UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

ANÁLISIS FUNCIONAL DEL PAPEL BIOLÓGICO DE RNAS LARGOS NO CODIFICANTES SOBRE-EXPRESADOS EN CÁNCER DE MAMA

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

M. EN C. RÍOS ROMERO MAGDALENA

TUTOR PRINCIPAL DE TESIS:

DR. ALFREDO HIDALGO MIRANDA
FACULTAD DE MEDICINA, UNAM

COMITÉ TUTOR:

DRA. ELIZABETH LANGLEY MCCARRON
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM
DRA. MARCELA LIZANO SOBERÓN
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

Ciudad Universitaria, CDMX, OCTUBRE, 2021

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

OFICIO CPCB/967/2021

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted qué en la reunión ordinaria del Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día **03 de mayo de 2021** se aprobó el siguiente jurado para el examen de grado de **DOCTORA EN CIENCIAS** de la estudiante **RÍOS ROMERO MAGDALENA** con número de cuenta **304625231** con la tesis titulada **"ANÁLISIS FUNCIONAL DEL PAPEL BIOLÓGICO DE RNAS LARGOS NO CODIFICANTES SOBRE-EXPRESADOS EN CÁNCER DE MAMA",** realizada bajo la dirección del **DR. ALFREDO HIDALGO MIRANDA,** quedando integrado de la siguiente manera:

Presidente:	DRA. SILVIA JIMÉNEZ MORALES
Vocal:	DRA. SARA HUERTA YEPEZ
Vocal:	DR. DIDDIER GIOVANNI PRADA ORTEGA
Vocal:	DRA. GISELA CEBALLOS CANCINO
Secretario:	DRA. MARCELA LIZANO SOBERÓN

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 11 de octubre de 2021

COORDINADOR DEL PROGRAMA

POSGRADO EN CIENCIAS BIOLOGICAS
COORDINACIÓN

DR. ADOLFO GERARDO NAVARRO SIGÜENZA

Agradecimientos

- Al Posgrado en Ciencias Biológicas, UNAM.
- A CONACYT, por el apoyo recibido para poder realizar mis estudios e investigación (número de beca 258936).
- Al Dr. Alfredo Hidalgo Miranda, mi Tutor de Doctorado.
- A la Dra. Elizabeth Langley McCarron.
- A la Dra. Marcela Lizano Soberón.

Dedicatorias

- **A mis padres.**
- **A Memo.**
- **A Django, Einstein y Valkiria.**

†A la memoria de Gisselle. Te fuiste muy pronto.

Índice

Resumen

El cáncer de mama (CaMa) es un problema de salud pública, ya que es el tumor maligno más frecuente en las mujeres de todo el mundo. El CaMa es una enfermedad molecularmente heterogénea, sobre todo a nivel de expresión transcriptómica (mRNA). Datos recientes demuestran que los RNAs codificantes (cRNA) representan sólo el 34% del transcriptoma total de una célula humana. El 66% restante son RNA no codificantes (ncRNA), por lo que se podría estar perdiendo información biológica, clínica o regulatoria relevante. En este trabajo de tesis doctoral, se identificaron dos RNAs largos no codificantes (lncRNAs) con importantes implicaciones biológicas y clínicas en el cáncer de mama, concretamente FAM83H AS1 y LINC00460.

Se identificaron nueve nuevos tipos de tumores derivados de la base de datos Tha Cancer Genome Atlas (TCGA) con desregulación de FAM83H AS1. Se utilizó un análisis de supervivencia para demostrar que la expresión de FAM83H AS1 es un marcador de mala supervivencia en pacientes con CaMa positivos para receptor de estrógenos (ER) y receptor de progesterona (PR) detectados por inmunohistoquímica (IHC). Adicionalmente, se encontró una correlación significativa entre la sobreexpresión de FAM83H AS1 y la resistencia al tamoxifeno. Los niveles de expresión de los receptores de estrógeno y progesterona interactúan con FAM83H-AS1 para potenciar su efecto en la predicción de la supervivencia global (OS). El silenciamiento de FAM83H-AS1 afecta a dos importantes vías relacionadas con el cáncer de mama: la migración y la muerte celular. Entre los genes blanco potenciales de FAM83H AS1, encontramos que los transcritos p63 y Claudina 1 (CLDN1) se desregulan tras el silenciamiento de FAM83H AS1. Utilizando un análisis de correlación, se mostró que FAM83H AS1 puede regular un conjunto de genes relacionados con el cáncer en múltiples tipos de tumores, incluyendo CaMa. Estas evidencias sugieren que FAM83H AS1 es un regulador maestro en diferentes tipos de cáncer, y CaMa en particular.

En cuanto a LINC00460, se identificaron dos nuevos tipos de tumores en la base de datos TCGA con desregulación del mencionado lncRNA. LINC00460 es un marcador cuya sobreexpresión predice una mala OS, baja supervivencia libre de recaídas (RFS) y baja supervivencia libre de metástasis a distancia (DMFS) en pacientes CaMa parecido a basal. La expresión de LINC00460 es un marcador potencial de fenotipos agresivos en

distintos tumores, incluyendo el cáncer de cabeza y cuello (HNSC) negativo al Virus de Papiloma Humano (VPH), cáncer de riñón de células claras (KIRC) en etapa IV, cáncer de pulmón localmente avanzado y CaMa parecido a basal. Se mostró que el efecto de la expresión pronóstica de LINC00460 es específico de cada tejido, ya que su regulación al alza puede predecir una mala OS en algunos tumores, pero también predice una mejor evolución clínica en los pacientes con CaMa. Se identificó que la expresión de LINC00460 está significativamente enriquecida en el subtipo Triple negativo (TNBC) pa-recido a basal 2 (BL2) y regula potencialmente la vía de diferenciación *Wingless* (*WNT).* LINC00460 también puede modular un conjunto de transcritos relacionados con la inmu-nidad innata en CaMa, como SFRP5, FOSL1, IFNK, CSF2, DUSP7 e IL1A. LINC00460 interactúa con miR-103-a-1, *in-silico*, el cual, a su vez, ya no puede unirse a WNT7A. Por último, la relación LINC00460:WNT7A constituye un marcador compuesto para la dismi-nución de la OS y la RFS en CaMa parecido a basal, y predice la respuesta al tratamiento con antraciclinas en pacientes con CaMa ER-. Estas evidencias confirman que LINC00460 es un regulador maestro de los circuitos moleculares del CaMa e influye de manera importante en el curso clínico de las pacientes.

Abstract

Breast cancer (BRCA) is a serious public health problem, as it is the most frequent ma-lignant tumor in women worldwide. BRCA is a molecularly heterogenic disease, particu-larly at gene expression (mRNAs) level. Recent evidence shows that coding RNAs rep-resent only 34% of the total transcriptome in a human cell. The remaining 66% RNAs are non-coding, so relevant biological, clinical or regulatory information could be lost. In this doctoral thesis report, we identified two long non coding RNAs (lncRNAs) with important biological and clinical implications in breast cancer, namely FAM83H-AS1 and LINC00460.

Nine novel tumor types have been identified from TCGA with FAM83H-AS1 deregulation. This study used survival analysis to demonstrate that FAM83H-AS1 expression is a marker for poor survival in IHC-detected ER and PR positive CaMa patients and found a significant correlation between FAM83H-AS1 overexpression and tamoxifen resistance. Estrogen and Progesterone receptor expression levels interact with FAM83H-AS1 to

potentiate its effect in OS prediction. FAM83H-AS1 silencing impairs two important breast cancer related pathways: cell migration and cell death. Among the most relevant potential FAM83H-AS1gene targets, there were found p63 and claudin 1 (*CLDN1*) to be deregulated after FAM83H-AS1 knockdown. Using correlation analysis, this study shows that FAM83H-AS1 can regulate a plethora of cancer-related genes across multiple tumor types, including BRCA. This evidence suggests that FAM83H-AS1 is a master regulator in different cancer types, and BRCA in particular.

Regarding LINC00460, two novel tumor types from TCGA were found with deregulation of the aforementioned lncRNA. The survival analysis was used to demonstrate that LINC00460 expression is a marker for poor overall (OS), relapse-free (RFS) and distant metastasis-free survival (DMFS) in basal-like BRCA patients. LINC00460 expression is a potential marker for aggressive phenotypes in distinct tumors, including HPV-negative HNSC, stage IV KIRC, locally advanced lung cancer and basal-like BRCA.

It was shown/proved that the LINC00460 prognostic expression effect is tissue-specific, since its upregulation can predict poor OS in some tumors, but also predicts an improved clinical course in CaMa patients. It was found that the LINC00460 expression is significantly enriched in the Basal-like 2 (BL2) TNBC subtype and potentially regulates the WNT differentiation pathway. LINC00460 can also modulate a plethora of immunogenic related genes in CaMa, such as SFRP5, FOSL1, IFNK, CSF2, DUSP7 and IL1A and interacts with miR-103-a-1, in-silico, which, in turn, can no longer target WNT7A. Finally, LINC00460:WNT7A ratio constitutes a composite marker for decreased OS and DMFS in Basal-like CaMa, and can predict anthracycline therapy response in ER-BRCA patients. This evidence confirms that LINC00460 is a master regulator in BRCA molecular circuits and influences clinical outcome.

1. INTRODUCCIÓN

1.1. Epidemiología del cáncer

El cáncer representa un reto en salud pública, ya que es la segunda causa de muerte a nivel mundial. En el año 2012, una de cada ocho muertes se debió a este conjunto de enfermedades, ocasionando 8.2 millones de defunciones. Actualmente 32.6 millones de personas viven con cáncer, siendo los tipos más comunes: pulmón, próstata y mama, este último es el más frecuente en mujeres de todo el mundo, con 1.67 millones de casos nuevos cada año y 520 mil fallecimientos, de los cuales el 70% ocurren en países en vías de desarrollo (Bray *et al.*, 2018).

1.2. Cáncer de mama

En México, el cáncer de mama (CaMa) es desde el 2006 una enfermedad de interés nacional, ya que el 24.8% de todos los cánceres femeninos corresponde a este tipo. El CaMa se posiciona como el primer lugar en enfermedades neoplásicas y la segunda causa de muerte en mujeres mexicanas de todos los estratos económicos (Figura 1. (Bray *et al.*, 2018; Cárdenas-Sánchez *et al.*, 2019).

Se estima que la incidencia y la mortalidad seguirán aumentando de manera importante debido al envejecimiento poblacional, a los cambios en los patrones reproductivos y a una mayor prevalencia de los factores de riesgo; convirtiéndose entonces en una prioridad para el sistema de salud en México (Cárdenas-Sánchez *et al.*, 2019).

El diagnóstico y las decisiones de tratamiento en pacientes con cáncer de mama se realizan con base en factores pronóstico y predictivos que incluyen variables clínicas, como la estadificación del tumor, tamaño tumoral, presencia de invasión en nódulos linfáticos y metástasis. Adicionalmente, se realizan estudios inmunohistoquímicos que caracterizan la presencia y nivel de expresión de 3 receptores: el receptor de estrógenos (RE), receptor de progesterona (RP) y el receptor 2 del factor de crecimiento epidérmico (HER2) (Allred *et al.*, 1998), los cuales son los tres principales biomarcadores utilizados en la práctica clínica para clasificar el CaMa (figura 2).

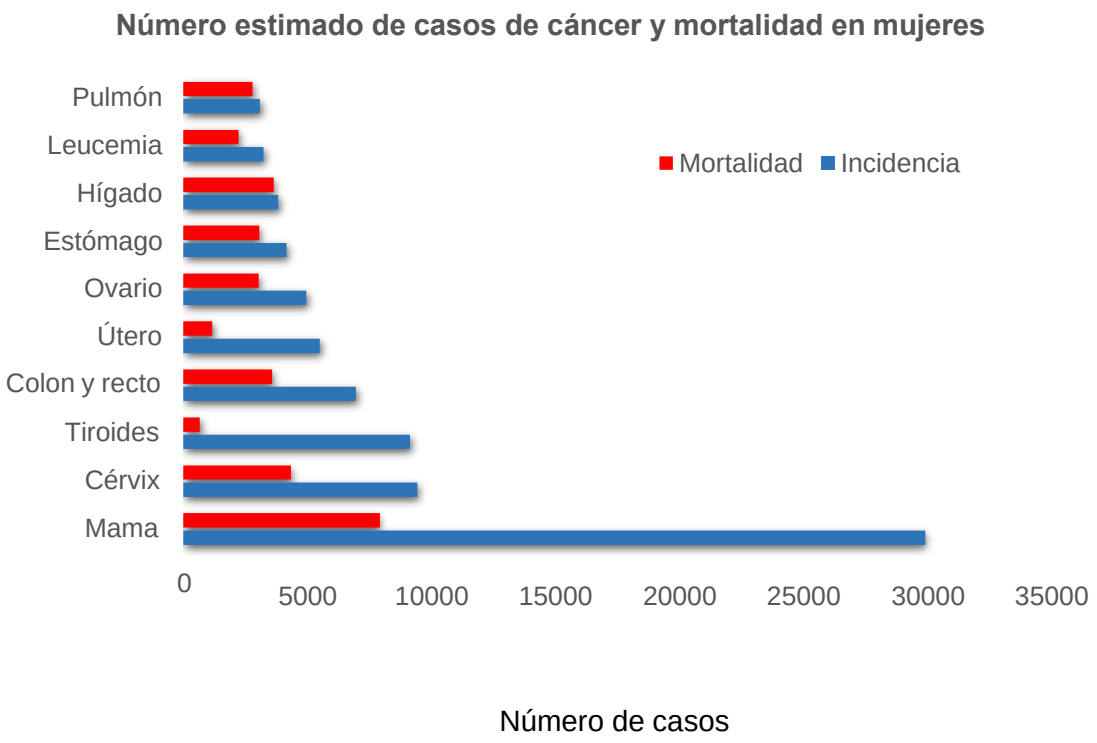

Figura 1. **Epidemiología del cáncer de mama (CaMa) y otros tumores en México.** El CaMa es el primer lugar en incidencia y mortalidad en mujeres en nuestro país. Modificado de (Bray *et al.*, 2018).

Estos factores, en conjunto, determinan el diagnóstico clínico, pronóstico y el esquema de tratamiento de las pacientes. No obstante, se ha observado que las pacientes presentan cursos clínicos distintos y responden diferencialmente a los tratamientos, aun cuando el diagnóstico inicial es similar. Adicionalmente, se estima que la tasa de mortalidad estandarizada (por 100,000 habitantes) es de 12.9 (Ghoncheh, Pournamdar y Salehiniya, 2016). La mortalidad en CaMa está relacionada con la recurrencia del tumor y la resistencia a los tratamientos disponibles. Estos fenómenos clínicos se asocian principalmente a que la clasificación fenotípica basada en receptores no considera la clasificación genómica del CaMa.

1.3. Genómica del cáncer de mama

En las últimas décadas, se ha realizado un esfuerzo considerable en cuanto a la caracterización de las alteraciones genómicas presentes en el genoma y transcriptoma de

los tumores de mama.

Uno de los abordajes genómicos lo constituye el análisis de cambios en los perfiles de expresión de RNA mensajero (mRNA) como herramienta de clasificación y evaluación de riesgo de recurrencia o muerte por actividad tumoral. En el año 2001, Perou y colaboradores analizaron la expresión de 8,102 mRNA en 78 tumores de mama, observando que 476 genes (perfil intrínseco de expresión) permiten subclasificar a

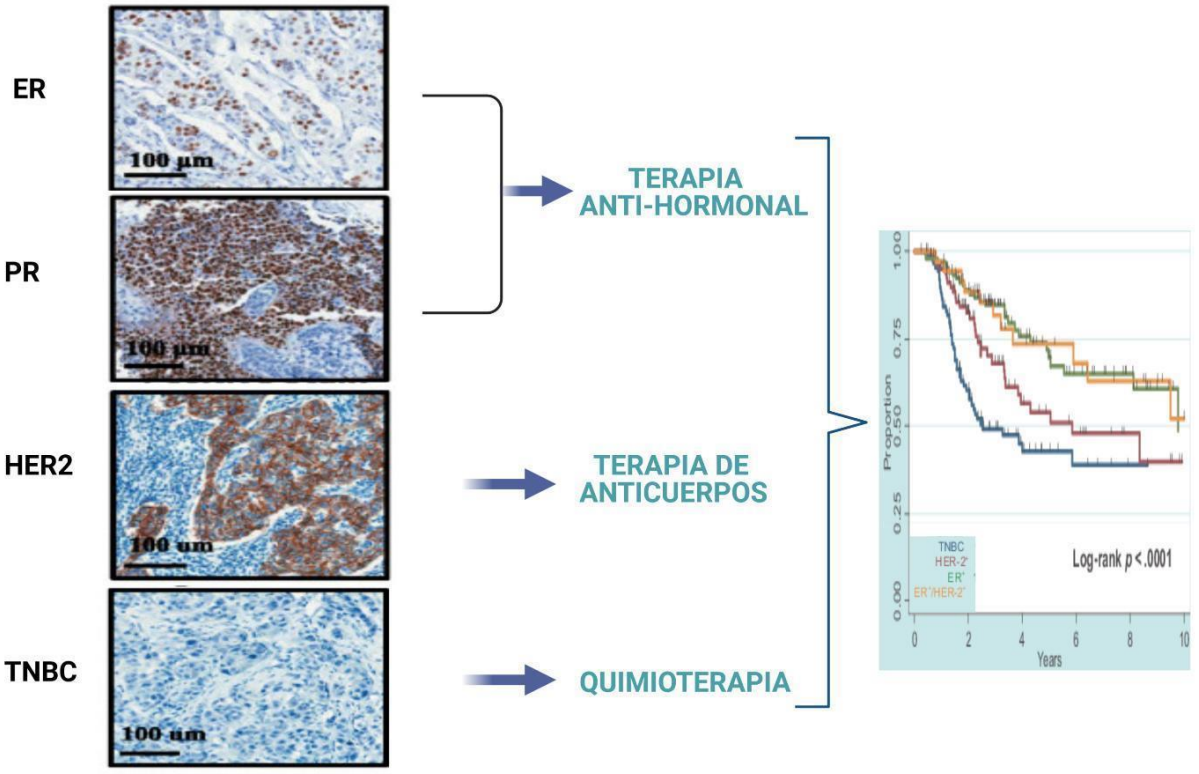

Figura 2. La expresión de receptores hormonales y HER2 determina la terapia de las pacientes y el curso clínico de éstas. Las biopsias de pacientes son analizadas por inmunohistoquímica para detectar ER, PR y HER2. Dependiendo de la presencia o ausencia de éstos, se decide el tratamiento de las pacientes. Esto, a su vez, determina su pronóstico clínico. Modificado de Sawe *et al.*(2016) y de Li *et al.*, (2011) en BioRender (BioRender.com).

los tumores de mama en 5 fenotipos moleculares: Luminal A, Luminal B, enriquecido en Her2, parecido a basal y parecido a normal, cada uno con firmas de expresión génica diferente (Sørlie *et al.*, 2001). En 2009, con el uso de 50 genes se conservó el agrupamiento de tumores por subtipo molecular (PAM50) y se refinó a 4 subtipos moleculares

intrínsecos, que son Luminal A, Luminal B, enriquecido en Her2 y parecido a basal. Este clasificador constituye un avance considerable en el campo, no solo en la comprensión de la biología del CaMa, sino también porque es uno de los primeros ejemplos de la traducción de una herramienta genómica en un producto de utilidad directa en la atención clínica. El PAM50 da información de pronóstico personalizado en cuanto a riesgo de recurrencia y eficacia de la terapia en las pacientes con CaMa (Parker *et al.*, 2009).

A continuación, se describen brevemente las características de los subtipos moleculares de CaMa identificados en el perfil de expresión PAM50(ver figura 3).

1.3.1. Luminal A

Es el subtipo más frecuente en las pacientes con CaMa, ya que representa el 50-60% de los casos. El 85% de los tumores Luminal A presentan bajo grado tumoral (de Kruijf *et al.*, 2014), niveles altos de expresión de ER y PR, así como expresión moderada de genes relacionados con la proliferación celular (Yersal y Barutca, 2014). Adicionalmente, estos tumores presentan una alta expresión de GATA Binding Protein 3 (GATA3), Hepatocyte Nuclear Factor 3a (HNF3-a), Solute Carrier Family 39 Member 6 (SLC39A6), entre otros (Dressman *et al.*, 2001).
Los subtipos moleculares intrínsecos pueden aproximarse a través de las técnicas de tinción inmunohistoquímica (Cheang *et al.*, 2009; Lips *et al.*, 2013; Prat *et al.*, 2013; Maisonneuve *et al.*, 2014). Para que una muestra sea considerada como Luminal A por aproximación inmunohistoquímica debe cumplir los siguientes parámetros: ER+, PR>20%, HER2-, Ki67<20% y un grado histológico bajo (1 ó 2) (Cárdenas-Sánchez *et al.*, 2019). Las pacientes Luminal A son las que presentan un mejor pronóstico clínico (Howlader *et al.*, 2018). La tasa de recaída es significativamente menor que la de otros subtipos (<5%) (Shim *et al.*, 2014). La metástasis es común en hueso (70%), mientras que, en el hígado, los pulmones y el cerebro, ocurre en menos del 10% de las pacientes (Kennecke *et al.*, 2010).

1.3.2. Luminal B

Estos tumores son los segundos más frecuentes de CaMa, ya que se presentan en el 15-20% de los casos. Los cánceres tipo Luminal B son considerados como agresivos, ya que presentan un grado tumoral más alto (93% de los casos) y un peor pronóstico clínico,

comparado con las pacientes Luminal A (Creighton, 2012; de Kruijf *et al.*, 2014; Howlader *et al.*, 2018). Los tumores subtipo Luminal B presentan una alta expresión de genes asociados a la proliferación. Pueden presentar un nivel de expresión de Ki67 alto (mayor a 14% de células positivas), pero una baja expresión de ciertos genes pertenecientes a las vías asociadas a hormonas, como el PR, aunque no ER (Prat *et al.*, 2013). Adicionalmente, se ha reportado que expresan altos niveles de los genes Homólogo de oncogen viral de mieloblastosis (*v-MYB*), y Ciclina E1 (*CCNE1),* entre otros.

Los tumores Luminal B se aproximan por inmunohistoquímica, mediante los siguientes criterios: ER+, PR <20%, HER2- o HER2+, Ki67>20% y un grado histológico alto (Cárdenas-Sánchez *et al.*, 2019). Adicionalmente, se ha observado un aumento en la expresión de HER2, del Receptor 1 del factor de crecimiento de fibroblastos (FGFR1) y Receptor 1 del factor de crecimiento similar a la insulina (IGF1R) (Reis-Filho *et al.*, 2010). Aproximadamente el 20%-30% de los tumores HER2+ son clasificados en el subtipo Luminal B (Tran y Bedard, 2011).

1.3.3. Enriquecido en HER2

Este subtipo comprende el 15-20% de los casos. Se caracteriza por la alta expresión de genes contenidos en el amplicón *HER2* La positividad a la proteína HER2 se asocia a un comportamiento biológico y clínico más agresivo (Yersal and Barutca, 2014; Howlader *et al.*, 2018). Morfológicamente, los tumores son muy proliferativos; el 75% de ellos tienen un alto grado tumoral y más del 40% presentan mutaciones en p53 (Tsutsui *et al.*, 2003). En ausencia de tratamiento, los tumores HER2 positivos se asocian con un pronóstico desfavorable, aun comparándolo con el subtipo parecido a basal (Yersal y Barutca, 2014; Cedro-Tanda *et al.*, 2020).

1.3.4. Parecido a basal

Representa del 8-37% de todos los casos. Se asocia con un alto grado tumoral (93% de los casos) (de Kruijf *et al.*, 2014) e índices mitóticos y proliferativos altos (Ki67>50%) (Rakha *et al.*, 2009; Wang *et al.*, 2016). Adicionalmente, presentan una alta expresión de citoqueratinas basales (CK5/6 y CK17), proteínas asociadas a la proliferación y lamininas, entre otros. La mayoría de estos tumores tienen un comportamiento clínico altamente agresivo (probabilidad de supervivencia, (74% a tres años) (Howlader *et al.*, 2018)

y alta tasa de metástasis a hueso (37%) y pulmones (32%) (Wu *et al.*, 2017). El 62% de los tumores parecidos a basal, presentan mutaciones en p53 (Shah *et al.*, 2012). Además, dentro de este subtipo, se ubican cerca del 75% de las pacientes con por lo menos una mutación en el gen *BRCA1* (Foulkes *et al.*, 2004).

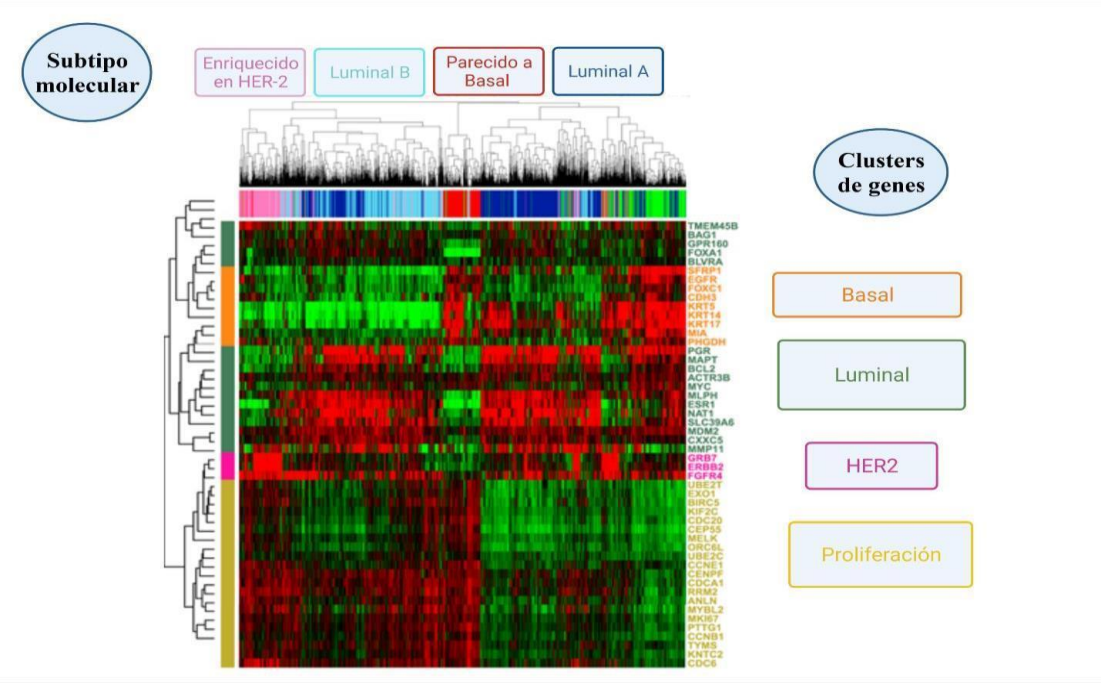

Figura 3. El cáncer de mama se clasifica en 4 subtipos moleculares intrínsecos Cada subtipo molecular (enriquecido en HER2, Luminal B, Parecido a basal y Luminal A) se caracteriza por firmas de expresión de genes particulares (*clusters* de genes), que se relacionan con un fenotipo celular Basal, Luminal o con procesos biológicos como la Proliferación. En rojo se muestran los transcritos sobreexpresados, mientras que en verde, los transcritos subexpresados. Modificado de (Prat *et al.*, 2013) en BioRender (BioRender.com).

1.3.5. Predicción de sobrevida y respuesta a terapia utilizando los subtipos moleculares

El principio sobre el cual se basa la capacidad predictiva del PAM50 está en que cada uno de los subtipos moleculares que identifica presenta un curso clínico diferente. En un estudio donde se estratificó a las pacientes con base en diferentes etapas clínicas (I- IV) en los subtipos moleculares se observó que las pacientes Luminal A y B presentan las mejores tasas de OS (97% a 45% a 4 años) respecto a los subtipos enriquecidos en Her2 (97% a 34% a 4 años) y parecidos a basal (95%-11% a 4 años) (Howlader *et al.*, 2018). Sin embargo, existe una mayor tasa de muerte en pacientes con subtipos Luminales B después de 5 años, siendo igual a la mortalidad asociada al subtipo parecido a basal.

En la actualidad, los factores clínico-patológicos y moleculares se han utilizado para evaluar el riesgo de muerte y la respuesta a tratamiento en pacientes con CaMa. Los biomarcadores utilizados en la práctica clínica son: el receptor de estrógenos (ER); el receptor de Progesterona (PR) y el HER2. Dependiendo de los niveles de expresión de estos marcadores, es posible evaluar el pronóstico y el tratamiento de las pacientes.

Como se mencionó previamente, Perou et al. (2001) fueron los primeros en describir los cuatro subtipos moleculares intrínsecos de CaMa, que se identificaron originalmente mediante la expresión genética diferencial de 1.753 genes. Estudios subsecuentes confirmaron la existencia de estos subgrupos y mostraron que son capaces de predecir la OS y la RFS (Sorlie, T., et al., 2001). Posteriormente, la lista de genes se limitó a los 50 genes clasificadores del ensayo PAM50, manteniendo la clasificación biológica y la precisión pronóstica de subtipos moleculares intrínsecos (van ’t Veer, L., et al., 2002).

El clasificador PAM50 basado en el método de Reacción en Cadena de la Polimerasa (PCR) es pronóstico en mujeres posmenopáusicas RE positivas tratadas solamente con tamoxifeno (Nielsen, T., et al., 2010). Además, este clasificador funciona como prueba pronóstica y predictiva del beneficio de la terapia con tamoxifeno en mujeres premenopáusicas tratadas con terapia hormonal adyuvante (Chia, S. K., et al., 2012). Los subtipos de PAM50 también predicen el beneficio del tratamiento con quimioterapia adyuvante (Martín, M., et al., 2013; Cheang, M. C. U., et al., 2012; Jørgensen, C. L. T., et al, 2014) y la respuesta patológica completa (Parker, J. S., et al., 2009). Adicionalmente, estos estudios mostraron que el PAM50 es un mejor predictor de la eficacia de la terapia y del pronóstico que los sustitutos aproximados basados en la IHQ (Nielsen, T., et al., 2010; Chia, S. K., et al., 2012; Martín, M., et al., 2013). En mujeres posmenopáusicas con CaMa temprano RE+, tratadas con tamoxifeno o tamoxifeno seguido de anastrozol, el índice de riesgo de recurrencia basado en PAM50 (ROR) es un predictor con validez clínica para predecir el riesgo de recurrencia en estas pacientes. Otro ensayo clínico con 2,137 pacientes mostró que el índice ROR es capaz de predecir el riesgo de recurrencia tardía (de 5 a 10 años después del diagnóstico) a distancia más eficientemente los marcadores clínicos convencionales (estado de los ganglios, tamaño del tumor, grado, edad y tratamiento) en todos los subgrupos de CaMa (Sestak, Ivana, et al., 2015). Por lo tanto, la clasificación por PAM50 y el índice ROR pueden usarse para separar a los pacientes en

grupos de riesgo que beneficiarse potencialmente de una terapia anti-hormonal a corto plazo o prolongada más allá de los 5 años de tratamiento.

Por su parte, Oncotype DX® Recurrence Score® consiste en una prueba de 21 genes que predice la probabilidad de beneficiarse de la quimioterapia y el riesgo de recurrencia distante a 10 años que ayuda a guiar las decisiones de tratamiento adyuvante en pacientes con CaMa previamente tratadas con terapia anti-hormonal en etapas tempranas (Albanell, J., et al., 2012).

Por lo tanto, el uso de marcadores moleculares, en combinación con los marcadores tradicionales, reduce significativamente los costos económicos de tratamiento. Esto se logra al estratificar adecuadamente a las pacientes que realmente responderán a determinados fármacos. Así, se evita el sobretratamiento de las pacientes, lo cual evita morbilidades y gastos innecesarios (Steuten *et al.*, 2019). Adicionalmente, el uso de marcadores moleculares de pronóstico permite identificar a las pacientes que tendrán un curso clínico más agresivo, por lo que auxilia al médico tratante para tomar decisiones sobre tratamientos más agresivos en esas pacientes. En México, un estudio mostró que el uso de Oncotype DX® Recurrence Score®, cambió el 32% de las decisiones de tratamiento que habían sido tomadas previamente (Bargallo et al., 2015), previniendo el sobretratamiento y los perjuicios asociados a este.

A pesar de los beneficios que generan, los paneles de marcadores presentan todavía algunas desventajas. Una de ellas es el alto costo y el moderado beneficio −ambos medidos en términos de calidad de vida y años de sobrevida− que reportan (Steuten *et al.*, 2019). En este sentido, las pruebas con marcadores individuales (un solo gen) han mostrado ser menos específicas, pero con menores costos asociados (Bargallo *et al.*, 2015).

Por lo tanto, se hace necesario identificar marcadores individuales efectivos de pronóstico clínico y predictivos de respuesta a tratamiento, más accesibles en costo a la mayoría de las pacientes. Adicionalmente, es importante que estos marcadores funcionen en distintos modelos tumorales, para así captar al mayor número de pacientes posible.

