7	< 2.2e-16
PCED1B	PCED1B-AS1	0.7	< 2.2e-16
SEL1L3	RP11-302F12.10	0.7	< 2.2e-16
PLEKHA2	CTD-2544N14.3	0.7	< 2.2e-16
ELMO1	ELMO1-AS1	0.7	< 2.2e-16
KCNA3	RP11-284N8.3	0.7	< 2.2e-16
TNFAIP3	RP11-10J5.1	0.7	0.03591851
GPR65	RP11-300J18.2	0.7	< 2.2e-16
RFX5	RP11-126K1.8	0.6	< 2.2e-16
IFNG	IFNG-AS1	0.6	< 2.2e-16
CD38	TCONS_00008443-XLOC_003894_	0.6	< 2.2e-16
RASGRP1	RP11-102L12.2	0.6	< 2.2e-16
CELF2	RP3-323N1.2	0.6	< 2.2e-16
DEXI	RP11-876N24.5	0.5	< 2.2e-16
SMARCA2	RP11-264I13.2	0.4	6.71E-08
PTPRJ	PTPRJ-AS1	0.4	9.70E-08
IDO1	RP11-44K6.4	0.4	7.26E-06

---

---

## Anexos

### Anexos 1. Protocolo de clonación y transformación para shRNAs (BLOCK-iT).

1. Generar “ds Oligo” (50µM).

	Vol.
Oligo Sentido (a 200 µM)	5 µL
Oligo Antisentido (a 200 µM)	5 µL
10x Oligo Annealing Buffer	2 µL
DNase/RNase water-free	8 µL
Vol.total	20 µL

#### Hibridación Y dilución de oligos

2. Hibridar los ds Oligos a 94°C por 4 minutos.
3. Dejar a temperar la reacción de 5-10 min (a temp. Ambiente).
4. Centrifugar por 5 segundos (spin) y mezclar cuidadosamente .
5. Diluir los oligos 10,000 mil veces a través de dos diluciones seriadas como se indica a continuación:
  - a) Tomar 1 µL (50µM) del mix de ds Oligo previamente hibridados y diluirlos en 99 µL de agua libre de nucleasas (suplido por el Kit).
  - b) Tomar 1 µL (500nM) del oligo del paso “a” y mezclarlo con 10µL de “10x Oligo Annealing Buffer” más 89 µL de agua libre de nucleasas (suplido por el Kit). El oligo ahora está a 5nM.
  - c) Hacer alícuotas del oligo ya diluido (3 de preferencia) y almacenar a -20°C. (estas alícuotas no lamacenar por mucho tiempo).

#### Ligación

6. Hacer el siguiente mix para la ligación.

	Vol.
5xLigation Buffer	4 µL
pENTR/U6(0.5ngµL)	2 µL
Ds Oligo (5nM; diluido 1:10000)	1 µL
DNase/RNase water-free	12 µL
T4 DNA ligase	1 µL
	20 µL

7. Mezclar e incubar por 5 mint. A temp. Ambiente
8. Colocar la reacción en hielo y proceder hacer la transformación en E. Coli (ONE SHOT TOP10 suplido por el Kit).

#### Transformación

9. Centrifugar la reacción de ligación brevemente y colocarlo en hielo.
10. Descongelar en hielo un vial de células de E. coli ONE SHOT® (50µL).
11. Colocar de 1-5 µL (se recomiendo 2µL) de cada ligación dentro de un vial de células competentes y mezclar suavemente con ligeros golpes al tubo. **NOTA: NO mezclar por pipeteo.** Almacenar los remanente de ligación nuevamente a -20°C.
12. Incubar las células en hielo de 5 a 30 minutos ( *recomiendo 10 minutos*).
13. Dar un choque térmico a las células por 30 segundos a 42 °C, sin mover la reacción.
14. Pasar el tubo de 42 °C a hielo (mantenerlo por 2 minutos) .

**Antes de la transformación asegurarse:**

- Baño Maria estabilizado a 42°C.
- Incubadora estabilizada a 37°C con agitador.
- Placas de medio LB
-

15. Agregar 250  $\mu\text{L}$  de medio S.O.C (atemperado, suplido por el kit).
16. Incubar a 37°C por 1 hora con agitación (200 rpm).
17. Sembrar de 20-100  $\mu\text{L}$  (se recomienda 50 $\mu\text{L}$ ) del cultivo bacteriano en placas de agar LB que contenga 50  $\mu\text{g}/\text{mL}$  de kanamicina. Nota: guardar a -80 el resto de las células competentes.
18. Incubar toda la noche a 37°C.

### Análisis de células transformadas

#### Por secuenciación

- a) Picar de 5-10 colonias resistentes a Kanamicina y cultivarlas toda la noche en medio LB (o medio SOB) con 50 $\mu\text{g}/\text{mL}$  de kanamicina.
- b) Aislar el DNA plasmídico (usar MidiPret.)
- c) Secuenciar el pENTR/U6 de cada construcción. Usar los oligos U6 Forward y el M13 reverse. (recomiendo secuenciar con el M13 para un mejor resultado de la lectura).
- d) Confirmar: la secuencia y correcta orientación de la secuencia (oligo ds).

#### PCR de colonia

- a) Preparar la reacción de PCR-TaqPlatinum.

94°C x 5 min  
 94°C x 30 seg  
 48°C x 30 seg     x 30  
 72°C x 30seg  
 72°C x 4 min

	Vol. Por 1rx	( ) Rx
<i>Buffer 10X</i>	2 $\mu\text{L}$	
<i>MgCl2 50mM</i>	0.6 $\mu\text{L}$	
<i>dNTPs 10mM</i>	0.4 $\mu\text{L}$	
<i>U6 Forward (100ng/<math>\mu\text{L}</math>)</i>	0.4 $\mu\text{L}$	
<i>M13 Reverse (100ng/<math>\mu\text{L}</math>)</i>	0.4 $\mu\text{L}$	
<i>TaqPlatinum</i>	0.1 $\mu\text{L}$	
<i>H<sub>2</sub>O</i>	16.1 $\mu\text{L}$	
<b>Vol.Total</b>	<b>20 <math>\mu\text{L}</math></b>	

- b) Correr 10  $\mu\text{L}$  de producto de PCR en gel de agarosa al 2%. Si se observa un amplicón de ~261 pb. Probablemente está el inserto de shRNA. Proceder a propagar la colonia y extraer el DNA plasmídico para su secuenciación capilar (Sanger).

#### Propagación y extracción de DNA plasmídico.

1. Tomar la colonia bacteriana de interés (positivas a PCR de colonia) y sembrar en 10 mL de **medio líquido LB** con kanamicina (50  $\mu\text{g}/\text{mL}$ ).
2. Incubar a 37°C de 14 -18 hrs en agitación a ~200 rpm (incubadora del cuarto de las líneas celulares). **Notas: no cultivar más de 20 hrs las bacterias.**



---

---

### Extracción MidiPret

3. Una vez incubado el tubo toda la noche, **centrifugar por 35 segundos a 12,000 rpm.**
4. **Lavar** el Pellet de bacterias **con agua libre de nucleasas.**
5. Resuspender el botón de bacterias en **600µL en agua libre de nucleasas.**
6. Agregar **150µL de 7X Lysis Buffer** (Azul) y dar **vortex por 1 minuto.** Esperar 2 minutos.  
(Nota: Después de agregar el buffer 7X lysis B deberá cambiar la tonalidad de un azul opaca a mas claro, indicando que la lisis de completó).
7. Agregar 400µL de **Neutralization Buffer (Amarillo) frío** y mezclar por **vortex por 2 minutos.** La muestra cambiará de amarillo a un color amarillento opaco con poco precipitado. Invertir la muestra 2-3 veces más para asegurar una neutralización completa.
8. Centrifugar de 13,000 rpm por 7 minutos.
9. Colocar la **columna Zymo-SpinIIN** en el tubo colector.
10. Transferir el sobrenadante (~900 µL) en la columna Zymo-Spin IIN y centrifugar por 15 segundos. Transferir el sobrenadante restante a la columna. **NOTA: no llevarse detritus del pellet.**
11. Decantar el fluido y agregar **200µL de Endo-Wash Buffer** a la columna. Centrifugar por 30 segundos.
12. Decantar y agregar **500µL de ZypPy Wash Buffer** a la columna. Centrifugar 1 minuto.
13. Secar la columna a 13,000 rpm por 2 minutos.
14. Transferir la columna a un tubo nuevo de 1.5mL. Agregar **30 µL de Agua libre de nucleasas** directamente a la columna e incubar durante 1 minuto a temperatura ambiente.  
**NOTA: NO USAR ZYPY ELUTION BUFFER para secuenciación.**
15. Centrifugar a 13,000rpm por 30 segundos para eluir el DNA.
16. Guardar a -20.

### Preparación de medio sólido LB (Lennox L agar) con 50µg/mL de kanamicina.

Para 1 litro:

- a) Disolver 32 gramos de medio LB agar en 1 litro de agua destilada.
- b) Esterilizar a calor húmedo (autoclave) a 121<sup>o</sup>C por 15 minutos.
- c) Dejar enfriar el medio esterilizado a 60<sup>o</sup>C y agregar 1 ml del antibiótico (concentración stock de 50mg/mL), para obtener una concentración final a  
a) 50µg/mL de kanamicina). **NOTA: usar mechero para evitar contaminación del medio estéril.**
- d) Vaciar ~15 mL del medio (antes que se solidifique) en placas de 10 cm de diámetro.
- e) Dejar solidificar el medio y hacer prueba de esterilidad dejando las placas a 37<sup>o</sup>C toda la noche.

---

---

f) Al día siguiente, evaluar y desechar las placas que tengan crecimiento bacteriano (USAR MECHERO).

**Preparación de medio sólido LB (Lennox L agar) con 100µg/mL de ampicilina.**

Para 1 litro:

- Disolver 32 gramos de medio LB agar en 1 litro de agua destilada.
- Esterilizar a calor húmedo (autoclave) a 121<sup>o</sup>C por 15 minutos.
- Dejar enfriar el medio esterilizado a 60<sup>o</sup>C y agregar 400 µL del antibiótico (concentración stock de 250mg/mL) de ampicilina, para obtener una concentración final a 100µg/mL). **NOTA: usar mechero para evitar contaminación del medio estéril.**
- Vaciar ~15 mL del medio (antes que se solidifique) en placas de 10 cm de diámetro.
- Dejar solidificar el medio y hacer prueba de esterilidad dejando las placas a 37<sup>o</sup>C toda la noche.
- Al día siguiente, evaluar y desechar las placas que tengan crecimiento bacteriano (USAR MECHERO).

**Preparación de medio líquido LB (Lennox L Broth base) 50µg/mL de kanamicina.**

Para 1 litro:

- Disolver 20 gramos de medio LB en 1 litro de agua destilada.
- Esterilizar a calor húmedo (autoclave) a 121<sup>o</sup>C por 15 minutos.
- Dejar enfriar el medio esterilizado a 60<sup>o</sup>C y agregar 1 ml del antibiótico (concentración stock de 50mg/mL), para obtener una concentración final a 50µg/mL). **NOTA: usar mechero para evitar contaminación del medio estéril.**

$$C_1V_1=C_2V_2$$

V1= Volumen necesario del antibiótico stock para obtener la C2 final.  
C1= Concentración stock del antibiótico (por ejemplo: 50mg/mL = 50,000 µg/mL).  
V2= Volumen total de medio a preparar( por ejemplo: 500 mL).  
C2= Concentración final del antibiótico (ejemplo: 50µg/mL).

Ejemplo:

$$V1 = \frac{C_2V_2}{C_1} = \frac{5000 \text{ mL} \cdot 50 \mu\text{g/mL}}{50000 \mu\text{g/mL}} = 1 \text{ mL}$$

∴ agregar 1 mL de antibiótico de concentración inicial de 50 mg/mL a 1Litro de medio LB para una concentración final de 50µg/mL de antibiótico.

---

---

## Anexo 2. Secuencias de shRNAs usadas en este estudio.

Nombre	Secuencia	Blanco
shRNA-1	5'-CACCGCCTCAGCCCAAGTCTTAACTCGAAAGTTAAGACTTGGGCTGAGGC-3'	lncKLHDC7B
shRNA-2	5'-CACCGCCCAAGTCTTAACTTCAGCTCGAAAGCTGAAGTTAAGACTTGGGC-3'	lncKLHDC7B
NC	5'-CACCGGAATTACGGAGTCTTCTTCGCGAACGAAGAAGACTCCGTAATTCC-3'	GPF

## Artículos publicados

1. **Expression of long non-coding RNA ENSG00000226738 (LncKLHDC7B) is enriched in the immunomodulatory triple-negative breast cancer subtype and its alteration promotes cell migration, invasion, and resistance to cell death.** Mol Oncol. 2019 Apr;13(4):909-927. doi: 10.1002/1878-0261.12446. Epub 2019 Feb 9. PMID: 30648789.
2. **Growth inhibition and transcriptional effects of ribavirin in lymphoma.** Oncol Rep. 2019 Sep;42(3):1248-1256. doi: 10.3892/or.2019.7240. Epub 2019 Jul 17
3. **New insights into radioresistance in breast cancer identify a dual function of miR-122 as a tumor suppressor and oncomiR.** Mol Oncol. 2019 May;13(5):1249-1267. doi: 10.1002/1878-0261.12483. Epub 2019 Apr 18. PMID: 30938061.
4. **Population and breast cancer patients' analysis reveals the diversity of genomic variation of the BRCA genes in the Mexican population.** Hum Genomics. 2019 Jan 10;13(1):3. doi: 10.1186/s40246-018-0188-9. PMID: 30630528.
5. **Functional Integrity and Gene Expression Profiles of Human Cord Blood-Derived Hematopoietic Stem and Progenitor Cells Generated In Vitro.** Stem Cells Transl Med. 2018 Aug;7(8):602-614. doi: 10.1002/sctm.18-0013. Epub 2018 Apr 26. PMID: 29701016.
6. **Insights into the Regulatory Role of Non-coding RNAs in Cancer Metabolism.** Front Physiol. 2016 Aug 8;7:342. doi: 10.3389/fphys.2016.00342. eCollection 2016. Review. PMID: 27551267.
7. **Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells.** Virology. 2016 Jan 15;488:187-95. doi: 10.1016/j.virol.2015.11.017. Epub 2015 Dec 2. PMID: 26655236.

# Expression of long non-coding RNA ENSG00000226738 (*LncKLHDC7B*) is enriched in the immunomodulatory triple-negative breast cancer subtype and its alteration promotes cell migration, invasion, and resistance to cell death

Fredy Omar Beltrán-Anaya<sup>1,2</sup>, Sandra Romero-Córdoba<sup>1,3</sup>, Rosa Rebollar-Vega<sup>1</sup>, Oscar Arrieta<sup>4</sup>, Verónica Bautista-Piña<sup>5</sup>, Carlos Dominguez-Reyes<sup>5</sup>, Felipe Villegas-Carlos<sup>5</sup>, Alberto Tenorio-Torres<sup>5</sup>, Luis Alfaro-Riuz<sup>1</sup>, Silvia Jiménez-Morales<sup>1</sup>, Alberto Cedro-Tanda<sup>1,2</sup>, Magdalena Ríos-Romero<sup>1</sup>, Juan Pablo Reyes-Grajeda<sup>6</sup>, Elda Tagliabue<sup>3</sup>, Marilena V. Iorio<sup>3</sup> and Alfredo Hidalgo-Miranda<sup>1</sup>

1 Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico

2 Programa de Doctorado en Ciencias Biomédicas, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico

3 Department of Experimental Oncology and Molecular Medicine, Istituto Nazionale dei Tumori, Milan, Italy

4 Thoracic Oncology Unit, Instituto Nacional de Cancerología (INCan), Mexico City, Mexico

5 Instituto de Enfermedades de la Mama FUCAM, Mexico City, Mexico

6 Unidad de Proteómica Médica, Instituto Nacional de Medicina Genómica, Mexico City, Mexico

## Keywords

*ENSG00000226738*; invasion; *LncKLHDC7B*; long non-coding RNA; migration; triple-negative breast cancer

## Correspondence

A. Hidalgo-Miranda, Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, CDMX, Mexico

Tel: +52 55 53501143

E-mail: ahidalgo@inmegen.gob.mx

(Received 19 June 2018, revised 29 November 2018, accepted 30 December 2018, available online 9 February 2019)

doi:10.1002/1878-0261.12446

## Abbreviations

ADAM9, ADAM metallopeptidase domain 9; AMOT, angiomin; APP, amyloid beta precursor protein; BL1, basa-like 1; BL2, basa-like 2; CAV1, caveolin 1; CCAR2, cell cycle and apoptosis regulator 2; CDH3, cadherin 3; CI, confidence interval; ECM, extracellular matrix; EPS8, epidermal growth factor receptor pathway substrate 8; ER, estrogen receptor; FDR, false discovery rate; FFPE, formalin-fixed paraffin-embedded; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GBA, guilt-by-association analysis; GEO, gene expression omnibus; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; HTA, human transcriptome array; IM, immunomodulatory; IPA, ingenuity pathway analysis; KLHDC7B, Kelch domain containing 7B; KLHL39, Kelch-like protein 39; LAR, luminal androgen receptor; *LncKLHDC7*, long non-coding *KLHDC7B*; *lncMAP*, *LncRNA* modulator atlas in pan-cancer; *lncRNA*, long non-coding RNA; MHC, major histocompatibility complex; M, mesenchymal; MMP2, matrix metallopeptidase 2; MSL, mesenchymal stem-like; PMAIP1, phorbol-12-myristate-13-acetate-induced protein 1; PR, progesterone receptor; QBS, 2-amino-N-quinolin-8-yl-benzenesulfonamide; RAD21, RAD21 cohesin complex component; shRNA, short hairpin RNA; TCGA, The Cancer Genome Atlas; TGF, transforming growth factor; TN, triple-negative; TNBC, triple-negative breast cancer; TNC, tenascin C; TNF, tumor necrosis factor; U87, small nucleolar RNA U87; UNS, unstable.