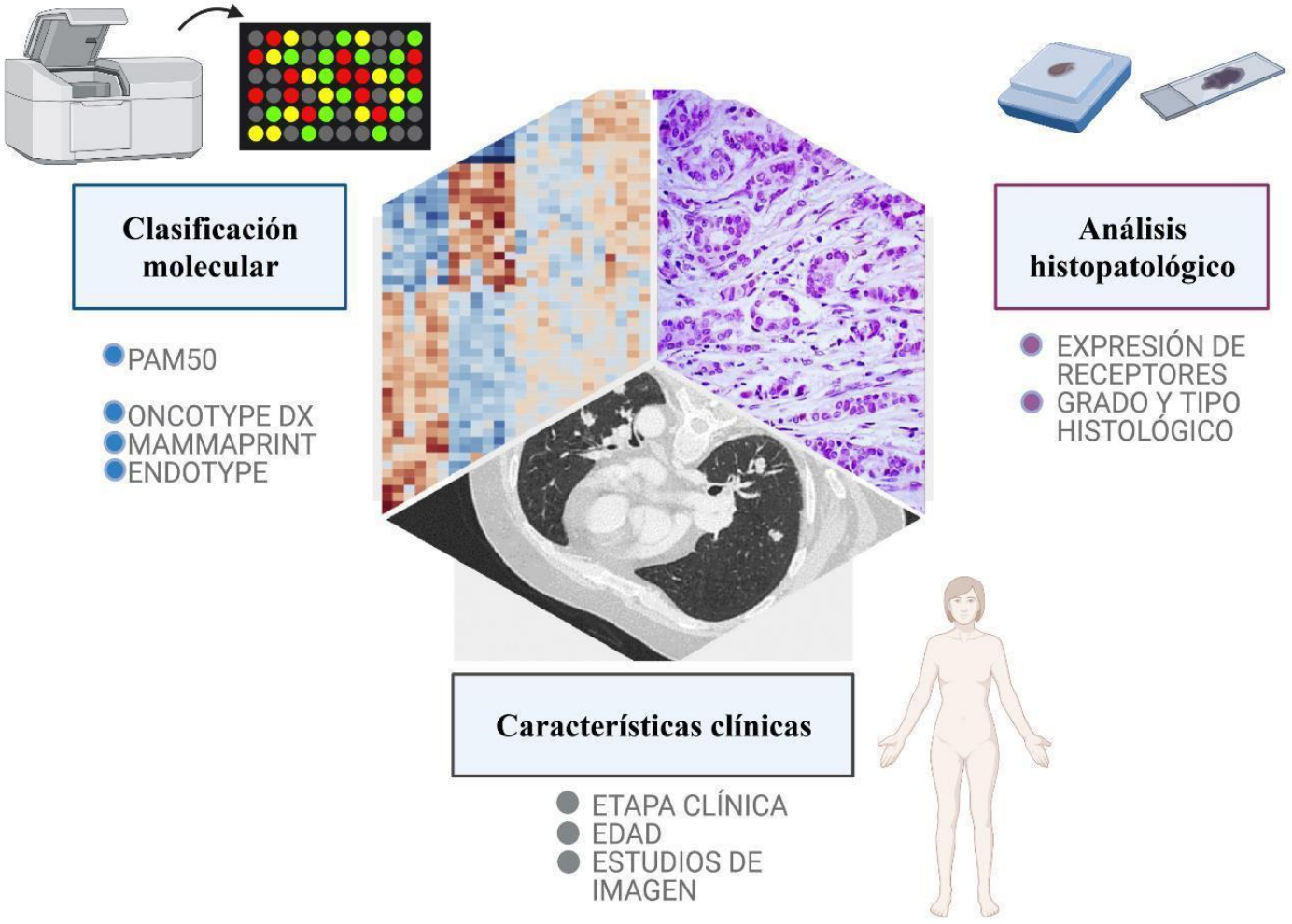

Figura 4. Diagnóstico integral de las pacientes con CaMa. En la actualidad, el diagnóstico se enfoca no solamente a las características clínicas (etapa clínica, edad estudios de imagen) e histopatológicas (grado, tipo histológico) de la paciente, sino también a la clasificación molecular basada en el PAM50. La expresión de receptores se utiliza para clasificar y dar tratamiento adecuado a las pacientes. Herramientas moleculares auxiliares en el pronóstico clínco son ONCOTYPE DX y MAMMAPRINT. Modificado de (Russnes *et al.*, 2017) con BioRender (BioRender.com).

1.4. El transcriptoma y los RNAs largos no codificantes

La caracterización del CaMa a nivel de transcripción génica refleja aspectos importantes de la biología de estos tumores y ha tenido un efecto importante en la práctica clínica. Sin embargo, estos clasificadores se basan en el perfil de expresión de RNAs mensajeros. Éstos constituyen solamente el 36.7% del transcriptoma humano; el resto está constituido por RNAs no codificantes (ncRNAs), los cuales tienen importantes funciones regulatorias (Derrien *et al.*, 2012).

Los ncRNAs se clasifican con base en su longitud. Los de tamaño menor a 200 nucleótidos se denominan ncRNAs pequeños. En este grupo encontramos a los microRNAs (miRNAs), RNAs pequeños nucleares (snRNAs), RNAs pequeños nucleolares (snoRNAs), RNAs asociados a Piwi (piRNAs) y RNAs pequeños derivados de tRNAs (tsRNAs) (Watson, Belli and Di Pietro, 2019). Por convención, los RNAs no codificantes cuya longitud excede los 200 nucleótidos, pertenecen a la categoría de RNAs largos no codificantes (lncRNAs). Los lncRNAs poseen una serie de características que los diferencian de los mRNAs: en general, los genes que transcriben lncRNAs presentan menos exones (pero de mayor longitud) que los mRNAs, sus promotores muestran marcas de histonas distintas, algunas clases son menos conservadas evolutivamente, su expresión es tejido-específica y su localización celular es variada (Derrien *et al.*, 2012). Los lncRNAs comparten algunos rasgos con los mRNAs, como que son mayoritariamente transcritos por la RNA polimerasa II (RNA pol II), sus transcritos son procesados para la adición de caperuza de 7-metil-guanosina (m7G) y en ambos grupos, la mayoría presentan cola de Poli adeninas (PolyA) (Derrien *et al.*, 2012; Derrien, Guigó and Johnson, 2012).

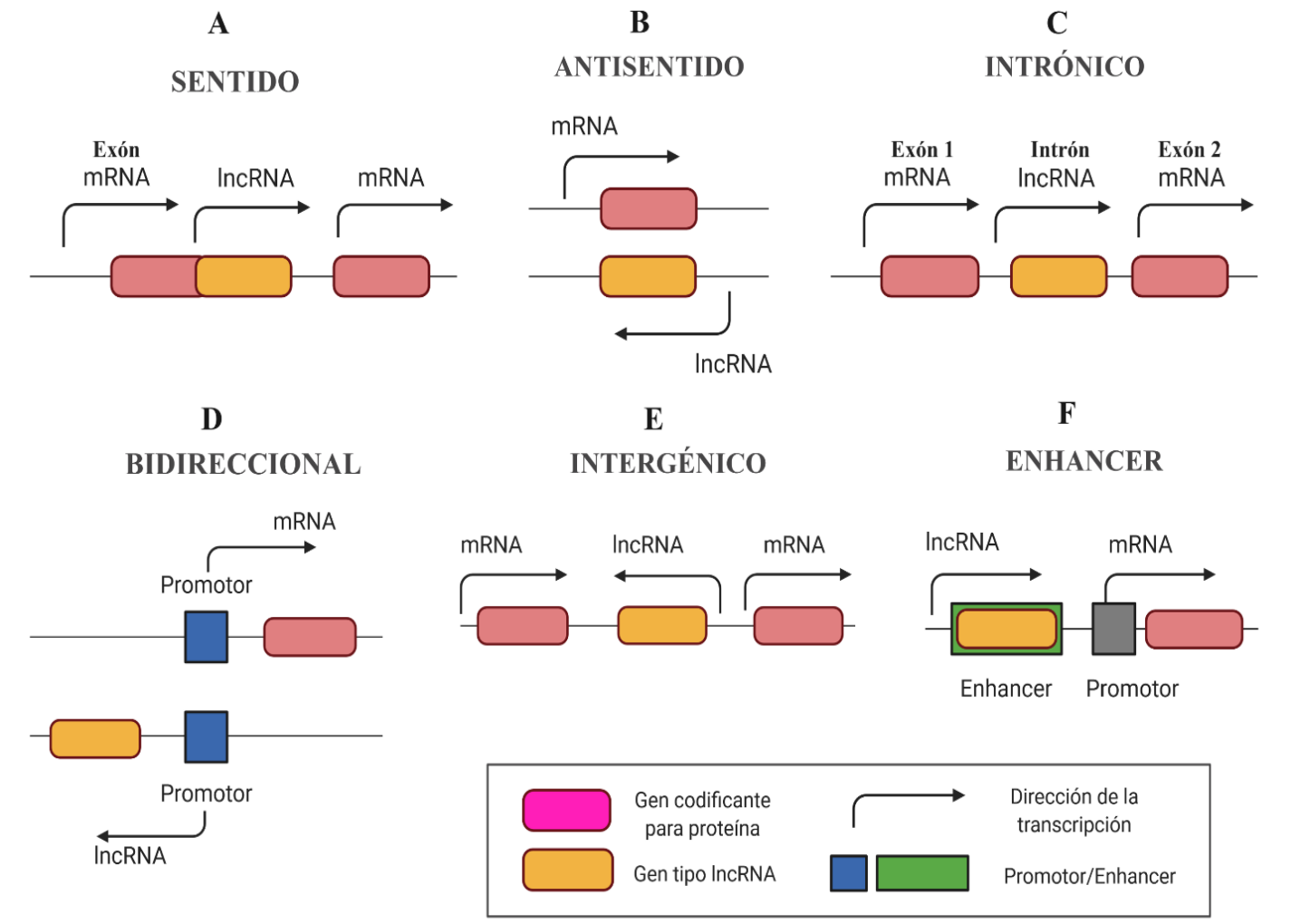

Figura 5. Clasificación de los lncRNAs con base en su posición genómica. **A)** Los lncRNAs sentido son transcritos que se ubican dentro de la misma cadena que el gen codificante para proteína y se pueden sobreponer a uno o más exones de este. **B)** Los transcritos de los lncRNAs antisentido se localizan en la cadena complementaria al gen codificante para proteína, y complementan parcial o totalmente dicha secuencia. **C)** Los lncRNAs intrónicos se generan a partir de los intrones de los genes codificantes. **D)** Los lncRNAs bidireccionales comparten la misma región promotora con los genes codificantes para proteínas, pero se transcriben en la dirección opuesta. **E)** Los lncRNAs intergénicos (lincRNAs) se localizan entre los genes codificantes para proteínas y se transcriben de forma independiente. **F)** Los lncRNAs *enhancers* o potenciadores (eRNAs) se producen a partir de las regiones *enhancers* de los genes codificantes para proteínas. Creado en BioRender (BioRender.com).

La diferencia principal entre los lncRNAs y los mRNAs es que los primeros no muestran potencial codificante para proteínas completas y funcionales (Housman y Ulitsky, 2016), aunque pueden presentar marcos de lectura abiertos cortos(Niazi y Valadkhan, 2012). Adicionalmente, hay algunos reportes de péptidos pequeños codificados por lncRNAs,que, aunque son traducidos, no dan lugar a proteínas completas (Andrews y Rothnagel,2014; Nelsonet al., 2016). Otra de las diferencias entre los mRNAs y los lncRNAs es su localizacióncelular.MientrasquelosmRNAs se restringen principalmente al citoplasma celular, los lncRNAs se encuentran en el núcleo, el citoplasma y en algunos casos en la mitocondria (Chen, 2016).

Se han estudiado los mecanismos y funciones celulares de los lncRNAs. Entre las funciones que se han caracterizado se encuentran procesos de regulación a todos los niveles: remodelación de la cromatina y marcaje de histonas, de actividad transcripcional, maduración, transporte y decaimiento de RNAs, señalización, regulación post-traduccional y síntesis de proteínas.

En la siguiente sección se describirán las funciones moleculares principales de los lncRNAs tanto a nivel de núcleo como de citoplasma.

1.4.1. Funciones generales de los lncRNAs en núcleo

Las funciones de los lncRNAs en el compartimento nuclear se realizan mediante las interacciones con la cromatina. Estas interacciones pueden clasificarse en dos tipos: (1) interacciones lncRNA-cromatina que regulan la expresión génica mediante modificaciones de histonas o del DNA sin afectar la estructura tridimensional de la cromatina y (2) lncRNAs que tienen como blanco elementos regulatorios y modulan la expresión génica iniciando o manteniendo horquillas de cromatina entre potenciadores (*enhancers* en inglés) y promotores en un espacio tridimensional (Jarroux, Morillon y Pinskaya, 2017).

En el primer grupo se encuentran algunos lnRNAs prototípicos y bien estudiados, como HOX Antisense Intergenic RNA (HOTAIR), HOXA Distal Transcript Antisense RNA (HOTTIP), Imprinted Maternally Expressed Transcript (H19) y Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT-1). Estos cuatro lncRNAs regulan la expresión génica mediante la modificación de histonas y de la arquitectura de la cromatina. En el segundo conjunto de lncRNAs, se encuentran X-inactive specific transcript (XIST), y Nuclear Enriched Abundant Transcript 1 (NEAT1), que cambian la estructura tridimensional de la cromatina para regular la expresión a nivel regional.

HOTAIR es un lncRNA cuyo gen se localiza en la cadena antisentido del locus 12q13.13, entre los genes *HOXC11 y HOXC12* (Rinn *et al*., 2007). La región 5´de HOTAIR interactúa directamente con el complejo Polycomb 2 (PRC2), en un mecanismo que silencia la expresión de los genes blanco de dicho complejo por medio de la trimetilación de histonas (particularmente, de la marca H3K27me3). Por otro lado, en la región 3´ HOTAIR interacciona con LSD1 (Desmetilasa de histonas específica de lisinas) y el complejo REST/CoREST (Factor de represión transcripcional RE-1), lo cual favorece la desmetilación de la histona 3 (marcas H3K4me2, H3K27me3) silenciando la actividad transcripcional en los genes blanco (Tsai et al., 2010). Por lo tanto, a nivel de núcleo, HOTAIR actúa

como un andamio molecular de complejos reguladores de la expresión génica, como PRC2 y LSD1/ REST/CoREST.

El lncRNA HOTTIP (HOXA transcript at the distal tip) es un transcrito que presenta diferentes mecanismos de acción. En el mecanismo más clásico, HOTTIP se acerca a sus blancos transcripcionales por medio de la formación de bucles de cromatina, se une directamente a la proteína WDR5 y posteriormente recluta al complejo MLL en el locus HOXA. Finalmente, promueve la metilación de histonas (marca H3K4) de los genes HOXA (Wang et al., 2011).

H19 es un lncRNA que presenta funciones tanto a nivel de núcleo como de citoplasma. En el núcleo, un ejemplo de mecanismo constituye el reclutamiento de EZH2, que, a su vez, promueve el silenciamiento de los genes pro-apoptóticos B IK (BCL-2 interacting killer) y NOXA (phorbol-12-myristate-13-acetate-induced protein 1) (Si et al., 2016a).

MALAT-1 (Metastasis Associated Lung Adenocarcinoma Transcript 1) es un lncRNA que fue originalmente relacionado con la regulación del corte y empalme (splicing) al-ternativo al interactuar directamente con factores de splicing y modular su distribución a las espículas nucleares. MALAT-1 regula el del corte y empalme alternativo de algunos pre-mRNAs al controlar los niveles funcionales (fosforilados) de los factores de splicing (Tripathi et al., 2010). Adicionalmente, MALAT-1 coordina la relocalización de genes de los sitios silenciados por PRC2 a sitios transcripcionalmente activos, respondiendo a la adición de suero. Estos mecanismos son mediados por MALAT-1, al unirse directamente con EZH2 y SUZ12 y, por lo tanto, regula la expresión de los genes blanco de PRC2 (Yang et al., 2011; Wang et al., 2015).

XIST fue uno de los primeros lncRNAs en ser descritos (Borsani et al., 1991; Brown et al., 1991) y es el principal factor que genera el mecanismo de compensación de dosis del cromosoma X. XIST coordina la inactivación de uno de los cromosomas X al azar y expande esa inactivación a lo largo de todo el cromosoma (Chaumeil et al., 2006; No-zawa et al., 2013). XIST recluta complejos remodeladores de la cromatina, metiltransfe-rasas de DNA, proteínas asociadas a RNA (Chu et al., 2015) y represores transcripcio-nales al cromosoma X que será inactivado (Engreitz et al., 2013; Simon et al., 2013). Durante este proceso, XIST produce un reacomodo tridimensional del cromosoma X, compactándolo

a gran escala (Chaumeil et al., 2006; Nozawa et al., 2013; McHugh et al., 2015) y anclándolo completamente a la lámina nu- clear (Chen et al., 2016). Por último, XIST regula el mantenimiento y la estabilidad del silenciamiento en el cromosoma X, estableciendo un compartimiento de silenciamiento tridimensional (Nora et al., 2012).

NEAT1 es un transcrito que pertenece al grupo de los RNAs de arquitectura celular (archRNAs). NEAT1 es fundamental y necesario para el ensamblaje y estructura de las paraspeckles nucleares (Clemson et al., 2009), que son estructuras sub-nucleares que se encuentran en el espacio inter-cromatínico de las células de mamíferos.
Las paras-peckles están conformadas por la interacción entre NEAT1 y diversos complejos proteicos, que pertenecen a la familia de DBHS (Drosophila Behavior Human Splicing, por sus siglas y nombre en inglés): P54NRB/NONO, PSPC1 y PSF/SFPQ (Fox and La-mond,2010). NEAT-1 secuestra a P54NRB/NONO, PSPC1 y PSF/SFPQ a través de la unión directa en tres dominios del RNA (Murthy and Rangarajan, 2010). A su vez, las proteínas PSPC1 y P54NRB/NONO regulan la edición Adenosina-Inosina (A-I) de mRNAs en el núcleo (Zhang and Carmichael, 2001).

La edición A-I es una marca de retención nuclear, mientras que la ausencia de esta edi-ción provoca la exportación al citoplasma. De esta manera, NEAT-1 regula la expresión de mRNAs editados. Por otro lado, se ha mostrado que P54NRB/NONO, PSPC1 y PSF/SFPQ actúan como factores transcripcionales en diferentes blancos a lo largo del genoma. Al secuestrarlos, NEAT-1 también modula indirectamente la expresión de di-chosgenes(Hiroseetal.,2014).

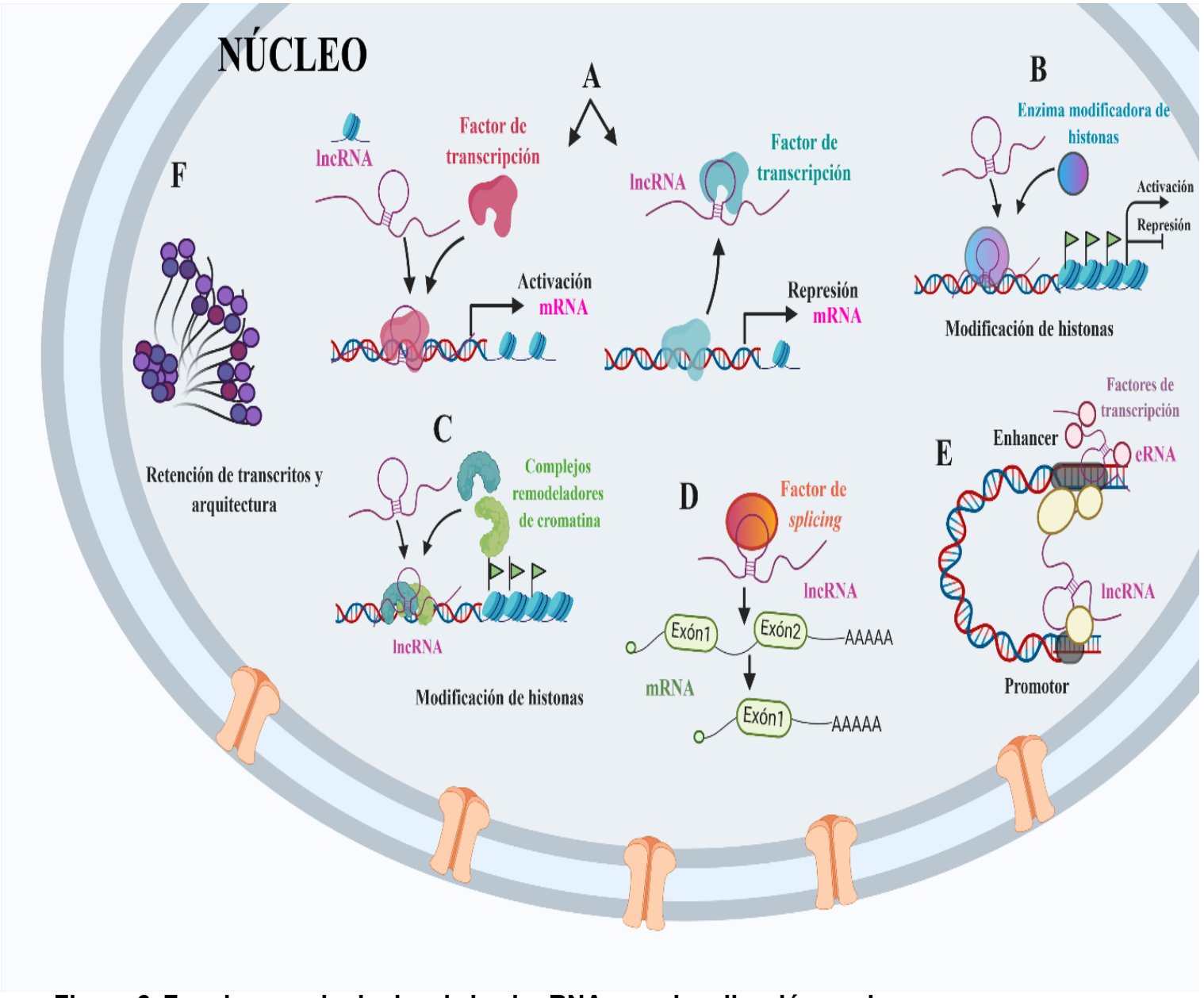

Figura 6. Funciones principales de los lncRNAs con localización nuclear.

A) Los lncRNAs se unena factores de transcripción para activar o reprimir la expresión génica.

B) Los lncRNAs interactúan con enzimas modificadoras de histonas para activar o reprimir la transcripción génica.

C) Muchos lncRNAs regulan la transcripción actuando como andamios moleculares para favorecer las interacciones proteína-proteína de los complejos remodeladores de cromatina.

D)Los lncRNAs interaccionan con factores de splicing para regular el procesamiento alternativo de los mRNAs.E)Los trans-critos de enhancers pueden generar lncRNAs enhancers (eRNAs) o lncRNAs asociados a enhancers; los primeros pueden captar factores de transcripción en el enhancer,mientras que los segundos regulan la transcripción en cis y en trans. Creado en BioRender (BioRender.com).

1.4.2. lncRNAs en el citoplasma: funciones múltiples

Los lncRNAs pueden ejercer funciones a nivel de citoplasma;inclusive hay algunos ejemplos de lncRNAs nucleares que se exportan al citoplasma para ejercer diferentes actividades. Algunos de estos transcritos se describirán en la siguiente sección.

Como ya se mencionó anteriormente, HOTAIR es un lncRNA que participa en la regulación de la expresión de cientos de genes en el núcleo, sin embargo, también actúa a nivel postranscripcional en el citoplasma celular. Se ha probado que induce la proteólisis de Ataxina 1 y Snurportina 1 mediada por marcas de ubiquitina, al interaccionar con Dzip3, and Mex3b (Ubiquitinas ligasa E3) (Yoon *et al.,* 2013). Adicionalmente, se ha reportado que HOTAIR interactúa con MDM2, incrementando así la estabilidad del receptor de Andrógenos (AR), al prevenir su degradación (Zhang *et al*., 2015).

Por otro lado, se ha demostrado que MALAT-1 se exporta al citoplasma y actúa como una esponja de miRNAs, como miR-125b (Han *et al*., 2013) y miR-101b en diferentes contextos celulares.

En el citoplasma, H19 presenta diferentes mecanismos de acción. Se ha mostrado que actúa como esponja para el microRNA let-7 (Kallen *et al.,* 2013) y de miR-675, que silencia al supresor tumoral PTEN. Además, H19 funciona como precursor de miRNAs (Cai y Cullen, 2007) como miR-675-5p y miR-675-3p, que, a su vez, tienen activida-des regulatorias relevantes (Vennin *et al.,* 2015).

El lincRNA-p21es otro ejemplo de un transcrito que presenta funciones regulatorias a nivel de núcleo como de citoplasma. A nivel nuclear regula la transcripción global de genes blanco de PRC2 (Dimitrova *et al*., 2014). En citoplasma, el lincRNA-p21 actúa como inhibidor de la traducción. En este mecanismo, el lincRNA-p21 se acumula en el citoplasma y se forman híbridos parciales a través de múltiples regiones de complementariedad con los mRNAs blanco, como CTNNB1 (que codifica a la proteína β-catenina) y JUNB (que codifica el factor de transcripción JUNB). Las regiones de separación parcial de bases en los segmentos del dúplex mRNA-lncRNA provocan el reclutamiento de los supresores traslacionales DDX6 (DEAD-Box Helicase 6) y FMRP (FMRP Translational Regulator 1), disminuyendo la tasa de traducción de CTNNB1 y JUNB (Chu y Rana, 2006; Yoon *et al.,* 2012).

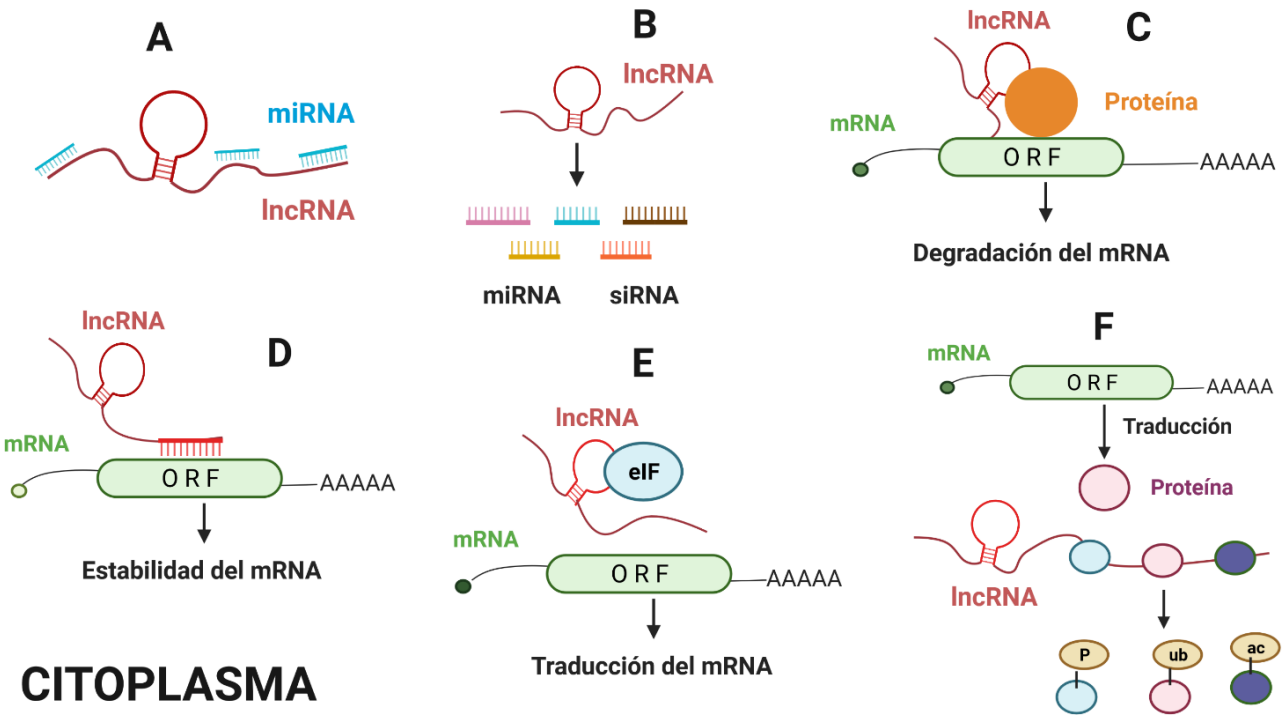

Figura 7. Mecanismos moleculares de los lncRNAs en el citoplasma.

A) Los lncRNAs actúan como esponjas de miR-NAs, secuestrando estas moléculas lejos de sus mRNAs blanco, y favoreciendo la regulación de la expresión de los genes. B) Los lncRNAs pueden ser precursores de miRNAs y siRNAs. C) Mediante la unión a proteínas, los lncR- NAs están implicados en la degradación de los mRNAs. D)Los lncRNAs se unen directamente al mRNA y regulan su estabilidad. E) Los lncRNAs se unen al complejo de iniciación de la traducción eucariota eIF para regular la traducción del mRNA. F) Los lncRNAs interactúan con las proteínas y controlan su fosforilación, acetilación y ubiquitinación, a nivel post- traduccional. Creado en BioRender (BioRender.com).

El lncRNA Growth Arrest Specific 5 (GAS5) es un transcrito citoplasmático que presenta una doble función en el citoplasma. En primer lugar, funciona como señuelo del receptor de glucocorticoides (GR). GAS5 contiene elementos de respuesta a glucocorticoides (GRE), lo cual promueve el secuestro de GR mediado por GAS5 en el citoplasma. Este mecanismo impide la movilización de GR al núcleo y suprime el programa transcripcional mediado por GR (Mourtada-Maarabouni et al., 2009). Por otro lado, GAS5 inhibe la traducción de la proteína MYC (c-Myc) sin afectar sus niveles de mRNA ni a la estabilidad de la proteína. GAS5 es reclutado en el complejo de iniciación de la traducción por medio de la unión directa al factor eIF4E. Se observó que c-Myc y GAS5 interactúan de manera

directa en la fracción no polisómica, lo cual sugiere que GAS5 disminuye la eficiencia de la iniciación de la traducción del mRNA c-Myc (Hu, Lou y Gupta, 2014).

El lncRNA citoplasmático half STAU1-binding site RNAs (1/2-sbsRNA) participa en el mecanismo de decaimiento de mRNAs mediado por Staufen1 (STAU1). Este mecanismo se genera a partir del apareamiento de bases entre un elemento Alu presente en la región 3′ no traducida (UTR) y una secuencia Aluparcialmente complementaria en el 1/2-sbsRNA. Como resultado, se forma un RNA de doble cadena, que a su vez recluta a la proteína de unión a RNA STAU1. Dicha proteína promueve la degradación de los mRNA blanco a través del mecanismo de decaimiento de mRNAs mediado por STAU1, que adicionalmente implica a la RNA helicasa UPF1 (Gong y Maquat, 2011).

El lncRNA Hepatocellular Carcinoma Up-Regulated Long Non-Coding RNA (HULC) actúa como un RNA señuelo o esponja para el miR-372 y, por tanto, potencia la traducción del mRNA Protein Kinase CAMP-Activated Catalytic Subunit Beta (PRKACB) el cual a su vez es blanco de inhibición por miR-372 (Wang *et al.,* 2010).

En los últimos años se han caracterizado una serie de funciones novedosas para los lncRNAs citoplásmicos. Un ejemplo de estos mecanismos es Long Intergenic Non-Protein Coding RNA 1139 (LINK-A), que está involucrado en la transducción de señales desencadenada por el factor de crecimiento epitelial ligado a la heparina (HB- EGF). LINK-A recluta a la cinasa BRK y al dímero receptor de EGF- glicoproteína de membrana NMB (EGFR-GPNMB), promoviendo la activación de BRK. LINK-A también interactúa con la cinasa Leucine Rich Repeat Kinase 2 (LRRK2), generando a su vez el cambio conformacional, la fosforilación y finalmente la activación de BRK. Esta señalización genera la estabilización del factor inducido por hipoxia 1α (HIF1α) (Lin *et al.,* 2016).
Adicionalmente,otro estudio demostró que LINK-A promueve la interacción de fosfatidil inositol trifosfato (PIP3) con la cinasa AKT y por lo tanto, la activación de esta proteína (Lin *et al.,* 2016).

1.5 Los lncRNAs y su asociación con el cáncer de mama

Algunos estudios han sugerido que los lncRNAs están involucrados en la carcinogénesis de distintos tumores. Además de que ciertos lncRNAs se encuentran aberrantemente expresados, también se ha mostrado que pueden funcionar como oncogenes o genes

supresores de tumores (Yan *et al.,* 2015). En algunos casos, un mismo lncRNA puede comportarse como oncogén y supresor de tumor, dependiendo del contexto celular (Aprile *et al.,* 2020).

En el CaMa se han caracterizado alteraciones en la expresión y función de algunos lncRNAs. Estas alteraciones se encuentran relacionadas con el desarrollo, agresividad y pronóstico clínico en pacientes con CaMa.

Los lncRNAs regulan procesos celulares asociados a la carcinogénesis como el ciclo celular, la diferenciación y la invasividad, por lo que el estudio de la sobreexpresión de lncRNAs ha cobrado un papel relevante en la actualidad. En este sentido, se realizó un estudio que analizó los niveles de expresión de lncRNAs cuya transcripción se altera en distintos tumores humanos. En dicho estudio los autores mostraron que existen 1,151 lncRNAs cuya expresión se altera en CaMa, comparado con tejido normal (Yan *et al*., 2015). Adicionalmente, otros estudios han mostrado la relevancia clínica de las alteraciones en la expresión de algunos lncRNAs, mismos que se describirán en la siguiente sección.