## Growth inhibition and transcriptional effects of ribavirin in lymphoma.

Dominguez-Gomez G<sup>1</sup>, Cortez-Pedroza D<sup>1</sup>, Chavez-Blanco A<sup>1</sup>, Taja-Chayeb L<sup>1</sup>, Hidalgo-Miranda A<sup>2</sup>, Cedro-Tanda A<sup>2</sup>, Beltran-Anaya F<sup>2</sup>, Diaz-Chavez J<sup>1</sup>, Schcolnik-Cabrera A<sup>1</sup>, Gonzalez-Fierro A<sup>1</sup>, Dueñas-Gonzalez A<sup>3</sup>.

### Author information

#### Abstract

Ribavirin exhibits inhibitory effects on the epigenetic enzyme enhancer of zeste homolog 2 (EZH2), which participates in lymphomagenesis. Additionally, preclinical and clinical studies have demonstrated the anti-lymphoma activity of this drug. To further investigate the potential of ribavirin as an anticancer treatment for lymphoma, the tumor-suppressive effects of ribavirin were analyzed in lymphoma cell lines. The effects of ribavirin on the viability and clonogenicity of the B-cell lymphoma cell line Pfeiffer (EZH2-mutant), Toledo (EZH2 wild-type) and cutaneous T-cell lymphoma Hut78 cell line were assessed. Expression of EZH2 and trimethylation status of histone 3, lysine 27 trimethylated (H3K27m3) was also determined in response to ribavirin. The transcriptional effects of ribavirin on Hut78 cells were analyzed by microarray expression and the results were validated by reverse transcription-quantitative polymerase chain reaction, western blotting and knockout of signal transducer and activator of transcription 1 (STAT1). The results of the present study demonstrated that ribavirin suppressed the growth and clonogenicity of cells in a dose-dependent manner. Ribavirin did not affect the expression of EZH2 nor altered its activity as evaluated by H3K27 trimethylation status. Furthermore, the results of transcriptome analysis indicated that the majority of the canonical pathways affected by ribavirin were associated with the immune system, including 'antigen presentation', 'communication between innate and adaptive immune cells' and 'cross-talk between dendritic and natural killer cells'. The results of gene expression analysis were confirmed, by demonstrating at the RNA and protein levels, downregulation of stearoyl-CoA desaturase and upregulation of STAT1. Depletion of STAT1, which was proposed as a key regulator of the aforementioned pathways, exerted growth inhibitory effects almost to the same extent as ribavirin. In conclusion, ribavirin was proposed to exert growth inhibitory effects on lymphoma cell lines, particularly Hut78 cells, a cutaneous T-cell lymphoma cell line. Of note, these effects may depend on, at least in part, the activation of canonical immune pathways regulated by the key factors STAT1 and interferon- $\gamma$ . Our results provide insight into the anti-lymphoma potential of ribavirin; however, further investigations in preclinical and clinical studies are required to determine the effectiveness of ribavirin as a therapeutic agent for treating lymphoma.

PMID: 31322273 DOI: [10.3892/or.2019.7240](https://doi.org/10.3892/or.2019.7240)

## New insights into radioresistance in breast cancer identify a dual function of miR-122 as a tumor suppressor and oncomiR

Isidro X. Perez-Añorve<sup>1,2</sup>, Claudia H. Gonzalez-De la Rosa<sup>2</sup>, Ernesto Soto-Reyes<sup>2</sup>, Fredy O. Beltran-Anaya<sup>3</sup>, Oscar Del Moral-Hernandez<sup>4</sup>, Marisol Salgado-Albarran<sup>2</sup>, Oscar Angeles-Zaragoza<sup>5</sup>, Juan A. Gonzalez-Barrios<sup>6</sup>, Daniel A. Landero-Huerta<sup>1,2,7</sup>, Margarita Chavez-Saldaña<sup>7</sup>, Alejandro Garcia-Carranca<sup>8</sup>, Nicolas Villegas-Sepulveda<sup>9</sup> and Elena Arechaga-Ocampo<sup>2</sup>

1 Posgrado en Ciencias Naturales e Ingeniería, División de Ciencias Naturales e Ingeniería, Universidad Autónoma Metropolitana, Mexico City, Mexico

2 Departamento de Ciencias Naturales, Universidad Autónoma Metropolitana, Unidad Cuajimalpa, Mexico City, Mexico

3 Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico

4 Laboratorio de Virología y Epigenética del Cáncer, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, Mexico

5 Unidad de Radioterapia, Instituto Nacional de Cancerología, Mexico City, Mexico

6 Laboratorio de Medicina Genómica, Hospital Regional "1° de Octubre", Mexico City, Mexico

7 Laboratorio de Biología de la Reproducción, Instituto Nacional de Pediatría Mexico City, Mexico

8 Unidad de Investigación Biomedica en Cáncer-Laboratorio de Virus y Cáncer, Instituto Nacional de Cancerología, Mexico City, Mexico

9 Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados (CINVESTAV), Mexico City, Mexico

### Keywords

breast cancer; microRNAs; miR-122; radioresistance

### Correspondence

E. Arechaga-Ocampo, Departamento de Ciencias Naturales, Universidad Autónoma Metropolitana, Unidad Cuajimalpa, Av. Vasco de Quiroga 4871, Col. Santa Fe Cuajimalpa de Morelos, D.F. C.P. 05348, Mexico City, Mexico  
Tel: +52 5558146500  
E-mails: earechaga@correo.cua.uam.mx; e.arechaga@dcniuamc.com

(Received 31 October 2018, revised 10 February 2019, accepted 13 March 2019, available online 18 April 2019)

doi:10.1002/1878-0261.12483

Radioresistance of tumor cells gives rise to local recurrence and disease progression in many patients. MicroRNAs (miRNAs) are master regulators of gene expression that control oncogenic pathways to modulate the radiotherapy response of cells. In the present study, differential expression profiling assays identified 16 deregulated miRNAs in acquired radioresistant breast cancer cells, of which miR-122 was observed to be up-regulated. Functional analysis revealed that miR-122 has a role as a tumor suppressor in parental cells by decreasing survival and promoting radiosensitivity. However, in radioresistant cells, miR-122 functions as an oncomiR by promoting survival. The transcriptomic landscape resulting from knockdown of miR-122 in radioresistant cells showed modulation of the *ZNF611*, *ZNF304*, *RIPK1*, *HRAS*, *DUSP8* and *TNFRSF21* genes. Moreover, miR-122 and the set of affected genes were prognostic factors in breast cancer patients treated with radiotherapy. Our data indicate that up-regulation of miR-122 promotes cell survival in acquired radioresistant breast cancer and also suggest that miR-122 differentially controls the response to radiotherapy by a dual function as a tumor suppressor and an oncomiR dependent on cell phenotype.

### Abbreviations

GO, Gene Ontology; miRNAs, microRNAs; oncomiRs, oncogenic miRNAs; qRT-PCR, quantitative reverse transcriptase-polymerase chain reaction; RFS, relapse-free survival; SF, surviving fraction; TCGA, The Cancer Genome Atlas; TLDA, TaqMan low-density arrays; TNBC, triple-negative breast cancer.




PRIMARY RESEARCH

Open Access



# Population and breast cancer patients' analysis reveals the diversity of genomic variation of the BRCA genes in the Mexican population

J. C. Fernández-Lopez<sup>1</sup>, S. Romero-Córdoba<sup>1</sup>, R. Rebollar-Vega<sup>1</sup>, L. A. Alfaro-Ruiz<sup>1</sup>, S. Jiménez-Morales<sup>1</sup>, F. Beltrán-Anaya<sup>1</sup>, R. Arellano-Llamas<sup>1</sup>, A. Cedro-Tanda<sup>1</sup>, M. Rios-Romero<sup>1</sup>, M. Ramirez-Florencio<sup>1</sup>, V. Bautista-Piña<sup>2</sup>, C. Dominguez-Reyes<sup>2</sup>, F. Villegas-Carlos<sup>2</sup>, A. Tenorio-Torres<sup>2</sup> and A. Hidalgo-Miranda<sup>1\*</sup> 

## Abstract

Interpretation of variants of unknown significance (VUS) in genetic tests is complicated in ethnically diverse populations, given the lack of information regarding the common spectrum of genetic variation in clinically relevant genes. Public availability of data obtained from high-throughput genotyping and/or exome massive parallel sequencing (MPS)-based projects from several thousands of outbred samples might become useful tools to evaluate the pathogenicity of a VUS, based on its frequency in different populations. In the case of the Mexican and other Latino populations, several thousands of samples have been genotyped or sequenced during the last few years as part of different efforts to identify common variants associated to common diseases. In this report, we analyzed Mexican population data from a sample of 3985 outbred individuals, and additional 66 hereditary breast cancer patients were analyzed in order to better define the spectrum of common genomic variation of the *BRCA1* and *BRCA2* genes. Our analyses identified the most common genetic variants in these clinically relevant genes as well as the presence and frequency of specific pathogenic mutations present in the Mexican population. Analysis of the 3985 population samples by MPS identified three pathogenic mutations in *BRCA1*, only one population sample showed a *BRCA1* exon 16–17 deletion by MLPA. This resulted in a basal prevalence of deleterious mutations of 0.10% (1:996) for *BRCA1* and 11 pathogenic mutations in *BRCA2*, resulting in a basal prevalence of deleterious mutations of 0.276% (1:362) for *BRCA2*, combined of 0.376% (1:265). Separate analysis of the breast cancer patients identified the presence of pathogenic mutations in 18% (12 pathogenic mutations in 66 patients) of the samples by MPS and 13 additional alterations by MLPA. These results will support a better interpretation of clinical studies focused on the detection of BRCA mutations in Mexican and Latino populations and will help to define the general prevalence of deleterious mutations within these populations.

**Keywords:** BRCA, Breast cancer, Populations, Germline, Genetic testing

\* Correspondence: ahidalgo@inmegenejob.mx

<sup>1</sup>Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Periférico Sur, 4809, Arenal Tepepan, 14610 Mexico City, CP, Mexico

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.





<sup>a</sup>Hematopoietic Stem Cells Laboratory, Oncology Research Unit, Oncology Hospital, <sup>c</sup>Immunochemistry Research Unit, Medical Specialties Hospital, IMSS National Medical Center, Mexico City, Mexico; <sup>b</sup>National Ministry of Health, National Institute of Genomic Medicine, Mexico City, Mexico; <sup>d</sup>Troncoso General Hospital, IMSS, Mexico City, Mexico

Correspondence: Hector Mayani, Ph.D., National Medical Center, IMSS, Tallo 2, D-102, San Pablo Tepetlapa, Coyoacan, Mexico D.F. 04620, Mexico. Telephone: 52-55-5627 6959; e-mail: hmayaniv@prodigy.net.mx

Received January 17, 2018; accepted for publication March 29, 2018; first published April 26, 2018.

<http://dx.doi.org/10.1002/sctm.18-0013>

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

This article was published online on 26 April 2018. The logo for the Cord Blood Association incorrectly appeared on the article and has now been removed because the authors are not members of the association. This notice is included in the online version to indicate this version was corrected on 7 May 2018.

## Functional Integrity and Gene Expression Profiles of Human Cord Blood-Derived Hematopoietic Stem and Progenitor Cells Generated In Vitro

ROBERTO DIRCIO-MALDONADO,<sup>a</sup> PATRICIA FLORES-GUZMAN,<sup>a</sup> JULIETA CORRAL-NAVARRO,<sup>a</sup> ILEANA MONDRAGÓN-GARCÍA,<sup>a</sup> ALFREDO HIDALGO-MIRANDA,<sup>b</sup> FREDY OMAR BELTRAN-ANAYA,<sup>b</sup> ALBERTO CEDRO-TANDA,<sup>b</sup> LOURDES ARRIAGA-PIZANO,<sup>c</sup> ODETTE BALVANERA-ORTIZ,<sup>d</sup> HECTOR MAYANI<sup>1b,a</sup>

**Key Words.** Cord blood • Ex vivo expansion • Gene expression profiles • Hematopoietic stem and progenitor cells

### ABSTRACT

To date, different experimental strategies have been developed for the ex vivo expansion of human hematopoietic stem (HSCs) and progenitor (HPCs) cells. This has resulted in significant advances on the use of such expanded cells in transplantation settings. To this day, however, it is still unclear to what extent those stem and progenitor cells generated in vitro retain the functional and genomic integrity of their freshly isolated counterparts. In trying to contribute to the solving of this issue, in the present study we have selected and purified three different hematopoietic cell populations: HSCs (CD34<sup>+</sup> CD38<sup>-</sup> CD45RA<sup>-</sup> CD71<sup>-</sup> Lin<sup>-</sup> cells), myeloid progenitor cells (CD34<sup>+</sup> CD38<sup>+</sup> CD45RA<sup>+</sup> CD71<sup>-</sup> Lin<sup>-</sup> cells), and erythroid progenitor cells (CD34<sup>+</sup> CD38<sup>+</sup> CD45RA<sup>-</sup> CD71<sup>+</sup> Lin<sup>-</sup> cells), obtained directly from fresh human umbilical cord blood (UCB) units or generated in vitro under particular culture conditions. We, then, compared their functional integrity in vitro and their gene expression profiles. Our results indicate that in spite of being immunophenotypically similar, fresh and in vitro generated cells showed significant differences, both in functional and genetic terms. As compared to their fresh counterparts, those HSCs generated in our culture system showed a deficient content of long-term culture-initiating cells, and a marked differentiation bias toward the myeloid lineage. In addition, in vitro generated HSCs and HPCs showed a limited expansion potential. Such functional alterations correlated with differences in their gene expression profiles. These observations are relevant in terms of HSC biology and may have implications in UCB expansion and transplantation. *STEM CELLS TRANSLATIONAL MEDICINE* 2018;7:602–614

### SIGNIFICANCE STATEMENT

This study presents evidence indicating that in vitro-generated hematopoietic stem and progenitor cells differ in functional and genetic terms from their freshly obtained counterparts. These findings have biologic and clinical implications.

### INTRODUCTION

Over the last three decades, umbilical cord blood (UCB) has become a major source of hematopoietic stem and progenitor cells (HSCs and HPCs, respectively) for hematopoietic transplantation [1, 2]. UCB not only contains clinically relevant amounts of primitive hematopoietic cells that possess great proliferation and expansion potentials [3–6], but such cells can be stored in public or private UCB banks [7–9]. To date, it is estimated that over 40,000 UCB transplants (UCBT) have been performed worldwide. In pediatric patients, UCB transplantation has become a first-line therapy for specific hematological conditions, with results

comparable to those obtained when using adult marrow or mobilized peripheral blood [10]. In contrast, most adult patients transplanted with UCB show prolonged periods of neutropenia and thrombocytopenia after transplant, which results in increased morbidity and mortality [11]. This seems to be due, at least in part, to the fact that the absolute numbers of HSCs and HPCs in a UCB unit are significantly lower than in a bone marrow or a mobilized peripheral blood unit [12].