Otra fuente reportó una firma de expresión basada en la expresión de 12 lncRNAs, la cual puede clasificar a las pacientes en grupos de alto o de bajo riesgo de muerte y recurrencia, así como predecir la OS de pacientes con CaMa. También se ha demostrado que la sobreexpresión de 4 lncRNAs está asociada con una pobre OS, y que la sobreexpresión de 9 lncRNAs está asociada a un riesgo mayor de recurrencia (Zhou *et al.,* 2016).

Otros lncRNAs cuyas alteraciones en la expresión se han asociado significativamente con el pronóstico clínico de las pacientes con CaMa son los siguientes:

LINC00978, cuya sobreexpresión se relaciona con una pobre OS comparada con las pacientes que presentan subexpresión del mismo y que además es un marcador independiente de los parámetros clínicos tradicionales (Deng et al., 2016) y. FGF14-AS2, cuya subexpresión predice una menor OS y aumenta el riesgo de progresión (Yang, Liu, *et al*., 2016).

1.5.1 DSCAM-AS1

DSCAM-AS1 es un lncRNA propuesto como un marcador de diagnóstico y resistencia a tratamiento anti-hormonal en CaMa (Tarighi *et al.*, 2021). DSCAM-AS1 fue descrito como un lncRNA regulado transcripcionalmente por REα en CaMa. Está involucrado en la progresión tumoral y en la resistencia a tamoxifén (Niknafs *et al.*, 2016). Se ha demostrado que DSCAM- AS1 actúa como esponja de miR-137 y regula el Epidermal Growth Factor Receptor Pathway Substrate 8 (EPS8), lo cual a su vez promueve la proliferación celular e inhibe la apoptosis en CaMa resistente a tamoxifén (Ma *et al.*, 2019). Mediante el reclutamiento de miR-204-5p y el aumento de la expresión de RRM2 (Liang *et al.*, 2019) se ha demostrado que DSCAM-AS1 promueve la regulación positiva de sus propios moduladores−ERα y FOXA1, y se activa transcripcionalmente por super potenciadores (*super-enhancers*) regulados por FOXA1; además, promueve la progresión del CaMa al interactuar con YBX1 y regular a su vez la expresión de FOXA1 y REα (Zhang *et al.*, 2020).

1.5.2. H19

H19 es precursor de diversos miRNAs. En células de CaMa T47D y MCF7H19 genera al miR-675-5p y al miR-675-3p, y contribuye al desarrollo de esta enfermedad. (Cordero *et al.*, 2015; Zhai *et al.*, 2015). La expresión ectópica de H19/miR-675 incrementa la agresividad de las células de CaMa. Esta expresión anormal aumenta la proliferación la migración, el crecimiento tumoral y la metástasis en este modelo (Matouk *et al.*, 2014; Vennin *et al.*, 2015; Wang *et al.*, 2017). Adicionalmente, H19 funciona como esponja de varios miRNAs, como miR- 200b/c y let-7b y regula la expresión de los genes blanco correspondientes, como *GIT2* y *CYTH3*. Esto a su vez, promueve la migración de las células de CaMa a través de la transición epitelio-mesénquima (EMT) y la diseminación de las células de CaMa (Zhou *et al.*, 2015). Por otro lado, la sobreexpresión de H19 influye en la regulación mediada por let-7b de ciertos genes promotores de la metástasis, incluyendo el oncogén *c-Myc* (Yan *et al.*, 2015).

1.5.3. MALAT1

El papel del lncRNA MALAT1 en CaMa sigue en controversia (Chen, Zhu y Jin, 2020). Algunos estudios han sugerido su papel como inductor de metástasis(Li *et al.*,2018), de la progresión (Feng *et al.*, 2016), la agresividad (Bamodu *et al.*, 2016), y promotor de la migración celular en CaMa (Chou *et al.*, 2016), entre otros. Sin embargo,otros estudios han mostrado que MALAT-1 inhibe la metástasis al demostrar que este transcrito se une a TEA domain transcription factor 1 (TEAD) −un factor de transcripción pro-metastásico y lo desactiva suprimiendo en consecuencia la metástasis del CaMa (Kim *et al.*, 2018).

1.5.4. HOTAIR

HOTAIR es uno de los primeros lncRNAs, cuya expresión aberrante fue asociada con la progresión y metástasis del CaMa (Gupta *et al.*, 2010; Sørensen *et al.*, 2013).
HOTAIR interactúa con algunos actores de las principales vías moleculares implicadas en el desarrollo de CaMa,como las vías de los receptores hormonales ER y PR. El estradiol regula la expresión de HOTAIR en células ER+ de CaMa por medio de diversos elementos de respuesta a estrógenos (EREs) en su promotor (Bhan *et al.*, 2013), además es regulado por agonistas de ER (Bhan *et al.*, 2014) se ha mostrado que HOTAIR, junto con BRCA1, se une a la subunidad de EZH2 de PRC2, coordinando la regulación epigenética del cromosoma (Wang *et al.*, 2013). Por otra parte, se tiene que la combinación de la sobreexpresión de HOTAIR y el estado de metilación genómico representa un importante predictor de mal pronóstico clínico (Lu *et al.*, 2012; Peng *et al.*, 2017).
La sobreexpresión de HOTAIR está involucrada en la respuesta a radioterapia (Zhou *et al.*, 2017; Xuguang Hu *et al.*, 2019), quimioterapia (Lu *et al.*, 2018; Tang *et al.*, 2019), terapia anti-hormonal y trastuzumab (Chen *et al.*, 2019). Dada la importancia biológica que tiene, se ha pro-puesto a HOTAIR como un potencial blanco terapéutico en pacientes con CaMa (Li *et al.*,2019; Ren *et al.*, 2019).

1.5. El RNA largo no codificante FAM83H-AS1

1.5.2. Aspectos generales y localización

El RNA FAM83H antisentido 1 (FAM83H-AS1) es un transcrito humano codificado por la región 8q24.3 y que mide 2,737 pares de bases (pb). FAM83H-AS1 está constituido por cuatro exones y se ha demostrado que no posee potencial codificante. Según la base de datos ENSEMBL, el gen FAM83H-AS1 presenta tres variantes transcripcionales reportadas hasta la fecha (ver figura 8).

1.5.3. Papel de FAM83H-AS1 en cáncer

En un estudio, se mostró que FAM83H-AS1 se encuentra asociado a una menor OS, particularmente en el subtipo luminal A, y es un marcador pronóstico independiente en estos tumores (Yang *et al.*, 2016). El silenciamiento del lncRNA FAM83H-AS1 está asociado con una disminución en la proliferación celular, apoptosis, migración e invasión en células de cáncer de pulmón (Zhang *et al.*, 2017).

FAM83H-AS1 es un lncRNA cuya expresión aberrante desregula vías relacionadas al cáncer como proliferación celular, migración, invasión y muerte celular en células de cáncer de pulmón, colorrectal (S. Lu *et al.*, 2018), glioma (Bi *et al.*, 2018), vejiga (Shan *et al.*, 2019), ovario (Dou *et al.*, 2019; Gong and Zou, 2019) y cérvix (Barr *et al.*, 2019). A nivel molecular, un reporte mostró que la señalización por MET/EGFR se regula mediante la acción de FAM83H-AS1 (Zhang *et al.*, 2017). Otro estudio mostró que FAM83H-AS1 regula el silenciamiento de *CDKN1A* por medio de la unión con EZH2 en células de glioma (Bi *et al.*, 2018). FAM83H-AS1 se encuentra significativamente sobre- expresado en muestras de cáncer de ovario en comparación con tejido normal. Una alta expresión del FAM83H-AS1 se asocia con una menor supervivencia global de los pacientes con cáncer de ovario así como a la radio resistencia en células de cáncer de ovario a través de la estabilización de la proteína HuR (Dou *et al.*, 2019). El silenciamiento de FAM83H-AS1 inhibe la proliferación y la invasión de las células de cáncer de ovario *in vitro* (Gong and Zou, 2019). FAM83H-AS1 se encontró sobreexpresado en líneas celulares de cáncer de cérvix positivas al Virus del Papiloma Humano 16 (VPH16). Esta expresión es dependiente de la oncoproteína viral E6, pero independiente de la regulación de p53 (Barr *et al.*, 2019). En este mismo estudio, los autores describen que la sobreexpresión de FAM83H-AS1 se asocia con una menor OS en pacientes con cáncer de cérvix.

Sorprendentemente, FAM83H-AS1 funciona como una esponja del miR-22-3p en muestras de pacientes con degeneración del disco intervertebral (Jiang and Chen, 2020). En otro estudio en cáncer gástrico, se identificaron altos niveles de FAM83H-AS1 en comparación con los de tejidos normales. También se observó que la sobreexpresión de FAM83H-AS1 es un factor de riesgo relacionado con una menor OS y sobrevida libre de enfermedad (Da *et al.*, 2019). Sin embargo, los autores no replicaron los resultados de su cohorte de pacientes con cohortes independientes. FAM83H-AS1 puede potenciar la carcinogénesis y se propone como un biomarcador de mal pronóstico para cáncer de colon y recto (Yang *et al.*, 2019).

Dos estudios afirman que FAM83H-AS1 se encuentra desregulado en muestras de CaMa (Yang*et al.*, 2016; Deva Magendhra Rao *et al.*, 2019). La sobreexpresión de FAM83H-AS1 indica un mal pronóstico clínico en pacientes con tumores de mama de tipo luminal A y con tumores poco avanzados. Estas evidencias, en su conjunto, sugieren que FAM83H-AS1 es un factor importante en la biología del cáncer, y del CaMa en particular.

1.6. El RNA largo no codificante intergénico LINC00460

1.6.2. Aspectos generales y localización

El RNA largo no codificante intergénico 460 (LINC00460) se localiza en el cromosoma 13q33.2 y mide 913 pb (GenBank, 2021) (ver figura 8). El LINC00460 está constituido por tres exones y se ha demostrado que carece de capacidad codificante de proteínas (Cao *et al.*, 2017; Wen *et al.*, 2019). Según la base de datos ENSEMBL, el gen humano LINC00460 tiene siete variantes transcripcionales reportadas hasta la fecha (Yates *et al.*, 2020).

1.6.3. Papel de LINC00460 en cáncer

Se ha observado que LINC00460 funciona como oncogén, actuando como esponja de miRNAs (Lian *et al.*, 2019; Xiaodong Hu *et al.*, 2019; Xie *et al.*, 2019; Zhang *et al.*, 2019; Zhu *et al.*, 2019; Cui *et al.*, 2020; Hong *et al.*, 2020; Wang *et al.*, 2020; Zou *et al.*, 2020). En estos modelos está bien descrita la actividad de esponja de LINC00460 y su asociación con procesos relacionados con el cáncer, como el aumento de la proliferación, la transición EMT, la migración, invasión y la metástasis (Lian *et al.*, 2018; Wang *et al.*,

2018; Jiang *et al.*, 2019). Por ejemplo, LINC00460 promueve la progresión del carcinoma hepatocelular actuando como esponja de miR-342-3p y, por tanto, aumentando el nivel de expresión de AGR2 (Anterior Gradient 2, Protein Disulphide Isomerase Family Member) (Yang *et al.*, 2020). Estudios en otros tumores, incluyendo cáncer colorectal, cáncer de cabeza y cuello, cáncer de tiroides, cáncer de pulmón, osteosarcoma, cáncer de páncreas, cáncer de cabeza y cuello, cáncer de vejiga y cáncer de mama han descrito mecanismos similares, donde LINC00460 actúa como esponja de distintos miRNAS(Li y Kong, 2019a; Lian *et al.*, 2019; Xiaodong Hu *et al.*, 2019; Cui *et al.*, 2020; Hong *et al.*, 2020; Wang *et al.*, 2020; Zou *et al.*, 2020).

Se ha estudiado la asociación del LINC00460 con diversas características clínicas del cáncer. La sobreexpresión de LINC00460 está asociada con tasas menores de supervivencia global en diferentes tumores, como pulmón, ovario, laringe, nasofaringe, cabeza y cuello, meningioma, riñón, tiroides, colorrectal, glioma, osteosarcoma, vejiga y cérvix (Wang *et al.*, 2016; Liang *et al.*, 2017; Huang *et al.*, 2018; Xing *et al.*, 2018; Liu *et al.*, 2018; Feng *et al.*, 2019; Li y Kong, 2019b; Nakano *et al.*, 2019; Wen *et al.*, 2019; Xu *et al.*, 2019; Li, Zhu y Wang, 2020). Por lo tanto, LINC00460 es un lncRNA esponja bien definido con un potencial pronóstico significativo en varios tumores; sin embargo, su relevancia clínica y biológica en CaMa es poco conocida (Zhu *et al.*, 2019) y se requieren estudios exhaustivos.

www.nature.com/scientificreports

SCIENTIFIC REPORTS

nature research

OPEN

FAM83H-AS1 is a potential modulator of cancer driver genes across different tumors and a prognostic marker for ER/PR + BRCA patients

Magdalena Ríos-Romero[1,3], Alberto Cedro-Tanda[1], Mónica Peña-Luna[1], Marco Antonio Mancera-Rodríguez[1], Lizbett Hidalgo-Pérez[1], Mireya Cisneros-Villanueva[1], Fredy Omar Beltrán-Anaya[1], Rocío Arellano-Llamas[1], Silvia Jiménez-Morales[1], Luis Alberto Alfaro-Ruíz[1], Alberto Tenorio-Torres[2], Carlos Domínguez-Reyes[2], Felipe Villegas-Carlos[2], Elsa Ochoa-Mendoza[1] & Alfredo Hidalgo-Miranda[1]✉

Breast cancer (BRCA) is a serious public health problem, as it is the most frequent malignant tumor in women worldwide. BRCA is a molecularly heterogenic disease, particularly at gene expression (mRNAs) level. Recent evidence shows that coding RNAs represent only 34% of the total transcriptome in a human cell. The rest of the 66% of RNAs are non-coding, so we might be missing relevant biological, clinical or regulatory information. In this report, we identified nine novel tumor types from TCGA with FAM83H-AS1 deregulation. We used survival analysis to demonstrate that FAM83H-AS1 expression is a marker for poor survival in IHC-detected ER and PR positive BRCA patients and found a significant correlation between FAM83H-AS1 overexpression and tamoxifen resistance. Estrogen and Progesterone receptor expression levels interact with FAM83H-AS1 to potentiate its effect in OS prediction. FAM83H-AS1 silencing impairs two important breast cancer related pathways: cell migration and cell death. Among the most relevant potential FAM83H-AS1 gene targets, we found p63 and claudin 1 (*CLDN1*) to be deregulated after FAM83H-AS1 knockdown. Using correlation analysis, we show that FAM83H-AS1 can regulate a plethora of cancer-related genes across multiple tumor types, including BRCA. This evidence suggests that FAM83H-AS1 is a master regulator in different cancer types, and BRCA in particular.

Breast cancer (BRCA) is a serious public health problem, as it is the most frequent malignant tumor in women worldwide. According to GLOBOCAN[1], at least 1.67 million new cases and a total of 522,000 deaths are reported globally.

BRCA is a phenotypically heterogenic disease; with well-defined histological types and protein markers, such as Estrogen receptor (ER), Progesterone Receptor (PR) and membrane receptor HER-2. The physical and phenotypical BRCA heterogeneity is also reflected at the molecular, particularly at gene expression (mRNAs) level. This heterogeneity has been extensively studied, and evidence shows that breast cancer comprises four intrinsic groups: Luminal A, Luminal B, HER-2 enriched and Basal-like tumors[2,3].

Molecular classification has been an important milestone in BRCA biology, as it has been used to differentiate aggressive and non-aggressive tumors, metastatic potential, clinical prognosis and survival, among other relevant cancer-related features[4]. Additionally, therapy response-associated expression profiles are now available. This expression profiles are able to predict if a patient can benefit from chemotherapy or anti-hormonal therapy[5].

[1]Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Periferico Sur 4809, Tlalpan, Arenal Tepepan, 14610 Ciudad de México (CDMX), Mexico. [2]Fundación de Cáncer de Mama, FUCAM, Ciudad de México, Mexico. [3]Programa de Doctorado de Ciencias Biológicas, Universidad Nacional Autónoma de México, Ciudad de México, Mexico. ✉email: ahidalgo@inmegen.gob.mx

These useful clinical advances are focused on coding RNAs profiles only; however, recent evidence show that coding (messenger) RNAs represent only 34% of the total transcriptome in a human cell[6]. The rest of the 66% of RNAs are non-coding, so we might be missing relevant biological, clinical or regulatory information if we only focus on messenger RNA.

In this regard, recent papers have focused on the role of long non coding RNAs (lncRNAs) in cancer biology[7–10] and in the role of specific lncRNAs in breast cancer.

FAM83H-AS1 is a lncRNA whose expression impairs important cancer-related pathways such as cell proliferation, migration, invasion and cell death in lung, colorectal, glial, bladder, ovarian and cervical cancer cells[11–16]. At the molecular level, one report showed that MET/EGFR signaling is regulated by FAM83H-AS1, and[16] showed that FAM83H-AS1 epigenetically silenced *CDKN1A* by binding to EZH2 in glioma cells.

In addition, two reports showed that FAM83H antisense RNA 1 (FAM83H-AS1) is deregulated BRCA samples. High expression of FAM83H-AS1 indicated an unfavorable prognosis in luminal type BRCA[12] and early-stage BRCA[16]. Altogether, this evidence shows that FAM83H-AS1 is an important actor in cancer biology. In this paper, we identified nine novel tumor types from TCGA with FAM83H-AS deregulation, and used a multivariate Cox regression analysis to demonstrate that FAM83H-AS1 expression is a marker for poor survival in Progesterone receptor (PR) positive BRCA. We found a significant correlation between FAM83H-AS1 overexpression and tamoxifen resistance in luminal BRCA patients. Using Kaplan–Meier and Cox regression analysis, we found that estrogen and progesterone receptor expression levels interact with FAM83H-AS1 to potentiate its effect in OS prediction. Using FAM83H-AS1 short hairpin knockdown coupled with microarray analysis, we demonstrate that FAM83H-AS1 silencing impairs two important breast cancer related pathways: cell migration and cell death. We further validate this phenotypic effect with in vitro migration and caspase 3 assays. Among the most relevant potential FAM83H-AS1 gene targets, we found p63 and claudin 1 (*CLDN1*) to be deregulated after FAM83H-AS1 knockdown. Using correlation analysis, we show that FAM83H-AS1 can regulate a plethora of cancer-related genes across multiple tumor types, including BRCA.

Results

FAM83H-AS1 is deregulated in multiple tumor types. Multiple studies have related FAM83H-AS1 high expression levels with different tumors, including luminal breast cancer[11–16]. These findings suggest an important role for FAM83H-AS1 in cancer tumor biology. We therefore screened FAM83H-AS1 expression levels in the TCGA database, which comprises data from 33 different tumor types and the correspondent normal tissues. As expected, we found significant FAM83H-AS1 expression deregulation in 16 different tumor types (Fig. 1A) (Log2FC > 1; $p < 0.01$). Some of these FAM83H-AS1 expression deregulation data has been reported previously[11,12,16], but we have also found significant deregulation of FAM83H-AS1 in nine additional tumor types (Fig. 1B) (Log2FC > 1; $p < 0.01$). Interestingly, FAM83H-AS1 was up-regulated in 15 different tumor types, but down-regulated in acute myeloid leukemia (LAML), suggesting a different mechanism for this particular malignancy (Fig. 1B).

FAM83H-AS1 expression level is enriched in BRCA locally-advanced tumors. It was reported that FAM83H-AS1 expression is a prognostic marker for luminal breast cancers[12]. We were interested to see if FAM83H-AS1 expression was more widely associated with BRCA tumors, since we and others found alterations for this lncRNA in a large number of malignancies. As shown in Fig. 1B, FAM83H-AS1 is significantly up-regulated in all BRCA patients, not only in the luminal subtype BRCA. FAM83H-AS1 over-expression is also marginally associated with BRCA locally advanced (II and III) clinical stages (one-way ANOVA; $p = 0.05$) (Fig. 1C).

FAM83H-AS1 is a prognostic marker for ER and PR positive BRCA and its expression is related with tamoxifen resistance. Altogether, these widespread alterations in FAM83H-AS1 expression suggested that its expression could be a prognostic biomarker for all BRCA subtypes: but as mentioned above, FAM83H-AS1 was previously reported to have particular prognostic association with the BRCA luminal subtype[12,16]. To test if FAM83H-AS1 is a widespread or a luminal specific prognostic marker in BRCA, we first screened FAM83H-AS1 expression as a prognostic marker for all BRCA tumors. We did not find significant association with poor OS in the Cox regression model (n = 743; 95% CI [0.442–1.15] Cox p value = 0.09); however, we observed a clear tendency in poor survival prognosis in the FAM83H-AS1 high expression group (see Fig. 2A). We further validated these results in an independent Mexican patient cohort (Fig. 2B) with all the BRCA subtypes. The general clinical features of this cohort are listed in Table 1.

We then tested if this effect was due to FAM83H-AS1 interacting with other significant clinical and survival-related variables, in particular, luminal type-related. FAM83H-AS1 predictive value was significant when interacting with Immunohistochemistry (IHC)-detected Progesterone receptor (n = 743; HR = 1.55; 95% CI [1.005–2.376] Cox p value = 0.047) (supplementary Table 1). Marginal, but not significant, association was observed with ER status (supplementary Table 2). Kaplan–Meier analysis of PR (logrank; $p = 0.014$) or ER positive patients (logrank; $p = 0.006$) showed significant association poor OS when FAM83H-AS1 was over-expressed (Fig. 2C, D). No effect in survival rate was seen when FAM83H-AS1 was over-expressed in PR and ER negative patients (supplementary Figs. 1 and 2).

This data strongly suggest that FAM83H-AS1 is an independent prognostic marker for OS in PR positive BRCA subtype and confirms with further statistical analyses, previous findings made by[12].

We found a significant association with IHC-detected PR and ER status in the survival context (Fig. 2 C, D; supplementary Table 1). We then analyzed tamoxifen treatment resistance or sensitivity in our independent cohort (Fig. 2E), and we found that FAM83H-AS1 overexpression was significantly related with poor tamoxifen initial response (n = 42; OR = 3.9; one-tailed F-exact test; $p = 0.045$).

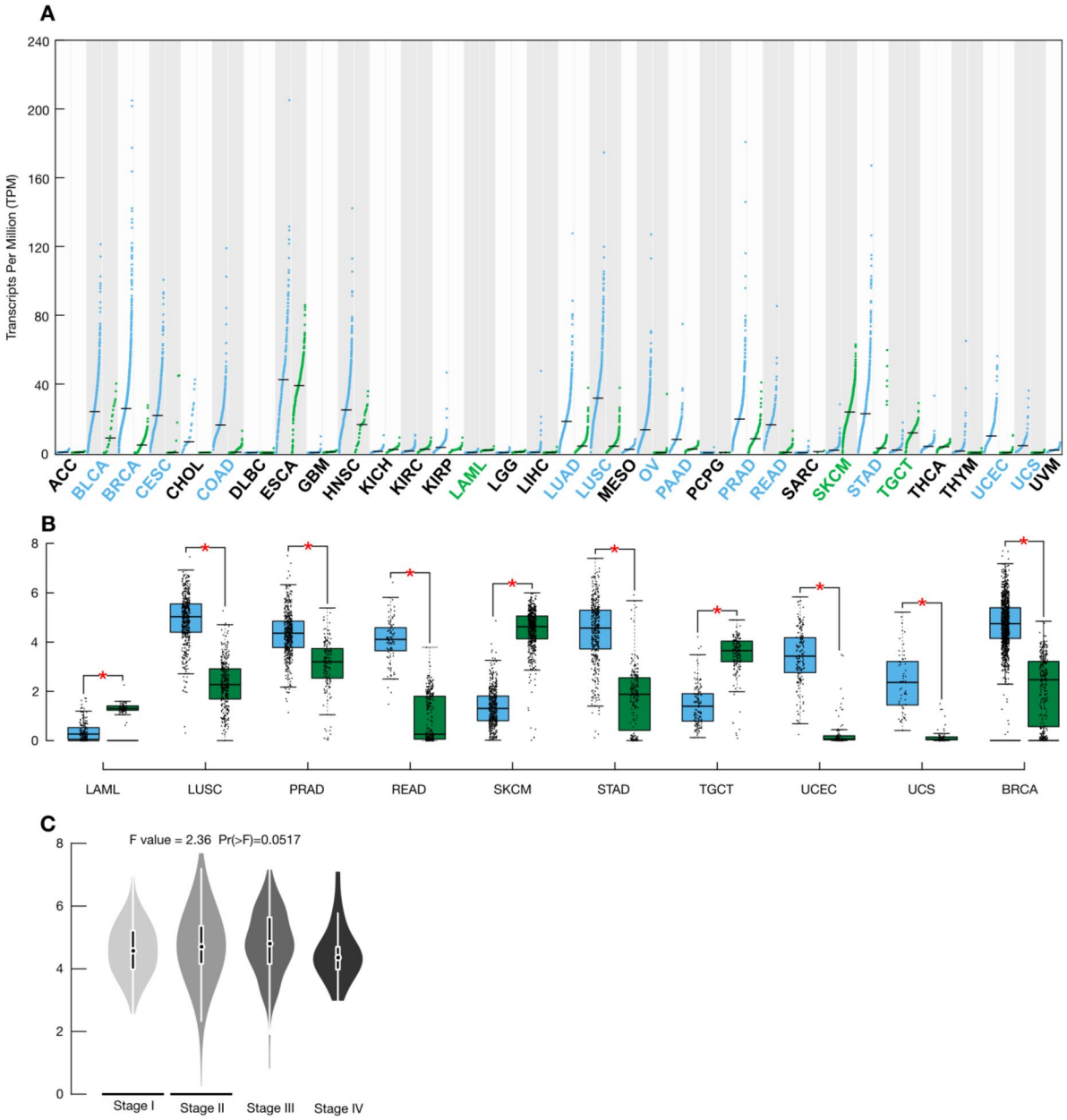

Figure 1. FAM83H-AS1 expression is altered in multiple human tumors. (**A**) FAM83H-AS1 expression levels (TPM) in 33 tumors from the TCGA database. In green are shown normal tissue samples, in blue tumor samples. (**B**) FAM83H-AS1 is aberrantly expressed in nine not previously reported tumors. Color code as in A. (**C**) FAM83H-AS1 is enriched in locally advanced BRCA clinical stages (II and III).

We did not find a significant prognosis association for 8 unreported tumors shown above in the Gene expression Profiling Interactive Analysis (GPIA) database (see methods). We could determine, however, a strong correlation between high FAM83H-AS1 expression and poor OS in skin cutaneous melanoma (SKCM) patients (n = 461; HR = 1.6; log-rank test, $p = 0.0003$) (Supplementary Fig. 3).

ER and PR expression levels potentiate FAM83H-AS1 prediction of survival in BRCA patients. In order to further characterize our previous finding regarding ER and PR status, and its association with FAM83H-AS1 in BRCA prognosis, we built a risk model taking into account ER, PR and FAM83H-AS1 expression levels in the same analysis. This model fundamentally displays ER, PR and FAM83H-AS1 interaction and potentiation of the poor OS prediction in BRCA.

We first calculated the Allred score[17] to identify IHC ER and PR level of positivity, *ie.*ER and PR expression levels from TCGA data. We then obtained FAM83H-AS1 expression levels from the TCGA cohort and divided it in four strata (quartiles). We then multiplied these two values (Allred score values and FAM83H-AS1 quartile values), and the product was a new risk score. We obtained four risk groups with the above described method, shown in Fig. 3. As shown here, combination of very high FAM83H-AS1 and ER/PR expression levels potentiates

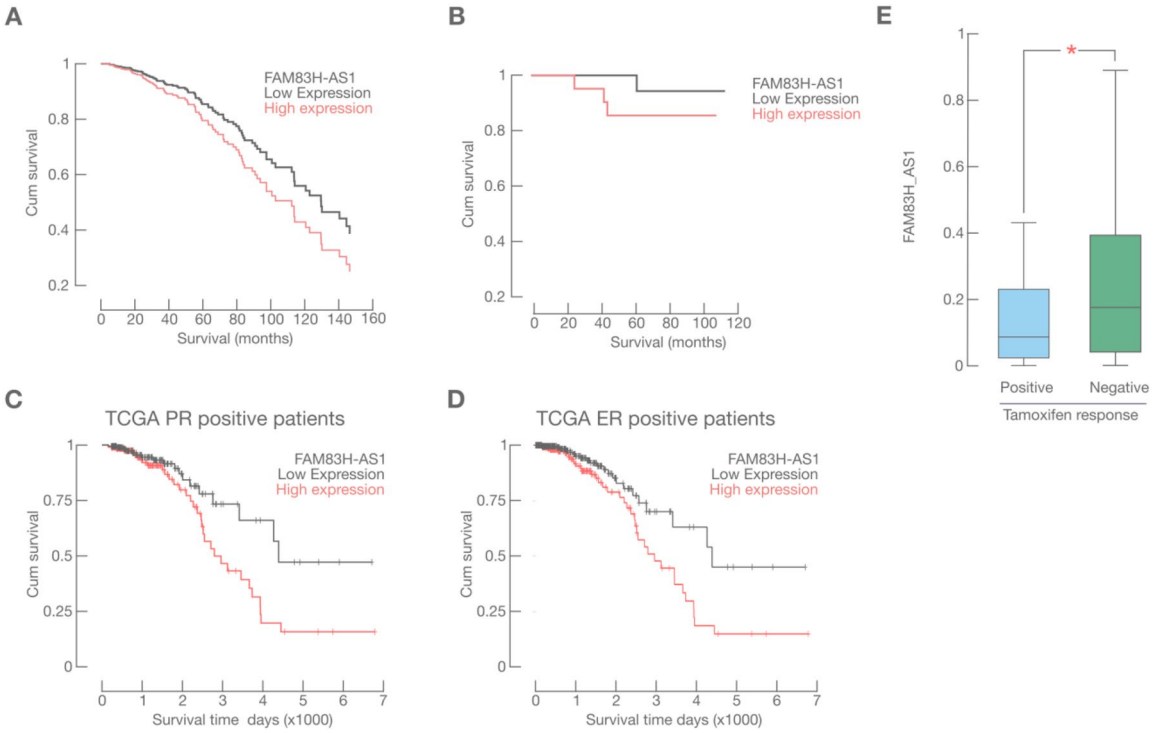

Figure 2. FAM83H-AS1 over-expression is a marker for poor prognosis in ER/PR BRCA patients. (**A**) Overall survival analysis (Cox Regression) for the BRCA TCGA cohort. (**B**) Kaplan–Meier analysis for our independent BRCA cohort. (**C**) Kaplan–Meier analysis for PR and **D**, ER positive BRCA patients from the TCGA cohort. (**E**) FAM83H-AS1 expression is associated with tamoxifen resistance in BRCA patients.

Characteristic	Class	Frequency	Percent
Age	0–50	19	45.2
	51–100	23	54.8
Tumor grade	I	3	7.1
	II	26	61.9
	III	9	21.4
	SV	4	9.5
Clin. stage	0, IA, IIA, IIB	37	88.1
	IIIA, IIIB, IIIC	5	11.9
ER	Positive	25	59.5
	Negative	17	40.5
PR	Positive	23	54.8
	Negative	19	45.2
HER2	Positive	13	31
	Negative	29	69
Lymph Nodes	Positive	18	42.9
	Negative	10	23.8
	NA	14	33.3
Recurrence	Positive	7	16.7
	Negative	35	83.3
Metastasis	Positive	7	16.7
	Negative	35	83.3

Table 1. Clinical-pathological characteristics of population (n = 42). In this Mexican cohort, none of these clinical variables were significantly correlated with FAM83H-AS1 expression level.

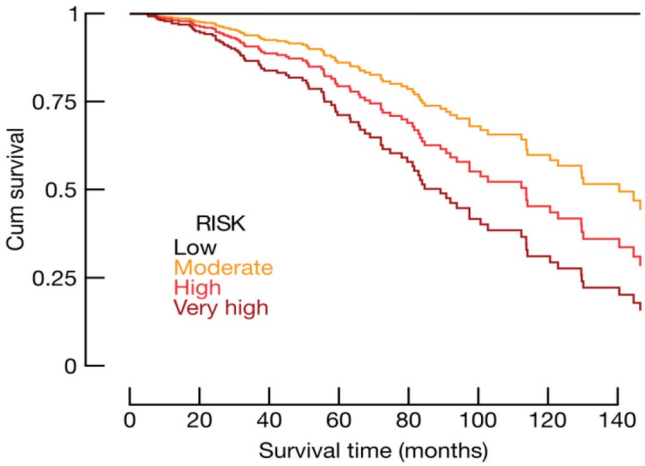

Figure 3. ER/PR expression potentiates FAM83H-AS1 death risk prediction. Expression levels from ER and PR were calculated using Allread scores. FAM83H-AS1 expression levels were calculated from the quartile distribution of the log2 values (TANRIC). We then multiplied Allread scores (17) and FAM83H-AS1 expression levels in order to define risk groups (see methods and main text).

high risk of decease in the TCGA cohort (Kaplan–Meier model, one tailed logrank; $p = 0.034$). Cox hazard proportional risks for decease in these four groups were: 0 for low risk, 25 for moderate risk, 39 for high and 57 for very high risk. This data confirms a strong interaction between ER, PR and FAM83H-AS1 expression levels in BRCA. Interaction between these three variables potentiates poor OS prediction in BRCA patients, and possibly other hormone-related human tumors.

FAM83H-AS1 potentially regulates a plethora of cancer-related genes. In order to better understand FAM83H-AS1 role in BRCA, we performed a differential expression analysis using the TCGA BRCA cohort. We compared high versus low FAM83H-AS1 RNA expression level samples (see methods). We found 2,668 differentially expressed genes between these two groups (Log2FC > 1.5 and − <1.5; *p*.adj-value < 0.01). The vast majority of these transcripts (98.6%; 2,631 RNAs) were found to be down-regulated in this analysis. These results might suggest candidate target genes for FAM83H-AS1 (Fig. 4A).

Among the down-regulated RNAs, we identified several cancer-related transcripts, such as: fibroblast growth factor 4 (*FGF4*); fibroblast growth factor 21 (*FGF21*); leptin (*LEP*), Claudin 17 (*CLDN17*); cadherin 9 (*CDH9*) Tumor Necrosis Factor receptor (*TNFRSF11B*); BCL2-associated X (*BAX*); Tumor protein p53 (*TP53*) and Phosphatase and tensin homolog (*PTEN*) (see Table 2).

After pathway enrichment analysis, we found a significant down-regulation of cellular migration, synthesis of steroids and lipid metabolism (Fig. 4B). Gene set enrichment analysis (GSEA) also showed alterations in apoptosis and p53-signalling pathways (Fig. 4C). Taken together, this evidence suggests an important regulatory role for FAM83H-AS1 in cancer-related pathways.

FAM83H-AS1 is present both in nucleus and cytoplasm in ER/PR BRCA cells. To further characterize FAM83H-AS1 functional role in BRCA, we measured its expression in nine breast cancer cell lines, including: MDA-MB-231, 468, 453, HCC1187, MCF7, SKBR3, BT20, Hs578, ZR75 and one non-transformed cell line, MCF10 (Fig. 5A). As expected, FAM83H-AS1 was up-regulated in transformed cell lines. We then performed cellular fractionation assays in MCF7 cells and detected its enrichment in the cytoplasmic fraction (67.3% of total input RNA) (Fig. 5B). LncATLAS screening further confirmed this observation, as MCF7 cells display both FAM83H-AS1 cytoplasmic and nuclear localization (Supplementary Fig. 4).

FAM83H-AS1 knockdown deregulates 415 transcripts expression in MCF7 cells. In order to gain further insight into the potential FAM83H-AS1 targets, we performed sh-mediated FAM83H-AS1 silencing experiments in MCF7 cells. As shown in Fig. 5C, we obtained 85% of silencing efficiency after 48 h of plasmid transfection. After knockdown, we performed microarray experiments in MCF7 cells. We identified 415 differentially-expressed genes in the FAM83H-AS1-silenced cells (FC > 1.5 and − <1.5; *p* value < 0.05) (Fig. 5D). Key cancer-related transcripts were identified, such as *CLDN1, TP63, FGF14, DDX60* and *DRAM* (see Table 3). 262 transcripts were found to be down-regulated whereas 153 were up-regulated (Fig. 5D; Table 3). Among the most up-regulated RNAs, we found *TP63* and *CLDN1*. These genes can also be potential candidate target genes for FAM83H-AS1 activity.

FAM83H-AS1 silencing impairs cellular migration and apoptosis. Pathway enrichment analysis showed that the most activated cellular processes in the sh-FAM83H-AS1 condition are: migration and cellular

Description	Pathway	Fold change (log2)	*p*adj-value
Myosin Light chain 2	Tight junction signaling	4.268825581	3.38E−52
Myosin heavy chain 2	WNT signaling pathway	3.267559072	9.51E−42
Actin alpha 1	WNT signaling pathway	2.247491528	1.06E−30
Leptin	Cell migration	−2.03868572	1.10E−08
MicroRNA 181	Cell migration	−1.757287497	8.07E−06
Nuclear receptor 1 H4	Cell migration	−1.923864676	1.47E−05
Fibroblast Growth Factor 4	MAPK signaling/Focal adhesion	−1.60587355	9.78E−08
Fibroblast Growth Factor 21	MAPK signaling/Focal adhesion	−1.813231804	3.74E−05
Claudin 17	Epithelial to mesenchymal transition	−1.609953043	0.00048284
Insulin like growth factor binding protein 1	PI3K pathway	−3.010911164	1.75E−11
Cadherin 9	Cadherin signaling	−1.579133916	2.44E−05
Hydroxy-delta-5-steroid dehydrogenase	Steroid synthesis	−1.961005522	1.57E−10
UDP glucuronosyltransferase family 1 member A1	Steroid synthesis	−1.554304254	0.001164245
Alcohol dehydrogenase 1A (class I) alpha polypeptide	Drug metabolism	−2.556929166	4.05E−09
Cytochrome P450 family 2 subfamily A member 6	Drug metabolism	−2.989324666	3.13E−09
Deathassociated protein kinase 3	Cell death regulation	−0.277	1.15E−04
Tumor protein p53	Cell death regulation	0.139	5.00E−02
TNF receptor superfamily member 11b	Cell death regulation	−0.126	4.22E−01
BCL2 associated X, apoptosis regulator	Apoptosis regulation	−0.136	6.64E−02
BH3-like motif containing cell death inducer	Cell death regulation	−2.779	1.07E−11
Phosphate and tensin homolog	Cell survival signaling	−0.19	1.46E−03

Table 2. Examples of differentially expressed genes in the FAM83H-AS1 high versus. low BRCA sample.

motility (29 altered molecules; p value range <0.01) and cellular death pathways (49 molecules; p value range <0.01) (Fig. 5E). These data are in accordance with our previous differential expression results in BRCA samples (Fig. 4), further supporting a master regulatory role for FAM83H-AS1 in breast cancer cells.

We then aimed to functionally validate that both cellular processes are altered after FAM83H-AS1 silencing. We thus performed Transwell migration assays and caspase 3 activity assays in MCF7 cells. As shown in Fig. 6A, MCF7 cells migration significantly increases after sh-FAM83H-AS1 transfection (one tailed T-test; $p=0.035$). We then performed matrigel-invasion assays and observed an increase 24 h after transfection, but this observation did not reach statistical significance. (Supplementary Fig. 5). Furthermore, we did not find significantly altered invasion related genes or pathways in the microarray assays nor the differential expression analysis in the TCGA cohort, further suggesting that FAM83H-AS1 is not involved in invasion in this model.

Caspase 3 assays identified a significantly increase in enzymatic activity (T-test; $p=0.032$) after 72 h of FAM83H-AS1 silencing, which corroborates its role in apoptosis mediated cell death. Caspase 3 is the primary activator of apoptotic DNA fragmentation[18]. Significant increase in caspase 3 activity in the sh-FAM83H-AS1 condition suggests FAM83H-AS1 regulation of late stage apoptosis (Fig. 6B).

Migration and cell death alterations are enriched in the FAM83H-AS1 low expression group in BRCA samples. To further validate our previous in vitro results, we performed single sample GSEA (ssGSEA) in a BRCA independent cohort (Gene Expression Omnibus dataset GSE115577) (see methods). In this approach, gene sets are ranked according to absolute expression values in every sample, rather than by a comparison with another sample. We first stratified the GSE115577 cohort onto two groups: samples with high levels of FAM83H-AS1 RNA, and samples with low FAM83H-AS1 levels, using the quartile approach described above. We then aimed to know if the gene sets corresponding to migration, apoptosis and other cell death processes, like necrosis, were significantly enriched in any of these two FAM83H-AS1 expression groups. As expected, we found a significant enrichment of the migration (Normalized Mutual Index [NMI] score = 0.16; AUC = 0.79; p value = 9.18e−15) and apoptosis (NMI score = 0.033; AUC = 0.61; p value = 0.001) processes in the FAM83H-AS1-low expression group (Fig. 6 B). Interestingly, we also found a strong enrichment of the necrosis pathway in this BRCA cohort (NMI score = 0.3; AUC = 0.83; p value = 1.3e−18).

FAM83H-AS1 and its potential target genes are co-deregulated across multiple tumor types. We found that FAM83H-AS1 is up-regulated not only in BRCA, but also in other tumors (see Fig. 1). We reasoned that, if up-regulated, FAM83H-AS1 may be exerting similar regulatory roles in other tumors as well. In order to show this, we correlated FAM83H-AS1 expression levels with its potential target genes (BAX, CLDN1, CLDN17, DRAM, DDX60, FGF4, FGF14, LEP, PTEN, TNFRSF11B, TP53 and TP63). As shown in Fig. 7A, FAM83H-AS1 is strongly correlated with these coding genes across multiple tumor types, namely BLCA, BRCA, CESC, COAD, LAML, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, TGCT, UCEC and UCS. We then calculated the hazard ratio (HR) of decease event related to FAM83H-AS1 and its potential targets expression in 16 different tumor types from TCGA (Fig. 7B). We found a significant association (log-rank

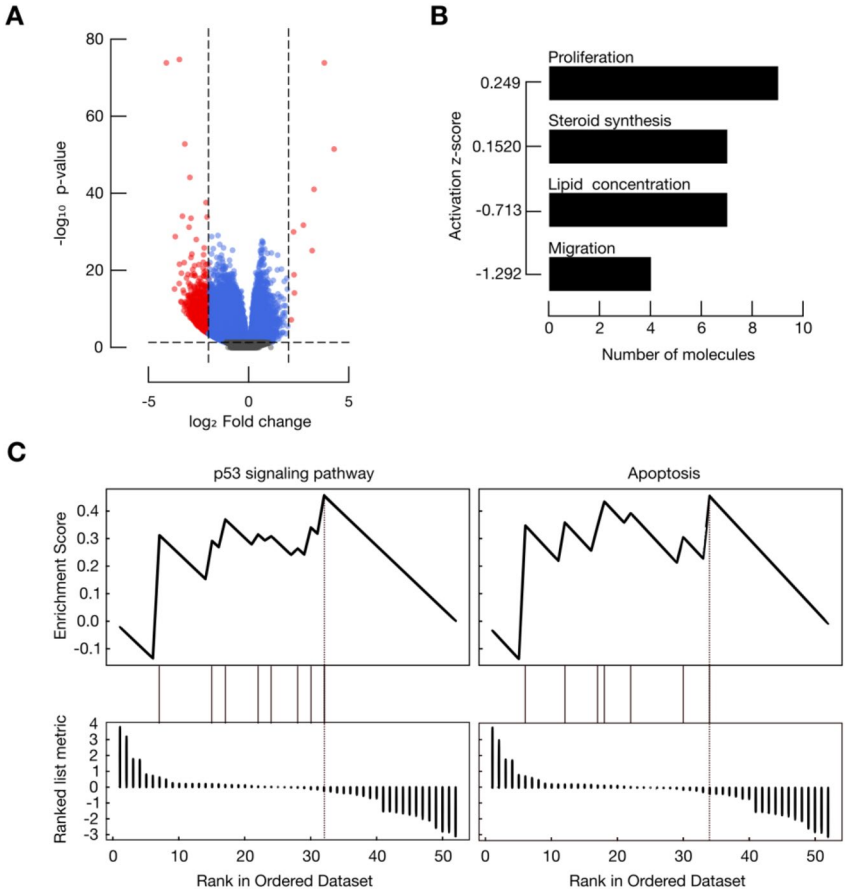

Figure 4. High FAM83H-AS1 levels in BRCA samples are correlated with down-regulation of cancer-related inhibitors. (**A**) Volcano plot depicting differentially expressed transcripts in high versus low FAM83H-AS1 levels. Points in grey = non-significant; blue = *p* value significant; red = fold change, and *p* value significant. (**B**) Migration-related and steroid metabolism genes are significantly down-regulated when FAM83H-AS1 is highly expressed in BRCA. (**C**) GSEA analysis showed enrichment for apoptosis cell death and p53-signalling pathways in the FAM83H-AS1-high samples.

test; $p < 0.05$) of FAM83H-AS1 overexpression and a high decease HR in BRCA, PAAD and SKCM (Fig. 7B). The expression of FAM83H-AS1 potential targets is also associated with high or low HR in these tumors. Some of the coding genes that we either found in the differential expression analysis (Fig. 4) or after FAM83H-AS1 knockdown in MCF7 cells (Fig. 5D) are also co-deregulated in other tumor types (Fig. 7C). Altogether, this evidence suggests a wide master regulatory role for FAM83H-AS1 not only in ER/ PRBRCA, but in other tumor types, such as PAAD and SKCM.

Discussion

Long non-coding RNAs are molecules that exert numerous roles in human cancers, as their biological activities involve regulation of cell proliferation, cell death, differentiation, migration and invasion. Deregulation in lncRNAs expression has also been associated with clinical outcome. LncRNAs can affect expression of thousands of genes, so they are regarded as key master regulators[7–10].

In this work, our aim was to investigate a wider de-regulation for FAM83H-AS1 expression in tumors, focusing on its functional and clinical role in breast cancer and the identification of potential FAM83H-AS1 targets. We found that FAM83H-AS1 was overexpressed in nine different tumor types in the TCGA database. In particular, FAM83H-AS1 is overexpressed and significantly correlated with a worse clinical outcome in PR positive (detected by immunohistochemistry) BRCA subtypes, in the TCGA breast cancer cohort. One previous report[12] shows that FAM83H-AS1 high expression indicated unfavorable prognosis in luminal breast cancer and was an independent prognostic indicator. To the best of our knowledge, this is the first report that suggests variable interaction between FAM83H-AS1 and IHC-detected PR and ER in the clinical outcome context. Furthermore, we demonstrate that ER and PR expression levels can act as potentiators of FAM83H-AS1 poor OS prediction. In particular, ER and PR high expression levels, together with FAM83H-AS1 over-expression, confers a very high risk (HR = 57) of decease in BRCA patients. This data suggest an important clinical role for FAM83H-AS1 in ER/PR positive breast cancer. It is currently unknown, however, if these statistical and clinical interactions are reflected at the biological or molecular level, and future studies must address this issue.

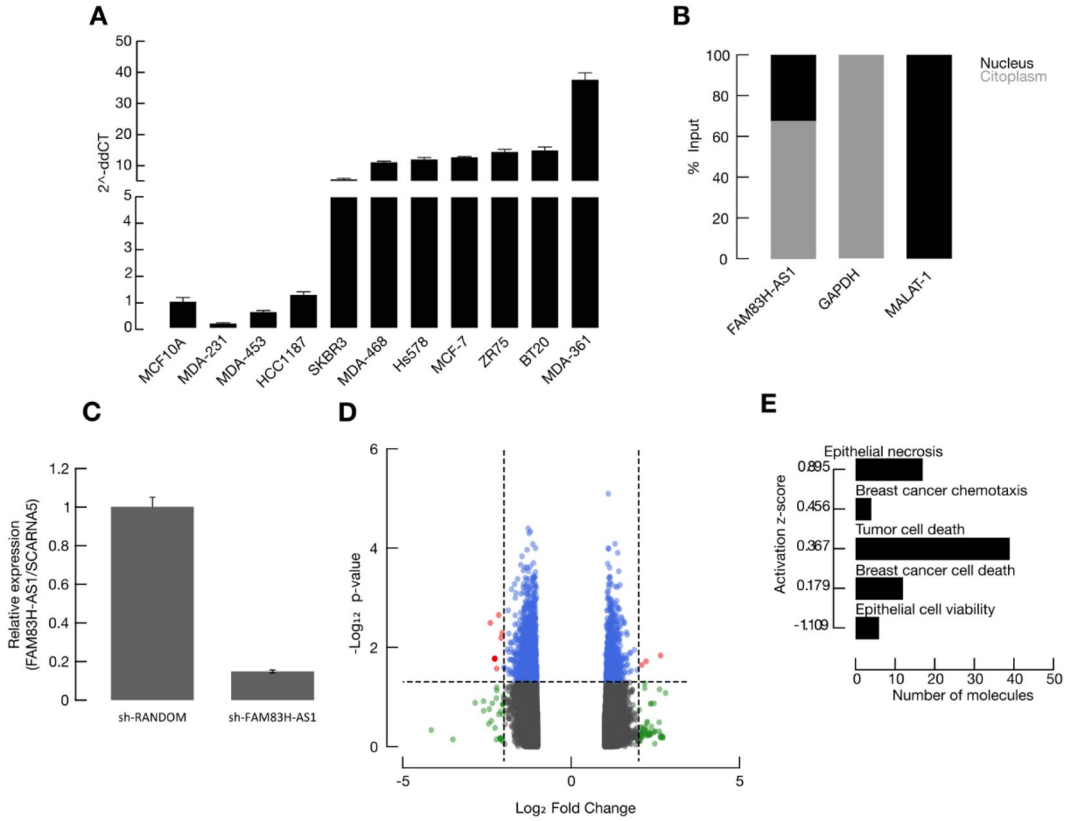

Figure 5. FAM83H-AS1 knockdown in MCF7 cells is associated with de-regulation of multiple cancer-related genes. (**A**) FAM83H-AS1 expression profile in breast cancer cell lines. (**B**) FAM83H-AS1 is localized both in nucleus and cytoplasm, but is enriched in cytoplasm in MCF7 cells. (**C**) short hairpin RNA silencing of FAM83H-AS1 in MCF7 cells. (**D**) sh-mediated silencing of FAM83H-AS1 induces differential expression in 415 genes in MCF7 cells. Grey = non-significant; blue = *p* value significant; green = fold change significant; red = *p* value, and fold change significant. **E,** Cellular migration and cell death are two significantly enriched pathways after FAM83H-AS1 silencing in MCF7 cells.

Gene symbol	Description	Pathway	Fold change (sh-FAM83H-AS1 vs. sh-RANDOM)	*p* value
CLDN1	Claudin 1	Epithelial to mesenchymal transition	1.64	0.023413
TP63	Tumor protein p63	Apoptosis signaling	1.56	0.017146
H3F3C	H3 histone, family 3C	DNA replication	1.51	0.033872
FGF18	Fibroblast growth factor 18	MAPK signaling/Focal adhesion	−1.43	0.000613
DDX60	DEAD (Asp–Glu–Ala–Asp) box polypeptide 60	Cell death regulation	−1.45	0.008851
IFI6	Interferon, alpha-inducible protein 6	Interferon signaling	−2.4	0.003207
DRAM1	DNA-damage regulated autophagy modulator 1	Cell death regulation	−1.54	0.018695
GPER1	G-protein coupled estrogen receptor 1	Endocrine resistance	−1.31	0.0286
PARP9	Poly(ADP-ribose) polymerase family member 9	DNA repair	−1.34	0.001424
GSTM3	Glutathione S + S transferase mu 3 (brain)	Drug metabolism	−1.49	0.012426

Table 3. Examples of differentially expressed genes after FAM83H-AS1 knockdown in MCF7 cells.

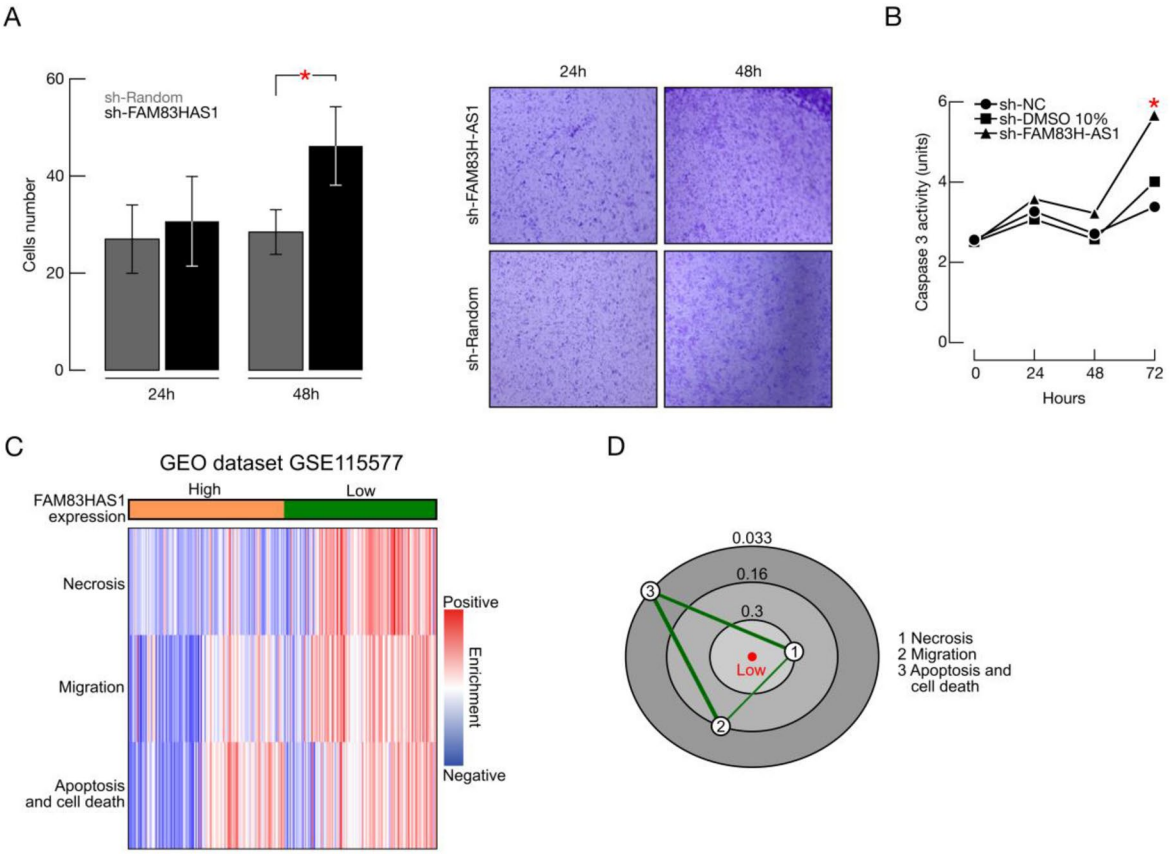

Figure 6. FAM83H-AS1 knockdown impairs cellular migration and induces cell death in breast cancer cells. (**A**) Transwell in vitro migration assays in MCF7 cells. (**B**) Caspase 3 activity assays in MCF7 cells. (**C**) Heatmap of the 3 gene sets enriched in the high FAM83H-AS1 expression group (blue) compared with that of FAM83H-AS1 low expression samples (green). (**D**) Constellation Map of the 3 gene sets. Three connected clusters of gene sets (migration, apoptosis and cell death and necrosis pathways) are detected in the low-FAM83H-AS1 group.

We were also able to find a significant correlation between FAM83H-AS1 high expression and poor tamoxifen response in BRCA patients. This association could partially explain the reduced clinical response in FAM83H-AS1 high expression group.

We also report that FAM83H-AS1 over-expression in TCGA breast cancer samples is associated with down-regulation of migration and cell death-related transcripts, like *FGF4, FGF21, LEP, CLDN17, TP53, BAX* and *TNFRSF11B.* Accordingly, we also found that FAM83H-AS1 knockdown significantly deregulates migration and apoptosis-related genes, such as *TP63* and *CLND1*. Transwell migration assays showed that indeed, cellular migration increases after FAM83H-AS1 silencing. LEP and CLDN1 had been both shown to induce cellular migration and epithelial to mesenchymal transition (EMT) in breast cancer cells[19–23], further suggesting a FAM83H-AS1 role in the early steps of migration. Taken together, this data might explain the underlying mechanisms related to FAM83H-AS1 cell migration impairment in breast cancer cells.

FAM83H-AS1 might play a dual role, probably due to cellular context. FAM83H-AS1 was involved in regulation of cell proliferation, migration and invasion processes that were decreased after FAM83H-AS1 knockdown in lung cancer cells. Further analysis indicated the cell cycle was arrested at the G2 phase after FAM83H-AS1 knockdown[11] In the same report, they found that MET/EGFR signaling was regulated by FAM83H-AS1. These conflicting results may be due to cellular context or specific regulation mechanisms, and henceforth, specific molecular targets.

We also identified that FAM83H-AS1 overexpression is associated with down-regulation of cellular death-related transcripts, like *BAX, TNFRSF11B* and *P53*. In vitro assays also show that FAM83H-AS1 silencing increases cellular death, possibly by up regulating genes like *p63.* One previous report[14] showed that cell death was markedly increased after with FAM83H-AS1 knockdown in colorectal cell lines. FAM83H-AS1, Notch1 and Hes1 were significantly increased in colorectal cancer samples and cell lines. Cell proliferation was inhibited with FAM83H-AS1 knockdown and this effect mediated by FAM83H-AS1 could be reversed by Notch1 regulators[14].

It is currently not clear, however, if FAM83H-AS1 has a direct or an indirect effect in gene regulation. In this regard, it has been shown that FAM83H-AS1 epigenetically silenced *CDKN1A* by binding to EZH2 in glioma cells[24]. In our differential expression analysis, we demonstrate that FAM83H-AS1 is mostly down-regulating gene expression. This data might suggest an inhibitory regulation role for FAM83H-AS1. Future studies must address

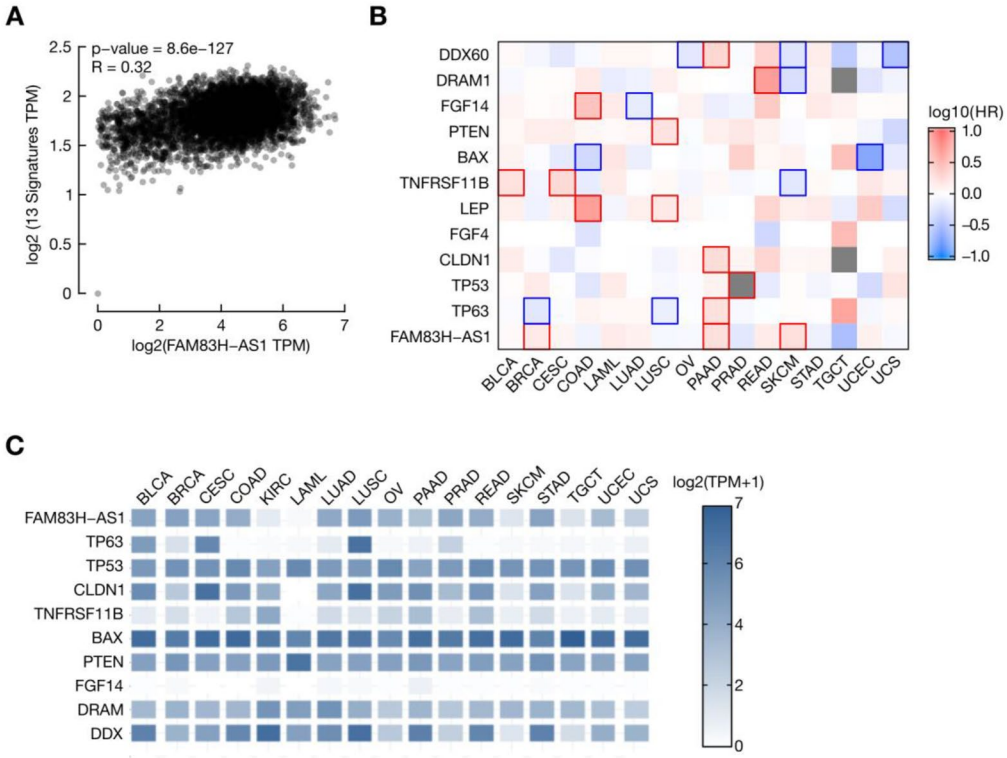

Figure 7. The expression of master regulators of cancer, such as p53 and p63, are dependent of FAM83H-AS1. (**A**) FAM83H-AS1 expression levels are strongly correlated with 12 potential target genes (TP53, TP63, BAX, CLDN1, CLDN17, CDH9, TNFRSF11B, PTEN, FGF4, FGF14, DRAM, DDX60) across 13 different tumors (BLCA, BRCA, CESC, COAD, LUAD, LUSC, PAAD, PRAD, READ, OV, STAD, UCEC, UCS, SKCM). (**B**) Overall survival heatmap, depicting that over-expression of FAM83H-AS1 confers high risk of death in BRCA, PAAD and SKCM patients. (**C**) FAM83H-AS1 and its potential target genes are deregulated in 17 different tumors.

these mechanisms; as we cannot discard that FAM83H-AS1 may regulate master gene expression via recruiting epigenetic complexes (e.g. EZH2). In addition, the exact role for FAM83H-AS1 in up-regulated genes remains obscure. We cannot discard a subtle, alternative role for this lncRNA in gene activation, and future studies must address this issue.

Our results also show that FAM83H-AS1 is present both in nucleus and cytoplasm of breast cancer cells. It is possible that this lncRNA is playing a different role in cytoplasm, and future studies must focus on this question.

In conclusion, FAM83H-AS1 is a lncRNA that is de-regulated in multiple cancers, and is a promising molecule that can perform as an independent prognostic factor in ER/ PR positive breast cancer. FAM83H-AS1 deregulation is associated with migration and cell death impairment in BRCA samples and breast cancer cells, and may regulate a plethora of cancer-related gene targets, such as *p63, BAX, LEP* and *CLDN1*.

Methods

The Cancer Genome Atlas (TCGA) and Gene expression omnibus (GEO) datasets. FAM83H-AS1 expression levels were screened in the 33 tumor datasets (see supplementary Table 3 for details) from TCGA and correspondent normal tissues using the Gene expression Profiling Interactive Analysis (GPIA) platform (https://gepia.cancer-pku.cn/). The 33 tumors included are enlisted as follows: Acute myeloid leukemia (LAML); Adrenocortical carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA); Brain Lower Grade Glioma (LGG); Breast invasive carcinoma (BRCA); Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC); Cholangiocarcinoma (CHOL); Chronic Myelogenous Leukemia (LMCL); Colon adenocarcinoma (COAD); Esophageal carcinoma (ESCA); Glioblastoma multiforme (GBM); Head and Neck squamous cell carcinoma (HNSC); Kidney Chromophobe (KICH); Kidney renal clear cell carcinoma (KIRC); Kidney renal papillary cell carcinoma (KIRP); Liver hepatocellular carcinoma (LIHC); Lung adenocarcinoma(LUAD);Lung squamous cell carcinoma (LUSC); Mesothelioma (MESO); Ovarian serous cystadenocarcinoma (OV); Pancreatic adenocarcinoma (PAAD); Prostate adenocarcinoma (PRAD); Rectum adenocarcinoma (READ); Sarcoma (SARC); Skin Cutaneous Melanoma (SKCM); Stomach adenocarcinoma (STAD); Testicular Germ Cell Tumors (TGCT); Thyroid carcinoma (THCA);Uterine Carcinosarcoma (UC);Uterine Corpus Endometrial Carcinoma (UCEC); Uveal Melanoma (UVM).

Potential target genes expression correlation, Hazard ratio (HR) map, co-expression map and BRCA stage plots were also generated in the GPIA platform. FAM83H-AS1 expression levels were considered significantly correlated with tumors when log2FoldChange > 1 and p value < 0.01.

Microarray generated expression data was downloaded from the GEO dataset GSE115577. This dataset includes RNA levels from 467 BRCA samples, analyzed with the Affymetrix HTA 2.0 platform. Downstream analysis is described below.

IHC-detected hormonal receptors and FAM83H-AS1 risk model. Clinical information of the BRCA patients was downloaded from the TCGA database (https://portal.gdc.cancer.gov/). FAM83H-AS1 expression levels were downloaded from the TANRIC tool (https://ibl.mdanderson.org/tanric/_design/basic/main.html). We first searched for the ER and PR status, and the numerical value for percent stained cells, also available in the clinical data. We then calculated the Allred score[17] (which measures the stain intensity and stain pattern) for ER and PR positivity levels in TCGA samples. We calculated the FAM83H-AS1 expression quartiles and stratified its levels of expression in this four groups (quartiles).

We then multiplied the Allred score with the FAM83H-AS1 expression levels (quartiles), and obtained nine patient groups: all the possible combinations for this particular model. We then performed two regression Cox models: in the first one, we established major risk groups. In the second one, we computed these major risk groups, shown in the results section.

Breast cancer samples differential expression analysis. Breast cancer RNAseq counts were downloaded from the TCGA Data Portal (https://portal.gdc.cancer.gov). After dataset preparation, we identified the FAM83H-AS1 ID (ENSG00000282685) and downloaded the expression counts. Transcripts with 10 counts or less were not included in the analysis. In order to generate the high and low FAM83H-AS1 expression groups, we calculated two percentiles from the count expression data. The first quartile (25) contains the lowest FAM83H-AS1 expression counts, and the upper quartile (75), contains the highest expression levels for this transcript. We then performed differential expression analysis with the DESeq2 module from the Gene Pattern platform (https://software.broadinstitute.org/cancer/software/genepattern/). Genes were considered differentially expressed when Log2Fold Change was > 1.5 and − <1.5 and p adjusted value < 0.05. Volcano plots were generated with the Enhanced Volcano R Package.

The upper (75) and lower quartile (25) approach to generate FAM83H-AS1 expression groups described above was also performed with the expression data from GSE115577. In this particular case, groups were generated in order to perform single sample Gene Set Enrichment Analysis (ssGSEA).

Breast cancer patients and biological samples. A total of 42 biological samples (biopsies) were collected from breast cancer patients attending to Fundación Cáncer de Mama (FUCAM) in Mexico City, Mexico. Adjacent normal samples were obtained from 2 cm above the surgical tumor margin. All patients signed a written informed consent before participating in this study. The study was approved by the Research Ethics Committee (INMEGEN) and the FUCAM Ethics Committee (Registration number: CE2009/11). All research methods were performed in accordance with relevant regulations and guidelines.

Biological samples were bisected; one portion was fixed in formaldehyde (10%), paraffin embedded (Paraplast Plus®; Sigma Aldrich ®, St Louis, Missouri, USA) and then submitted to haematoxylin and eosin staining for histopathological examination by an expert pathologist. Tumor stage was assessed, according to International standards. The second portion the sample was used for RNA extraction and functional downstream analysis. All tissues were liquid nitrogen-frozen and stored at − 80 °C.