In the last few years, different strategies have been explored to improve UCBT in adult patients. Among them, three approaches have already been taken into the clinic, namely, double-unit transplants, intra bone injection of





# Insights into the Regulatory Role of Non-coding RNAs in Cancer Metabolism

Fredy O. Beltrán-Anaya, Alberto Cedro-Tanda, Alfredo Hidalgo-Miranda\* and Sandra L. Romero-Cordoba\*

Cancer Genomics Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico

## OPEN ACCESS

### Edited by:

Firas H. Kobeissy,  
University of Florida, USA

### Reviewed by:

Guanglong Jiang,  
Indiana University School of Medicine,  
USA

Naniko Hiroi,  
Keio University, Japan

### \*Correspondence:

Alfredo Hidalgo-Miranda  
ahidalgo@inmegen.gob.mx  
Sandra L. Romero-Cordoba  
sromero\_cordoba@hotmail.com

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Physiology

Received: 25 May 2016

Accepted: 25 July 2016

Published: 08 August 2016

### Citation:

Beltrán-Anaya FO, Cedro-Tanda A,  
Hidalgo-Miranda A and  
Romero-Cordoba SL (2016) Insights  
into the Regulatory Role of  
Non-coding RNAs in Cancer  
Metabolism. *Front. Physiol.* 7:342.  
doi: 10.3389/fphys.2016.00342

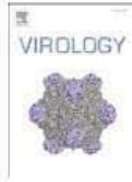
Cancer represents a complex disease originated from alterations in several genes leading to disturbances in important signaling pathways in tumor biology, favoring heterogeneity that promotes adaptability and pharmacological resistance of tumor cells. Metabolic reprogramming has emerged as an important hallmark of cancer characterized by the presence of aerobic glycolysis, increased glutaminolysis and fatty acid biosynthesis, as well as an altered mitochondrial energy production. The metabolic switches that support energetic requirements of cancer cells are closely related to either activation of oncogenes or down-modulation of tumor-suppressor genes, finally leading to dysregulation of cell proliferation, metastasis and drug resistance signals. Non-coding RNAs (ncRNAs) have emerged as one important kind of molecules that can regulate altered genes contributing, to the establishment of metabolic reprogramming. Moreover, diverse metabolic signals can regulate ncRNA expression and activity at genetic, transcriptional, or epigenetic levels. The regulatory landscape of ncRNAs may provide a new approach for understanding and treatment of different types of malignancies. In this review we discuss the regulatory role exerted by ncRNAs on metabolic enzymes and pathways involved in glucose, lipid, and amino acid metabolism. We also review how metabolic stress conditions and tumoral microenvironment influence ncRNA expression and activity. Furthermore, we comment on the therapeutic potential of metabolism-related ncRNAs in cancer.

**Keywords:** metabolic reprogramming, cancer metabolism, ncRNA regulation, miRNAs

## METABOLIC REPROGRAMMING: CANCER METABOLISM CHANGING THE ENERGETIC STATE TO FULFILL CELLULAR REQUIREMENTS

Deregulation of cellular energetics has been pointed out as one of the emerging hallmarks of cancer, both during early cellular transformation and as a driving phenotype of several tumorigenic programs (Kroemer and Pouyssegur, 2008; Munoz-Pinedo et al., 2012). Under physiological conditions, cells maintain regulated and complex metabolic homeostasis by diverse signaling pathways that function as energetic sensors. Metabolic sensors act under a network of cooperative signaling cascades, not only to fulfill the energetic requirements of the cells, but also to influence cellular pathways like cell growth, proliferation, and death (Dumortier et al., 2013). In contrast,





## Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells



Ana Elvira Zacapala-Gómez<sup>a</sup>, Oscar Del Moral-Hernández<sup>a</sup>, Nicolás Villegas-Sepúlveda<sup>b</sup>, Alfredo Hidalgo-Miranda<sup>c</sup>, Sandra Lorena Romero-Córdoba<sup>c</sup>, Fredy Omar Beltrán-Anaya<sup>c</sup>, Marco Antonio Leyva-Vázquez<sup>a</sup>, Luz del Carmen Alarcón-Romero<sup>d</sup>, Berenice Illades-Aguilar<sup>a,\*</sup>

<sup>a</sup> Laboratorio de Biomedicina Molecular, Unidad Académica de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, Gro., México

<sup>b</sup> Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), México, D.F., México

<sup>c</sup> Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica (INMEGEN), México, D.F., México

<sup>d</sup> Laboratorio de Investigación en Citopatología e Histoquímica, Unidad Académica de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, Gro., México

### ARTICLE INFO

#### Article history:

Received 25 August 2015

Returned to author for revisions

12 November 2015

Accepted 16 November 2015

Available online 2 December 2015

#### Keywords:

HPV 16 variants

HPV 16 E6 oncoprotein variants

C33-A cells

Global gene expression

Cervical carcinoma

### ABSTRACT

We analyzed the effects of the expression of HPV 16 E6 oncoprotein variants (AA-a, AA-c, E-A176/G350, E-C188/G350, E-G350), and the E-Prototype in global gene expression profiles in an in vitro model. E6 gene was cloned into an expression vector fused to GFP and was transfected in C33-A cells. Affymetrix GeneChip Human Transcriptome Array 2.0 platform was used to analyze the expression of over 245,000 coding transcripts. We found that HPV16 E6 variants altered the expression of 387 different genes in comparison with E-Prototype. The altered genes are involved in cellular processes related to the development of cervical carcinoma, such as adhesion, angiogenesis, apoptosis, differentiation, cell cycle, proliferation, transcription and protein translation. Our results show that polymorphic changes in HPV16 E6 natural variants are sufficient to alter the overall gene expression profile in C33-A cells, explaining in part the observed differences in oncogenic potential of HPV16 variants.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Introduction

It has been widely shown that chronic infection by High Risk Human Papillomavirus (HR-HPV) is the main risk factor for developing cervical carcinoma (CC) and its precursor lesions (Münger et al., 2004). Genotype 16 of the HR-HPV is the most common in cervical carcinoma worldwide (Crosbie et al., 2013; Obeidat et al., 2013), and its oncogenic potential is mainly related to the action of oncoproteins E6 and E7. These oncoproteins have the ability to bind several cellular proteins and interfere with fundamental processes which promote cervical carcinogenesis, such as apoptosis, cell

differentiation, adhesion, cell cycle, and immune response (Wise-Draper and Wells, 2008; Narisawa-Saito and Kiyono, 2007; Pim and Banks, 2010; Moody and Laimins, 2010).

Several studies have suggested that HPV 16 variants may contribute to cancer development (Xi et al., 2007). Genetic studies have shown natural amino acid variants of HPV16 that are classified according to the polymorphic changes in the E6 gene and the LCR region (Cornet et al., 2012). The reference HPV 16 genome was sequenced in 1985 (Seedorf et al., 1985) and many variants have been found since that time. HPV 16 variants have been classified into four major lineages (European-Asian, African 1, African 2 and Asian-American/North-American) and nine sublineages (Cornet et al., 2012). The presence of polymorphisms that generate amino acid changes in the E6 oncoprotein has been shown. T350G, a common polymorphism in European variants, causes a leucine to valine change (L83V); T178G, common in Asian variants, leads D25E; G145T, common in Asian American and North American variants, leads Q14H; C143G, common in African variants, leads Q14D (Huertas-Salgado et al., 2011). Polymorphisms frequency depends on the population studied (Xi et al., 2007, 2006; Cornet et al., 2012; Tornesello et al., 2011; Pillai et al., 2009).

\* Correspondence to: Av. Lázaro Cárdenas S/N, Chilpancingo, Gro, México, 39090.

E-mail addresses: [zak\\_ana@yahoo.com.mx](mailto:zak_ana@yahoo.com.mx) (A.E. Zacapala-Gómez),

[odelmoralh@gmail.com](mailto:odelmoralh@gmail.com) (O. Del Moral-Hernández),

[nvillega@cinvestav.mx](mailto:nvillega@cinvestav.mx) (N. Villegas-Sepúlveda),

[ahidalgo@inmegen.gob.mx](mailto:ahidalgo@inmegen.gob.mx) (A. Hidalgo-Miranda),

[sromero\\_cordoba@hotmail.com](mailto:sromero_cordoba@hotmail.com) (S.L. Romero-Córdoba),

[frebeltran@hotmail.com](mailto:frebeltran@hotmail.com) (F.O. Beltrán-Anaya),

[leyvamarco13@gmail.com](mailto:leyvamarco13@gmail.com) (M.A. Leyva-Vázquez),

[uzdelcarmen14@gmail.com](mailto:uzdelcarmen14@gmail.com) (L.C. Alarcón-Romero),

[b.illadesaguilar@gmail.com](mailto:b.illadesaguilar@gmail.com) (B. Illades-Aguilar).

<http://dx.doi.org/10.1016/j.viro.2015.11.017>

0042-6822/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

# Expression of long non-coding RNA *ENSG00000226738* (*LncKLHDC7B*) is enriched in the immunomodulatory triple-negative breast cancer subtype and its alteration promotes cell migration, invasion, and resistance to cell death

Fredy Omar Beltrán-Anaya<sup>1,2</sup>, Sandra Romero-Córdoba<sup>1,3</sup>, Rosa Rebollar-Vega<sup>1</sup>, Oscar Arrieta<sup>4</sup>, Verónica Bautista-Piña<sup>5</sup>, Carlos Dominguez-Reyes<sup>5</sup>, Felipe Villegas-Carlos<sup>5</sup>, Alberto Tenorio-Torres<sup>5</sup>, Luis Alfaro-Riuz<sup>1</sup>, Silvia Jiménez-Morales<sup>1</sup>, Alberto Cedro-Tanda<sup>1,2</sup>, Magdalena Ríos-Romero<sup>1</sup>, Juan Pablo Reyes-Grajeda<sup>6</sup>, Elda Tagliabue<sup>3</sup>, Marilena V. Iorio<sup>3</sup> and Alfredo Hidalgo-Miranda<sup>1</sup>

1 Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico

2 Programa de Doctorado en Ciencias Biomédicas, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico

3 Department of Experimental Oncology and Molecular Medicine, Istituto Nazionale dei Tumori, Milan, Italy

4 Thoracic Oncology Unit, Instituto Nacional de Cancerología (INCan), Mexico City, Mexico

5 Instituto de Enfermedades de la Mama FUCAM, Mexico City, Mexico

6 Unidad de Proteómica Médica, Instituto Nacional de Medicina Genómica, Mexico City, Mexico

## Keywords

*ENSG00000226738*; invasion; *LncKLHDC7B*; long non-coding RNA; migration; triple-negative breast cancer

## Correspondence

A. Hidalgo-Miranda, Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, CDMX, Mexico

Tel: +52 55 53501143

E-mail: ahidalgo@inmegen.gob.mx

(Received 19 June 2018, revised 29 November 2018, accepted 30 December 2018, available online 9 February 2019)

doi:10.1002/1878-0261.12446

Triple negative breast cancer (TNBC) represents an aggressive phenotype with poor prognosis compared with ER, PR, and HER2-positive tumors. TNBC is a heterogeneous disease, and gene expression analysis has identified seven molecular subtypes. Accumulating evidence demonstrates that long non-coding RNA (lncRNA) are involved in regulation of gene expression and cancer biology, contributing to essential cancer cell functions. In this study, we analyzed the expression profile of lncRNA in TNBC subtypes from 156 TNBC samples, and then characterized the functional role of *LncKLHDC7B* (*ENSG00000226738*). A total of 710 lncRNA were found to be differentially expressed between TNBC subtypes, and a subset of these altered lncRNA were independently validated. We discovered that *LncKLHDC7B* (*ENSG00000226738*) acts as a transcriptional modulator of its neighboring coding gene *KLHDC7B* in the immunomodulatory subtype. Furthermore, *LncKLHDC7B* knockdown enhanced migration and invasion, and promoted resistance to cellular death. Our findings confirmed the contribution of *LncKLHDC7B* to induction of apoptosis and inhibition of cell

## Abbreviations

ADAM9, ADAM metalloproteinase domain 9; AMOT, angiomin; APP, amyloid beta precursor protein; BL1, basal-like 1; BL2, basal-like 2; CAV1, caveolin 1; CCAR2, cell cycle and apoptosis regulator 2; CDH3, cadherin 3; CI, confidence interval; ECM, extracellular matrix; EPS8, epidermal growth factor receptor pathway substrate 8; ER, estrogen receptor; FDR, false discovery rate; FFPE, formalin-fixed paraffin-embedded; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GBA, guilt-by-association analysis; GEO, gene expression omnibus; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; HTA, human transcriptome array; IM, immunomodulatory; IPA, ingenuity pathway analysis; KLHDC7B, Kelch domain containing 7B; KLHL39, Kelch-like protein 39; LAR, luminal androgen receptor; *LncKLHDC7*, long non-coding *KLHDC7B*; lncMAP, lncRNA modulator atlas in pan-cancer; lncRNA, long non-coding RNA; MHC, major histocompatibility complex; M, mesenchymal; MMP2, matrix metalloproteinase 2; MSL, mesenchymal stem-like; PMAIP1, phorbol-12-myristate-13-acetate-induced protein 1; PR, progesterone receptor; QBS, 2-amino-N-quinolin-8-yl-benzenesulfonamide; RAD21, RAD21 cohesin complex component; shRNA, short hairpin RNA; TCGA, The Cancer Genome Atlas; TGF, transforming growth factor; TN, triple-negative; TNBC, triple-negative breast cancer; TNC, tenascin C; TNF, tumor necrosis factor; U87, small nucleolar RNA U87; UNS, unstable.



migration and invasion, suggesting that TNBC tumors with enrichment of *LncKLHDC7B* may exhibit distinct regulatory activity, or that this may be a generalized process in breast cancer. Additionally, *in silico* analysis confirmed for the first time that the low expression of *KLHDC7B* and *LncKLHDC7B* is associated with poor prognosis in patients with breast cancer.

## 1. Introduction

Triple negative breast cancer (TNBC), defined in the clinic by the negative expression of estrogen receptor (ER), progesterone receptor (PR), and lack of overexpression of human epidermal growth factor receptor 2 (HER2), is a heterogeneous tumor that represents ~14–23% of all breast cancers (BC) (Lara-Medina *et al.*, 2011; Martinez *et al.*, 2013; Perez-Rodriguez, 2015). TNBC currently lacks targeted treatment, is more aggressive, has worse global survival in metastatic disease, affects young women, and a higher proportion of African American and Hispanic women (Bauer *et al.*, 2007; Dent *et al.*, 2007; Foulkes *et al.*, 2010; Rastelli *et al.*, 2010). With the development and advancement of high throughput technologies, TNBC has been classified into intrinsic subtypes based on coding gene expression [messenger (m) RNA] and ontology analysis into the following groups: two basal-like (BL1 and BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL), luminal androgen receptor (LAR), and unstable (UNS) subtypes (Lehmann *et al.*, 2011). Subtyping of TNBC based on non-coding RNA (ncRNA) expression and research on its therapeutic implications are still being refined (Lehmann *et al.*, 2016).

During the development of cancer, major transcriptional alterations can be induced by ncRNA, leading to changes in global gene expression patterns and genomic instability (Wapinski and Chang, 2011). A large proportion of RNAs with no coding capacity (Consortium, 2012), such as long non-coding RNA (lncRNA) (Martin and Chang, 2012), have emerged as key players in several biological processes. LncRNA have > 200 nucleotides, transcribed by RNA polymerase II, mainly located within the nucleus but also found in the cytosolic compartment, lack protein-coding potential, and show lower expression compared with mRNA (Derrien *et al.*, 2012). The lncRNA species are also considered to be new protagonists in the development of cancer, with potential roles in oncogenic and tumor suppressor pathways (Gibb *et al.*, 2011; Li and Chen, 2013; Wapinski and Chang, 2011).

To analyze the differences in lncRNA expression landscapes in TNBC subtypes, we evaluated lncRNA expression profiles of 156 TNBC samples. TNBC

subtypes were first defined by mRNA microarray expression profiling, and differentially expressed lncRNA were identified between subtypes. A significant number of differentially expressed lncRNA were identified between all subtypes, and some of them were validated in an independent cohort of TNBC samples. *LncKLHDC7B*, an lncRNA that was over-expressed in the immunomodulatory subtype, was selected for further investigation of its functional role in TNBC through modulation of its expression in cell lines. Down-modulation of *LncKLHDC7B* resulted in down-regulation of its coding gene (*KLHDC7B*) and in the alteration of the expression of several other genes, as well as in an increase in migration and invasion, in addition to resistance to cell death not only in TNBC cell lines, but also in other breast tumor subtypes, suggesting its potential role in breast oncogenesis.

## 2. Materials and methods

### 2.1. Collection and processing of samples

A total of 156 formalin-fixed, paraffin-embedded (FFPE) tissues were collected from the Instituto de Enfermedades de la Mama FUCAM (Mexico) and the Istituto Nazionale dei Tumori (Italy) by retrospective collection from 2007 to 2015. Samples were defined as triple-negative using the current international histopathological criteria, less than 1% of positive cells for ER and PR, and HER2. Immunohistochemical detection of the three markers was carried out using the following antibodies: ER (clone 1D5, Dako, Denmark, Carpinteria, CA, USA), PR (clone PgR636, Dako, Denmark) and HER2 (K5204, Dako). Cytokeratin 5/6 and Ki67 (clone D5/16B4, Dako, Denmark) were also evaluated to rule out potential false-negative immunohistochemical reactions. All procedures were performed according to the Declaration of Helsinki, and were reviewed and approved by the Ethics and Research Committees of the National Institute of Genomic Medicine (approval number CEI2016/13).