In all 42 cases, demographic (age, sex), clinical (date of diagnosis, therapy received), pathologic (stage, grade, histological type) and prognostic data (recurrence, progression and overall survival) were available and correlated with FAM83H-AS1 expression status.

Cell culture. ER/PR-positive human breast cancer cell line MCF-7 (ATCC HTB-22) cells were purchased from American Type Culture Collection (ATCC, Manassas, Virginia, USA) and cultured in Dulbecco´s Modified Eagle Medium F12 (DMEM-F12; Corning® Inc, N.Y, USA) with 10% Fetal Bovine Serum (Corning® Inc, N.Y, USA). Cells were grown in 75 cm^3 cell culture bottles (Corning® Inc, N.Y, USA) at 37 °C with an atmosphere 95%/5% of air/CO_2.

Plasmids and transfection. Control and target plasmids were cloned using the BLOCK-iT™ U6 RNAi Entry Vector Kit (ThermoFisher™ Scientific, Waltham, Massachusetts, USA), according to the manufacturer´s recommendations. Short hairpin RNA oligos were designed using the Invitrogen Block-iT™ RNAi designer tool. Oligos were designed to target FAM83H-AS1 Exon 1 sequence (Oligo sh Top sequence: CACCGAAGAACA TCCCAGATTACCCGCGAACGGGTAATCTGGGATGTTCTTTT; Bottom sequence: AAAAGAACATCC CAGATTACCCGTTCGCGGGTAATCTGGGATGTTCTTC). This double stranded oligo was then annealed and cloned into the entry vector pENTR™/U6 (ThermoFisher™ Scientific, Waltham, Massachusetts, USA). The plasmids were then introduced onto *E. coli* TOP10 competent cells, which were grown in LB/agar medium at 37 °C.

Plasmids were purified using GeneJet Plasmid Miniprep Kit (ThermoFisher™ Scientific, Waltham, Massachusetts, USA) and sequenced to confirm insert integrity.

Transfection experiments were performed using Xfect™ Transfection Reagent (Clontech Laboratories Inc., Mountain View, California, USA) following the manufacturer´s instructions. Briefly, 100,000 cells were cultured 24 h prior to transfection in 24-well plates. 750 ng of the random plasmid and 750 ng of sh-FAM83H-AS1 plasmid

were diluted and then added to each well. Transfection reaction was incubated for 24 h; medium was removed and replaced with fresh complete medium.

RNA extraction. Genomic RNA was extracted using the commercial kit AllPrep® DNA/RNA FFPE (Qiagen® Inc, Valencia, CA) following manufacturer´s instructions. Briefly, the tissues were deparaffinized, disrupted and lysed. RNA was then precipitated, washed, purified and suspended in RNAse free water. RNA concentration was evaluated by spectrofotometry (NanoDrop Technologies, Wilmington, Delaware, USA). RNA integrity was analized using the BioAnalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). Samples were stored at − 80 °C.

Quantitative reverse transcription polymerase chain reaction (qRT-PCR). cDNA was synthetized using SuperScript III RT-PCR (Invitrogen, ThermoFisher™ Scientific, Waltham, Massachusetts, USA) following the manufacturer's instructions. Briefly, 100 ng of total RNA from cell lines or breast cancer samples were used to synthesize cDNA in a final reaction volume of 20 µL. The PCR reaction contained 1 µL of cDNA, 5 µL 2X TaqMan Universal Master Mix (Applied Biosystems, ThermoFisher™ Scientific, Waltham, Massachusetts, USA), 0.5 µL TaqMan probes (AIS09YL custom for FAM83H-AS1) and 3.5 µL of nuclease-free water. Both primers and reporter were designed to target FAM83H-AS1 exon 1 (Forward primer: ATCCCAGTTGATATC AGGGCAATC; reverse primer: TGTAAGCCCTTGATATTGG; reporter: TCCTGGCTGTTTTCC). GAPDH (Hs99999905) and SCARNA5 (Hs03391742_cn) transcripts were used as endogenous controls.

Subcellular fractionation and validation. MCF7 cellular fractionation (cytoplasm and nucleus) assays were performed using the Protein and RNA Isolation System (PARIS™) Kit (ThermoFisher™ Scientific, Waltham, MA, USA) according to manufacturer's instructions. Total and fractionated RNA was purified and then analyzed by RT-PCR, as described above. RNA percentages of each transcript over total RNA were calculated. GAPDH (Hs99999905) and MALAT-1 were used as cytoplasmic and nuclear controls, respectively.

The lncATLAS database (https://lncatlas.crg.eu/) was used to validate FAM83H-AS1 subcellular localization in MCF7 cells.

Microarray expression analysis. Global expression analysis was performed using Human Transcriptome Array 2.0® (Affymetrix® Inc, Santa Clara, CA, USA). This array covers 44,699 coding RNAs and 22,829 non-coding RNAs. A total of 200 nanograms of RNA were processed in each assay. All samples were processed using WT Plus Reagent Kit and Affymetrix hybridization kits, according to Affymetrix® recommendations.

Gene expression profiles. Affymetrix HTA 2.0 dataset analysis was performed using the Affymetrix® Expression Console and Transcriptome Analysis Console®. Normalized intensities from the sh-RANDOM condition was compared to normalized intensities from the sh-FAM83H-AS1 condition using one-way ANOVA. Genes were considered differentially expressed when fold change was > 1.5 and − <1.5 and p value < 0.05.

Pathway enrichment analysis. Pathway enrichment analysis was performed with Ingenuity Pathway Analysis® (IPA) software. Z-scores and p values were also computed using this platform. Only differential statistically significant genes were included in this analysis (see criteria above).

Gene set enrichment analysis (GSEA) was performed with the Web-based Gene SeT Analysis Toolkit (WebGestalt) platform (www.webgestalt.org). Non-significant pathway-enriched genes were included in this analysis as a priori set of genes.

ssGSEA was performed with the ssGSEA projection Gene Pattern module (https://software.broadinstitute.org/cancer/software/genepattern/). Graphic visualization was constructed using the Constellation Map module[25], also available in the Gene Pattern platform. Migration and cell death complete pathways that were generated in the IPA software, as a result of the microarray expression assays, were used as a priori set of genes. GSE115577 dataset (BRCA samples) was used as the input cohort to perform enrichment analysis.

Transwell migration and invasion assays. For migration assays, MCF-7 cells were seeded in 24-well plates (100,000 cells per well) and transfected with sh-RANDOM or sh-FAM83H-AS1 plasmids. 24 h after transfection, cells were tripsinized and cultured in Transwell® Chambers (8.0 µm) (Corning® Inc, N.Y, USA). Cells were incubated for 24 and 48 h after tripsinization and then fixed with cold 70% Ethanol (J.T. Baker®, Fisher Scientific, Waltham, Massachusetts, USA). Cells were stained with 1X SRB Staining Solution (VitroSure™ SRB Viability/Cytotoxicity Assay Kit, GeneCopoeia™, Rockville, Maryland, USA). In vitro invasion assays were done in the same fashion, but adding Matrigel Matrix (Corning® Inc, NY, USA) in each Transwell® Chamber. Invasion was analyzed after 24 and 48 h after seeding.

Cell migration and invasion was evaluated by double-blind manual counting and image analysis using the software ImageJ.

Caspase 3 activity assays. Cell death induction was evaluated using the Caspase-3 Colorimetric Activity Assay Kit (Merck Millipore®, Burlington MA, USA), following the manufacturer´s instructions. Briefly, MCF7 cells were transfected with sh-RANDOM, sh-FAM83H-AS1 or treated with dimethyl sulfoxide (DMSO) 10%. Caspase 3 activity was measured 24 h, 48 h and 72 h after transfection or DMSO treatment.

Statistical analysis. Kaplan Meier survival analysis for FAM83H-AS1 associated tumors (except for BRCA) were performed in the GPIA platform.

Overall survival (OS) of the BRCA TCGA patients and our independent cohort was analyzed with the Kaplan–Meier model and the multivariable Cox´s regression model. This analysis was performed with the PASW statistics software (SPSS, IBM®, Quarry Bay, Hong Kong). Fischer´s exact test was calculated in order to correlate clinical variables with FAM83H-AS1 expression level. Student´s T-tests were performed to calculate statistical differences in functional in vitro assays. For all statistical tests, the level of significance was < 0.05.

Ethics approval and consent to participate. This study was approved by the Research and Ethics Committee of National Institute of Genomic Medicine and the Institute of Breast Diseases, FUCAM (Registration number: CE2009/11). Written informed consent was obtained from each patient before any procedure.

Data availability

Gene Expression Omnibus (GEO): data submitted.

Received: 4 July 2019; Accepted: 10 August 2020
Published online: 24 August 2020

References

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. https://doi.org/10.3322/caac.21492 (2018).
2. Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167. https://doi.org/10.1200/JCO.2008.18.1370 (2009).
3. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752. https://doi.org/10.1038/35021093 (2000).
4. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* **98**, 10869–10874. https://doi.org/10.1073/pnas.191367098 (2001).
5. Solin, L. J. *et al.* A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. *J. Natl. Cancer Inst.* **105**, 701–710. https://doi.org/10.1093/jnci/djt067 (2013).
6. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773. https://doi.org/10.1093/nar/gky955 (2019).
7. Su, X. *et al.* Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. *Oncotarget* **5**, 9864–9876. https://doi.org/10.18632/oncotarget.2454 (2014).
8. Saleembhasha, A. & Mishra, S. Long non-coding RNAs as pan-cancer master gene regulators of associated protein-coding genes: A systems biology approach. *PeerJ* **7**, e6388. https://doi.org/10.7717/peerj.6388 (2019).
9. Beltran-Anaya, F. O. *et al.* 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.* **13**, 909–927. https://doi.org/10.1002/1878-0261.12446 (2019).
10. Hanly, D. J., Esteller, M. & Berdasco, M. Interplay between long non-coding RNAs and epigenetic machinery: Emerging targets in cancer?. *Trans. R Soc Lond. B Biol. Sci Philos* https://doi.org/10.1098/rstb.2017.0074 (2018).
11. Zhang, J. *et al.* Overexpression of FAM83H-AS1 indicates poor patient survival and knockdown impairs cell proliferation and invasion via MET/EGFR signaling in lung cancer. *Sci. Rep.* **7**, 42819. https://doi.org/10.1038/srep42819 (2017).
12. Yang, F. *et al.* Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer. *Oncol. Targets Ther.* **9**, 7039–7045. https://doi.org/10.2147/OTT.S110055 (2016).
13. Shan, H. *et al.* FAM83H-AS1 is associated with clinical progression and modulates cell proliferation, migration, and invasion in bladder cancer. *J. Cell Biochem.* **120**, 4687–4693. https://doi.org/10.1002/jcb.27758 (2019).
14. Lu, S., Dong, W., Zhao, P. & Liu, Z. lncRNA FAM83H-AS1 is associated with the prognosis of colorectal carcinoma and promotes cell proliferation by targeting the Notch signaling pathway. *Oncol. Lett.* **15**, 1861–1868. https://doi.org/10.3892/ol.2017.7520 (2018).
15. Dou, Q. *et al.* LncRNA FAM83H-AS1 contributes to the radioresistance, proliferation, and metastasis in ovarian cancer through stabilizing HuR protein. *Eur. J. Pharmacol.* **852**, 134–141. https://doi.org/10.1016/j.ejphar.2019.03.002 (2019).
16. Deva Magendhra Rao, A. K. *et al.* Identification of lncRNAs associated with early-stage breast cancer and their prognostic implications. *Oncol. Mol.* https://doi.org/10.1002/1878-0261.12489 (2019).
17. Allred, D. C., Harvey, J. M., Berardo, M. & Clark, G. M. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod. Pathol.* **11**, 155–168 (1998).
18. Wolf, B. B., Schuler, M., Echeverri, F. & Green, D. R. Caspase-3 is the primary activator of apoptotic DNA fragmentation via DNA fragmentation factor-45/inhibitor of caspase-activated DNase inactivation. *J. Biol. Chem.* **274**, 30651–30656. https://doi.org/10.1074/jbc.274.43.30651 (1999).
19. Zhao, X. *et al.* Lentiviral vector mediated claudin1 silencing inhibits epithelial to mesenchymal transition in breast cancer cells. *Viruses* **7**, 2965–2979. https://doi.org/10.3390/v7062755 (2015).
20. Wei, L. *et al.* Leptin promotes epithelial-mesenchymal transition of breast cancer via the upregulation of pyruvate kinase M2. *J. Exp. Clin. Cancer Res.* **35**, 166. https://doi.org/10.1186/s13046-016-0446-4 (2016).
21. Strong, A. L. *et al.* Leptin produced by obese adipose stromal/stem cells enhances proliferation and metastasis of estrogen receptor positive breast cancers. *Breast Cancer Res.* **17**, 112. https://doi.org/10.1186/s13058-015-0622-z (2015).
22. Li, K. *et al.* Leptin promotes breast cancer cell migration and invasion via IL-18 expression and secretion. *Int. J. Oncol.* **48**, 2479–2487. https://doi.org/10.3892/ijo.2016.3483 (2016).
23. Huang, Y. *et al.* Leptin promotes the migration and invasion of breast cancer cells by upregulating ACAT2. *Cell Oncol. Disord.* **40**, 537–547. https://doi.org/10.1007/s13402-017-0342-8 (2017).
24. Bi, Y. Y., Shen, G., Quan, Y., Jiang, W. & Xu, F. Long noncoding RNA FAM83H-AS1 exerts an oncogenic role in glioma through epigenetically silencing CDKN1A (p21). *J. Cell Physiol.* **233**, 8896–8907. https://doi.org/10.1002/jcp.26813 (2018).
25. Tan, Y., Wu, F., Tamayo, P., Haining, W. N. & Mesirov, J. P. Constellation map: Downstream visualization and interpretation of gene set enrichment results. *F1000Research* **4**, 167. https://doi.org/10.12688/f1000research.6644.1 (2015).

Acknowledgements

We thank Dr. Camilo Rios, Dr. Araceli Ruiz and M.C. Marcela Islas for valuable and critic comments regarding the apoptosis experimental design.

Author contributions

A.H-M. and M.R-R. conceived and designed the study; M.R.-R., A.C.-T., M.P-L., F.O.B.-A. and M.C.-V. performed the in vitro experiments; M.A.M.-R., L.H.-P., E.O.-M. tested the expression levels in biological samples; S.J.-M., R.A.-L., L.A.A.-R. handled the samples and clinical database; A.T.-T., C.D.-R., F.V.-C. and L.H.-P. did the clinical and pathological evaluation of the samples; M.R.-R. and A.C.-T. performed statistical analysis and differential expression analysis; M.R.-R. and A.H.-M. interpreted and discussed the data.

Funding

This work was funded by the Mexican National Council of Science and Technology (CONACYT) 480751/282036 grant (Scholarship), Mexican National Council of Science and Technology Basic Science grant (CONACYT Grant Number 258936) and Frontiers in Science Grant (Number 1285).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-71062-2.

Correspondence and requests for materials should be addressed to A.H.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

frontiers in Oncology

ORIGINAL RESEARCH
published: 12 April 2021
doi: 10.3389/fonc.2021.628027

LINC00460 Is a Dual Biomarker That Acts as a Predictor for Increased Prognosis in Basal-Like Breast Cancer and Potentially Regulates Immunogenic and Differentiation-Related Genes

Mireya Cisneros-Villanueva[1,2†], Lizbett Hidalgo-Pérez[1,3†], Alberto Cedro-Tanda[1], Mónica Peña-Luna[1], Marco Antonio Mancera-Rodríguez[1], Eduardo Hurtado-Cordova[1], Irene Rivera-Salgado[4], Alejandro Martínez-Aguirre[4], Silvia Jiménez-Morales[1], Luis Alberto Alfaro-Ruiz[1], Rocío Arellano-Llamas[1], Alberto Tenorio-Torres[5], Carlos Domínguez-Reyes[5], Felipe Villegas-Carlos[5], Magdalena Ríos-Romero[1,6] and Alfredo Hidalgo-Miranda[1*]*

[1] Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica (INMEGEN), Ciudad de México, México, [2] Laboratorio de Epigenética del Cáncer, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, Chilpancingo de los Bravo, Mexico, [3] Programa de Doctorado en Ciencias Biomédicas, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, Mexico, [4] Departamento de Anatomía Patológica, Hospital Central Sur de Alta Especialidad, Petróleos Mexicanos, Ciudad de México, México, [5] Instituto de Enfermedades de la Mama, FUCAM, Ciudad de México, Mexico, [6] Posgrado en Ciencias Biológicas, Unidad de Posgrado, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, México

OPEN ACCESS

Edited by:
Wenwen Zhang,
Nanjing Medical University, China

Reviewed by:
Haruhiko Sugimura,
Hamamatsu University School of Medicine, Japan
Kamran Ghaedi,
University of Isfahan, Iran
Ruoxi Yuan,
Hospital for Special Surgery, United States

***Correspondence:**
Alfredo Hidalgo-Miranda
ahidalgo@inmegen.gob.mx
Magdalena Ríos-Romero
magdarios@ciencias.unam.mx

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *10 November 2020*
Accepted: *10 March 2021*
Published: *12 April 2021*

Citation:
Cisneros-Villanueva M, Hidalgo-Pérez L, Cedro-Tanda A, Peña-Luna M, Mancera-Rodríguez MA, Hurtado-Cordova E, Rivera-Salgado I, Martínez-Aguirre A, Jiménez-Morales S, Alfaro-Ruiz LA, Arellano-Llamas R, Tenorio-Torres A, Domínguez-Reyes C, Villegas-Carlos F, Ríos-Romero M and Hidalgo-Miranda A (2021) LINC00460 Is a Dual Biomarker That Acts as a Predictor for Increased Prognosis in Basal-Like BreastCancer and Potentially Regulates Immunogenic and Differentiation-Related Genes. Front. Oncol. 11:628027. doi: 10.3389/fonc.2021.628027

Breast cancer (BRCA) is a serious public health problem, as it is the most frequent malignant tumor in women worldwide. BRCA is a molecularly heterogeneous disease, particularly at gene expression (mRNAs) level. Recent evidence shows that coding RNAs represent only 34% of the total transcriptome in a human cell. The rest of the 66% of RNAs are non–coding, so we might be missing relevant biological, clinical or regulatory information. In this report, we identified two novel tumor types from TCGA with LINC00460 deregulation. We used survival analysis to demonstrate that LINC00460 expression is a marker for poor overall (OS), relapse-free (RFS) and distant metastasis-free survival (DMFS) in basal-like BRCA patients. LINC00460 expression is a potential marker for aggressive phenotypes in distinct tumors, including HPV-negative HNSC, stage IV KIRC, locally advanced lung cancer and basal-like BRCA. We show that the LINC00460 prognostic expression effect is tissue-specific, since its upregulation can predict poor OS in some tumors, but also predicts an improved clinical course in BRCA patients. We found that the LINC00460 expression is significantly enriched in the Basal-like 2 (BL2) TNBC subtype and potentially regulates the WNT differentiation pathway. LINC00460 can also modulate a plethora of immunogenic related genes in BRCA, such as *SFRP5, FOSL1, IFNK, CSF2, DUSP7* and *IL1A* and interacts with miR-103-a-1, *in-silico*, which, in turn, can no longer target WNT7A. Finally, LINC00460:WNT7A ratio constitutes a composite

marker for decreased OS and DMFS in Basal-like BRCA, and can predict anthracycline therapy response in ER-BRCA patients. This evidence confirms that LINC00460 is a master regulator in BRCA molecular circuits and influences clinical outcome.

Keywords: LINC00460, breast cancer, basal-like, biomarker, increased prognosis, mir-103a, sponge, WNT7A

INTRODUCTION

Breast cancer (BRCA) is a major public health problem, as it is the most frequent malignant tumor in women worldwide. According to GLOBOCAN (2018), at least 2.088 million new cases and a total of 626,679 deaths were reported globally (1). Breast cancer is a phenotypically heterogeneous disease, with well-defined histological types. In 2000, Perou et al. (2), suggested that this physical heterogeneity is also reflected on molecular level, particularly on transcriptome, where same histological types can display a variety of differentially expressed genes. Four BRCA subgroups are differentiated by the expression of 50 genes (the PAM50 signature): Luminal A, Luminal B, HER2-enriched and Basal-like tumors. This molecular classification has been used to discern between aggressive and non-aggressive tumors, evaluate metastatic potential, establish clinical prognosis and estimate survival, among other relevant cancer-related features. Furthermore, patients´ therapy response-associated expression profiles can be subjected to the same classification, allowing us to predict a type of therapy a patient can benefit from, e.g. chemotherapy or anti-hormonal therapy (3). All these clinical advances are only focused on coding RNAs profiles; however, more recent reports have shown that coding (messenger) RNAs represent only 2% of the total transcriptome in a human cell (4). The remaining 98% of transcriptome are non-coding RNAs, that might nevertheless carry relevant biological and clinical information.

Long non-coding RNAs (lncRNAs) are non-translated transcripts with a length of 200 nucleotides or superior. In recent years, several papers have focused on their role in different types of cancer showing their contribution to critical biological processes including carcinogenesis, apoptosis, differentiation, proliferation, invasion and metastasis among others (5–8). LncRNAs display multiple regulatory functions, as they can act as modulators of transcription and chromatin remodeling (9, 10), as splicing factors (11, 12), as regulators of mRNA decay (13), mRNA stability (14), as protein decoys (15) and microRNA (miRNA) sponges (16, 17). In this sponging effect, a lncRNA competes with a miRNA to release the inhibition of other genes (18, 19). Several reports have shown that sponge lncRNAs play a pivotal role in various cancer types (20–24), including BRCA (25, 26) and abnormal expression of lncRNAs can significantly contribute to BRCA initiation and progression (25, 27, 28).

The long intergenic non-coding RNA 460 (LINC00460) is a human lncRNA gene, transcribed from chromosome 13q33.2 and measuring 913 bp (29). The LINC00460 is constituted by three exons and it has been shown that it lacks coding capability (30, 31). According to the ENSEMBL database, the human LINC00460 gene has seven splice variant transcripts reported to date (32). It has been observed that LINC00460 functions as an oncogene, acting as a miRNA sponge (17, 33–43). In these models, LINC00460 sponging activity and its association with cancer-related processes such as increase in proliferation, epithelial to mesenchymal transition (EMT), migration, invasion and metastasis (44–46) is well described. For example, LINC00460 promotes hepatocellular carcinoma progression by sponging miR-342-3p and hence increasing AGR2 expression level (47). Similar LINC00460 sponging and cancer-related mechanisms across distinct tumors are described in several reports (38–43, 48).

The LINC00460 role in clinical cancer features has also been studied. LINC00460 over-expression is strongly associated with poor survival rates in different tumors, such as lung, ovary, larynx, nasopharynx, head and neck, meningioma, kidney, thyroid, colorectal, glioma, osteosarcoma, bladder and cervix (31, 35, 48–56). Thus, LINC00460 is a well-defined sponge lncRNA with significant prognostic potential in several tumors, however its clinical and biological relevance in BRCA is poorly understood (33), and thorough studies are needed.

In this study, we aimed to elucidate the potential role of LINC00460 in BRCA. For this purpose, we performed a series of analysis to infer its biological relevance and verify the underlying role of LINC00460 in BRCA. We identified two novel (not previously reported) tumor types from the TCGA cohort with LINC00460 deregulation. In addition, we used a multivariate Cox regression analysis to demonstrate that LINC00460 expression is related to poor prognosis in three different tumors. We show here that LINC00460 upregulation is significantly associated with improved survival in BRCA in three independent cohorts. This increased survival effect is replicated in Basal-like BRCA and furthermore, we demonstrate that LINC00460 is significantly enriched in the Basal-like 2 (BL2) Triple Negative Breast Cancer (TNBC) subtype. Using differential expression and correlation analysis, we show that LINC00460 overexpression impairs several breast cancer-related pathways: WNT signaling pathway, cytokine inflammatory response and DNA damage response. Among the most relevant potential LINC00460 gene targets, we found *WNT7A, SFRP5, FOSL1* and *IFNK* to be deregulated in the LINC00460-high condition. Using *in-silico* prediction analysis, we show that LINC00460 could interact with miR-103-a, which in turn, potentially regulates *WNT7A*. Finally, we demonstrate that LINC00460: WNT7A ratio is a composite marker for increased OS and DMFS in Basal-like BRCA, and their combined expression can predict anthracycline therapy response in ER-BRCA patients, which further pinpoints the biological and clinical role of these transcripts in TNBC.

MATERIALS AND METHODS

The Cancer Genome Atlas (TCGA) Datasets

LINC00460 expression levels were screened in 31 TCGA tumor datasets and their corresponding GTEx normal tissues using the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) platform (http://gepia2.cancer-pku.cn/#index) (57). The 31 tumors included in the analysis are summarized in **Supplementary Table 1** and enlisted as follows: Acute myeloid leukemia (LAML); Adrenocortical carcinoma (ACC); Bladder urothelial carcinoma (BLCA); Brain lower grade glioma (LGG); Breast invasive carcinoma (BRCA); Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC); Cholangiocarcinoma (CHOL); Colon adenocarcinoma (COAD); Esophageal carcinoma (ESCA); Glioblastoma multiforme (GBM); Head and neck squamous cell carcinoma (HNSC); Kidney chromophobe (KICH); Kidney renal clear cell carcinoma (KIRC); Kidney renal papillary cell carcinoma (KIRP); Liver hepatocellular carcinoma (LIHC); Lung adenocarcinoma (LUAD); Lung squamous cell carcinoma (LUSC); Lymphoid neoplasm diffuse large B-cell lymphoma (DLBC); Ovarian serous cystadenocarcinoma (OV); Pancreatic adenocarcinoma (PAAD); Pheochromocytoma and Paraganglioma (PCPG); Prostate adenocarcinoma (PRAD); Rectum adenocarcinoma (READ); Sarcoma (SARC); Skin cutaneous melanoma (SKCM); Stomach adenocarcinoma (STAD); Testicular germ cell tumors (TGCT); Thymoma (THYM); Thyroid carcinoma (THCA); Uterine carcinosarcoma (UCS); Uterine corpus endometrial carcinoma (UCEC). Mesothelioma (MESO) and Uveal melanoma (UVM) were removed from the analysis, since the correspondent normal tissues were not available in the GEPIA2 platform.

Additional analysis with TCGA breast cancer, normal tissues and HNSC HPV status, was performed using data retrieved from the TANRIC platform (https://ibl.mdanderson.org/tanric/_design/basic/main.html) (58). Violin plots, box plots and notch plots were constructed using the ggplot2 R package. Validation of LINC00460 over-expression in BRCA versus normal tissues was performed using data retrieved from the GSE29431 dataset (https://www.ncbi.nlm.nih.gov/geo/) and analyzed using the lnCAR platform (https://lncar.renlab.org/) (59).

LUAD, LUSC and KIRC tumor stage plots were generated in the GEPIA2 platform. LINC00460 expression levels were considered significantly correlated with tumors when log2 fold change (Log2FC)>1 and p-value<0.01.

Breast Cancer Samples Differential Expression Analysis

Breast cancer RNA seq counts were obtained from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov). After dataset preparation, we identified the LINC00460 ID (ENSG00000233532), which is located in chromosome 13: 106,376,563-106,378,217. We then downloaded the LINC00460 expression counts and performed filtering of transcripts with 10 counts or less. In order to generate the high and low LINC00460 expression groups, we calculated two percentiles from the count expression data. The first percentile (25) contains the lowest LINC00460 expression counts, and the upper percentile (60), contains the highest expression levels for this transcript. We then performed differential expression analysis with the DESeq2 tool from the Gene Pattern platform, using the default parameters (http://software.broadinstitute.org/cancer/software/genepattern/) (61).

Genes were considered differentially expressed when Log2FC was >1.5 and -<1.5 and p- adjusted value <0.05. Volcano plots were generated with the ggplot2 and ggrepel R packages.

Breast Cancer Patients and Biological Samples

A total of 74 biological samples (frozen tissues) were collected from breast cancer patients attending Fundación de Cáncer de Mama A.C. (FUCAM) in Mexico City, Mexico. Formalin fixed paraffin embedded (FFPE) breast cancer samples (n=19) were collected from the Hospital Central Sur de Alta Especialidad, Petróleos Mexicanos (PEMEX). None of the patients received neoadjuvant therapy. All patients signed a written informed consent. The studies involving human participants were reviewed and approved by the Research Ethics Committee (INMEGEN) and the FUCAM Ethics Committee (Registration number: CE2009/11).

Biological samples were bisected; one portion was fixed in formaldehyde (10%), paraffin embedded (FFPE) (Paraplast Plus®; Sigma Aldrich ®, St Louis, Missouri, USA) and then submitted to Hematoxylin and Eosin (H&E) staining for histopathological examination by a certified pathologist. Tumor stage was assessed, according to international standards. The second portion of the sample was used for RNA extraction and functional downstream analysis. All tissues were liquid nitrogen-frozen and stored at -80°C. Samples with more than 80% tumor cells were included in the analysis, otherwise were discarded.

An expert pathologist performed the standardized evaluation of stromal tumor-infiltrating lymphocytes (TILs) based on H&E -stained slides of breast tumoral tissue. Briefly, TILs assessment was performed as follows: a) only TILs within the borders of the invasive tumors were evaluated; b) the invasive edge was included in the evaluation; c) only mononuclear infiltrates were included; d) immune infiltrates in adjacent normal tissue or areas of central necrosis or fibrosis were not included; e) the average TILs of the stromal area were reported. For the purposes of this research, the cut-off points of TILs were defined in less than 50% and more than 50% in the stromal area.

In all cases, demographic (age, sex), clinical (date of diagnosis, therapy received), pathologic (stage, grade, histological type) and prognostic data (recurrence, progression and overall survival) were available and correlated with LINC00460 expression status.

RNA Extraction

Total RNA was extracted using the commercial kit AllPrep® DNA/RNA FFPE (Qiagen® Inc, Valencia, CA) following manufacturer's instructions. Briefly, the tissues were deparaffinized, disrupted

and lysed. RNA was then precipitated, washed, purified and suspended in RNAse free water. RNA concentration was evaluated by spectrophotometry (NanoDrop Technologies, Wilmington, Delaware, USA). RNA integrity was analyzed using the BioAnalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA), only high-quality samples were used. Samples were stored at -80°C until processing.

Microarray Re-Analysis

Using the BRCA cohort previously reported by our group in (62) (n = 74; Luminal A = 24, Luminal B = 23, HER2 = 14, Basal = 13) in which samples were analyzed using Human Transcriptome Array 2.0 (Affymetrix, Inc, Santa Clara, CA), we were able to identify LINC00460 expression levels across BRCA subtypes. We used the Robust Multi-chip Analysis (RMA) algorithm to minimize the effect of probe-specific affinity differences and to normalize samples (63). Log 2 relative fluorescent signal intensities were computed using the Transcriptome Analysis Console (Affymetrix, Inc, Santa Clara, CA).

Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

cDNA was synthetized using SuperScript III RT-PCR (Invitrogen, ThermoFisher™ Scientific, Waltham, Massachusetts, USA) and High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems™, Foster City, California, USA), following the manufacturer's instructions. Briefly, 100 ng of total RNA from cell lines or breast cancer samples were used to synthesize cDNA in a final reaction volume of 20μL. The PCR reaction contained 1μL of cDNA, 5 μL 2X TaqMan Universal Master Mix (Applied Biosystems, ThermoFisher™ Scientific, Waltham, Massachusetts, USA), 0.5 μL TaqMan probes (custom-made for LINC00460) and 3.5 μL of nuclease-free water. Both primers and the reporter were designed to target LINC00460, exon 2, transcript variant 1 (NCBI Reference Sequence: NR_034119.2). Forward primer: CCTGGATGAACCACCATTGC; reverse primer: ATGAGAACGAAGGTTACGACCATT; reporter: ATGTTGCAGCTTTCCCA). GAPDH (Hs99999905) and SCARNA5 (Hs03391742_cn) transcripts were used as endogenous controls.

Pathway Enrichment Analysis

Pathway enrichment analysis was performed using Ingenuity Pathway Analysis® (IPA) software. Z-scores and p-values were also computed using this platform. Only differential statistically significant genes were included in this analysis (see criteria above).

Overrepresentation Enrichment Analysis (ORA) was performed with the web-based Gene SeT Analysis Toolkit (WebGestalt) platform (www.webgestalt.org) (64). LINC00460 significantly correlated genes were included in this analysis.

BRCA and HNSC Datasets for Validation

Survival analyses from independent BRCA Gene Expression Omnibus (GEO) cohorts were performed using the Kaplan-Meier plotter (KM plotter) site (http://kmplot.com/analysis/) (65). These GEO datasets contain gene expression data from 6,234 BRCA samples analyzed with multiple microarray platforms, including Human Genome U133 and Human Genome U133 Plus 2.0 Arrays (Affymetrix). We used GEO-derived cohorts, the GSE16446 and GSE21653 datasets, or a combination of several datasets. The GSE16446 dataset was selected, since it contains 120 microarray experiments from primary ER-negative breast tumors of anthracycline-treated patients (66). The GSE21653 dataset contains data of 266 invasive breast adenocarcinomas, with all BRCA subtypes included (67, 68). We used the Affymetrix probe ID 1558930_at, that targets LINC00460, and the Affymetrix probe ID 210248_at, which targets WNT7A. In addition, we performed survival analysis of hsa-miR-103a expression levels, using the mirPower tool (69), included in the KM plotter website.

LINC00460:WNT7A ratio for TNBC survival analyses was also performed in the KM plotter tool. LINC00460 and WNT7A signature expression validation as anthracycline predictive markers were analyzed with the ROC Plotter platform (http://www.rocplot.org/site/index) (70).

For HNSC validation, we downloaded the GEO dataset GSE3292. This cohort has 36 freshly frozen HNSC samples, stratified by HPV status (negative or positive) and analyzed by gene expression microarrays (Affymetrix Human Genome U133 Plus 2.0 Array).

For further TNBC validation, we downloaded the GEO dataset GSE76250, which includes 165 samples. We then analyzed the expression profile of these samples, and classified them with the TNBCtype tool to obtain the Lehman subtypes (http://cbc.mc.vanderbilt.edu/tnbc) (71). We then re-classified these samples into the TNBCtype4 re-assigning IM and MSL subtypes to the second highest correlated centroid, as described in (72). For LINC00460 expression in TNBC cell lines validation, we screened the Cancer Cell Line Encyclopedia (CCLE) online tool (https://portals.broadinstitute.org/ccle) (73).

The validation datasets analyzed for this study can be found in the Gene Expression Omnibus (GEO) web site (https://www.ncbi.nlm.nih.gov/geo/).

miRNA Interaction Prediction Analysis

We used the miRcode prediction tool (http://www.mircode.org/) (74) in order to identify potential LINC00460 miRNA targets. We further corroborated these potential interactions using the lncTAR tool (http://www.cuilab.cn/lnctar) (75). Interaction prediction was considered valid when normalized delta G (ndG) values reached the -0.1 cutoff.

We then performed a third *in-silico* prediction analysis, using the miRPathDB tool (https://mpd.bioinf.uni-sb.de/) (60). We searched the top 15 correlated mRNAs with LINC00460 expression levels (see **Table 4**) and downloaded the miRNAs interaction list. We then compared these new miRNAs list with the previous one described above, and identified potential miRNA-mRNA-linc00460 interactions.

Statistical Analysis

It has been previously shown that cancer gene expression profiles are not normally-distributed, either on the complete-experiment or on the individual-gene level (76). Thus, LINC00460 expression distributions from the TCGA and validation BRCA, KIRC, LUAD and HSNC datasets were first tested for normality

distribution using the Kolmogorov-Smirnov and the Shapiro-Wilk tests. After results computation, we selected comparison statistical tests accordingly.

Overall Survival (OS) using all patients from CESC, GBM, HNSC, LGG, LUAD, PAAD and SARC and the Relapse Free Survival (RFS) analysis using all patients from KIRC, LUAD, READ and SARC was performed in the GEPIA2 platform. Distant Metastasis Free Survival (DMFS) was computed using the KM plotter tool described above.

OS of the BRCA TCGA patients and our independent cohort was analyzed with the Kaplan-Meier model and the multivariable Cox regression model. Relative Risk was also calculated. This analysis was performed with the PASW statistics software (SPSS, IBM®, Quarry Bay, Hong Kong). Chi-square tests were calculated in order to correlate clinical variables status with LINC00460 expression level (high or low expression).

For all statistical tests, the level of significance was <0.05.

RESULTS

LINC00460 Expression Is Deregulated in Multiple Tumors

LINC00460 expression was evaluated in 31 tumor types and normal tissues included in TCGA database. We found that LINC00460 is overexpressed in five different epithelial cancers, namely: Breast cancer (BRCA) (**Figure 1B**); Colon adenocarcinoma (COAD); Head and Neck Squamous cell carcinoma (HNSC), Pancreatic adenocarcinoma (PAAD) and Rectum adenocarcinoma (READ) (Log2FC>1; p<0.05) (**Figure 1A**). Interestingly, we have also detected two central nervous system cancers with LINC00460 low expression, comparing with its normal tissue counterparts: Glioblastoma Multiforme (GBM) and Low-Grade Glioma (LGG) (Log2FC<1; p<0.01) (**Supplementary Figure S1**). These observations potentially suggest a different role for LINC00460 in central nervous system tumors, which can further be analyzed in future studies.

Among the seven LINC00460-deregulated tumors from the TCGA cohort, we identified two novel (not previously reported) tumors, namely LGG and GBM (**Supplementary Figure S1**). We have also confirmed previous observations regarding LINC00460 overexpression in BRCA (33), comparing with normal tissue, using two independent cohorts, namely TCGA (Tanric data, T test; p=1.6845E-14) and the GEO cohort GSE29431 (**Figure 1** and **Supplementary Figure S2**). These observations suggest a ubiquitous role for LINC00460 in cancer biology.

LINC00460 Expression Is Associated With Advanced Clinical Stages and Aggressive Phenotypes in Different Cancers

LINC00460 over-expression was significantly associated with advanced and locally advanced tumor clinical stages in three

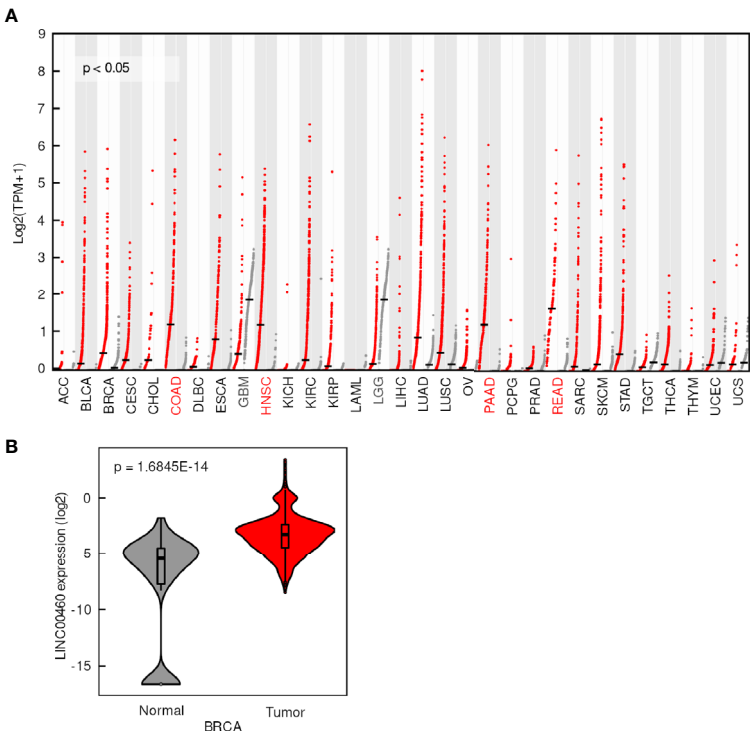

FIGURE 1 | LINC00460 is aberrantly expressed in different tumors. **(A)** LINC00460 expression levels in 33 tumors (red dots) and its correspondent normal tissues (grey dots). Tumors with overexpression and down expression of LINC00460 have red and grey color abbreviations, respectively. Data was obtained from GEPIA2 and modified. **(B)** LINC00460 is over-expressed in breast cancer compared with normal tissues. Data was retrieved from TANRIC and then re-plotted.

distinct cancers: LUSC, LUAD and KIRC (ANOVA; p<0.05; **Supplementary Figures S3A–C**). In the KIRC model, LINC00460 over-expression is significantly associated with high histological grade (Mann Whitney U-test; p=0.0001; **Supplementary Figure S3D**). In addition, LINC00460 overexpression is correlated with further aggressive tumor phenotypes, such as HPV negative HNSC (TCGA: Mann-Whitney U, p< 0.05. GSE3292: one-way ANOVA, p= 0.02) and basal-like BRCA (TCGA: Kruskal-Wallis test, p=3.14E-13. Mexican cohort: Kruskal-Wallis test, p=0.012), using two independent cohorts for each cancer type (**Supplementary Figure S4** and **Figures 2A and B**). This last observation was further validated in a panel of BRCA cell lines, where we detected that LINC00460 is mainly expressed in basal TNBC cells (**Figure 2C**). CCLE screening confirmed TNBC LINC00460 enrichment (**Supplementary Figure S6**). As LINC00460 expression is significantly related to Basal-like BRCA, we reasoned that LINC00460 over-expression might also be associated with Immunohistochemistry (IHC)-detected Estrogen Receptor (ER) and Progesterone Receptor (PR) status in BRCA patients. As expected, LINC00460 expression is enriched in IHC-detected negative Estrogen receptor (ER) (Mann-Whitney U; p=0.000079) and Progesterone Receptor (PR) (Mann-Whitney U; p= 0.024) TCGA BRCA patients (**Figure 2D**). These observations were validated in an independent mexican cohort (Chi squared test ER p-value= 0.028; Chi squared test PR p-value=0.001) (**Figure 2E**).

LINC00460 Is Related With Poor Prognosis in Eight Different Tumors, but Increased Survival Rate in BRCA

After demonstrating that LINC00460 is significantly related to aggressiveness markers in different tumors, we then aimed to know if its deregulation is also related to overall survival (OS) and relapse free survival (RFS) in these cancers, using the TCGA cohorts. As shown in **Figure 3**, LINC00460 overexpression is significantly related to high risk of death in eight different tumors, analyzed in a single model (CESC, GBM, HNSC, KIRC, LGG, LUAD, PAAD and SARC) (log2HR>1; p<0.05) (**Figure 3A**), and with high risk of relapse in four cancers (KIRC, LUAD, READ and SARC) (log2HR>1; p<0.05) (**Figure 3B**). Taken together with previous findings reported in the field (30, 34, 47, 48, 50–52, 54, 55, 57), this data suggests a relevant role for LINC00460 in clinical cancer biology.

An unexpected observation in this regard is the effect of LINC00460 over-expression in BRCA, which is associated with improved OS (**Figure 4A**), compared with the LINC00460 low expression group in two independent cohorts, namely TCGA (through Tanric tool, n=743; logrank p=0.01) and mexican (n=93; HR=1.655; 95% CI [1.038-26.41]; logrank p=0.045) (**Figure 4B**). The clinical and pathological characteristics of the TCGA and mexican BRCA patients are described in **Table 1**.

In addition, we observed that LINC00460 expression level statistically interacts with PR, HER2 status, patient age, and

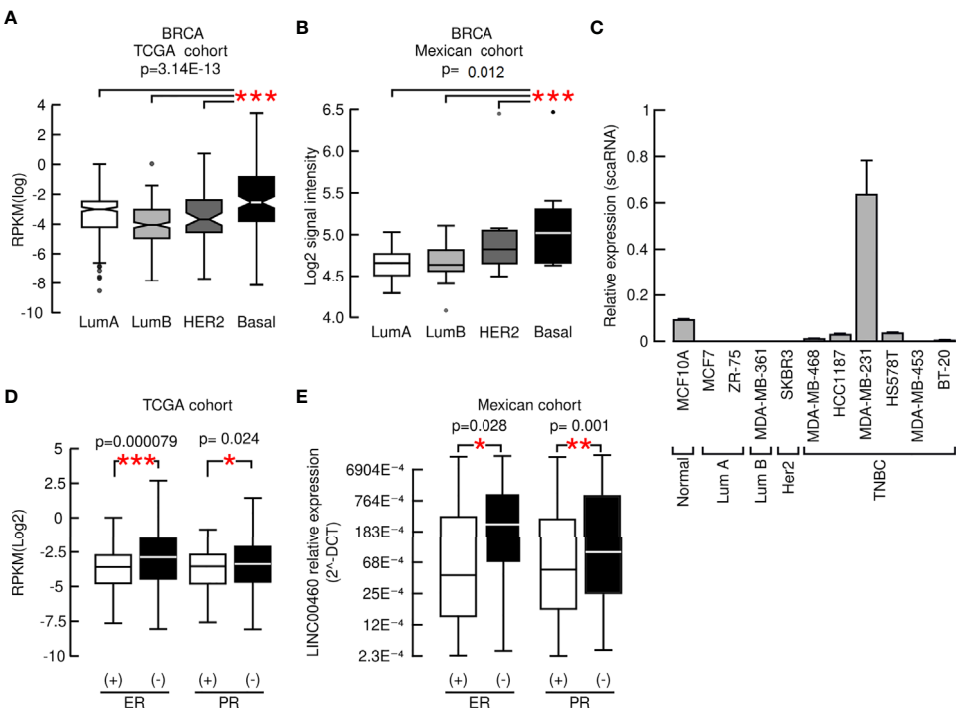

FIGURE 2 | LINC00460 high expression is associated with aggressive phenotypes in BRCA. **(A)** LINC00460 expression is significantly related with Basal-like BRCA in the TCGA cohort, **(B)** in a Mexican independent cohort and **(C)** enriched in TNBC cell lines, **(D)** LINC00460 is significantly enriched in ER- and PR- BRCA in the TCGA and **(E)** an independent Mexican BRCA cohort. *p<0.05, **p<0.005, ***p<0.0005.

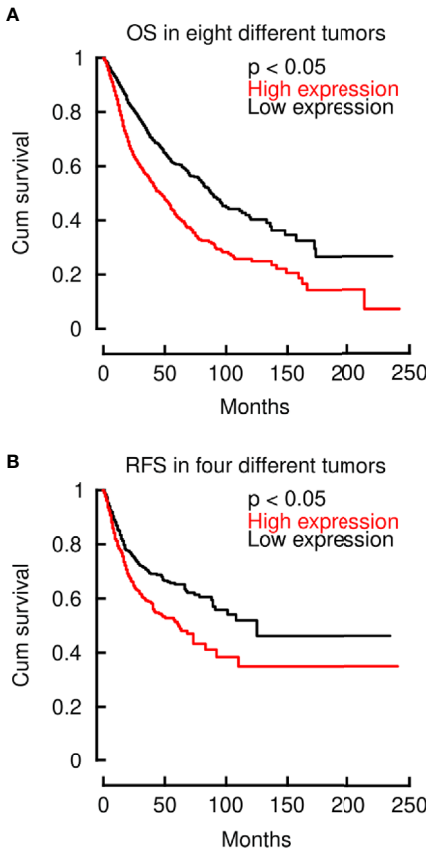

FIGURE 3 | LINC00460 deregulation is related to high risk of decease and relapse across multiple cancers. **(A)** LINC00460 overexpression is significantly associated with elevated HR of decease in eight different tumors (CESC, GBM, HNSC, KIRC, LGG, LUAD, PAAD and SARC). **(B)** High LINC00460 expression is associated with increased risk of relapse in four different tumors (KIRC, LUAD, READ and SARC).

TABLE 1 | Clinicopathological characteristics of the mexican BRCA cohort (n = 107).

Variable	Stratification	Frequency	n		
Age	<50	>50	NA		
	31	69	7		
ER	Negative	Positive	NA		
	35	68	4		
PR	Negative	Positive	NA		
	47	56	4		
HER2	Negative	Positive	NA		
	79	23	5		
Clinical Stage	I	II	III	IV	
	27	64	9	7	
Grade	I	II	III	NA	
	14	64	20	9	
Defunction	Positive	Negative	NA		
	11	94	2		
Metastasis	Positive	Negative	NA		
	9	26	72		
Survival (months)	< 60	>60	NA		
	16	91	0		
Tumor size	<20mm	>21 a 49 mm	>50 mm	NA	
	33	49	9	16	
Molecular subtype (IHC)	Luminal A	Luminal B	Her2	Basal	NA
	44	25	11	17	10
Lymphocyte infiltration	≤50	≥50	NA		
	61	1	22		

tumor grade in the survival Cox regression analysis of the Mexican cohort (**Supplementary Table S2**). These interactions suggest that LINC00460 is involved in important cancer-related processes like tumor differentiation, hormonal status and HER2 expression in mexican BRCA patients, although the exact related mechanisms remain unclear.

We have also assessed RFS in the GSE21653 dataset (n=240; HR = 0.59; 95% CI [0.38 – 0.94]; logrank p = 0.024) and DMFS (n=120; HR=0.78; 95% CI [0.6-1.02]; logrank p= 0.062) in BRCA patients in the GEO-derived cohorts (**Figure 4C** and **Supplementary Figure S5**). We found a corresponding significant association with improved RFS for LINC00460-high patients. Although non-significant, we observed a solid tendency to a higher DMFS in the BRCA LINC00460-overexpressed group (**Supplementary Figure S5**).

These data strongly suggest that LINC00460 might play a dual prognostic role across different tumors, as high LINC00460 expression predicts an increased OS, RFS and DMFS in the BRCA model, but it is also a marker for poor prognosis in at least

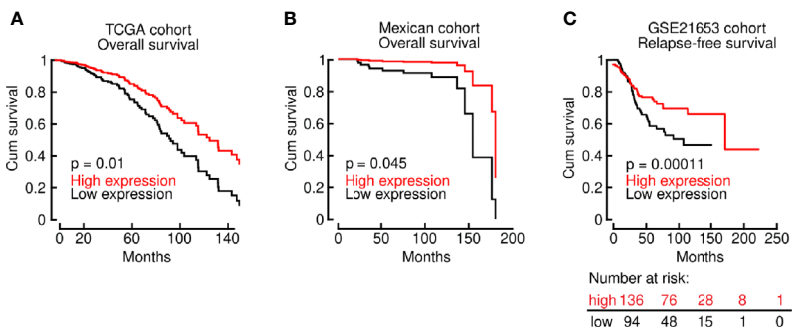

FIGURE 4 | LINC00460 up-regulation is associated with improved OS and RFS in multiple BRCA cohorts. **(A)** LINC00460 over-expression is significantly associated with improved OS in the BRCA TCGA cohort; **(B)** in an independent mexican patient cohort and **(C)** the GSE21653 GEO dataset.

eight distinct solid tumors (see **Figure 3C**). These intriguing findings aimed us to further investigate the role of LINC00460 in BRCA.

LINC00460 Is Significantly Enriched in Basal-Like 2 TNBC and Its Overexpression Predicts a Favorable Clinical Course

We then directed our efforts to elucidate the role of LINC00460 in OS prediction of the aggressive Basal-like BRCA model, a subtype which comprises the majority of TNBC cases (77). Interestingly, the improved OS effect shown in all BRCA patients (**Figures 4A, B**), is reproduced when we analyzed two independent cohorts of Basal-like BRCA tumors only, namely the TCGA cohort (TCGA *via* Tanric tool, n=139; logrank p=0.042) and GEO derived cohort GSE16446 (using basal-like samples only; n= 76; HR=0.23[0.08 −0.63]; logrank p = 0.0019) (**Figures 5A, B**). RFS/DMFS (GSE16446; n= 76; HR = 0.27; [0.11 − 0.68]; logrank p = 0.0027) were also significantly improved in LINC00460 over-expressed BRCA basal-like samples (**Figure 5C**). Analysis with LINC00460 predicted an increased OS in the TNBC GEO derived cohort GSE16446 (using all TNBC samples; n= 107; HR = 0.26; 95% CI [0.09 − 0.72]; logrank p = 0.0053) (**Figure 5D**).

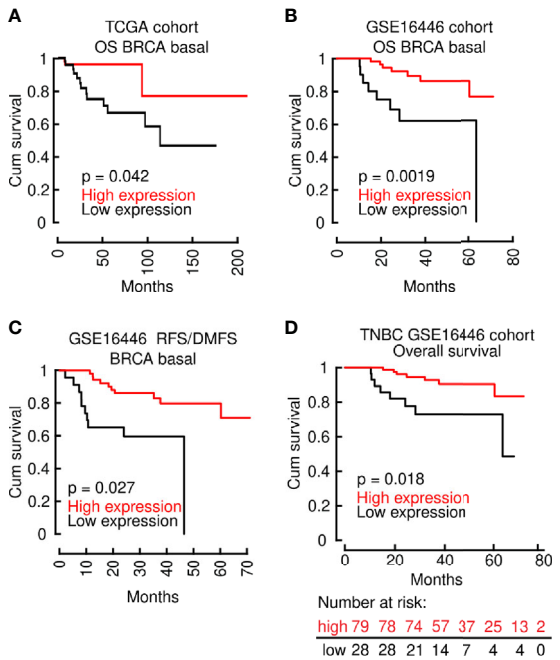

FIGURE 5 | The effect of improved OS, RFS and DMFS in the LINC00460-high group is maintained in Basal type BRCA. **(A)** OS in the TCGA BRCA Basal cohort; **(B)** OS in the GSE16446 BRCA Basal cohort; **(C)** RFS/DMFS in the GSE16446 BRCA Basal cohort; **(D)** OS in the GEO TNBC cohort.

To further characterize LINC00460 association with aggressive BRCA, we analyzed its expression in Lehman refined triple negative breast cancer classification (TNBCtype4) (**Figure 6A**). We observed a significant LINC00460 enrichment in the Basal-like 2 (BL2) subtype (n= 135, Jonckheere-Terpstra test for ordered variables; p=0.026) in a Chinese BRCA cohort (GSE76250). LINC00460 high expression level is also able to predict a favorable clinical response in BL2 triple negative BRCA GEO cohorts, both in OS (n=28; logrank, p=0.0062; **Figure 6B**) and RFS (n= 52; logrank, p=0.04; **Figure 6C**).

Altogether, these observations suggest that LINC00460 expression is generally related to intrinsically aggressive tumor phenotypes, as shown for HNSC, LUSC, LUAD, KIRC and basal-like BRCA (**Figure 2** and **Supplementary Figures S3, 4**). These findings are further corroborated when we showed that LINC00460 is enriched in the TNBCtype4 BL2 subtype in an independent cohort (**Figure 6A**). Interestingly, in aggressive BRCA subtypes, high LINC00460 expression is able to predict a favorable clinical course, further strengthening the dual role for this lncRNA in OS and RFS prediction in cancer.

LINC00460 Potentially Regulate a Plethora of Cancer-Related Genes in BRCA Involved in Proliferation, Cell Cycle and Migration

In order to elucidate the intriguing role of LINC00460 in BRCA, we aimed to identify potential LINC00460 expression targets. Differential expression analysis was performed in TCGA BRCA

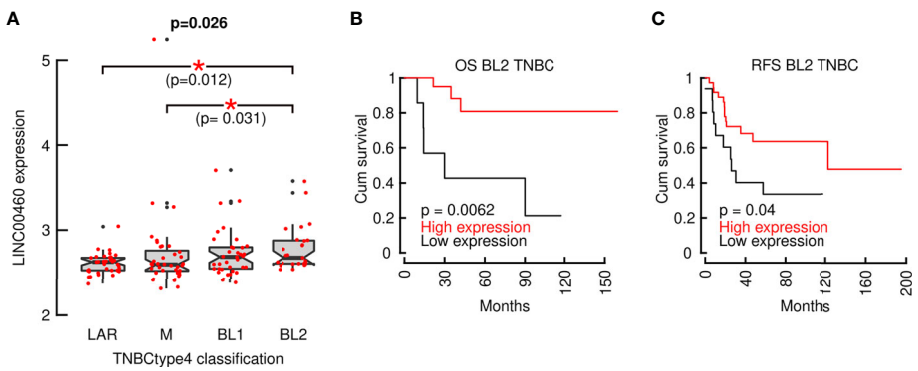

FIGURE 6 | LINC00460 is enriched in Basal-like 2 TNBC (TNBCtype4) and its expression can predict improved OS and RFS. **(A)** LINC00460 is significantly enriched in BL2 TNBC; **(B)** LINC00460 over-expression in associated with increased OS and **(C)** RFS in BL2 TNBC. *p<0.05.

samples between the LINC00460 high expression group and the LINC00460 low-group (see methods for details). This approach revealed 874 significantly deregulated transcripts (FC<1.5 and <-1.5, p.adj value<0.01; **Figure 7A**). Of those, 73% (638 RNAs) were up-regulated, and the remaining were found down-regulated (27%; 236 transcripts) between groups. Among the up-regulated RNAs, we identified several cancer-related transcripts, such as *HOXD13, CXCL1, CXCL5, FOXG1, SERPINB4, CLDN6* and *CLDN10* (see **Table 2**). Down-regulated transcripts list includes *EMX1, CYP2G1P, AMELX* and *SOX5-AS1* (**Table 3**). Furthermore, Ingenuity Pathway Analysis revealed that the proliferation process was negatively enriched in the high LINC00460 group (z score= -0.513). Cellular migration, adhesion, cell cycle progression and proliferation were also enriched in these BRCA samples, although no z-score enrichment was detected in these pathways (see **Figure 7B**).

This data pinpoints the role of LINC00460 as a potential regulator of transcripts and cellular cancer-related processes like proliferation, migration and cell proliferation in BRCA.

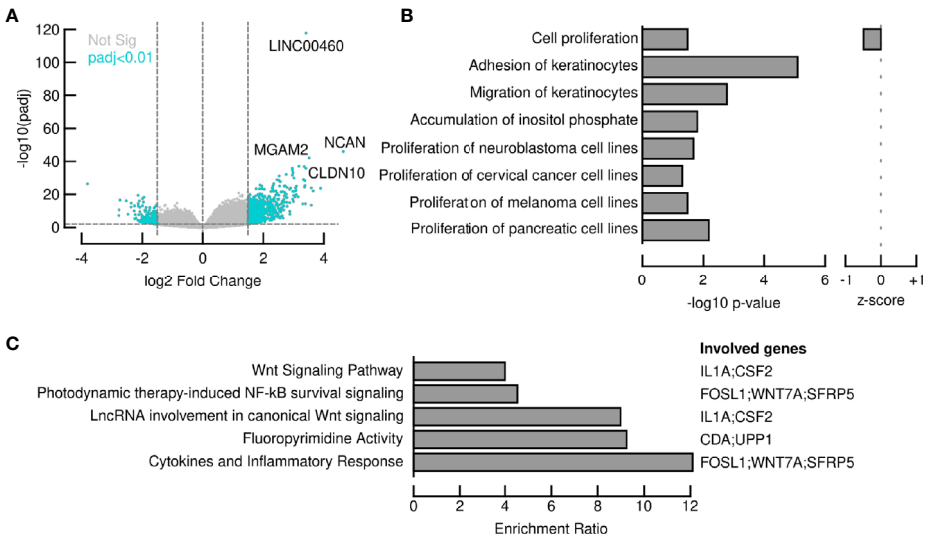

FIGURE 7 | LINC00460 potentially regulates WNT signaling pathway and the cytokine inflammatory response in BRCA. **(A)** 874 transcripts were significantly deregulated after comparing LINC00460-high samples vs LINC00460-low samples; **(B)** Several cancer-related pathways, such as cell proliferation, cellular migration, adhesion, cell cycle progression and carbohydrate metabolism are potentially altered in LINC00460-high BRCA samples; **(C)** ORA revealed LINC0460 involvement in WNT signaling and cytokine inflammatory response among other pathways in BRCA samples.

TABLE 2 | LINC00460 high vs low up-regulated genes.

	Gene symbol	Gene name	log2FC	padj-value
ENSG00000130287	NCAN	Neurocan	4.7	5.41E-50
ENSG00000103355	PRSS33	Serine Protease 33	3.7	1.22E-24
ENSG00000140522	RLBP1	Retinaldehyde binding protein 1	3.7	4.03E-27
ENSG00000275216	AL161431.1		3.7	8.28E-15
ENSG00000112238	PRDM13	PR Domain-containing protein 13	3.5	9.31E-17
ENSG00000177359	AC024940.2		3.4	2.06E-38
ENSG00000233532	LINC00460		3.4	1.00E-126
ENSG00000134760	DSG1	Desmoglein 1	3.4	4.23E-31
ENSG00000224887	AL513318.1		3.4	4.13E-16
ENSG00000166535	A2ML1	Alpha-2-Macroglobulin-Like 1	3.3	4.23E-31
ENSG00000206073	SERPINB4	Serpin Family B member 4	3.3	7.15E-15
ENSG00000181617	FDCSP	Follicular dendritic cell secreted protein	3.2	3.44E-21
ENSG00000229921	KIF25-AS1	KIF25 antisense RNA 1	3.2	2.85E-24
ENSG00000163735	CXCL5	C-X-C motif chemokine ligand 5	3.2	8.62E-34
ENSG00000047936	ROS1	ROS proto-oncogene 1	3.2	8.15E-28
ENSG00000134873	CLDN10	Claudin 10	3.2	2.36E-39
ENSG00000128714	HOXD13	Homebox D13	3.0	7.62E-27
ENSG00000163739	CXCL1	C-X-C motif chemokine ligand 1	3.0	1.66E-36
ENSG00000184697	CLDN6	Claudin 6	2.9	1.62E-25
ENSG00000176165	FOXG1	Forkhead box G1	2.7	2.22E-13

TABLE 3 | LINC00460 high vs low down-regulated genes.

	Gene symbol	Gene name	log2FC	padj-value
ENSG00000234398	AC134915.1		-1,5	0,00659454174
ENSG00000236532	LINC01695		-1,5	6,95E-09
ENSG00000137948	BRDT	Bromodomain Testis associated	-1,5	1,66E-07
ENSG00000238391	RNA5SP233	5S ribosomal pseudogene 233	-1,5	0,003359924242
ENSG00000275251		-1,5	0,004665490892	
ENSG00000249691	AC026117.1		-1,5	0,003145246539
ENSG00000130612	CYP2G1P	Cytochrome P450 family 2 subfamily G member 1, pseudogene	-1,5	7,07E-07
ENSG00000125363	AMELX	Amelogenin X-Linkes	-1,5	0,005986652631
ENSG00000135638	EMX1	Empty spiracles homebox 1	-1,5	2,26E-07
ENSG00000256120	SOX5-AS1	GRIK1 antisense-RNA1	-1,5	0,009555483978
ENSG00000255155	AP004371.1		-1,6	1,77E-06
ENSG00000225795	AC006463.1		-1,6	0,004581123508
ENSG00000256612	CYP2B7P	Cytochrome P450 family 2 subfamily B member 7	-1,6	2,64E-06
ENSG00000265595	MIR4756		-1,6	0,001510008292
ENSG00000186732	MPPED1	Metallophosphoesterase domain containing 1	-1,6	1,43E-08
ENSG00000248350	AC010265.1		-1,6	0,004980797457
ENSG00000139352	ASCL1	Achaete-scute family bHLH transcription factor 1	-1,6	3,46E-05
ENSG00000233420	AC002127.2		-1,6	0,001874852522
ENSG00000250223	LINC01216		-1,6	0,007490194855

LINC00460 Expression Is Significantly Correlated With the WNT Pathway and Cytokine Inflammatory Response Genes

LINC00460 expression level is correlated with at least 100 coding transcripts in BRCA (PCC>0.45, p<0.05; shown in **Table 4**). Some of these transcripts are classic cancer-related, such as *WNT7A, SFRP5, FOSL1, IFNK, CSF2, DUSP7* and *IL1A*. Overrepresentation Enrichment Analysis (ORA) showed that WNT signaling pathway, cytokine inflammatory response, fluoropyrimidine activity pathway and photodynamic therapy-induced NF-kB survival signaling are enriched in LINC00460 over-expressed BRCA samples (**Figure 7C**). In addition, we observed that 73% (61/84) of the BRCA Mexican cohort samples displayed high levels of stromal TILs (see **Table 1**).

TABLE 4 | LINC00460 correlated genes.

	Gene symbol	Gene name	PCC*
ENSG00000179046.8	TRIML2	Tipartite motif family like 2	0.61
ENSG00000175592.8	FOSL1	FOS like 1, AP-1 transcription factor subunit	0.6
ENSG00000115008.5	IL1A	Interleukin 1 alpha	0.57
ENSG00000163915.7	IGF2BP2-AS1		0.56
ENSG00000164400.5	CSF2	Colony stimulating factor 2	0.55
ENSG00000240476.1	LINC00973		0.55
ENSG00000147896.3	IFNK	Interferon kappa	0.54
ENSG00000154764.5	WNT7A	Wnt family member 7A	0.54
ENSG00000261780.2	CTD-2354A18.1		0.53
ENSG00000254403.1	OR10Y1P	Olfactory receptor family 10 subfamily Y member 1 pseudogene	0.53
ENSG00000158825.5	CDA	Cytidine deaminase	0.53
ENSG00000104327.7	CALB1	Calbindin 1	0.51
ENSG00000203783.4	PRR9	Proline rich 9	0.51
ENSG00000176797.3	DEFB103A	Defensin beta 103A	0.48
ENSG00000182591.5	KRTAP11-1	Keratin associated protein 11-1	0.48

**PCC, Pearson correlation coefficient.*

Analyzing the LINC00460 co-expressed genes list, we observed that similar genes had been previously identified as enriched in Lehman's TNBC subtypes (77) (**Supplementary Table S3**). Altogether, these enrichment data suggests that LINC00460 is potentially related with inflammatory pathways and might partially explain its effect in good prognosis prediction in BRCA.

LINC00460 Is an In-Silico Predicted miR-103-a-1 Sponge

LINC00460 has been described as a miRNA sponge lncRNA in different tumors (38, 44, 60, 75). With this evidence in mind, we aimed to further characterized its biological role in BRCA. We used an *in-silico* approach to predict LINC00460 interaction with novel miRNAs, using the miRcode tool (**Table 4**).

In the former section, we found that a group of mRNAs are significantly correlated with LINC00460 levels (see **Table 4**). We then reasoned that, if LINC00460 is acting as a miRNA sponge as previously reported, then some of these correlated mRNAs may be miRNAs targets as well. We used two *in-silico* tools to predict mRNA-miRNA binding, namely miRcode and mirPATHDB. As shown in **Table 5**, the 10 LINC00460 most correlated mRNAs have potential miRNA interactants. Furthermore, we observed that some of these mRNAs, such as WNT7A and KRTAP11-1, can potentially bind to the same miRNA: miR-103-a-1 (**Tables 5** and **6**). These findings suggest a role for LINC00460 as a miR-103-a-1 interactant and as a potential regulator of WNT7A and KRTAP11-1 expression.