A pathologist identified regions of the sample with the highest amount of cancer tissue, and tissue cores were obtained with a 2-mm-wide needle with a tissue arrayer. Tissue cores were used for RNA extraction

using the AllPrep® nucleic acid purification FFPE kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Subsequently, RNA was quantified by spectrophotometry (Nanodrop, Thermo Fisher Scientific, Waltham, MA, USA) and stored at  $-80^{\circ}\text{C}$  until further processing.

## 2.2. Microarray processing and data analysis

Transcriptional profiles were analyzed using the Affymetrix Human Transcriptome Array V2.0, following the manufacturer's instructions. Briefly,  $\sim 200$  ng of total RNA was converted into complementary (c)DNA and labeled with the SensationPlus™ FFPE Amplification and WT Labeling® kit (Affymetrix, Santa Clara, CA, USA) and hybridized on the array, which detects both mRNA and lncRNA. Arrays were washed, stained, and scanned using a Genechip Scanner 3000 7G (Affymetrix). The data were analyzed with the Robust Multichip Analysis (RMA) algorithm using Affymetrix default analysis settings and global scaling as normalization method on a Affymetrix Transcriptome Analysis Console (TAC) V3.0 [Data accessibility gene expression omnibus (GEO): GSE86948]. The ComBat function was used to remove batch effects using the SVA package. The boxplots, heat maps, and correlation analyses were conducted in R software (<http://cran.r-project.org>; <http://www.bioconductor.org>; ggplots [cran.r-project.org](http://cran.r-project.org)) using tools from the Bioconductor project (<http://www.bioconductor.org>) (Wickham, 2009).

## 2.3. Messenger RNA subtyping of triple-negative breast cancer samples and lncRNA expression profiles

The Web-based TNBCtype algorithm (<http://cbc.mc.vanderbilt.edu/tnbc/>) was used to identify the mRNA-based subtypes of TNBC in our cases (Chen *et al.*, 2012). Four of 160 TNBC samples were excluded from lncRNA expression profile analysis, according to the bimodal filter of the TNBCtype algorithm (Chen *et al.*, 2012). After mRNA subtyping, we analyzed the lncRNA expression profiles between subgroups using the Affymetrix Transcriptome Analysis Console (TAC) v.3.0, through comparison of the samples belonging to a specific subtype against to the rest of the samples (e.g. IM vs other subtypes). A  $\geq 1.5$ -fold change, ANOVA *P*-value less than or equal to 0.05, and false discovery rate (FDR)  $< 0.05$  were considered as significant to detect expression changes between the TNBC subtypes, except for BL2 where FDR was  $< 0.5$ . A heatmap with the differentially expressed lncRNA is shown in Fig. 1. TNBC samples from the GEO

database (Data accessibility GEO: GSE76250) were used for validation of differential gene expression in an independent cohort (Liu *et al.*, 2016). Gene ontology and cellular pathway analysis were carried out using DAVID (Frederick, MD, USA) (Huang da *et al.*, 2009). Venn diagrams were used to show the relationship between groups of differentially expressed lncRNA (Heberle *et al.*, 2015).

## 2.4. Guilt-by-association analysis

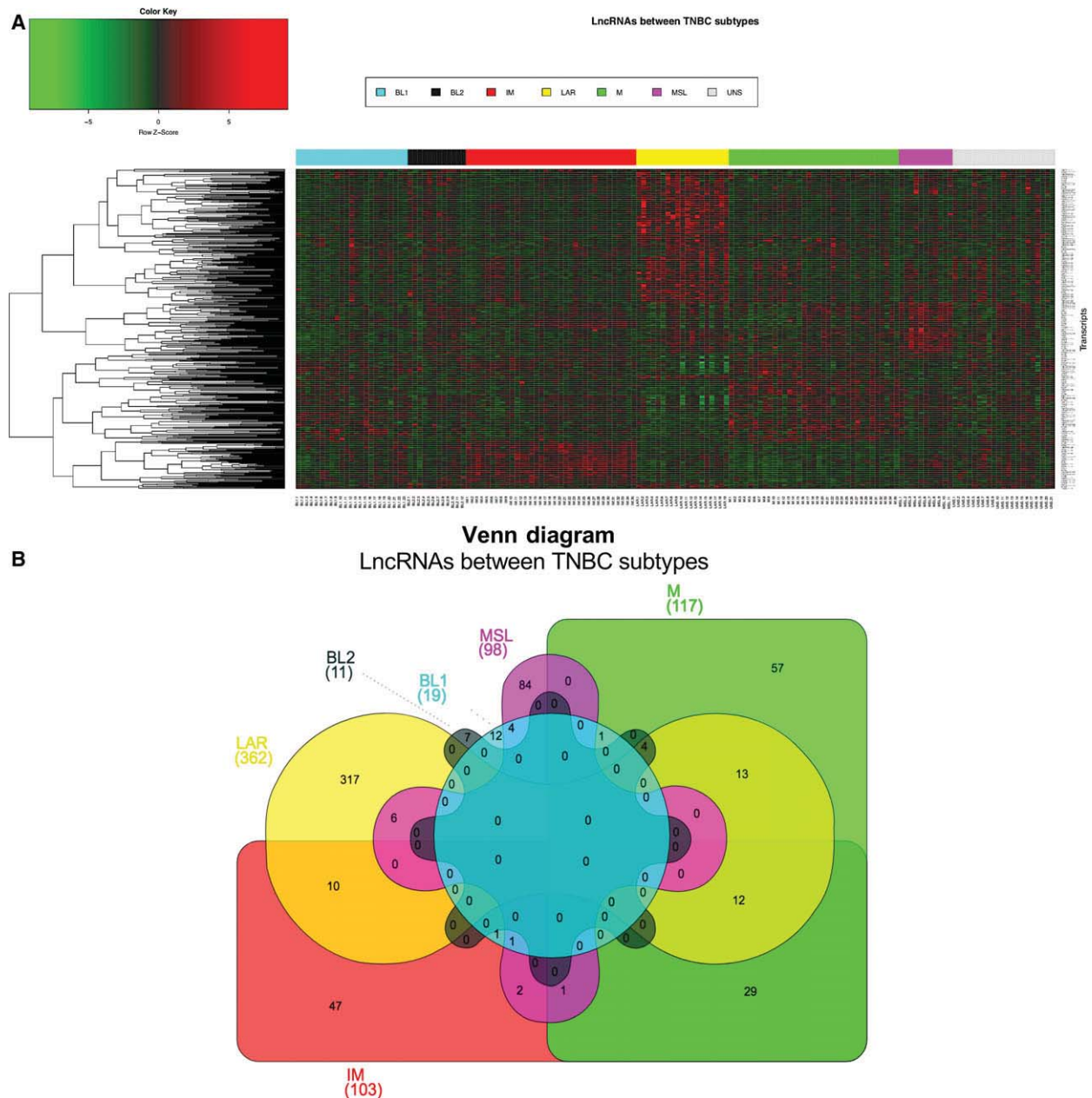
Enrichment pathway analysis was carried out by a guilt-by-association analysis (Guttman *et al.*, 2009). Briefly, normalized gene expression of mRNA and lncRNA of the 156 profiled tumors were used to build a correlation matrix including only the differentially expressed lncRNA and mRNA for each subtype. For each lncRNA, mRNA were ranked according to their Spearman correlation coefficient and *P* value. A significant correlation between lncRNA and mRNA was considered as follows: Spearman correlation  $> 30\%$  and *P* value  $< 0.05$ . The lncRNA-mRNA co-expression was computed to define candidate lncRNA prioritization by identifying which transcripts show a coordinated expression pattern across a group of samples. The significant correlated mRNA were then evaluated in an enrichment pathway analysis to define their impact on the signaling process, using ENRICH (New York, NY, USA) (Kuleshov *et al.*, 2016) to associate the altered lncRNA with pathways and signatures that were significantly enriched. The co-expression correlations were the basis for constructing an enrichment pathway analysis in which each term represents a set of transcripts with a significant strength co-expression relationship. To compute this weighted correlation analysis, some functions of the WGCNA package were applied (Los Angeles, CA, USA) (Langfelder and Horvath, 2008).

## 2.5. Immunoscore prediction

Gene expression profile of four immune-related classes (effector cells, immunosuppressive cells, MHC molecules, and selected immunomodulators) were evaluated to determine the Immunophenoscore\* using the available r-script deposited on GitHub (<https://github.com/mui-icbi/Immunophenogram>).

## 2.6. Analysis of lncKLHDC7B expression in breast cancer cell lines

MCF10A, BT20, MDA-MB-468, MDA-MB-231, Hs578-T, MCF-7, and HCC1187 cell lines were grown under standard conditions (DMEM or RPMI-1640



**Fig. 1.** Differential expression of lncRNA across TNBC subtypes. (A) Supervised hierarchical clustering of the differentially expressed lncRNA in TNBC subtypes. Red represents up-regulation and green down-regulation. Rows correspond to transcripts and columns to samples, color bar represents each TNBC subtype. (B) Venn diagram shows the common and differential lncRNA between TNBC subtypes.

medium, ATCC®, Manassas, VA, USA) according to each line, supplemented with 10% FBS (ATCC®) and incubated at ~ 37 °C with 5% CO<sub>2</sub>. Total RNA was isolated using the Trizol reagent (Invitrogen, Carlsbad, CA, USA) according the manufacturer's recommendation. Subsequently, RNA was quantified by spectrophotometry (Nanodrop) and stored at -80 °C until further processing.

## 2.7. Reverse transcription polymerase chain reaction

Complementary DNA was synthesized using SuperScript III RT-PCR (Invitrogen) following the manufacturer's recommendations. Briefly, 100 ng of total RNA was used to synthesize cDNA in a final reaction volume of 20 µL. The real time PCR mixture contained 1 µL of cDNA,



5  $\mu$ L 2 $\times$  TaqMan Universal Master Mix (Applied Biosystems), 0.5  $\mu$ L TaqMan® probes (Hs00536653\_s1, custom from tcon\_00029630), and 3.5  $\mu$ L of nuclease-free water. *GAPDH* (Hs99999905\_m1) and *U87* (Hs03298717\_s1) were used as endogenous controls for coding and non-coding genes, respectively. The fold change for each gene (*KLHDC7B* or *LncKLHDC7B*) in treated cells relative to control cells was calculated using the Ct ( $2^{-\Delta\Delta C_t}$ ) method, where  $\Delta\Delta C_t = \Delta C_t$  treated cell (shRNA-1 or -2) –  $\Delta C_t$  control cell (NC);  $\Delta C_t = C_t$  coding gene –  $C_t$  *GAPDH* or  $\Delta C_t = C_t$  non-coding gene –  $C_t$  *U87*.

## 2.8. Long non-coding RNA down-regulation with shRNA

Short hairpin RNA were generated using the BLOCK-iT™ U6 RNAi Entry Vector Kit (Invitrogen) following the manufacturer's instructions. Briefly, pairs of cDNA oligos were designed containing four nucleotide overhangs necessary for directional cloning. We generated the double-stranded oligo (ds-oligo) and subsequently performed the ligation reaction of ds-oligo into the pENTR™/U6 vector, which was used to transform competent *Escherichia coli* One Shot® TOP10 cells. Sanger sequencing was used to corroborate the presence and correct orientations of the ds-oligo insert. Two shRNA were used to silence *LncKLHDC7B* (shRNA-1, 5'-CACCGCCTCAGCCCAAGTCTTAACTCGAAAGTTAAGACTTGGGCTGAGGC-3' and shRNA-2, 5'-CACCGCCCAAGTCTTAACTTCAGCTCGAAAGCTGAAGTTAAGACTTGGGC-3') and negative control (NC, 5'-CACCGGAATTACGGAGTCTTCTTCGCGAACGAAGAAGACTCCGTAATTC-3'), with a sequence that does not target any mRNA in the human genome *GFP*.

Silencing of *LncKLHDC7B* was performed in three cell lines: HCC1187 (TNBC), MCF-7 (Luminal A) and BT-20 (TNBC) using two shRNA (shRNA-1 and shRNA-2), which were transfected using Xfect™ Transfection Reagent Protocol (Clontech, Palo Alto, CA, USA). Briefly, HCC1187 cells ( $6.5 \times 10^5$  cells per well), BT-20 ( $5 \times 10^5$  cells per well) and MCF-7 ( $5 \times 10^5$  cells per well) were seeded in 6-well plates and 3  $\mu$ g of plasmid was transfected during 24 h. Nanoparticle complexes were removed and replaced with complete growth medium. After 48 h post-transfection, expression of *LncKLHDC7B* was evaluated by RT-PCR. The RNA was extracted using Trizol according the manufacturer's recommendations and stored at  $-80^\circ\text{C}$  until processing. A > 70% decrease of lncRNA expression was achieved after 48 h post-transfection using 2–3.5  $\mu$ g of plasmid. The

experiments were carried out in triplicate. The effect of silencing *LncKLHDC7B* on the whole-genome transcriptional landscape in HCC1187 was analyzed with the HTA 2.0 microarray (Affymetrix; Data accessibility GEO: GSE114468), as described above. Genes with a fold change of > 2 and < –2, and a *P* value < 0.05 were considered significant and selected for biological pathway analysis using Ingenuity Pathway Analysis (IPA) software and GENETRAIL2 (Stockel *et al.*, 2016).

## 2.9. Migration and invasion assay

Cell migration and invasion assays were performed using Transwell migration chambers with 8- $\mu$ m pores (Corning, NY, USA). Transfected cells were harvested, re-suspended in RPMI without FBS at a concentration of  $5 \times 10^4$  cells in 100  $\mu$ L, and seeded into the upper chamber of the 24-well plate. The lower chamber was filled with 600  $\mu$ L RPMI containing 10% FBS. For the invasion assay,  $1 \times 10^5$  cells were plated on chambers precoated with 1.6 mg Matrigel (Corning). Cells were incubated (HCC1187 and MCF-7 for 24 h, BT-20 for 6 h) for migration assay and (HCC1187 for 36 h, BT-20 and MCF-7 for 48 h) for invasion assay at 37  $^\circ\text{C}$ , in 5%  $\text{CO}_2$ . Each experiment was done in triplicate. At the end of the experiments, cells that migrated to the reverse side of the Transwell membrane were fixed with paraformaldehyde (3.7%), stained with crystal violet, and counted using IMAGEJ software (Madison, WI, USA) (Schneider *et al.*, 2012).

## 2.10. Apoptosis assay

Apoptosis was induced with  $\sim 50$   $\mu\text{M}$  of 2-amino-N-quinolin-8-yl-benzenesulfonamide (QBS; A3105, Sigma-Aldrich, St. Louis, MO, USA) for 24 h before the end of 48 h post-transfection. The cells were then immediately harvested by trypsinization and washed with PBS. The cells were re-suspended in binding buffer and stained with Annexin V and PI (FITC Annexin V/ Dead cell Apoptosis kit, Invitrogen) for 15 min in the dark at room temperature. The stained cells were examined by cytometry. The cells were categorized into early apoptotic cells and late apoptotic cells (bounded in red lines).

## 2.11. Statistical analysis

Statistical significance was analyzed using GRAPHPAD PRISM (version 6, San Diego, CA, USA) and STATA (Version 12, College Station, TX, USA) software. Variances in Kruskal–Wallis tests, ANOVA test, Student's *t* test, and chi-square tests were performed for all comparisons involving categorical variables. Correlation between

variables was determined by Spearman's correlation coefficient. The Kaplan–Meier method and survival differences among groups were assessed by log-rank test. A *P* value < 0.05 was considered significant (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001).

### 3. Results

#### 3.1. Population characteristics and TNBC subtypes distribution

A total of 156 samples derived from patients with TNBC were included in this study. The mean age of patients was 53 years old (range 26–88) and infiltrating ductal carcinoma was the most frequent type of tumor. Clinical and pathological information on TNBC patients included in this study is shown in Table 1. TNBC subtypes were identified using the TNBCtype algorithm, which assigned samples to the six triple-negative breast cancer subtypes. Analyzing our TNBC cohort, the majority of tumors were clustered as IM and M (22.4%, *n* = 35/156), followed by BL1 (14.7%, *n* = 23/156) and LAR subtypes (12.2%, *n* = 19/156), BL2 (7.7%, *n* = 12/156), MSL subtype (7%, *n* = 11/156), and UNS (13.5%, *n* = 21/156) (Fig. S1a). The distribution of the TNBC subtypes in our

cohort was similar to previous studies (Lehmann *et al.*, 2011; Liu *et al.*, 2016; Masuda *et al.*, 2013) (Fig. S1b). DAVID enrichment analysis (Huang *et al.*, 2009) was applied to each of the TNBC subtypes to corroborate previously reported signaling pathways enriched in each TNBC subtype (Table S1, Fig. S2).