The LINC00460: WNT7A Ratio Is a Composite Marker for Increased OS and DMFS in Basal-Like BRCA and Can Predict Anthracycline Response in ER- BRCA Patients

To further demonstrate the role of LINC00460 in potential regulation of WNT7A and its combined role in BRCA prognosis, we computed the LINC00460:WNT7A ratio to construct an OS and

TABLE 5 | Matching sites and predicted miRNAs interacting with LINC00460.

microRNA family	Seed position	Seed type
miR-503	chr13:107029280	7-mer-m8
miR-143/1721/4770	chr13:107028970	7-mer-m8
miR-150/5127	chr13:107029727	7-mer-m8
miR-1ab/206/613	chr13:107030346	7-mer-A1
miR-200bc/429/548a	chr13:107029706	7-mer-A1
miR-221/222/222ab/1928	chr13:107029417	7-mer-A1
miR-23abc/23b-3p	chr13:107029820	7-mer-A1
miR-24/24ab/24-3p	chr13:107029217	7-mer-A1
miR-24/24ab/24-3p	chr13:107029855	7-mer-A1
miR-24/24ab/24-3p	chr13:107029953	8-mer
miR-103a/107/107ab	chr13:107030479	8-mer
miR-338/338-3p	chr13:107029282	7-mer-A1
miR-338/338-3p	chr13:107029426	7-mer-A1
miR-425/425-5p/489	chr13:107029977	7-mer-A1
miR-129-5p/129ab-5p	chr13:107030286	7-mer-A1

TABLE 6 | Predicted miRNAs that are potentially interacting with LINC00460 and some coexpressed mRNAs.

miRNA	mRNA	PCC (mRNA vs LINC00460)
miR-133a-3p	TRIML2	0.61
miR-130a	FOSL1	0.6
miR-544b	CSF2	0.55
miR-216b	IFNK	0.54
miR-103a-1	WNT7A	0.54
miR-34a	CDA	0.53
miR-140-5p	CALB1	0.51
miR-124-1	DEFB103A	0.48
miR-103a-1	KRTAP11-1	0.48
miR-455-5p	FERMT1	0.47

DMFS models for Basal-like BRCA. As shown in **Figure 8**, the LINC00460:WNT7A ratio is able to predict an increased OS (n=153, logrank p=0.028) and DMFS (n=145, logrank p=0.0057) in Basal-like BRCA, using GEO cohorts. In contrast, analysis with LINC00460:KRTAP11-1 ratio did not retrieve any significant survival effect (data not shown). Survival analysis using the mature sequence of miR-103-a-1 (hsa-miR-103a), showed a marginal, non-significant association between overexpression of the of miR-103-a-1 and decreased survival in TCGA TNBC cohort (n=97, logrank p=0.059; **Supplementary Figure S7**).

Taken together, these data show that the expression ratio of two genes, LINC00460:WNT7A is a composite marker that accurately predicts Basal-like BRCA OS and DMFS. Furthermore, we identified that the combination of LINC00460 and WNT7A over-expression is significantly associated with pathological complete response (pCR) after anthracycline therapy in ER- BRCA patients (n= 665, Mann-Whitney U test p= 0.0047) (**Figure 8C**). These evidences clearly indicate that both transcripts exert a central and beneficial role to basal-like, ER- BRCA patients.

DISCUSSION

Long non-coding RNAs exert numerous roles in human cancers, as their biological activities involve regulation of cell proliferation, cell death, differentiation, migration, invasion and metastasis. Deregulation in lncRNAs expression has also been associated with clinical outcome. LncRNAs can affect expression of thousands of genes, so they are regarded as key master regulators (78).

In this work, our aim was to investigate if LINC00460 expression was deregulated in different tumors, and if it was associated with clinical and pathological characteristics in these tumors. We then focused on its clinical role in aggressive (basal-like) breast cancer and the identification of potential LINC00460 targets in this model. We aimed to know if LINC00460 can target miRNAS *in-silico*, which can in turn bind to relevant mRNAs. Finally, we sought to investigate if some of the potential candidate genes would have a combinatorial role for OS and therapy response in basal-like BRCA.

We found that LINC00460 was deregulated in seven different tumor types in the TCGA database, including two not previously reported tumors, namely LGG and GBM. We have also confirmed LINC00460 deregulation in BRCA using two independent cohorts. Interestingly, LINC00460 expression is associated with clinically aggressive tumors, such advanced stage LUAD, LUSC and KIRC, high histology grade in KIRC, HPV-negative HNSC and basal-like BRCA, suggesting an important role of LINC00460 in the progression or intrinsically aggressiveness of these tumors. Indeed, it has been previously shown that LINC00460 expression can promote cancer progression (34, 35, 38, 50), metastasis (36) and influences therapy response (79).

Several reports show that LINC00460 is a marker for poor OS prognosis across different tumors, such as CESC (48), HNSC (80), KIRC (81), LUAD (49) and PAAD (82). In addition, we describe here that LINC00460 high expression is significantly associated with poor survival in three different tumors (GBM, LGG and SARC) but related with a favorable survival rate in BRCA, *i.e.*, its association to clinical outcome varies between tumors. This data suggests that the expression levels and its impact on OS, RFS or DMFS may be tissue-specific. Indeed, it has previously been shown that the same lncRNA may exert dual prognostic roles in distinct tumors. For example, high MALAT-1 expression has been reported as a marker for poor prognosis in various tumors, including COAD, NSCLC, STAD, PAAD, ESCA, among others (83), but also as a good prognosis factor for BRCA, acting as a metastasis suppressor (84). Another interesting example of this dual phenomenon is the expression of XIST. It has been shown that high expression of this lncRNA is related to poor clinical outcome in different cancers (85), but in another study, authors demonstrate that high XIST expression is related to an increased brain metastasis–free survival in BRCA patients (6). We suggest that LINC00460 can perform as a dual tissue-specific prognostic marker, similar to MALAT-1 and XIST; although we cannot discard alternative mechanisms, such as differences in the splicing variants measured between studies. These variant transcripts expression patterns should be taken into account to evaluate lncRNA-based predictive biomarkers (86). This latter possibility must be addressed in future studies.

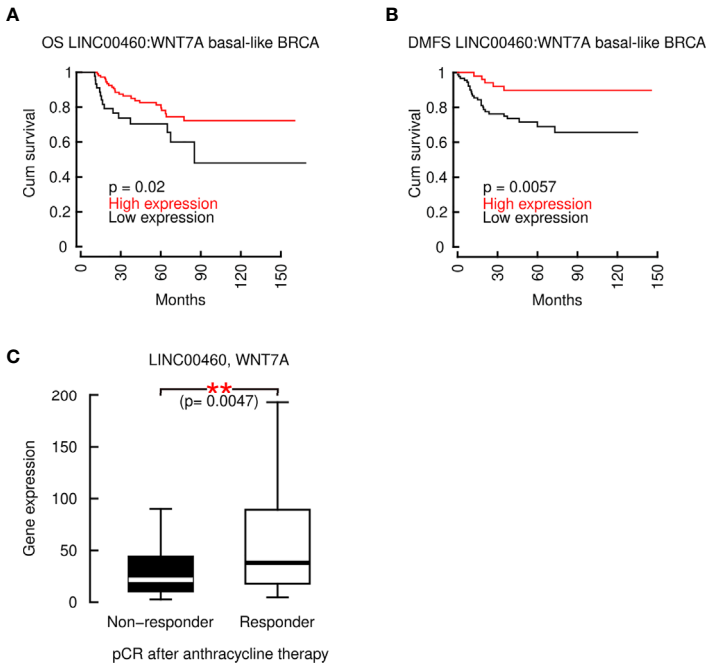

FIGURE 8 | The composite marker LINC00460:WNT7A is predicts an increased OS and RFS in basal-like BRCA and its median is associated with a complete pathological response to chemotherapy in ER- BRCA patients. **(A)** LINC00460:WNT7A ratio predicts a significantly improved OS and **(B)** RFS in basal-like BRCA; **(C)** LINC00460 and WNT7A combined expression is significantly associated with pCR after anthracycline-based therapy in ER- patients. **p<0.005.

In the BRCA model, LINC00460 expression is associated with the phenotypic makeup of the tumor, where overexpression of LINC00460 is associated with a negative result to hormone receptors (ER and PR) when subject to IHC. It was also observed that in the most aggressive phenotype, basal-like BRCA, it actually favored the clinical outcome (**Figure 6**), even when being subject to analysis in the different TNBC subtypes. Its similar behavior through different BRCA subtypes and BRCA different cohorts, led us to propose LINC00460 as a potential biomarker for improved OS, RFS and conceivably DMFS prediction in basal-like BRCA (**Figures 4–6**). This behavior can potentially be explained through the candidate co-expressed genes found in the differential expression analyses and ORA (**Figure 7**). We propose five main mechanisms that potentially explain the role of LINC00460 in the increased prognosis of BRCA patients: (1) regulation/co-expression of good prognosis-related genes, (2) co-expression/modulation of genes that promote an immunogenic niche, (3) decrease of tumor cell proliferation, (4) regulation of WNT7A through sponging of miR-103-a-1 and (5) promoting chemotherapy (anthracycline) complete pathologic response. As the majority of these interactions are *in-silico* predicted, or associated with clinical features in patient-based cohorts, experimental validation is needed to support these hypotheses.

In this regard, we observed that the LINC00460 expression is significantly enriched in the BL2 TNBC subtype. This finding is of particular interest, since the BL2 subtype displays a variety of gene ontologies enriched in components and pathways involved in cell proliferation, growth, survival and cell differentiation pathways, such as the WNT pathway (77). In accordance, we identified important members of the WNT pathway, such as WNT7A, as being potentially regulated by LINC00460 (see **Figure 7C** and the discussion below). These data strengthen the role of LINC00460 in BL2 TNBC, and further clarify its mechanism of action in these tumors.

We detected that some of the LINC00460 co-expressed genes, such as *SFRP5, HOXD13*, are also related to increased survival rate in different tumors. High expression of *SFRP5* is significantly associated with a better prognosis in PAAD (87) and BRCA (88). HOXD13 protein levels are related to an increased OS in BRCA (89). Thus, this LINC00460 candidate target genes potentially contribute to the increased survival effect observed in BRCA patients.

On the other hand, we found that several LINC00460 co-expressed genes (*TRIML2, SFRP5, FOSL1, IFNK, CSF2, DUSP7, DEFB103A* and *IL1AA)* are immunogenic-related. Immune response pathways are clinically relevant, as it has been previously described that a highly immunogenic niche in a tumor may improve the outcome of the disease (90). These genes and pathways are frequently enriched in TNBC (77, 91). Furthermore, TNBC also display enrichment of tumor-infiltrating lymphocytes (TILs) (92). We suggest that LINC00460 is related with good prognosis in TNBC/basal-like due to its potential relation with these genes and

with the presence of tumor-infiltrating lymphocytes (TILs), as shown in our samples (see **Table 1**). It has been previously shown that the presence of TILs improves prognosis as it modulates cancer progression and enhances chemotherapy response in TNBC, conferring a protective immunity in these patients (93).

Therefore, we suggest that the correlation between these immunogenic genes and LINC00460 can partially explain the clinical behavior in the breast cancer cohorts, as the overexpression of LINC00460 is associated with upregulation of immunogenic factors that in turn, permit the migration of components of the immune system. These mechanisms can promote an immunogenic tumor environment and thus favor tumor cell death. Future studies are needed to experimentally validate these data, as it is relevant for clinical outcome in aggressive TNBC.

In addition to promoting the immunogenic niche, LINC00460 could decrease proliferation of the tumor cells, as we observed that the proliferation pathway is negatively enriched in the high LINC00460 group in BRCA samples. This could also explain the increased survival rates of the BRCA patients. Indeed, multiple lines of evidence suggest that LINC00460 can modulate cell proliferation, cell death, migration and invasion and EMT, through its sponging activity and targeting various key transcripts in several cancer types (31, 37, 43, 94–96).

We predicted that LINC00460 potentially binds to miR-103-a-1 *in-silico*, which, in turn, can target WNT7A. In this regard, although there have been reports that suggest that WNT7A is an oncoprotein (97), it has also been shown that loss of WNT7A expression is significantly associated with poor RFS in BRCA (98) and it is also involved in tumor cell differentiation (99). Thus, the exact role of WNT7A in BRCA is currently unclear. Our results might suggest that WNT7A potentially play an important clinical role in the BRCA, as LINC00460 could sponge miR-103-a-1 and henceforth, liberate WNT7A. Further research must validate these predictions experimentally.

We observed that the LINC00460:WNT7A ratio is a composite marker that can predict a favorable OS and DMFS in TNBC. These results highlight the clinical and biological role of LINC00460 and WNT7A transcripts in TNBC and constitute valuable data, as simple ratios of gene expression levels can be used to accurately diagnose (100) and predict cancer outcomes (101, 102), while circumventing many of the limitations that preclude the use of microarray techniques in extensive clinical applications (103, 104). In accordance with our observations regarding immunogenic factor potential upregulation and previous findings in the field (92, 93, 105), we have also identified a significant association between the expression enrichment of LINC00460 and WNT7A with anthracycline responsive ER- patients. This finding further strengthens the beneficial role of both transcripts in patient's prediction and prognosis as, it has been demonstrated that ER-negative breast cancers with high levels of TILs have heightened sensitivity to anthracycline-based chemotherapy (106), and that TILs are an independent predictor of good response to anthracycline/taxane neoadjuvant chemotherapy (105). All these observations, however, are limited by the size of GEO patient cohorts and will require validation in a larger independent cohort.

Regarding the role of miR-103-a-1, we identified a marginal association with poor OS in BRCA. This is in accordance with previous findings, as it has been shown that miR-103-a-1 acts as an oncogene to promote TNBC cells migration and invasion (107). In another report, authors show that serum miR-103 over-expression was significantly correlated with worse clinical factors, as well as poorer recurrence-free survival or overall survival in colorectal cancer (108). MiR-103/107 expression displays stemness-promoting functions, and a signature of miR-103/107 high and Axin2 low expression profile correlates with poor prognosis in colorectal cancer patients (109). In gastric cancer patients, high expression of miR-103 was significantly associated with poor overall survival and disease-free survival and is a key factor that contributes to tumor progression (110). Altogether, these data suggest that of miR-103-a-1 is a marker for poor prognosis in several tumors, including BRCA. To the best of our knowledge, this is the first report that suggests a potential connection between LINC00460, WNT7A and miR-103-a-1. Future research must elucidate the exact mechanisms involved in this potential 3-gene network, and its impact in basal-like BRCA biology.

In conclusion, LINC00460 expression is a dual potential marker for aggressive phenotypes and poor clinical outcome in distinct tumors, including HNSC, KIRC LUSC and LUAD, that is also associated with increased prognosis in basal-like BRCA. LINC00460 is enriched in BL2 TNBC, and potentially regulates the WNT differentiation pathway. LINC00460 can also modulate a plethora of immunogenic related genes in BRCA, such as *SFRP5, FOSL1, IFNK, CSF2, DUSP7* and *IL1A* and interacts with miR-103-a-1, *in-silico*, which, in turn, can no longer target WNT7A. LINC00460:WNT7A ratio is a composite marker that can predict a favorable OS and DMFS in TNBC, and combination of LINC00460 and WNT7A over-expression is associated with complete pathological response (pCR) after anthracycline therapy in ER- BRCA patients. This data confirms that LINC00460 is a master regulator in BRCA molecular circuits and influences clinical outcome.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research and Ethics Committee of National Institute of Genomic Medicine and the Institute of Breast

Diseases, FUCAM (Registration number: CE2009/11). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AH-M, MR-R, LH-P and MC-V conceived and designed the study. MM-R and LH-P handled the samples, constructed the clinical database and tested the expression levels in biological samples. SJ-M, RA-L, LA-R, and LH-P handled the samples and clinical database. AT-T, CD-R, FV-C, IR-S, and AM-A obtained the samples and performed the clinical evaluation and following of the patients. LH-P did the pathological evaluation of the samples. AC-T, MC-V, MP-L, RA-L, and EH-C performed all the *in-silico* validation cohort-related analysis and the *in-silico* prediction of the sponge regulation. MR-R, MC-V, and AC-T performed statistical analysis, survival analysis, differential expression analysis, and risk analysis. All authors interpreted and discussed the data. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Mexican National Council of Science and Technology 480751/282036 grant (Scholarship), and Mexican National Council of Science and Technology Basic Science grant (CONACYT grant number 258936).

ACKNOWLEDGMENTS

We want to thank all the patients that accepted to participate in our research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.628027/full#supplementary-material

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
2. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* (2000) 406 (6797):747–52. doi: 10.1038/35021093
3. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, Yoshizawa C, et al. A Multigene Expression Assay to Predict Local Recurrence Risk for Ductal Carcinoma In Situ of the Breast. *JNCI J Natl Cancer Institute* (2013) 105 (10):701–10. doi: 10.1093/jnci/djt067
4. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, et al. Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum Mol Genet* (2014) 23(22):5866–78. doi: 10.1093/hmg/ddu309
5. Wu B-Q, Jiang Y, Zhu F, Sun D-L, He X-Z. Long Noncoding RNA PVT1 Promotes EMT and Cell Proliferation and Migration Through Downregulating p21 in Pancreatic Cancer Cells. *Technol Cancer Res Treat* (2017) 16(6):819–27. doi: 10.1177/1533034617700559
6. Xing F, Liu Y, Wu S-Y, Wu K, Sharma S, Mo Y-Y, et al. Loss of XIST in Breast Cancer Activates MSN-c-Met and Reprograms Microglia via Exosomal miRNA to Promote Brain Metastasis. *Cancer Res* (2018) 78 (15):4316–30. doi: 10.1158/0008-5472.CAN-18-1102
7. Dimitrova N, Zamudio JR, Jong RM, Soukup D, Resnick R, Sarma K, et al. LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint. *Mol Cell* (2014) 54 (5):777–90. doi: 10.1016/j.molcel.2014.04.025
8. Sanchez Calle A, Kawamura Y, Yamamoto Y, Takeshita F, Ochiya T. Emerging roles of long non-coding RNA in cancer. *Cancer Sci* (2018) 109 (7):2093–100. doi: 10.1111/cas.13642
9. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* (2007) 129(7):1311–23. doi: 10.1016/j.cell.2007.05.022
10. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, et al. Kcnq1ot1 Antisense Noncoding RNA Mediates Lineage-Specific Transcriptional Silencing through Chromatin-Level Regulation. *Mol Cell* (2008) 32(2):232–46. doi: 10.1016/j.molcel.2008.08.022
11. Jiang K, Patel NA, Watson JE, Apostolatos H, Kleiman E, Hanson O, et al. Akt2 Regulation of Cdc2-Like Kinases (Clk/Sty), Serine/Arginine-Rich (SR) Protein Phosphorylation, and Insulin-Induced Alternative Splicing of PKCβII Messenger Ribonucleic Acid. *Endocrinology* (2009) 150(5):2087–97. doi: 10.1210/en.2008-0818
12. Kong J, Sun W, Li C, Wan L, Wang S, Wu Y, et al. Long non-coding RNA LINC01133 inhibits epithelial–mesenchymal transition and metastasis in colorectal cancer by interacting with SRSF6. *Cancer Lett* (2016) 380(2):476–84. doi: 10.1016/j.canlet.2016.07.015
13. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* (2011) 470(7333):284–8. doi: 10.1038/nature09701
14. Zhao Y, Liu Y, Lin L, Huang Q, He W, Zhang S, et al. The lncRNA MACC1-AS1 promotes gastric cancer cell metabolic plasticity via AMPK/Lin28 mediated mRNA stability of MACC1. *Mol Cancer* (2018) 17(1):69. doi: 10.1186/s12943-018-0820-2. Available from.
15. Garabedian MJ, Logan SK. Glucocorticoid Receptor DNA Binding Decoy Is a Gas. *Sci Signaling* (2010) 3(108):pe5–5. doi: 10.1126/scisignal.3108pe5
16. Zhang X, Zhou Y, Chen S, Li W, Chen W, Gu W. LncRNA MACC1-AS1 sponges multiple miRNAs and RNA-binding protein PTBP1. *Oncogenesis* (2019) 8(12):73. doi: 10.1038/s41389-019-0182-7
17. Tu J, Zhao Z, Xu M, Chen M, Weng Q, Ji J. LINC00460 promotes hepatocellular carcinoma development through sponging miR-485-5p to up-regulate PAK1. *Biomed Pharmacother* (2019) 118:109213. doi: 10.1016/j.biopha.2019.109213
18. Sun B, Liu C, Li H, Zhang L, Luo G, Liang S, et al. Research progress on the interactions between long non−coding RNAs and microRNAs in human cancer (Review). *Oncol Lett* (2019) 19(1):595–605. doi: 10.3892/ol.2019.11182
19. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA Hypothesis: The Rosetta Stone of a Hidden RNA Language? *Cell* (2011) 146(3):353–8. doi: 10.1016/j.cell.2011.07.014
20. Wu X, Wang F, Li H, Hu Y, Jiang L, Zhang F, et al. LncRNA- PAGBC acts as a microRNA sponge and promotes gallbladder tumorigenesis. *EMBO Rep* (2017) 18(10):1837–53. doi: 10.15252/embr.201744147
21. Shan Y, Ma J, Pan Y, Hu J, Liu B, Jia L. LncRNA SNHG7 sponges miR-216b to promote proliferation and liver metastasis of colorectal cancer through upregulating GALNT1. *Cell Death Dis* (2018) 9(7):722. doi: 10.1038/s41419-018-0759-7
22. Cao F, Wang Z, Feng Y, Zhu H, Yang M, Zhang S, et al. lncRNA TPTEP1 competitively sponges miR−328−5p to inhibit the proliferation of non

−small cell lung cancer cells. *Oncol Rep* (2020) 43:1606–18. doi: 10.3892/or.2020.7522
23. Zhou X, Ye F, Yin C, Zhuang Y, Yue G, Zhang G. The Interaction Between MiR-141 and lncRNA-H19 in Regulating Cell Proliferation and Migration in Gastric Cancer. *Cell Physiol Biochem* (2015) 36(4):1440–52. doi: 10.1159/000430309
24. Xu F, Zhang J. Long non-coding RNA HOTAIR functions as miRNA sponge to promote the epithelial to mesenchymal transition in esophageal cancer. *Biomed Pharmacother* (2017) 90:888–96. doi: 10.1016/j.biopha.2017.03.103
25. Jin C, Yan b, Lu Q, Lin Y, Ma L. Reciprocal regulation of Hsa-miR-1 and long noncoding RNA MALAT1 promotes triple-negative breast cancer development. *Tumor Biol* (2016) 37(6):7383–94. doi: 10.1007/s13277-015-4605-6
26. Cui Y, Fan Y, Zhao G, Zhang Q, Bao Y, Cui Y, et al. Novel lncRNA PSMG3−AS1 functions as a miR−143−3p sponge to increase the proliferation and migration of breast cancer cells. *Oncol Rep* (2019) 43:229–39. doi: 10.3892/or.2019.7390
27. Li W, Zhang Z, Liu X, Cheng X, Zhang Y, Han X, et al. The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. *J Clin Invest* (2017) 127(9):3421–40. doi: 10.1172/JCI94233
28. Li Z, Hou P, Fan D, Dong M, Ma M, Li H, et al. The degradation of EZH2 mediated by lncRNA ANCR attenuated the invasion and metastasis of breast cancer. *Cell Death Differ* (2017) 24(1):59–71. doi: 10.1038/cdd.2016.95
29. GenBank. *Homo sapiens long intergenic non-protein coding RNA 460 (LINC00460), transcript variant 1, long non-coding RNA*. Available at: https://www.ncbi.nlm.nih.gov/nuccore/1636409762/, [cited 2020 Nov 10].
30. Cao W, Liu J, Liu Z, Wang X, Han Z-G, Ji T, et al. A three-lncRNA signature derived from the Atlas of ncRNA in cancer (TANRIC) database predicts the survival of patients with head and neck squamous cell carcinoma. *Oral Oncol* (2017) 65:94–101. doi: 10.1016/j.oraloncology.2016.12.017
31. Wen L, Zhang X, Bian J, Han L, Huang H, He M, et al. The long non−coding RNA LINC00460 predicts the prognosis and promotes the proliferation and migration of cells in bladder urothelial carcinoma. *Oncol Lett* (2019) 17:3874–80. doi: 10.3892/ol.2019.10023
32. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. *Nucleic Acids Res* (2020) 48(D1):D682–8. doi: 10.1093/nar/gkz966
33. Zhu Y, Yang L, Chong Q-Y, Yan H, Zhang W, Qian W, et al. Long noncoding RNA Linc00460 promotes breast cancer progression by regulating the miR-489-5p/FGF7/AKT axis. *CMAR* (2019) 11:5983–6001. doi: 10.2147/CMAR.S207084
34. Xie X, Xiong G, Wang Q, Ge Y, Cui X. Long non-coding RNA LINC00460 promotes head and neck squamous cell carcinoma cell progression by sponging miR-612 to up-regulate AKT2. *Am J Transl Res* (2019) 11(10):6326–40.
35. Li G, Kong Q. LncRNA LINC00460 promotes the papillary thyroid cancer progression by regulating the LINC00460/miR-485-5p/Raf1 axis. *Biol Res* (2019) 5261(1):1–12. doi: 10.1186/s40659-019-0269-9
36. Zhang Y, Liu X, Li Q, Zhang Y. lncRNA LINC00460 promoted colorectal cancer cells metastasis via miR-939-5p sponging. *CMAR* (2019) 11:1779–89. doi: 10.2147/CMAR.S192452
37. Wang F, Liang S, Liu X, Han L, Wang J, Du Q. LINC00460 modulates KDM2A to promote cell proliferation and migration by targeting miR-342-3p in gastric cancer. *OTT* (2018) 11:6383–94. doi: 10.2147/OTT.S169307
38. Lian H, Xie P, Yin N, Zhang J, Zhang X, Li J, et al. Linc00460 promotes osteosarcoma progression via miR-1224-5p/FADS1 axis. *Life Sci* (2019) 233:116757. doi: 10.1016/j.lfs.2019.116757
39. Hu X, Liu W, Jiang X, Wang B, Li L, Wang J, et al. Long noncoding RNA LINC00460 aggravates invasion and metastasis by targeting miR-30a-3p/Rap1A in nasopharyngeal carcinoma. *Hum Cell* (2019) 32(4):465–76. doi: 10.1007/s13577-019-00262-4
40. Hong W, Ying H, Lin F, Ding R, Wang W, Zhang M. lncRNA LINC00460 Silencing Represses EMT in Colon Cancer through Downregulation of ANXA2 via Upregulating miR-433-3p. *Mol Ther Nucleic Acids* (2020) 19:1209–18. doi: 10.1016/j.omtn.2019.12.006
41. Wang H-X, Kang L-J, Qin X, Xu J, Fei J-W. LINC00460 promotes proliferation and inhibits apoptosis of non-small cell lung cancer cells through targeted regulation of miR-539. *Eur Rev Med Pharmacol Sci* (2020) 24(12):6752–8. doi: 10.26355/eurrev_202006_21663
42. Cui Y, Zhang C, Lian H, Xie L, Xue J, Yin N, et al. LncRNA linc00460 sponges miR-1224-5p to promote esophageal cancer metastatic potential and epithelial-mesenchymal transition. *Pathol Res Pract* (2020) 216 (7):153026. doi: 10.1016/j.prp.2020.153026
43. Zou X, Guo ZH, Li Q, Wang PS. Long Noncoding RNA LINC00460 Modulates MMP-9 to Promote Cell Proliferation, Invasion and Apoptosis by Targeting miR-539 in Papillary Thyroid Cancer. *Cancer Manag Res* (2020) 12:199–207. doi: 10.2147/CMAR.S222085
44. Jiang Y, Cao W, Wu K, Qin X, Wang X, Li Y, et al. LncRNA LINC00460 promotes EMT in head and neck squamous cell carcinoma by facilitating peroxiredoxin-1 into the nucleus. *J Exp Clin Cancer Res* (2019) 38(1):365. doi: 10.1186/s13046-019-1364-z
45. Lian Y, Yan C, Xu H, Yang J, Yu Y, Zhou J, et al. A Novel lncRNA, LINC00460, Affects Cell Proliferation and Apoptosis by Regulating KLF2 and CUL4A Expression in Colorectal Cancer. *Mol Ther Nucleic Acids* (2018) 12:684–97. doi: 10.1016/j.omtn.2018.06.012
46. Wang X, Mo F-M, Bo H, Xiao L, Chen G-Y, Zeng P-W, et al. Upregulated Expression of Long Non-Coding RNA, LINC00460, Suppresses Proliferation of Colorectal Cancer. *J Cancer* (2018) 9(16):2834–43. doi: 10.7150/jca.26046
47. Yang J, Li K, Chen J, Hu X, Wang H, Zhu X. Long Noncoding RNA LINC00460 Promotes Hepatocellular Carcinoma Progression via Regulation of miR-342-3p/AGR2 Axis. *OTT* (2020) 13:1979–91. doi: 10.2147/OTT.S239258
48. Li F, Zhu W, Wang Z. Long noncoding RNA LINC00460 promotes the progression of cervical cancer via regulation of the miR-361-3p/Gli1 axis. *Hum Cell* (2020) 34229–37. doi: 10.1007/s13577-020-00447-2
49. Nakano Y, Isobe K, Kobayashi H, Kaburaki K, Isshiki T, Sakamoto S, et al. Clinical importance of long non−coding RNA LINC00460 expression in EGFR−mutant lung adenocarcinoma. *Int J Oncol* (2019) 56:243–57. doi: 10.3892/ijo.2019.4919
50. Liu X, Wen J, Wang H, Wang Y. Long non-coding RNA LINC00460 promotes epithelial ovarian cancer progression by regulating microRNA-338-3p. *Biomed Pharmacother* (2018) 108:1022–8. doi: 10.1016/j.biopha.2018.09.103
51. Liang Y, Wu Y, Chen X, Zhang S, Wang K, Guan X, et al. A novel long noncoding RNA linc00460 up-regulated by CBP/P300 promotes carcinogenesis in esophageal squamous cell carcinoma. *Biosci Rep* (2017) 37(5):BSR20171019. doi: 10.1042/BSR20171019
52. Xing H, Wang S, Li Q, Ma Y, Sun P. Long noncoding RNA LINC00460 targets miR-539/MMP-9 to promote meningioma progression and metastasis. *Biomed Pharmacother* (2018) 105:677–82. doi: 10.1016/j.biopha.2018.06.005
53. Feng L, Rao M, Zhou Y, Zhang Y, Zhu Y. Long noncoding RNA 00460 (LINC00460) promotes glioma progression by negatively regulating miR-320a. *J Cell Biochem* (2019) 120(6):9556–63. doi: 10.1002/jcb.28232
54. Huang G-W, Xue Y-J, Wu Z-Y, Xu X-E, Wu J-Y, Cao H-H, et al. A three-lncRNA signature predicts overall survival and disease-free survival in patients with esophageal squamous cell carcinoma. *BMC Cancer* (2018) 18 (1):147. doi: 10.1186/s12885-018-4058-6
55. Xu Z, Wang C, Xiang X, Li J, Huang J. Characterization of mRNA Expression and Endogenous RNA Profiles in Bladder Cancer Based on The Cancer Genome Atlas (TCGA) Database. *Med Sci Monit* (2019) 25:3041–60. doi: 10.12659/MSM.915487
56. Wang Z-L, Li B, Piccolo SR, Zhang X-Q, Li J-H, Zhou H, et al. Integrative analysis reveals clinical phenotypes and oncogenic potentials of long non-coding RNAs across 15 cancer types. *Oncotarget* (2016) 7(23):35044–55. doi: 10.18632/oncotarget.9037
57. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* (2019) 47(W1):W556–60. doi: 10.1093/nar/gkz430
58. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, et al. TANRIC: An Interactive Open Platform to Explore the Function of lncRNAs in Cancer. *Cancer Res* (2015) 75(18):3728–37. doi: 10.1158/0008-5472.CAN-15-0273
59. Zheng Y, Xu Q, Liu M, Hu H, Xie Y, Zuo Z, et al. lnCAR: A Comprehensive Resource for lncRNAs from Cancer Arrays. *Cancer Res* (2019) 79(8):2076–83. doi: 10.1158/0008-5472.CAN-18-2169
60. Backes C, Kehl T, Stöckel D, Fehlmann T, Schneider L, Meese E, et al. miRPathDB: a new dictionary on microRNAs and target pathways. *Nucleic Acids Res* (2017) 45(D1):D90–6. doi: 10.1093/nar/gkw926
61. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet* (2006) 38(5):500–1. doi: 10.1038/ng0506-500