#### 3.2. Unique lncRNA expression profiles in TNBC subtypes

Once our tumor cohort was classified into TNBC subtypes, we identified expression patterns of lncRNA among the tumor subclasses (Fig. 1A). In particular, 710 lncRNA showed differential expression: 84 were altered in at least two of the TNBC subtypes, whereas 524 were only altered in a particular tumor subtype (Fig. 1B, Tables 2 and S2). Comparison between the LAR subtype and all other subtypes yielded the highest number of differentially expressed lncRNA (50.9%, *n* = 362/710) followed by M subtype (16.5%, *n* = 117/710) and IM subtype (14.5%, *n* = 130/710). The MLS subtype showed 98 altered lncRNA, whereas the basal subtypes (BL1 and BL2) had a lower number of deregulated lncRNA (19 and 11 transcripts, respectively) (Fig. 1B, Tables 2 and S2).

**Table 1.** Clinical-pathological characteristics of the population.

Characteristics	Number (Total = 156)	BL1 ( <i>n</i> = 23)	BL2 ( <i>n</i> = 12)	IM ( <i>n</i> = 35)	M ( <i>n</i> = 35)	MSL ( <i>n</i> = 11)	LAR ( <i>n</i> = 19)	UNS ( <i>n</i> = 21)	<i>P</i> -value <sup>a</sup>
Age (Mean ± SD)	145	47.7±10.4	59±14.6	50.1±12.6	55.2±14.2	53.3±10.4	59.1±15.6	51.9±14.6	0.0835 <sup>b</sup>
Tumor grade									
≤II	26	1 (3.8)	1 (2.1)	3 (5.5)	7 (5.9)	1 (1.9)	8 (3.1)	5 (3.6)	
>II	110	19 (16.2)	10 (8.9)	26 (23.5)	24 (25.1)	9 (8.1)	8 (12.9)	14 (15.4)	<b>0.0148<sup>c</sup></b>
Unknown	20								
Tumor size (cm)									
≤ 2 cm	64	6 (9.3)	6 (5.1)	12 (14.4)	15 (13.9)	7 (4.6)	9 (7.4)	9 (9.3)	
> 2 cm	74	14 (10.7)	5 (5.9)	19 (16.6)	15 (16.1)	3 (5.4)	7 (8.6)	11 (10.7)	0.398 <sup>c</sup>
Unknown	18								
Type histology									
IDC	117	17 (15.9)	6 (9.2)	29 (26.7)	28 (26.7)	8 (8.4)	13 (13.4)	16 (16.7)	
Other	23	2 (3.1)	5 (1.8)	3 (5.3)	4 (5.3)	2 (1.6)	3 (2.6)	4 (3.3)	0.168 <sup>c</sup>
Unknown	16								
Follow up, month (mean)	55.8	50.5	56.6	55	48	65.8	59.8	62.1	0.562 <sup>d</sup>
CI	50.7–60.9	34.8–66.2	37.3–75.8	43.9–66.1	37.6–58.4	41.6–90.1	42–77.6	49.2–75.1	
Dead event									
Yes	30	8 (4.2)	4 (2.3)	9 (6.8)	5 (6.8)	1 (2.3)	1 (3.4)	2 (4.2)	0.064 <sup>c</sup>
No	112	12 (15.8)	7 (8.7)	23 (25.2)	27 (25.2)	10 (8.7)	15 (12.6)	18 (15.8)	
Unknown	14								

BL1, basal-like 1; BL2, basal-like 2; CI, 95% confidence interval; IDC, infiltrating ductal carcinoma; IM, immunomodulatory; LAR, luminal androgen receptor; M, mesenchymal; MSL, mesenchymal stem-like; SD, standard deviation; UNS, undetermined.

<sup>a</sup>Unknown data were not included for the statistical significance.

<sup>b</sup>ANOVA test.

<sup>c</sup>Chi-square test *P* < 0.05 (in bold).

<sup>d</sup>Kruskal–Wallis test.

**Table 2.** Up-regulated and down-regulated lncRNA with higher rate of change between TNBC subtypes.

lncRNA name	Fold change <sup>a</sup>	ANOVA P-value	FDR P-value	TNBC subtype
RP11-532E4.2	16.98	0.000133	0.015499	BL1
LINC01956	3.26	6.72E-13	4.54E-08	BL1
LINC01123	1.88	7.11E-07	0.000572	BL1
RP11-619J20.1	-1.65	0.000054	0.009449	BL1
RP11-815J21.4	-1.65	0.000489	0.031038	BL1
RP1-142L7.5	-1.82	0.000059	0.00983	BL1
RP11-116G8.5	2.73	0.000015	0.060613	BL2
RP11-3K16.2	2.22	0.000036	0.089263	BL2
LINC01133	1.69	0.000655	0.307124	BL2
LINC02095	-7.91	0.000565	0.300484	BL2
RP4-620F22.2	9.87	0.00E+00	0.00E+00	IM
RP5-1171110.5	1.9	1.71E-14	2.95E-12	IM
ELMO1-AS1	1.71	7.48E-10	5.31E-08	IM
RP13-455A7.1	-1.79	0.000545	0.014289	IM
CTD-2033D15.1	-2.98	0.000041	0.001466	IM
IFNG-AS1	1.82	6.93E-12	6.63E-10	IM
RP11-206M11.7	80.48	1.78E-15	1.08E-12	LAR
LINC00993	4.24	8.76E-07	0.000048	LAR
PCAT18	2.56	1.11E-16	8.24E-14	LAR
AF178030.2	-7.68	9.13E-10	1.41E-07	LAR
LINC01152	-2.27	0.000181	0.002271	LAR
VIM-AS1	-1.83	3.63E-07	0.000024	LAR
GAS5	-1.63	0.000023	0.000538	LAR
LINC02095	20.31	1.11E-16	1.50E-12	M
RP11-84E24.2	7.5	0.00E+00	0.00E+00	M
SOX9-AS1	13.71	0.00E+00	0.00E+00	M
H19	2.99	0.000066	0.002324	M
ANO1-AS1	2.13	0.000059	0.00212	M
CASC15	1.82	9.47E-08	0.00001	M
RP11-553K8.5	-3.39	6.12E-10	2.27E-07	M
RP11-989E6.8	-2.97	4.34E-07	0.000037	M
ANKRD44-IT1	-1.86	8.62E-08	0.00001	M
MEG3	4.21	5.29E-11	4.35E-08	MSL
AC133106.2	3.97	1.11E-09	6.08E-07	MSL
DNM3OS	3.73	0.000039	0.003966	MSL
VCAN-AS1	3.22	0.000019	0.002127	MSL
CARMN	2.44	2.25E-11	2.03E-08	MSL
TCONS_I2_00019097	-1.59	0.000161	0.0125	MSL
LINC00302	-1.64	0.000782	0.042797	MSL
TCONS_I2_00003601	-1.89	0.000845	0.045414	MSL

<sup>a</sup>Results from TAC software (Affymetrix).

To validate the lncRNA expression patterns in the TNBC subtypes, we corroborated the expression of some candidates using an independent cohort containing lncRNA profiles. Consistent with our data, a set of lncRNA was differentially expressed in the comparisons between BL1, BL2, M, MSL, IM, and LAR tumors in our *in silico* analysis (Fig. S3). Validation of the expression of this set of lncRNA in an independent cohort suggests that they might play important biological roles in the establishment of TNBC subtypes.

### 3.3. Identifying biological pathways associated with lncRNA in TNBC subtypes

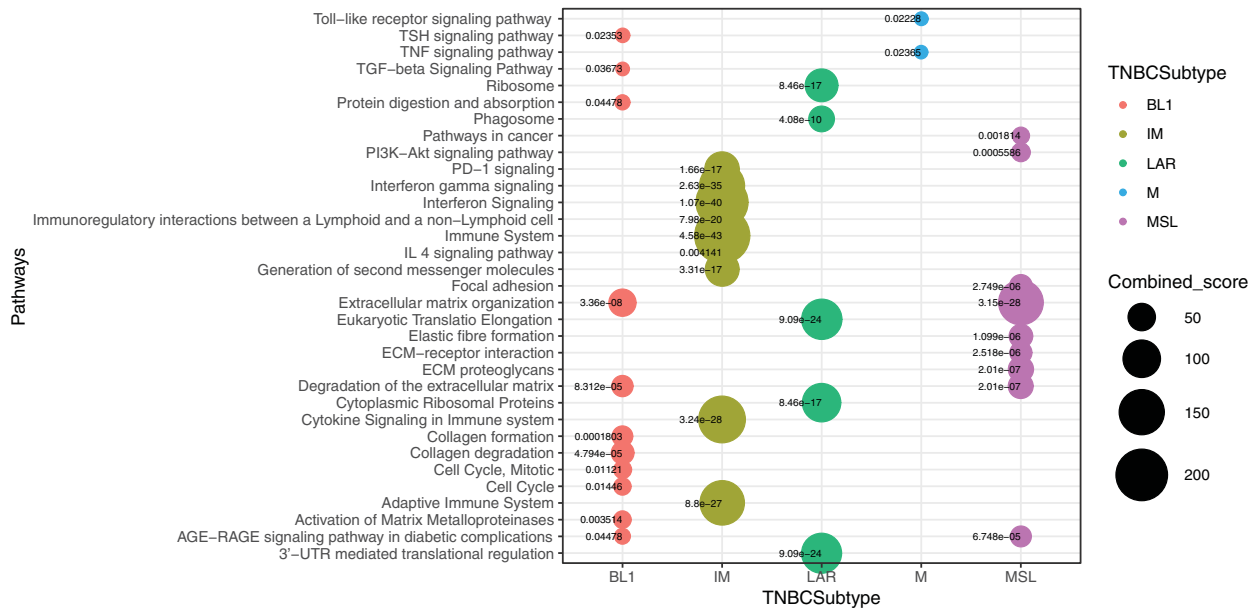
To gain insight into the biological relevance of the altered lncRNA, we used a guilt-by-association approach to investigate their relationship to dysregulated mRNA and their possible impact on different pathways. This analysis revealed a significant association between some of the differentially expressed lncRNA and mRNA, as well as the enrichment in key breast cancer-related pathways of associated coding genes for each TNBC subtype. lncRNA altered in BL1 correlated most strongly with extracellular matrix organization, cell cycle and transforming growth factor-beta (TGF- $\beta$ ) signaling. IM lncRNA are most strongly related to antigen processing, interferon signaling and immune system. LAR contains lncRNA that are potentially involved in translation regulation. Altered lncRNA in M tumors are associated with toll-like receptor and TNF signaling. Finally, the co-expression correlation network of altered lncRNA-mRNA in the MSL subtype was enriched in focal adhesion, ECM receptor interaction, and PI3K-Akt signaling pathway (Fig. 2, Table S3). With the exception of the BL2 subtype, where we did not find any significant association, guilt-by-association analyses suggested functions for a set of altered lncRNA by their weighted co-expression with related mRNA that defined a group of pathways with relevant correlations with lncRNA expression portraits in human tumors.

### 3.4. Co-expression of lncRNA and coding genes in IM subtypes

Long non-coding RNAs can function as enhancers of the expression of both adjacent and distant genes. To prioritize the most biologically relevant lncRNA in the IM subtypes (one of the most frequent subtypes), as a first approach we identified lncRNA (intergenic or intragenic) with possible regulatory functions over their neighboring genes; thus we analyzed lncRNA-mRNA pairs positively co-expressed (Fig. S4, Table S4). This approach defined the positive correlation of *KLHDC7B* and *ENSG00000226738*.

### 3.5. *LncKLHDC7B* (*ENSG00000226738*) and *KLHDC7B* show a correlated expression pattern in the immunomodulatory subtype

The lncRNA *ENSG00000226738* (which we call *LncKLHDC7B*) and its coding gene *KLHDC7B* were among one the most correlated pairs evaluated ( $r^2 = 0.9$ ,



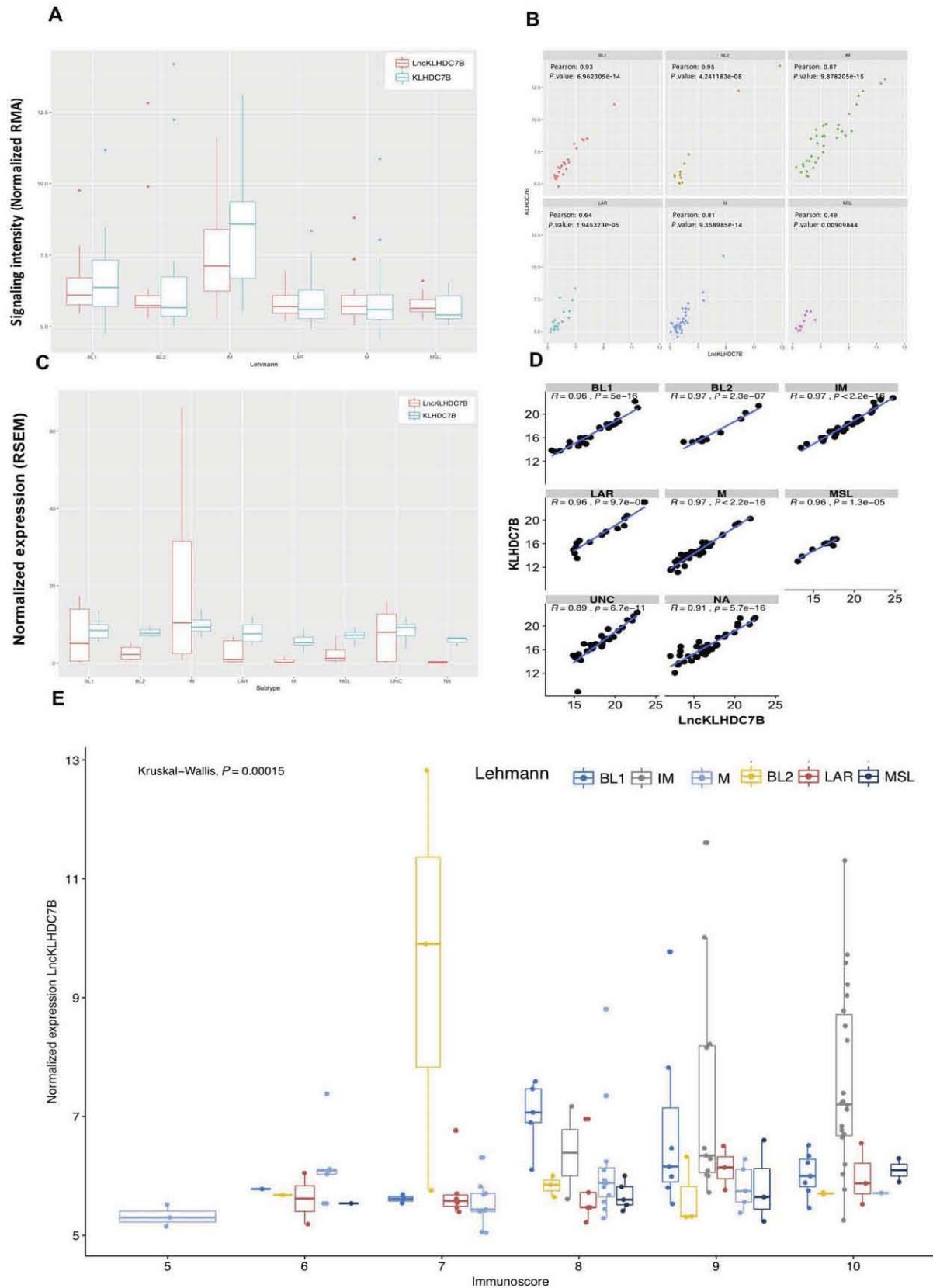
**Fig. 2.** Overview of the biological pathways by guilt-by-association analysis across TNBC subtypes by lncRNA-mRNA co-expression. Guilt-by-association analysis showing the significant enriched pathways of the resulted co-expression mRNA-lncRNA. Pathway enrichment analysis resulted in significant association with cancer-related signaling. The y-axis label represents pathways and the x-axis label represents TNBC subtypes. Bubble chart shows biological pathways enrichment by differential expression of lncRNA for each TNBC subtype. Size and color of the bubble represent the score each pathway and TNBC subtypes, respectively.

$P < 0.0001$ ; Figs S4 and 3A–D). The overexpression of the pair *LncKLHDC7B*/*KLHDC7B* as well as their correlation in the IM subtype was also validated in TCGA data (Fig. 3C,D, Table S4). We evaluated the expression of both transcripts in a panel of breast cancer cell lines, which showed increased expression of both genes in many tumor cell lines, including IM cell line (HCC1187), in contrast to a non-tumorigenic epithelial cell line MCF10A (Fig. 4A). Interestingly, these observations were reproducible in breast cancer tissues compared with normal tissue using databases from TCGA (Fig. S6). We noticed a distinct expression pattern of *KLHDC7B* and *LncKLHDC7B* among the different triple-negative subgroups, both in tumors and cell line models (Figs 3A–D, 4A and S6). To further validate the relationship of *LncKLHDC7B* with the immunophenotype, we calculated the immunophenoscore of our profiled cohort. Our findings showed distinct expression patterns across the scores, where the highest scores (9 and 10) present a

significant enrichment of *LncKLHDC7B*, particularly in the IM subtype (Fig. 3E). Given these results, we selected the HCC1187 cell line to investigate further the functional role of *LncKLHDC7B* in the immunomodulatory subtype.

We then chose the *LncKLHDC7B*/*KLHDC7B* pair for further analysis for the following reasons: (1) *LncKLHDC7B* may be a subtype tumor marker of the IM subtype, but its function is not well described in the literature; (2) *LncKLHDC7B* presents a significant association with its coding gene (localized in the same locus), which indicates a possible regulatory activity; (3) our guilt-by-association analysis highlighted correlations between the lncRNA and key cancer-related pathways such as antigen presentation, natural killer-mediated cytotoxicity, interferon signaling, and cell adhesion (Fig. S5). All these data suggest a potential role for *LncKLHDC7B* in the TNBC subtype establishment and biological regulatory programs.

**Fig. 3.** Evaluation of *LncKLHDC7B* and *KLHDC7B* expression in TNBC subtypes. (A) *LncKLHDC7B* and *KLHDC7B* expression levels in our cohort and (C) TCGA database. (B) Correlation between gene expression level of *LncKLHDC7B* and *KLHDC7B* across TNBC subtypes in our dataset, and TCGA database (D). (E) Relation across the immunoscore with the expression of *LncKLHDC7B* in our cohort. Kruskal–Wallis test was performed to determine significance. The y-axis label represents the normalized expression of the *LncKLHDC7B* and the x-axis label represents the immunoscore.

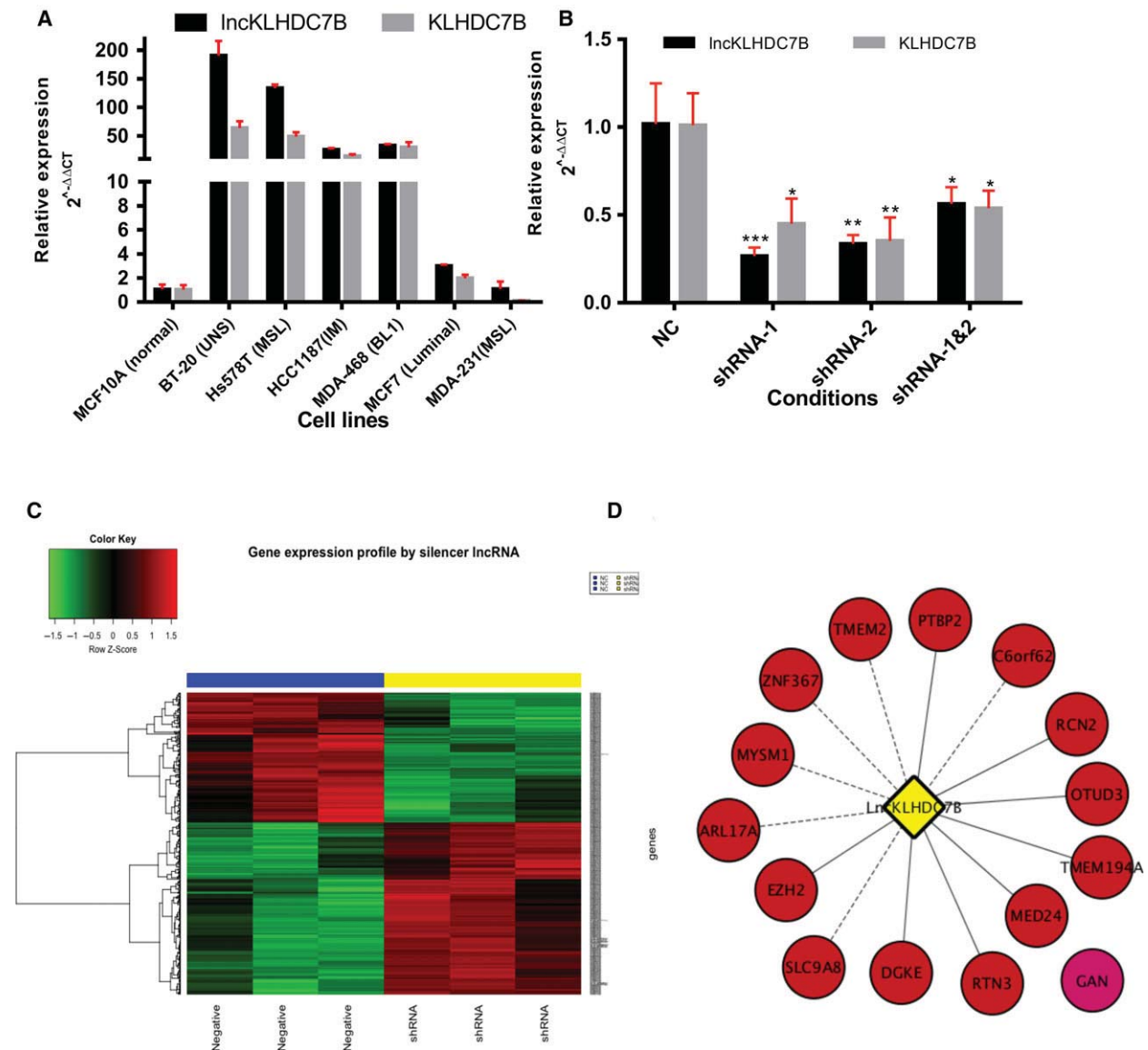




### 3.6. Silencing of *LncKLHDC7B* down-regulates *KLHDC7B* and impacts the expression of a large number of genes

We evaluated the role of *LncKLHDC7B* on oncogenic phenotypes in the IM breast cancer HCC1187 cell line model. The knockdown of *LncKLHDC7B* was achieved with ~75% silencing efficiency (Fig. 4B) and a 50%

inhibition of the expression of *KLHDC7B* gene (Fig. 4B). Furthermore, a complete genomic analysis of the transcriptomic state after silencing of *LncKLHDC7B* revealed an impact on the global expression of 1265 transcripts (Fig. 4C, Table S5), including several transcripts from the Kelch family. These data demonstrated that *LncKLHDC7B* might regulate gene expression both in *cis*, mediating the transcriptional activation of the



**Fig. 4.** Impact of *LncKLHDC7B* silencing on the TNBC immunomodulatory phenotype. (A) Validation of *LncKLHDC7B* and *KLHDC7B* expression in a panel of breast cancer cell lines. (B) Expression levels of *LncKLHDC7B* in immunomodulatory phenotype (HCC1187) of breast cancer cell line after silencing by shRNA system. *KLHDC7B* coding gene expression is affected by *LncKLHDC7B* silencing. All data are shown as the mean  $\pm$  SD of at least three independent experiments. Student's *t* test was performed to determine significance \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  of NC vs shRNA. (C) Heatmap from transcripts altered by *LncKLHDC7B* silenced. The microarray was performed in triplicate for the NC condition (blue bar) and shRNA-1 (yellow bar). Red represents elevated and green down-regulated expression. (D) Prediction of interactions between lncRNA and mRNA targets sub-expressed by the silencing of the *LncKLHDC7B*. Continuous and dotted lines represent a medium and high interaction, respectively.

*KLHDC7B* gene, and in *trans*, by regulating distant genes located throughout the genome.

Long non-coding RNAs show a diversity of functions through diverse mechanisms, interactions with other RNA molecules being some of the most representative. We applied an mRNA-lncRNA interaction prediction with a nearest-neighbor method based on thermodynamic parameters using the LncTar tool (Li *et al.*, 2015). Genes most significantly down-regulated upon *LncKLHDC7B* silencing were evaluated. Our results indicate a high and medium direct interaction between the mRNA and *LncKLHDC7B* (ndG value  $\leq -0.1$  indicates a true predicted interaction that could not be considered a chance event). Thus, their down-modulation might be a direct consequence of *LncKLHDC7B* knockdown. The regulator-signaling network mediated by *LncKLHDC7B* in the IM cell line model is described in Fig. 4D and Table S6. This approach provided relevant information to identify valuable lncRNA–mRNA interactions for predicting RNA targets and figuring out the possible biological function of the lncRNA.

### 3.7. *LncKLHDC7B* modulates the expression of genes associated with relevant hallmarks of cancer

To investigate the global impact of *LncKLHDC7B* inhibition in the IM cell line model, we performed an Ingenuity Pathway Analysis on up- or down-modulated targets after *LncKLHDC7B* reduction, which revealed a significant over-representation of cancer-relevant pathways related to cell migration, cellular death, apoptosis, and invasion (Fig. 5A).

### 3.8. Silencing of *LncKLHDC7B* increases cellular migration and invasion, and decreases cell death

*LncKLHDC7B* knockdown significantly increased the ability of cells to migrate in comparison with control HCC1187 cells (Fig. 5B,C). This oncogenic phenotype could be achieved by the observed up-modulation of *TNC*, *CDH3*, *APP*, *EPS8*, and *CAVI*, as previously reported (Baek *et al.*, 2010; Chen *et al.*, 2015a; Diaz *et al.*, 2014; Lim *et al.*, 2014; Zhang *et al.*, 2016). As shown in Fig. 5C, the invasion ability of HCC1187 cells transfected with shRNA was significantly enhanced compared with the control group. This phenotype might be related to the up-modulation (Table S5) of genes that have been previously involved in the invasion process, such as *MMP2* (Fagan-Solis *et al.*, 2013; Kim and Rhee, 2016), *ADAM9* (Micocci *et al.*, 2013), and *AMOT* (Zhang and Fan, 2015). Additionally, the depletion of *LncKLHDC7B* decreased the number of apoptotic cells, reaching

~ 23% of reduction in HCC1187 cells going through apoptotic death (Fig. 5D). This effect might be explained by the down-regulation of genes related to the induction of apoptosis, such as *PMAIP1* (Zhao *et al.*, 2014), *RAD21* (Chen *et al.*, 2002; Pati *et al.*, 2002), and *CCAR2* (known as *DBC1*) (Kim *et al.*, 2008), that we observed to be under-expressed after *LncKLHDC7B* silencing (Table S5).

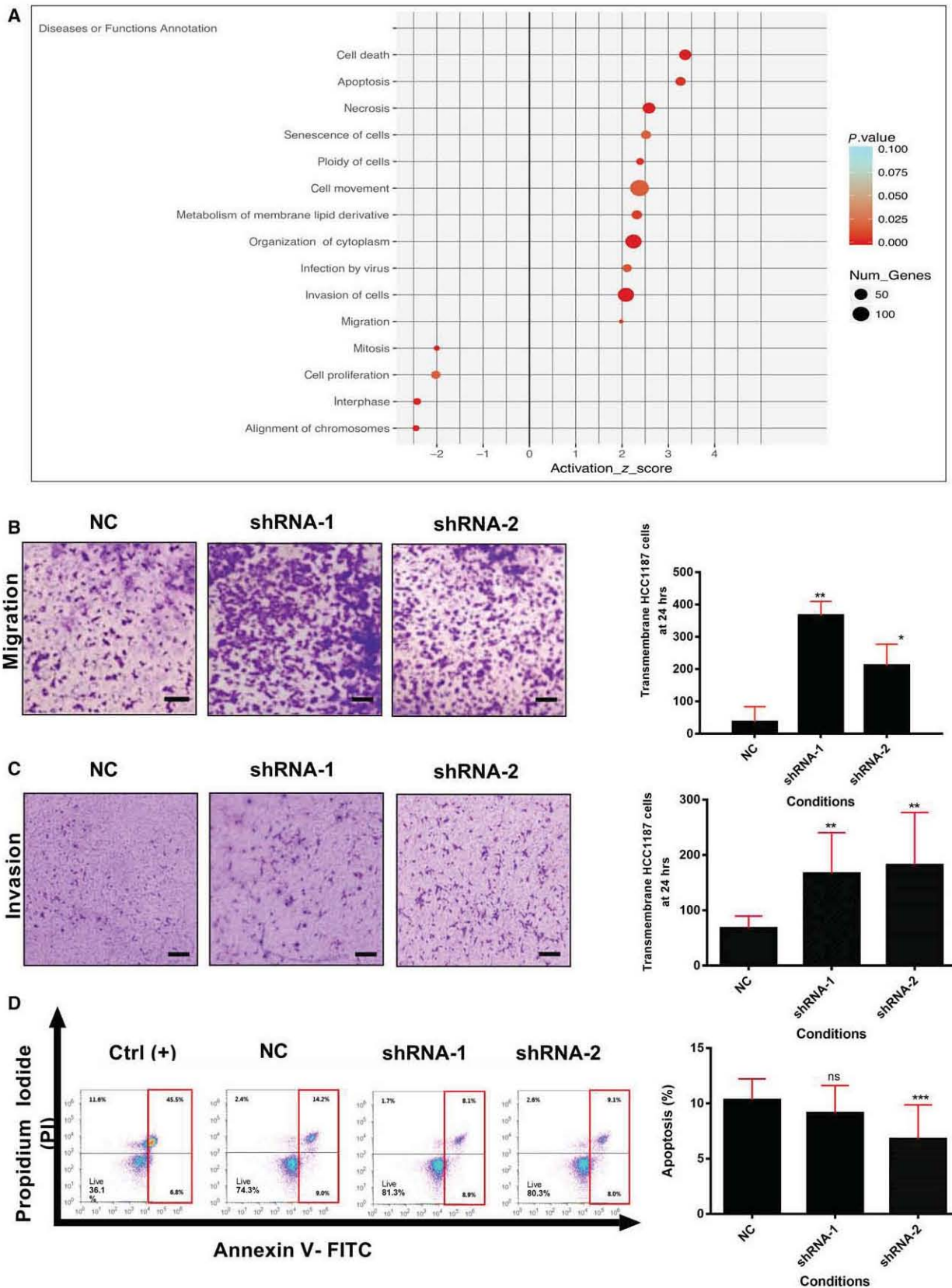
To determine whether the silencing effect of *LncKLHDC7B* on cellular migration, invasion, and apoptosis events are exclusive for triple-negative breast cancer, we silenced lncRNA in two additional cell lines (BT-20 and MCF-7). Interestingly, the down-regulation of *LncKLHDC7B* in the MCF-7 luminal cell line affected *KLHDC7B* mRNA expression (Fig. 6A). Moreover, silencing promoted migration and invasion (Fig. 6B,C) of MCF-7 cells, as well as a resistance to apoptosis (18–24%), compared with the control group (Fig. 6D).

However, in the triple-negative cell line BT-20, we were only able to obtain a 35% of inhibition of *LncKLHDC7B*, and we only observed a 34% down-regulation of *KLHDC7B* (Fig. S7a,b). In addition, the migration, invasion, and resistance to apoptosis were enhanced, albeit non-significantly, by an shRNA (shRNA-1) that had the greatest effect on the silencing the lncRNA in this particular cell line (Fig. S7b-d).

### 3.9. Clinical implications of *LncKLHDC7B* and *KLHDC7B* expression in triple-negative breast cancer

To investigate the potential clinical implication of down-regulating *LncKLHDC7B* and *KLHDC7B* in triple-negative breast cancer, we identified and analyzed public data (Anaya, 2016; Curtis *et al.*, 2012; Lanczky *et al.*, 2016) with available mRNA or lncRNA expression profiles and clinical information (disease-free survival or overall survival). The down-modulation of *KLHDC7B* was associated with an increased probability of a recurrent or metastatic event ( $n = 161$ ,  $P = 0.0053$ , Fig. 7A), but not with risk of death ( $n = 133$ ,  $P = 0.212$ ; Fig. 7B) compared with samples showing a higher expression of the gene. We did not detect a significant probability of death in tumors with low expression of the lncRNA in triple-negative breast cancer samples (Fig. 7C, but a generalized analysis in breast cancer showed that the low *KLHDC7B* and *LncKLHDC7B* expression was significantly associated with lower survival (HR 1.910, 95% CI 1.147–3.233, log-rank test,  $P = 0.0137$  and HR 1.933, 95% CI 1.161–3.273, log-rank test,  $P = 0.0120$ , respectively) compared with patients who presented higher expression of both transcripts (Fig. S8).





**Fig. 5.** Enrichment and functional analysis for the silencing of *LncKLHDC7B* in HCC1187 cell line. (A) Enrichment of biological process from transcriptional alteration by *LncKLHDC7B* silencing in HCC1187 cell line (IPA analysis). The size of the circle represents the number of genes involved in the biological process; the color indicates the statistical significance. (B) Transwell migration and (C) invasion assay showed that *LncKLHDC7B* silencing increases the migration and invasion of HCC1187. Representative images are shown on the left and quantification on the right. Scale bar: 50  $\mu\text{m}$ . (D) Flow cytometric analysis of apoptosis (early and late) in HCC1187 cell transfected with control and shRNA-1 and -2 after Annexin V/PI staining. All data are shown as the mean  $\pm$  SD of at least three independent experiments. Student's *t* test was performed to determine significance of NC vs shRNA: \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001.

Previous studies and our results shown above indicated that lncRNA is differentially expressed between TN breast cancer subtypes, thus lncRNA may be explored as a new class of clinical biomarkers for therapy response or targets for drug discovery (Arun *et al.*, 2018; Prabhakar *et al.*, 2017; Wu and Du, 2017). To identify drug-related lncRNA in breast cancer, we explored the association of the *LncKLHDC7B* levels and drug activity with the lncMAP tool (Li *et al.*, 2018) (<http://bio-bigdata.hrbmu.edu.cn/LncMAP/index.jsp>), which computes the Spearman correlation coefficient between lncRNA expression and the half maximal inhibitory concentration ( $\text{IC}_{50}$ ) values of 24 drugs approved for oncology purposes in TNBC cell lines. We observed a significant negative correlation between the expression of *LncKLHDC7B* and concentrations of lapatinib and paclitaxel (Fig. 7D), which are being evaluated in various clinical trials (<https://clinicaltrials.gov/ct2/home>).

#### 4. Discussion

Triple-negative breast cancer is a heterogeneous disease at the cellular, molecular, and clinical levels (Lehmann *et al.*, 2011). Long non-coding RNA have recently emerged as important players implicated in relevant biological processes (Martin and Chang, 2012) as well as potential cancer biomarkers or therapeutic targets (Arun *et al.*, 2018; Chandra Gupta and Nandan Tripathi, 2017; Prabhakar *et al.*, 2017; Wan *et al.*, 2016; Wu and Du, 2017). In the present study, we classified TNBC using mRNA expression profiles (Chen *et al.*, 2012) and identified several lncRNA expressed in specific tumor subtypes. The most frequent subtypes in our sample collection were mesenchymal-like and immunomodulatory, consistent with previous reports (Lehmann *et al.*, 2011; Liu *et al.*, 2016; Masuda *et al.*, 2013). Regarding tumor subgrouping and clinical characteristics, we did not detect any significant association between TNBC subtypes and tumor size, histology or survival data, except for tumor grade.

A class comparison analysis identified a profile of 710 differentially expressed lncRNA among the 156 samples belonging to different TNBC subtypes. A guilt-by-association analysis revealed the possible biological pathways associated with the lncRNA

expressed for each of the TNBC subtypes except for BL2, possibly due to lack of correlation between the expression of mRNA and lncRNA.

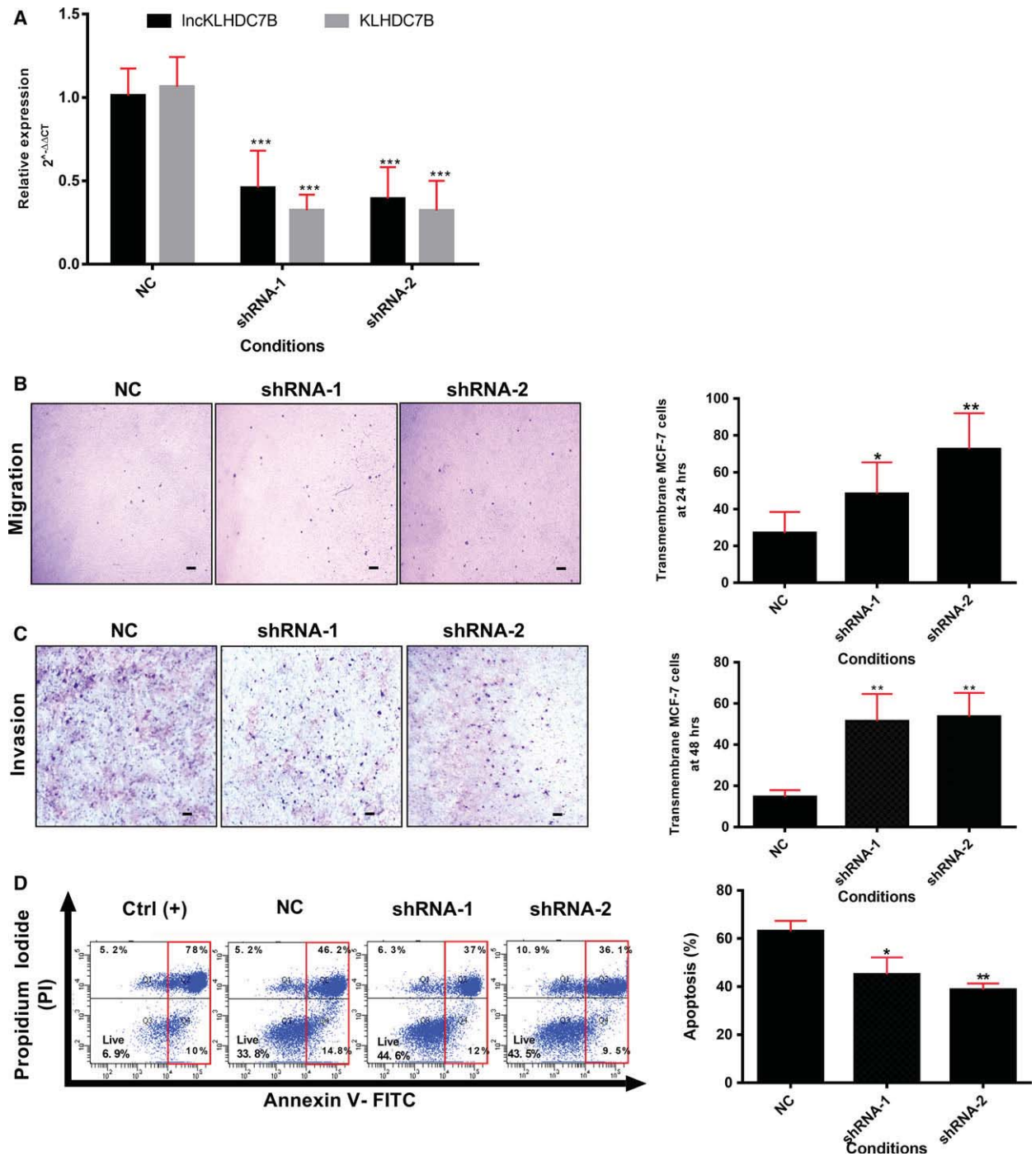
Overall, our findings indicate that lncRNA expression patterns can distinguish TNBC subtypes and that lncRNA might be involved in important pathways related to tumor biology, contributing to the regulatory circuits between lncRNA and mRNA. There was a quite interesting relationship between lncRNA and the immunomodulatory subtype, one of the altered tumor groups in the context of lncRNA transcriptional landscape, in which lncRNA activity seems to impact an important number of pathways involved in the establishment of the immune infiltrated phenotype. Given these results, we decided to further characterize the IM subtype.

In the IM subtype, *LncKLHDC7B* and *KLHDC7B* were up-regulated with positive correlated expression ( $r^2 = 0.9$ ,  $P < 0.0001$ ; Fig. S4), suggesting that the antisense lncRNA might regulate the expression of the *KLHDC7B* coding gene.

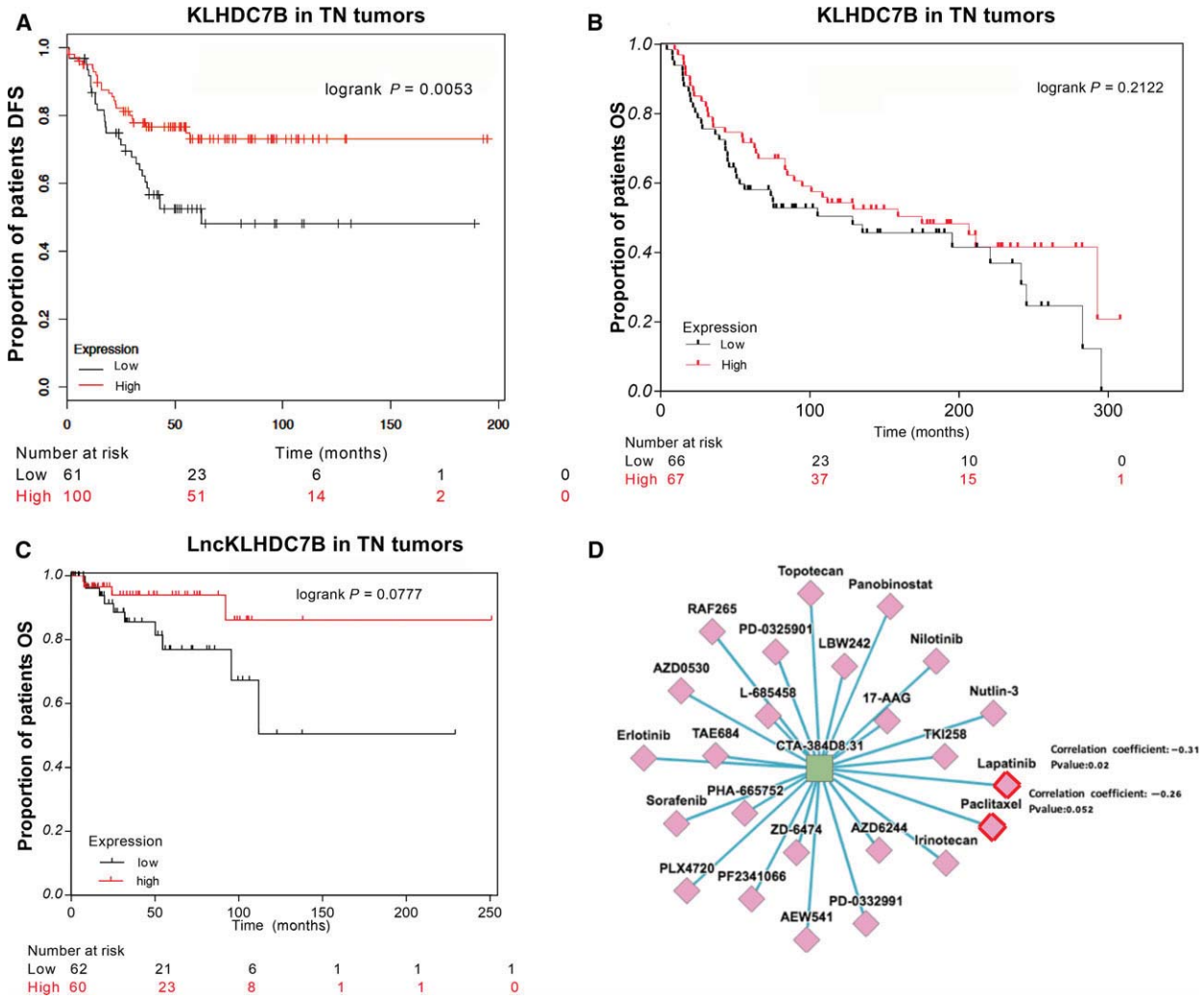
Several lines of evidence indicated the importance of the *LncKLHDC7B*/*KLHDC7B* pair of transcripts, including that *LncKLHDC7B* expression is specific for the IM subtype (or, based on our analyses, on breast cancer in a generalized way). *LncKLHDC7B* expression levels are significantly correlated with its coding gene (localized in the same locus) and our guilt-by-association analysis highlighted a correlation between lncRNA and key cancer-related pathways.

*KLHDC7B* (Kelch Domain Containing 7B) is a protein member of the Kelch superfamily, proteins related to cellular processes such as cytoskeletal rearrangement and protein degradation (Adams *et al.*, 2000). Alterations in this protein superfamily have been associated with various types of cancer, including leukemia, lung, prostate, brain, and Hodgkin's disease (Gupta and Beggs, 2014). To the best of our knowledge, the only report that associates this gene with breast cancer, reported the hypermethylation and over-regulation of this gene in mammary tumors, suggesting a possible role as an epigenetic marker (Kim *et al.*, 2010). However, the role of its non-coding antisense *LncKLHDC7B* was unknown.

*LncKLHDC7B* is a 260-base transcript and has two exons. At the genomic level *LncKLHDC7B* is an



**Fig. 6.** The silencing of *LncKLHDC7B* modulates breast cancer cell migration and invasion as well as resistance to apoptosis in MCF-7. (A) The expression of *LncKLHDC7B* and *KLHDC7B* in MCF-7 cells transfected with NC or shRNA-1 or -2 was determined by qRT-PCR. (B) Transwell migration and (C) invasion assay showed that *LncKLHDC7B* silencing increases the migration and invasion of MCF-7. Representative images are shown on the left and quantification on the right. Scale bar: 25  $\mu$ m. (D) Flow cytometric analysis of apoptosis (early and late) in MCF-7 cell transfected with control and shRNA-1 and -2 after Annexin V/PI staining. All data are shown as the mean  $\pm$  SD of at least three independent experiments. Student's *t* test was performed to determine significance of NC vs shRNA: \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001.



**Fig. 7.** Clinical implications of the sub-expression of *LncKLHDC7B* and *KLHDC7B* in TNBC. (A) Kaplan–Meier analysis curve of the disease-free survival (DFS) of TNBC according to *KLHDC7B* expression. (B) Kaplan–Meier analysis of overall survival (OS) according to *KLHDC7B* expression. (C) Kaplan–Meier analysis of prognostic relevance of *LncKLHDC7B* expression from public data. (D) *In silico* prediction of susceptibility to USFDA-approved treatment drugs related with *LncKLHDC7B* expression in triple-negative breast cancer cell line.

antisense lncRNA of ~ 1.26 Kb, located on the long arm of chromosome 22q13.33, sharing a locus with the *KLHDC7B* gene (Kent *et al.*, 2002; Volders *et al.*, 2015). Different approaches have demonstrated the potential role of lncRNA as regulators of the expression of genes in *cis* or in *trans* (Engreitz *et al.*, 2016; Gupta *et al.*, 2010). We first corroborated the positive correlation in the expression of *LncKLHDC7B* and *KLHDC7B* in independent tumor datasets and in cell lines. This over-expression was corroborated in the immunomodulatory HCC1187 cell line. Subsequently, we silenced the *LncKLHDC7B* using shRNA, which resulted in the repression of the *KLHDC7B* coding gene. Thus, *LncKLHDC7B* might be acting as a *cis* transcriptional regulator of its coding gene *KLHDC7B*.

We also analyzed the effects of *LncKLHDC7B* knock-down on the global transcriptional landscape in HCC11887 cells. We found that expression of several genes was modified after silencing the lncRNA, suggesting its potential role as a transcriptional regulator in *trans*. Enrichment pathways analysis identified a significant up-modulation of cell death, necrosis, apoptosis, cell movement, migration, invasion, and organization of cytoplasm upon lncRNA silencing. Interestingly, we observed the deregulation of several genes of the Kelch family both upwards and downwards, some of them related to processes such as cell migration and invasion, as well as apoptosis (Lian *et al.*, 2016; Ohta *et al.*, 2010).

Previous studies showed that silencing *KLHL39*, another member of the Kelch family, increased cell



migration and invasion, and decreased cell death by anoikis (Chen *et al.*, 2015b). In accordance with these data, we showed that silencing *LncKLHDC7B* significantly increased the cellular migration and invasion, as well as the resistance to apoptosis in HCC1187 and MCF-7 cells. These observations are also supported by our analysis of the BT-20 cell line (which showed higher expression of *LncKLHDC7B* and *KLHDC7B*), where a ~ 30% silencing was obtained by one of the shRNA, although in a non-significant way, of migration and cellular invasion with a tendency toward resistance to apoptosis. In general, these observations suggest that the effect of *LncKLHDC7B* and *KLHDC7B* inhibition might be applicable beyond the triple-negative subtype.

Previous studies suggest that the overexpression of protein members of the Kelch family, such as *KLHL2*, increases apoptosis (Tseng and Bixby, 2011) or, in the case of silencing *KLHDC7B*, favors death resistance, as previously published (Jeong *et al.*, 2018). Here, we show that the silencing of *LncKLHDC7B* led to resistance to cell death. Similarly, it has been reported that silencing of members of the Kelch family leads to death resistance by anoikis and is associated with lower survival in colorectal cancer (Chen *et al.*, 2015b).

These results support the potential role of *LncKLHDC7B* as a gene expression regulator of *KLHDC7B*, and its association with cytoskeletal rearrangement, cell migration, and invasion. Further evidence of the potential role of the *LncKLHDC7B*–*KLHDC7B* correlated expression in cancer comes from the recent description of a novel 5' alternative splicing site in the *KLHDC7B* gene in cervical squamous cell carcinoma. This novel splice site coincides with the location of a lncRNA (Tcons\_00029745) and was related to cellular differentiation and tumor size in this disease (Guo *et al.*, 2015).

Our data indicate that *LncKLHDC7B* is required for cell migration and invasion inhibition, and contributes to apoptosis induction, suggesting its important regulatory activity in TNBC tumors with enrichment of *LncKLHDC7B*, or maybe a generalized process in breast cancer.

To explore further the altered expression of *LncKLHDC7B* and *KLHDC7B* and their clinical implication in breast cancer, we evaluated the correlation between their expression and survival probability in a public database. Down-regulation of *KLHDC7B* was associated with lower disease-free survival and *LncKLHDC7B* was closely associated with poor clinical outcome in triple-negative tumors and breast cancer in general. As far as we know, this is the first

report of the association of *KLHDC7B* or *LncKLHDC7B* expression with poor prognosis in breast cancer. In addition, we analyzed the clinical implication of *LncKLHDC7B* expression and oncological drugs. Our results suggest a higher sensitivity to lapatinib and paclitaxel in cells over-expressing the *LncKLHDC7B*; however, the limitation of this approach was the lack of *in vitro* validation.

## 5. Conclusions

We explored the lncRNA expression landscape of the very complex and heterogeneous disease represented by triple-negative breast cancer, and identified several lncRNA with subtype-specific expression. Through this analysis we unveiled a correlation in the expression of diverse lncRNA with their corresponding coding genes in the IM subtype, including *LncKLHDC7B*–*KLHDC7B*. Silencing of *LncKLHDC7B* led to altered expression of *KLHDC7B*, a member of the Kelch family, and increased cell migration and invasion with decreased apoptosis, indicating that *LncKLHDC7B* may modulate the aggressiveness of breast cancer IM subtype at least in part by regulating the expression of its host gene *KLHDC7B*. We observed a consistent trend toward shorter survival for patients with low expression levels of *LncKLHDC7B* and *KLHDC7B*. Although the clinical relevance of *LncKLHDC7B* expression remains incompletely understood, our results reaffirm the down-modulation of these transcripts in patients with aggressive clinical TN tumors and suggest their association with poor outcomes. These data must be further characterized to elucidate the exact mechanism involved in these observations.

## Acknowledgements

Fredy Omar Beltrán Anaya is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), and received fellowship (CVU 416163) from CONACYT.

We would like to thank all the patients who participated in this study, and all the medical and nursing staff of FUCAM for their support. We appreciate the support of M.Sc. Raul Mojica from the INMEGEN microarray core lab, M.Sc. L. Nelly Patiño of the cytometry unit, and Dra. Diana I. Aparicio-Baustista and Marco González for their help. This work was funded by the Mexican National Council of Science and Technology Basic Science grant (grant number 258936) and Frontiers in Science grant (number 1285).

## Author contributions

B.A.: performed most of the experimental work and data analysis. R.C., C.T., R.R.: data analysis and discussion. R.V., A.R., J.M., R.G.: sample collection and processing. A.O.: data analysis. B.P.: histopathological review of the cases. D.R., V.C., T.T.: patient identification and clinical follow up. T.E., I.M., B.A., H.M.: project coordination and leadership, data analysis, and study design. All authors contributed to the writing of the manuscript.

## Conflict of interest

A.H.M. has received grants from Astra Zeneca, but these are not related to this particular project. All other authors declare no conflict of interest.

## References

- Adams J, Kelso R and Cooley L (2000) The Kelch repeat superfamily of proteins: propellers of cell function. *Trends Cell Biol* **10**, 17–24.
- Anaya J (2016) OncoLnc: linking TCGA survival data to mRNA, miRNAs, and lncRNAs. *PeerJ Comput Sci* **2**: e67, 2e:67.
- Arun G, Diermeier SD and Spector DL (2018) Therapeutic targeting of long non-coding RNAs in cancer. *Trends Mol Med* **24**, 257–277.
- Baek S, Lee YW, Yoon S, Baek SY, Kim BS and Oh SO (2010) CDH3/P-Cadherin regulates migration of HuCCT1 cholangiocarcinoma cells. *Anat Cell Biol* **43**, 110–117.
- Bauer KR, Brown M, Cress RD, Parise CA and Caggiano V (2007) Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* **109**, 1721–1728.
- Chandra Gupta S, Nandan Tripathi Y (2017) Potential of long non-coding RNAs in cancer patients: from biomarkers to therapeutic targets. *Int J Cancer* **140**, 1955–1967.
- Chen HY, Hu JY, Chen TH, Lin YC, Liu X, Lin MY, Lang YD, Yen Y and Chen RH (2015b) KLHL39 suppresses colon cancer metastasis by blocking KLHL20-mediated PML and DAPK ubiquitination. *Oncogene* **34**, 5141–5151.
- Chen F, Kamradt M, Mulcahy M, Byun Y, Xu H, McKay MJ and Cryns VL (2002) Caspase proteolysis of the cohesin component RAD21 promotes apoptosis. *J Biol Chem* **277**, 16775–16781.
- Chen X, Li J, Gray WH, Lehmann BD, Bauer JA, Shyr Y and Pietenpol JA (2012) TNBCtype: a subtyping tool for triple-negative breast cancer. *Cancer Informat* **11**, 147–156.
- Chen C, Liang Z, Huang W, Li X, Zhou F, Hu X, Han M, Ding X and Xiang S (2015a) Eps8 regulates cellular proliferation and migration of breast cancer. *Int J Oncol* **46**, 205–214.
- Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74.
- Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y *et al.* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352.
- Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P and Narod SA (2007) Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* **13**, 4429–4434.
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775–1789.
- Diaz J, Mendoza P, Ortiz R, Diaz N, Leyton L, Stupack D, Quest AF and Torres VA (2014) Rab5 is required in metastatic cancer cells for Caveolin-1-enhanced Rac1 activation, migration and invasion. *J Cell Sci* **127**, 2401–2406.
- Engreitz JM, Haines JE, Perez EM, Munson G, Chen J, Kane M, McDonel PE, Guttman M and Lander ES (2016) Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455.
- Fagan-Solis KD, Schneider SS, Pentecost BT, Bentley BA, Otis CN, Gierthy JF and Arcaro KF (2013) The RhoA pathway mediates MMP-2 and MMP-9-independent invasive behavior in a triple-negative breast cancer cell line. *J Cell Biochem* **114**, 1385–1394.
- Foulkes WD, Smith IE and Reis-Filho JS (2010) Triple-negative breast cancer. *N Engl J Med* **363**, 1938–1948.
- Gibb EA, Brown CJ and Lam WL (2011) The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* **10**, 38.
- Guo P, Wang D, Wu J, Yang J, Ren T, Zhu B and Xiang Y (2015) The landscape of alternative splicing in cervical squamous cell carcinoma. *Oncotargets Ther* **8**, 73–79.
- Gupta VA and Beggs AH (2014) Kelch proteins: emerging roles in skeletal muscle development and diseases. *Skelet Muscle* **4**, 11.
- Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL *et al.* (2010) Long non-coding RNA HOTAIR reprograms

- chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP *et al.* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227.
- Heberle H, Meirelles GV, da Silva FR, Telles GP and Minghim R (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform* **16**, 169.
- da Huang W, Sherman BT and Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocol* **4**, 44–57.
- Jeong G, Bae H, Jeong D, Ham J, Park S, Kim HW, Kang HS and Kim SJ (2018) A Kelch domain-containing KLHDC7B and a long non-coding RNA ST8SIA6-AS1 act oppositely on breast cancer cell proliferation via the interferon signaling pathway. *Sci Rep* **8**, 12922.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM and Haussler D (2002) The human genome browser at UCSC. *Genome Res* **12**, 996–1006.
- Kim D and Rhee S (2016) Matrix metalloproteinase2 regulates MDAMB231 breast cancer cell invasion induced by active mammalian diaphanous-related formin 1. *Mol Med Rep* **14**, 277–282.
- Kim JE, Chen J and Lou Z (2008) DBC1 is a negative regulator of SIRT1. *Nature* **451**, 583–586.
- Kim TW, Kim YJ, Lee HJ, Min SY, Kang HS and Kim SJ (2010) Hs.137007 is a novel epigenetic marker hypermethylated and up-regulated in breast cancer. *Int J Oncol* **36**, 1105–1111.
- Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90–W97.
- Lanczyk A, Nagy A, Bottai G, Munkacsy G, Szabo A, Santarpia L and Gyorffy B (2016) miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat* **160**, 439–446.
- Langfelder P and Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* **9**, 559.
- Lara-Medina F, Perez-Sanchez V, Saavedra-Perez D, Blake-Cerda M, Arce C, Motola-Kuba D, Villarreal-Garza C, Gonzalez-Angulo AM, Bargallo E, Aguilar JL *et al.* (2011) Triple-negative breast cancer in Hispanic patients: high prevalence, poor prognosis, and association with menopausal status, body mass index, and parity. *Cancer* **117**, 3658–3669.
- Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y and Pietenpol JA (2011) Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **121**, 2750–2767.
- Lehmann BD, Jovanovic B, Chen X, Estrada MV, Johnson KN, Shyr Y, Moses HL, Sanders ME and Pietenpol JA (2016) Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. *PLoS ONE* **11**, e0157368.
- Li CH and Chen Y (2013) Targeting long non-coding RNAs in cancers: progress and prospects. *Int J Biochem Cell Biol* **45**, 1895–1910.
- Li J, Ma W, Zeng P, Wang J, Geng B, Yang J and Cui Q (2015) LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief Bioinform* **16**, 806–812.
- Li Y, Li L, Wang Z, Pan T, Sahni N, Jin X, Wang G, Li J, Zheng X, Zhang Y *et al.* (2018) LncMAP: Pan-cancer atlas of long noncoding RNA-mediated transcriptional network perturbations. *Nucleic Acids Res* **46**, 1113–1123.
- Lian YF, Yuan J, Cui Q, Feng QS, Xu M, Bei JX, Zeng YX and Feng L (2016) Upregulation of KLHDC4 predicts a poor prognosis in human nasopharyngeal carcinoma. *PLoS ONE* **11**, e0152820.
- Lim S, Yoo BK, Kim HS, Gilmore HL, Lee Y, Lee HP, Kim SJ, Letterio J and Lee HG (2014) Amyloid-beta precursor protein promotes cell proliferation and motility of advanced breast cancer. *BMC Cancer* **14**, 928.
- Liu YR, Jiang YZ, Xu XE, Yu KD, Jin X, Hu X, Zuo WJ, Hao S, Wu J, Liu GY *et al.* (2016) Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. *Breast Cancer Res* **18**, 33.
- Martin L and Chang HY (2012) Uncovering the role of genomic ‘dark matter’ in human disease. *J Clin Invest* **122**, 1589–1595.
- Martinez ME, Wertheim BC, Natarajan L, Schwab R, Bondy M, Daneri-Navarro A, Meza-Montenegro MM, Gutierrez-Millan LE, Brewster A, Komenaka IK *et al.* (2013) Reproductive factors, heterogeneity, and breast tumor subtypes in women of Mexican descent. *Cancer Epidemiol Biomarkers Prev* **22**, 1853–1861.
- Masuda H, Baggerly KA, Wang Y, Zhang Y, Gonzalez-Angulo AM, Meric-Bernstam F, Valero V, Lehmann BD, Pietenpol JA, Hortobagyi GN *et al.* (2013) Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res* **19**, 5533–5540.
- Micocci KC, Martin AC, Montenegro Cde F, Durante AC, Pouliot N, Cominetti MR and Selistre-de-Araujo HS (2013) ADAM9 silencing inhibits breast tumor cell invasion in vitro. *Biochimie* **95**, 1371–1378.



- Ohta Y, Fujimura L, Nishio S, Arima M, Sakamoto A, Shimada H, Ochiai T, Tokuhisa T and Hatano M (2010) A Kelch family protein Nd1-L functions as a metastasis suppressor in cancer cells via Rho family proteins mediated mechanism. *Int J Oncol* **36**, 427–434.
- Pati D, Zhang N and Plon SE (2002) Linking sister chromatid cohesion and apoptosis: role of Rad21. *Mol Cell Biol* **22**, 8267–8277.
- Perez-Rodriguez G (2015) Prevalence of breast cancer subtypes by immunohistochemistry in patients in the Regional General Hospital 72, Instituto Mexicano del Seguro Social. *Cir Cir* **83**, 193–198.
- Prabhakar B, Zhong XB and Rasmussen TP (2017) Exploiting long noncoding RNAs as pharmacological targets to modulate epigenetic diseases. *Yale J Biol Med* **90**, 73–86.
- Rastelli F, Biancanelli S, Falzetta A, Martignetti A, Casi C, Bascioni R, Giustini L and Crispino S (2010) Triple-negative breast cancer: current state of the art. *Tumori* **96**, 875–888.
- Schneider CA, Rasband WS and Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675.
- Stockel D, Kehl T, Trampert P, Schneider L, Backes C, Ludwig N, Gerasch A, Kaufmann M, Gessler M, Graf N *et al.* (2016) Multi-omics enrichment analysis using the GeneTrail2 web service. *Bioinformatics* **32**, 1502–1508.
- Tseng LA and Bixby JL (2011) Interaction of an intracellular pentraxin with a BTB-Kelch protein is associated with ubiquitylation, aggregation and neuronal apoptosis. *Mol Cell Neurosci* **47**, 254–264.
- Volders PJ, Verheggen K, Menschaert G, Vandepoele K, Martens L, Vandesompele J and Mestdagh P (2015) An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Res* **43**, D174–D180.
- Wan X, Huang W, Yang S, Zhang Y, Pu H, Fu F, Huang Y, Wu H, Li T and Li Y (2016) Identification of androgen-responsive lncRNAs as diagnostic and prognostic markers for prostate cancer. *Oncotarget* **7**, 60503–60518.
- Wapinski O and Chang HY (2011) Long noncoding RNAs and human disease. *Trends Cell Biol* **21**, 354–361.
- Wickham H (2009) Ggplot2 Elegant Graphics for Data Analysis, Use R! Dordrecht: Springer, pp. 1 online resource (viii, 212 pages) illustrations (some color).
- Wu T and Du Y (2017) lncRNAs: From basic research to medical application. *Int J Biol Sci* **13**, 295–307.
- Zhang H and Fan Q (2015) MicroRNA-205 inhibits the proliferation and invasion of breast cancer by regulating AMOT expression. *Oncol Rep* **34**, 2163–2170.
- Zhang Z, Yu B, Gu Y, Zhou S, Qian T, Wang Y, Ding G, Ding F and Gu X (2016) Fibroblast-derived tenascin-C promotes Schwann cell migration through beta1-integrin dependent pathway during peripheral nerve regeneration. *Glia* **64**, 374–385.
- Zhao X, Liu X and Su L (2014) Parthenolide induces apoptosis via TNFRSF10B and PMAIP1 pathways in human lung cancer cells. *J Exp Clin Cancer Res* **33**, 3.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** Gene ontology and KEGG pathway enrichment analysis from differentially expressed genes from TNBC subtypes.

**Table S2.** Differentially expressed lncRNA in the TNBC subtypes.

**Table S3.** *In silico* prediction of biological pathways by guilt-by-association analysis across lncRNA-mRNA co-expressed in the TNBC subtypes.

**Table S4.** Correlation analysis of coding and non-coding genes co-expressed positively in the immunomodulatory phenotype.

**Table S5.** Differentially expressed genes that are significantly modulated after *lncKLHDC7B* silencing in HCC1187 cells.

**Table S6.** Prediction lncRNA-mRNA interactions from down-regulated genes after silencing *lncKLHDC7B*.

**Fig. S1.** (a) Frequency of TNBC subtypes in this study and (b) other studies (Cancer Genome Atlas, 2012; Curtis *et al.*, 2012; Lehmann *et al.*, 2011; Liu *et al.*, 2016; Masuda *et al.*, 2013).

**Fig. S2.** Gene ontology from FFPE TNBC samples in this study.

**Fig. S3.** Independent validation of some lncRNA in TNBC subtypes.

**Fig. S4.** Correlation analysis of coding and non-coding genes co-expressed positively in the immunomodulatory phenotype.

**Fig. S5.** Guilt-by-association analysis.

**Fig. S6.** Up-modulation of *lncKLHDC7B* and *KLHDC7B* in tumor samples of breast cancer.

**Fig. S7.** Functional analysis by the silencing of *lncKLHDC7B* in BT-20 cell line.

**Fig. S8.** Kaplan–Meier curve of overall survival (OS).