62. Cedro-Tanda A, Ríos-Romero M, Romero-Córdoba S, Cisneros-Villanueva M, Rebollar-Vega RG, Alfaro-Ruiz LA, et al. A lncRNA landscape in breast cancer reveals a potential role for AC009283.1 in proliferation and apoptosis in HER2-enriched subtype. *Sci Rep* (2020) 10(1):13146. doi: 10.1038/s41598-020-69905-z
63. Irizarry RA. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* (2003) 4(2):249–64. doi: 10.1093/biostatistics/4.2.249
64. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* (2005) 33(Web Server issue):W741–748. doi: 10.1093/nar/gki475
65. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* (2010) 123(3):725–31. doi: 10.1007/s10549-009-0674-9
66. Desmedt C, Di Leo A, de Azambuja E, Larsimont D, Haibe-Kains B, Selleslags J, et al. Multifactorial approach to predicting resistance to anthracyclines. *J Clin Oncol* (2011) Apr 2029(12):1578–86. doi: 10.1200/JCO.2010.31.2231
67. Finetti P, Cervera N, Charafe-Jauffret E, Chabannon C, Charpin C, Chaffanet M, et al. Sixteen–Kinase Gene Expression Identifies Luminal Breast Cancers with Poor Prognosis. *Cancer Res* (2008) 68(3):767–76. doi: 10.1158/0008-5472.CAN-07-5516
68. Sabatier R, Finetti P, Cervera N, Lambaudie E, Esterni B, Mamessier E, et al. A gene expression signature identifies two prognostic subgroups of basal breast cancer. *Breast Cancer Res Treat* (2011) 126(2):407–20. doi: 10.1007/s10549-010-0897-9
69. Lánczky A, Nagy Á, Bottai G, Munkácsy G, Szabó A, Santarpia L, et al. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat* (2016) Dec160(3):439–46. doi: 10.1007/s10549-016-4013-7
70. Fekete JT, Győrffy B. ROCplot.org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. *Int J Cancer* (2019) 145(11):3140–51. doi: 10.1002/ijc.32369
71. Chen X, Li J, Gray WH, Lehmann BD, Bauer JA, Shyr Y, et al. TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer Inform* (2012) 11: CIN.S9983. doi: 10.4137/CIN.S9983
72. Lehmann BD, Jovanović B, Chen X, Estrada MV, Johnson KN, Shyr Y, et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. Sapino A, editor. *PloS One* (2016) 11(6):e0157368. doi: 10.1371/journal.pone.0157368
73. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* (2019) 569(7757):503–8. doi: 10.1093/bioinformatics/bts344
74. Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* (2012) 28 (15):2062–3. doi: 10.1093/bioinformatics/bts344
75. Li J, Ma W, Zeng P, Wang J, Geng B, Yang J, et al. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief Bioinform* (2015) Sep16(5):806–12. doi: 10.1093/bib/bbu048
76. Marko NF, Weil RJ. Non-Gaussian Distributions Affect Identification of Expression Patterns, Functional Annotation, and Prospective Classification in Human Cancer Genomes. Coleman WB, editor. *PloS One* (2012) 7(10): e46935. doi: 10.1371/journal.pone.0046935
77. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* (2011) 121 (7):2750–67. doi: 10.1172/JCI45014
78. Chiu H-S, Somvanshi S, Patel E, Chen T-W, Singh VP, Zorman B, et al. Pan-Cancer Analysis of lncRNA Regulation Supports Their Targeting of Cancer Genes in Each Tumor Context. *Cell Rep* (2018) 23(1):297–312.e12. doi: 10.26355/eurrev_201907_18420
79. Ma G, Zhu J, Liu F, Yang Y. Long Noncoding RNA LINC00460 Promotes the Gefitinib Resistance of Nonsmall Cell Lung Cancer Through Epidermal Growth Factor Receptor by Sponging miR-769-5p. *DNA Cell Biol* (2019) 38 (2):176–83. doi: 10.1089/dna.2018.4462
80. Zhang C, Cao W, Wang J, Liu J, Liu J, Wu H, et al. A prognostic long non-coding RNA-associated competing endogenous RNA network in head and neck squamous cell carcinoma. *PeerJ* (2020) 8:e9701. doi: 10.7717/peerj.9701
81. Zhang D, Zeng S, Hu X. Identification of a three-long noncoding RNA prognostic model involved competitive endogenous RNA in kidney renal clear cell carcinoma. *Cancer Cell Int* (2020) 20(1):319. doi: 10.1186/s12935-020-01423-4
82. Sun J, Yang J, Lv K, Guan J. Long non–coding RNA LINC00460 predicts poor survival and promotes cell viability in pancreatic cancer. *Oncol Lett* (2020) 20(2):1369–75. doi: 10.3892/ol.2020.11652
83. Zhu L, Liu J, Ma S, Zhang S. Long Noncoding RNA MALAT-1 Can Predict Metastasis and a Poor Prognosis: a Meta-Analysis. *Pathol Oncol Res* (2015) 21(4):1259–64. doi: 10.1007/s12253-015-9960-5
84. Kim J, Piao H-L, Kim B-J, Yao F, Han Z, Wang Y, et al. Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat Genet* (2018) 50 (12):1705–15. doi: 10.1038/s41588-018-0252-3
85. Zhou Q, Hu W, Zhu W, Zhang F, Lin-lin L, Liu C, et al. Long non coding RNA XIST as a prognostic cancer marker – A meta-analysis. *Clin Chim Acta* (2018) 482:1–7. doi: 10.1016/j.cca.2018.03.016
86. Meseure D, Vacher S, Lallemand F, Alsibai KD, Hatem R, Chemlali W, et al. Prognostic value of a newly identified MALAT1 alternatively spliced transcript in breast cancer. *Br J Cancer* (2016) 114(12):1395–404. doi: 10.1038/bjc.2016.123
87. Zhou W, Tian, Ming, Hu J, Li L, He Y. SFRP5 as a prognostic biomarker for patients with pancreatic ductal adenocarcinoma. *Int J Clin Exp Pathol* (2018) 9(3):3442–7. doi: 10.1016/j.cca.2018.03.016
88. Wu Z-H, Zhang Y, Yue J-X, Zhou T. Comprehensive Analysis of the Expression and Prognosis for *SFRPs* in Breast Carcinoma. *Cell Transplant* (2020) 29:096368972096247. doi: 10.1177/0963689720962479
89. Zhong Z-B, Shan M, Qian C, Liu T, Shi Q-Y, Wang J, et al. Prognostic significance of HOXD13 expression in human breast cancer. *Int J Clin Exp Pathol* (2015) 8(9):11407–13.
90. Giraldo NA, Becht E, Vano Y, Petitprez F, Lacroix L, Validire P, et al. Tumor-Infiltrating and Peripheral Blood T-cell Immunophenotypes Predict Early Relapse in Localized Clear Cell Renal Cell Carcinoma. *Clin Cancer Res* (2017) 23(15):4416–28. doi: 10.1158/1078-0432.CCR-16-2848
91. Rody A, Karn T, Liedtke C, Pusztai L, Ruckhaeberle E, Hanker L, et al. A clinically relevant gene signature in triple negative and basal-like breast cancer. *Breast Cancer Res* (2011) 13(5):R97. doi: 10.1186/bcr3035
92. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *J Immunother Cancer* (2016) 4:59. doi: 10.1186/s40425-016-0165-6
93. Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol* (2018) 19(1):40–50. doi: 10.1016/S1470-2045(17)30904-X
94. Dong Y, Quan H-Y. Downregulated LINC00460 inhibits cell proliferation and promotes cell apoptosis in prostate cancer. *Eur Rev Med Pharmacol Sci* (2019) 23(14):6070–8. doi: 10.26355/eurrev_201907_18420
95. Li K, Sun D, Gou Q, Ke X, Gong Y, Zuo Y, et al. Long non-coding RNA linc00460 promotes epithelial-mesenchymal transition and cell migration in lung cancer cells. *Cancer Lett* (2018) 420:80–90. doi: 10.1016/j.canlet.2018.01.060
96. Yuan B, Yang J, Gu H, Ma C. Down-Regulation of LINC00460 Represses Metastasis of Colorectal Cancer via WWC2. *Dig Dis Sci* (2020) 65(2):442–56. doi: 10.1007/s10620-019-05801-5
97. Avgustinova A, Iravani M, Robertson D, Fearns A, Gao Q, Klingbeil P, et al. Tumour cell-derived Wnt7a recruits and activates fibroblasts to promote tumour aggressiveness. *Nat Commun* (2016) 7(1):10305. doi: 10.1038/ncomms10305
98. Yi K, Min K-W, Wi YC, Kim Y, Shin S-J, Chung MS, et al. Wnt7a Deficiency Could Predict Worse Disease-Free and Overall Survival in Estrogen Receptor-Positive Breast Cancer. *J Breast Cancer* (2017) 20(4):361. doi: 10.4048/jbc.2017.20.4.361
99. Liu Y, Meng F, Xu Y, Yang S, Xiao M, Chen X, et al. Overexpression of Wnt7a Is Associated With Tumor Progression and Unfavorable Prognosis in Endometrial Cancer. *Int J Gynecol Cancer* (2013) 23(2):304–11. doi: 10.1097/IGC.0b013e31827c7708
100. Gordon GJ, Jensen RV, Hsiao L-L, Gullans SR, Blumenstock JE, Ramaswamy S, et al. Translation of microarray data into clinically relevant cancer

diagnostic tests using gene expression ratios in lung cancer and mesothelioma. *Cancer Res* (2002) Sep 162(17):4963–7.
101. Ma X-J, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* (2004) 5(6):607–16. doi: 10.1016/j.ccr.2004.05.015
102. Gordon GJ, Jensen RV, Hsiao L-L, Gullans SR, Blumenstock JE, Richards WG, et al. Using Gene Expression Ratios to Predict Outcome Among Patients With Mesothelioma. *JNCI J Natl Cancer Institute* (2003) 95 (8):598–605. doi: 10.1093/jnci/95.8.598
103. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RCT, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* (2002) 8(1):68–74. doi: 10.1038/nm0102-68
104. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* (1999) 286(5439):531–7. doi: 10.1126/science.286.5439.531
105. Denkert C, Loibl S, Noske A, Roller M, Müller BM, Komor M, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* (2010) 28 (1):105–13. doi: 10.1200/JCO.2009.23.7370
106. West NR, Milne K, Truong PT, Macpherson N, Nelson BH, Watson PH. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. *Breast Cancer Res* (2011) 13(6):R126. doi: 10.1186/bcr3072
107. Xiong B, Lei X, Zhang L, Fu J. miR-103 regulates triple negative breast cancer cells migration and invasion through targeting olfactomedin 4. *Biomed Pharmacother* (2017) 89:1401–8. doi: 10.1016/j.biopha.2017.02.028
108. Wang D-S, Zhong B, Zhang M-S, Gao Y. Upregulation of serum miR-103 predicts unfavorable prognosis in patients with colorectal cancer. *Eur Rev Med Pharmacol Sci* (2018) 22(14):4518–23. doi: 10.26355/eurrev_201807_15506
109. Chen H-Y, Lang Y-D, Lin H-N, Liu Y-R, Liao C-C, Nana AW, et al. miR-103/107 prolong Wnt/β-catenin signaling and colorectal cancer stemness by targeting Axin2. *Sci Rep* (2019) 9(1):9687. doi: 10.1038/s41598-019-41053-z
110. Zheng J, Liu Y, Qiao Y, Zhang L, Lu S. miR-103 Promotes Proliferation and Metastasis by Targeting KLF4 in Gastric Cancer. *Int J Mol Sci* (2017) 18 (5):1–13. doi: 10.3390/ijms18050910

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

4. OTRO ARTÍCULO PUBLICADO DURANTE LA TESIS DOCTORAL

SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

A lncRNA landscape in breast cancer reveals a potential role for AC009283.1 in proliferation and apoptosis in HER2-enriched subtype

Alberto Cedro-Tanda[1,3], Magdalena Ríos-Romero[1,6], Sandra Romero-Córdoba[1,5], Mireya Cisneros-Villanueva[1], Rosa Gloria Rebollar-Vega[4], Luis Alberto Alfaro-Ruiz[1], Silvia Jiménez-Morales[1], Carlos Domínguez-Reyes[2], Felipe Villegas-Carlos[2], Alberto Tenorio-Torres[2], Veronica Bautista-Piña[2], Fredy Omar Beltrán-Anaya[1,3] & Alfredo Hidalgo-Miranda[1]✉

Breast cancer is the most commonly diagnosed neoplasm in women worldwide with a well-recognized heterogeneous pathology, classified into four molecular subtypes: Luminal A, Luminal B, HER2-enriched and Basal-like, each one with different biological and clinical characteristics. Long non-coding RNAs (lncRNAs) represent 33% of the human transcriptome and play critical roles in breast carcinogenesis, but most of their functions are still unknown. Therefore, cancer research could benefit from continued exploration into the biology of lncRNAs in this neoplasm. We characterized lncRNA expression portraits in 74 breast tumors belonging to the four molecular subtypes using transcriptome microarrays. To infer the biological role of the deregulated lncRNAs in the molecular subtypes, we performed co-expression analysis of lncRNA–mRNA and gene ontology analysis. We identified 307 deregulated lncRNAs in tumor compared to normal tissue and 354 deregulated lncRNAs among the different molecular subtypes. Through co-expression analysis between lncRNAs and protein-coding genes, along with gene enrichment analysis, we inferred the potential function of the most deregulated lncRNAs in each molecular subtype, and independently validated our results taking advantage of TCGA data. Overexpression of the AC009283.1 was observed in the HER2-enriched subtype and it is localized in an amplification zone at chromosome 17q12, suggesting it to be a potential tumorigenic lncRNA. The functional role of lncRNA AC009283.1 was examined through loss of function assays in vitro and determining its impact on global gene expression. These studies revealed that AC009283.1 regulates genes involved in proliferation, cell cycle and apoptosis in a HER2 cellular model. We further confirmed these findings through ssGSEA and CEMITool analysis in an independent HER2-amplified breast cancer cohort. Our findings suggest a wide range of biological functions for lncRNAs in each breast cancer molecular subtype and provide a basis for their biological and functional study, as was conducted for AC009283.1, showing it to be a potential regulator of proliferation and apoptosis in the HER2-enriched subtype.

[1]Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico. [2]FUCAM, Instituto de Enfermedades de la Mama, Mexico City, Mexico. [3]Programa de Doctorado en Ciencias Biomédicas, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico. [4]Genomics Laboratory, Red de Apoyo a la Investigación, Universidad Nacional Autónoma de México, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico. [5]Biochemistry Department, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico. [6]Programa de Doctorado de Ciencias Biológicas (Biomedicina), Universidad Nacional Autónoma de Mexico, Ciudad de México, Mexico. ✉email: ahidalgo@inmegen.gob.mx

SCIENTIFIC REPORTS | (2020) 10:13146 | https://doi.org/10.1038/s41598-020-69905-z

Breast cancer (BRCA) is one of the most common tumors in women around the world. Incidence of this disease is increasing, particularly in countries with emerging economies. In Mexico, BRCA affects about 27,830 women and represents the first cause of cancer related deaths, with approximately 6,884 deaths per year[1]. BRCA is a complex and heterogeneous disease. In the clinical setting, breast tumors are classified into clinically relevant groups based on the expression of three main immunohistochemical (IHC) markers: estrogen and progesterone receptors and (Human Epidermal Growth Factor Receptor 2) HER2,triple negative tumors are defined by the absence of the expression of these markers[2].

Breast tumors can be further subdivided based on messenger RNA (mRNA) expression signatures. One of these, the intrinsic breast tumor signature, identifies sub-groups based on the expression of 50 genes (the PAM50 signature): Luminal A, Luminal B, HER2-enriched, and Basal-like tumors[3,4]. More importantly, each of these tumor-subtypes has been associated with particular biological and clinical behaviors. Compared to luminal subtypes, Basal-like and HER2-enriched subtypes present worse five-year prognosis, and the HER2-enriched subtype has the worst relapse-free survival[5]. Several lines of evidence indicate that the most representative breast tumor subtype phenotypes identified by mRNA intrinsic signatures reflect not only the alteration of mRNA expression levels, but also the genetic and epigenetic alterations that contribute to the establishment and maintenance of different cancer pathways[6].

Most of the transcriptional differences between breast tumor subtypes focus on mRNA expression analysis. However, coding transcripts only represent around 39% of the human transcriptome, the rest are non-coding RNAs, which can be divided into two classes: small non-coding RNAs (less than 200 nucleotides, this category includes microRNAs) and long non-coding RNAs (lncRNAs), with more than 200 nucleotides[7].

LncRNAs have several important biological roles, regulating gene expression at the epigenetic, transcriptional and post-transcriptional levels, and their deregulation has been associated with cancer[8]. In BRCA, lncRNAs are emerging as master regulators in tumor biology, and have oncogenic and tumor suppressor functions related to initiation and cancer progression, some examples are HOTAIR, MALAT-1, lincRNAp21, and GAS5. Deregulation of these lncRNAs is associated with biological functions such as invasion, proliferation, apoptosis, cell cycle, and clinical features such as survival, progression and risk of metastasis[9,10].

Analyses of lncRNA expression patterns in BRCA have identified differences between tumor groups defined by IHC markers, including HER2 positive[11] and triple-negative tumors[12]. Other studies have evaluated the aberrant expression of lncRNAs in gene expression-based tumor subtypes, identifying specific lncRNA expression patterns for each subtype[13,14]. Nevertheless, there is still little information about the biological roles of specific differentially expressed lncRNAs in each tumor molecular subtype, particularly in those with worse prognosis, such as HER2-enriched.

The aim of the present study was to explore the transcriptional lncRNA landscape across the four molecular subtypes in 74 Mexican breast tumors. We further validated altered lncRNA expression in the Cancer Genome Atlas (TCGA) cohort. We identified aberrantly expressed lncRNAs in each BRCA molecular subtype, and then selected a group of lncRNAs with exclusive deregulation in each subtype and inferred their potential biological relevance through in silico analysis. Focusing on the HER2-enriched subtype, which has the worst relapse-free and overall survival rates, we identified AC009283.1 as the most up-regulated lncRNA, with a previously unknown biological role. Knockdown of AC009283.1 showed reduced cell proliferation, promoted S phase arrest, and enhanced apoptosis in SKBR3 cells (HER2-enriched). These biological processes were also enriched in HER2-enriched breast cancer tumors with high AC009283.1 expression levels, suggesting that AC009283.1 plays a role in promoting HER2-enriched tumors. The influence of lncRNAs on deregulated pathways in different breast tumor subtypes is relevant and supports the need for new approaches to understand the biology of BRCA.

Results

lncRNA expression is deregulated in breast cancer. The transcriptional portrait among tumors and adjacent tissues revealed 307 differentially expressed lncRNAs (27 up-regulated and 280 down-regulated) (fold change > 2.0 and < − 2.0 and FDR < 0.05) (Supplementary data 1.1). Hierarchical clustering analysis based on differentially expressed lncRNAs was able to optimally discriminate tumor breast tissues from normal adjacent samples, as shown in the heat map (Fig. 1A). The altered lncRNAs were classified into ncRNA categories as follows: lincRNAs (45.6%), antisense (27.36%), sense-intronic (20.52%), sense-overlapping (5.86%) and 3′-overlapping (0.65%) (Fig. 1B).

Several differentially expressed lncRNAs identified in our study have been previously reported to be related to BRCA, such as HOTAIR, HOTAIRM1, XIST, PANDAR, EPB41L4A-AS1, BC040587, FGF14-AS2, DSCAM-AS1, LINC00472, MIR31HG and FAM83H-AS1 (Supplementary data 1.2).

A set of differentially expressed lncRNAs was independently validated using whole transcriptome RNA-seq analysis data from TCGA. Some of the lncRNAs described as altered in our profile did not significantly change in the TCGA dataset while others were not detected by the sequencing experiment (Supplementary data 1.3).

Molecular subtypes are associated with distinct clinical outcomes. A total of 74 breast tumor samples and 12 adjacent tissue samples were included in the PAM50 intrinsic subtype analysis. Incidence of PAM50 subtypes were reported as follows: 32.4% Luminal A (n = 24), 31.1% Luminal B (n = 23), 18.9% HER2-enriched (n = 14), and 17.6% Basal-like (n = 13) subtypes. Clinical behavior differs among the molecular subtypes. For instance, in our in-home profiled cohort, the HER2-enriched subtype showed the highest rate of metastasis (log-rank p = 0.034). Similarly, lower overall-survival was observed in the METABRIC data set (n = 1,463) (95% CI [1.417–2.368] Cox p-value = 2.08E−10) as well as in the TCGA data set (n = 743) (95% CI [0.97–4.01] Cox p-value = 2.19E−7) (Supplementary Figure 1). Clinical and pathological features of our cohort are summarized in Supplementary Table 1.

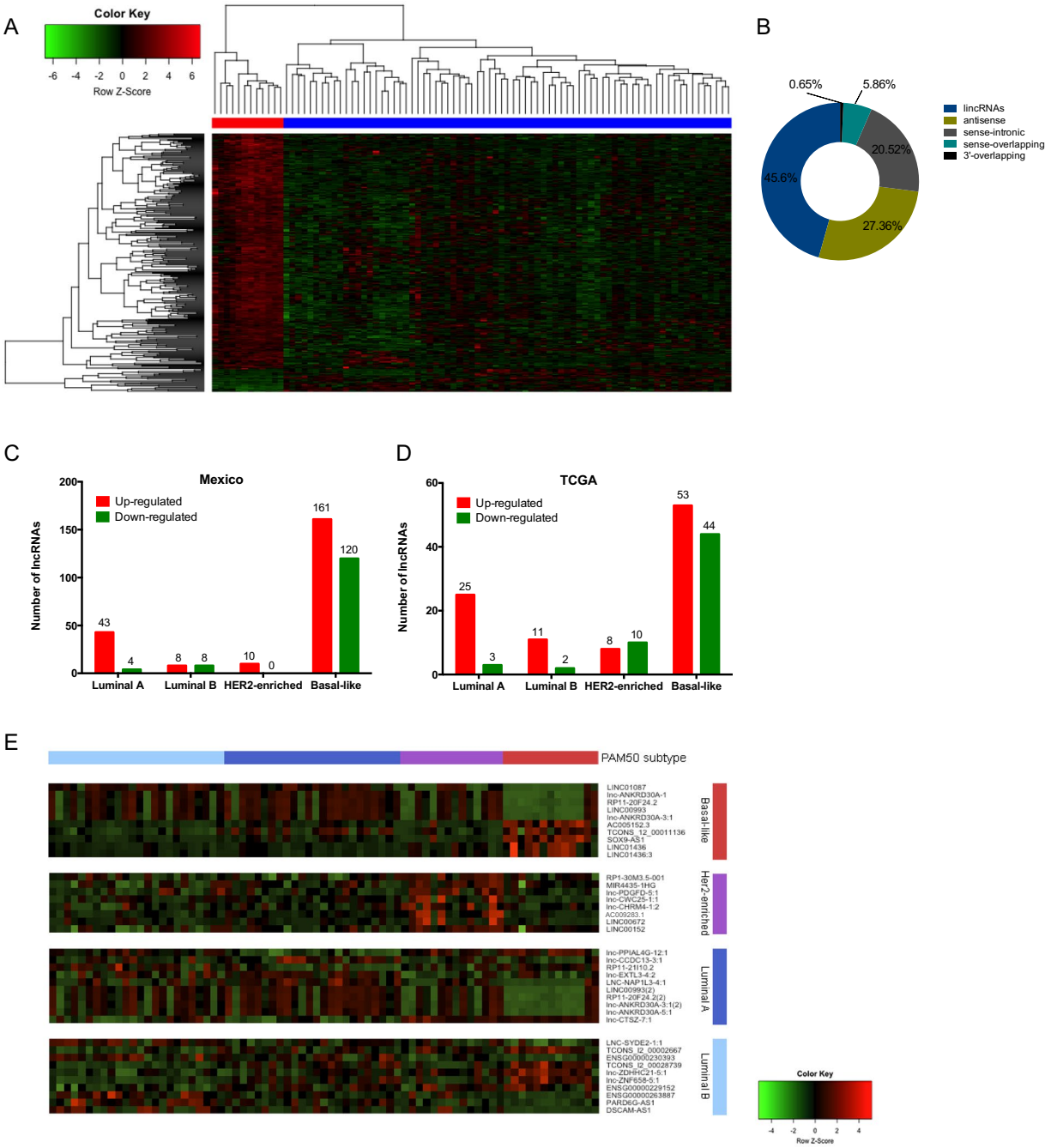

Figure 1. (**A**) Supervised hierarchical cluster analysis of the expression of the 307 lncRNAs deregulated in breast tumor (blue) and normal tissue (red). (**B**) Classification of lncRNAs deregulated between tumor and adjacent tissues, grouped based on their genomic localization by intersection with protein-coding genes. (**C**) lncRNAs deregulated across molecular subtypes in the Mexican cohort (fold change > 1.5, < − 1.5, FDR < 0.05) and (**D**) in the TCGA project (Fold change log 1.0. adj. p-value < 0.05). (**E**) Top ten lncRNAs deregulated across molecular subtypes in the Mexican cohort. Basal-like in red, Her2-enriched in purple, Luminal A in blue and Luminal B in cyan. Each column represents a sample and each row a lncRNA.

Expression profile of lncRNAs in breast cancer subtypes. To understand the biology of the intrinsic breast cancer subtypes, we explored the transcriptional landscape of the lncRNAs along the PAM50 subtypes. 354 differentially expressed lncRNAs were identified. In Luminal A tumors, we observed 47 altered lncRNAs, 16 in Luminal B, 10 in HER2-enriched, and 281 in Basal-like (Fig. 1C) (Supplementary data 2.1). TCGA data showed a similar proportion of lncRNAs deregulated in each molecular subtype (Fig. 1D) (Supplementary data 2.2). Figure 1E shows the expression of the top deregulated lncRNAs (5 up-regulated and 5 down-regulated) in Basal, Luminal A, Luminal B, and eight up-regulated in HER2-enriched tumors in our cohort.

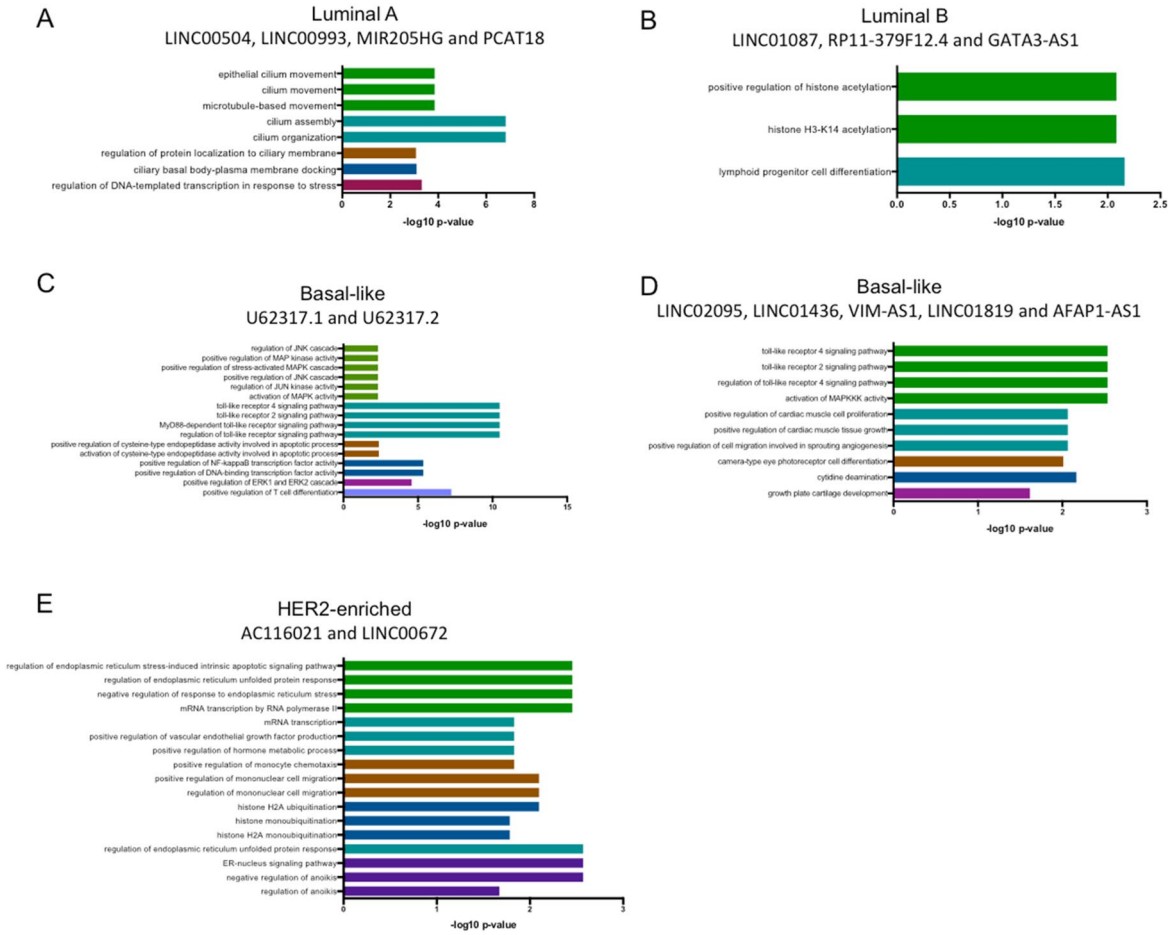

Figure 2. Gene ontology (GO) analysis of lncRNAs up-regulated in our cohort and TCGA. GO was performed using mRNAs that correlate with lncRNAs in each subtype using ClueGo in Cytoscape with PValue Corrected < 0.05.

lncRNA expression portraits in Luminal A tumors. In our cohort, we identified 47 lncRNAs differentially-expressed in the Luminal A subtype (43 up-regulated and 4 down-regulated) (Supplementary data 2.1). To define robust altered lncRNAs, we examined the TCGA cohort. Four lncRNAs were commonly up-regulated in Mexican and TCGA cohorts: LINC00504, LINC00993, MIR205HG and PCAT18. To get insight into their biological role as transcriptional regulators we investigated co-expression lncRNA–mRNA patterns of these four lncRNAs. 340 mRNAs were significantly correlated with LINC00504, LINC00993, MIR205HG and PCAT18 in both cohorts ($R > 0.3$ p-value < 0.05) (Supplementary data 3.1). Gene ontology (GO) analysis of the correlated mRNAs identified positive regulation of DNA-templated transcription in response to stress, microtubule-based movement and processes related to ciliary docking and organization (Fig. 2A).

lncRNA expression portraits in Luminal B tumors. We identified 16 deregulated lncRNAs in the Luminal B subtype in our cohort (8 up-regulated and 8 down-regulated) (Supplementary data 2.1). Up-regulation of LINC01087, RP11-379F12.4 and GATA3-AS1 was validated in the TCGA data set. Correlation analysis of the altered lncRNA and mRNAs, detected 70 mRNAs co-expressed in both cohorts (Supplementary data 3.2). According to GO analysis, the correlated genes participate in the regulation of histone H3-K14 acetylation and lymphoid progenitor cell differentiation (Fig. 2B). Moreover, in the TCGA cohort, the over-expression of these lncRNAs (LINC01087, RP11-379F12.4) was significantly associated with increased survival rates (log-rank test < 0.05), compared with tumors with low expression levels (Supplementary figure 2).

lncRNA expression portraits in basal-like tumors. The basal-like subtype showed the highest number of differentially expressed lncRNAs, with 161 up-regulated and 120 down-regulated transcripts in our cohort (Supplementary data 2.1). Regarding the up-regulated lncRNAs, only a low number of them have been previously implicated in cancer, such as the up-regulated CASC15, LINC00662, LINC01272, LINC00857, LINC00092, LINC00673 and the down-regulated LINC00657, LINC01087 and LINC00993. LncRNAs U62317.1 and U62317.2 are adjacent lncRNAs localized in chromosome 22q13.33, and both are up-regulated in our samples and in the TCGA cohort. U62317.1 and U62317.2 expression was correlated with 419 mRNAs in both cohorts (Supplementary data 3.3). GO analysis showed that these mRNAs are implicated in positive regula-

tion of JNK cascade, activation of cysteine-type endopeptidase involved in apoptotic process, positive regulation of NF-kappaB, ERK1 (MAPK3) and ERK2 (MAPK1), and positive regulation of T cell differentiation (Fig. 2C). High expression of U62317.2 was associated with higher OS in patients with basal subtype (log-rank p-value = 0.04378) (Supplementary Figure 3). Moreover, another five lncRNAs were common between our cohort and TCGA: LINC02095, LINC01436, VIM-AS1, LINC01819 and AFAP1-AS1. Co-expression analysis revealed 103 correlated mRNAs in both cohorts (Supplementary data 3.3), which are associated with the activation of MAPKKK and positive regulation of cell proliferation and differentiation (Fig. 2D).

lncRNA expression portraits in HER2-enriched tumors. We detected 8 novel up-regulated lncRNAs in the HER2-enriched subtype, not reported previously in BRCA (Supplementary data 2.1). Two of them (AC116021 y LINC00672) are up-regulated in both cohorts (our study and TCGA), and their expression is correlated with 90 mRNAs in both cohorts (Supplementary data 3.4). Gene ontology analysis revealed their involvement in the positive regulation of RNA Pol II transcription, response to endoplasmic reticulum stress, positive regulation of VEGF production, positive regulation of monocyte chemotaxis, histone monoubiquitination, negative regulation of anoikis, among other processes (Fig. 2E).

As described above, HER2-enriched breast cancer patients showed a significantly decreased global survival rate, both in TCGA and METABRIC cohorts. Therefore, we endeavored to characterize altered lncRNAs in this molecular subtype. Interestingly, AC009283.1 presented the highest expression level in HER2-enriched tumors compared with other BRCA intrinsic molecular subtypes, in both analyzed datasets (Mexican cohort: FC 4.02, p value < 0.0001, TCGA: FC 2.46 and p-value < 0.0001) (Fig. 3A), without any previous functional report (Supplementary data 2.1). Therefore, AC009283.1 was selected for further in vitro evaluation.

AC009283.1 is up-regulated and amplified in HER2-enriched tumors. AC009283.1 is an intergenic lncRNA localized in chr17:39,566,915–39,567,559 (GRCh38/hg38) annotated in the ENSEMBL database as ENSG00000273576. Some alternative names have been reported in other databases such as HSALNG0116248, NONHSAG021701, lnc-CDK12-1 and RP11-390P24.1**.** The Coding potential (CP) of AC009283.1 is 0.113, according to the Coding Potential Calculator, similar to other well-known lncRNAs, such as HOTAIR, MALAT-1 and PANDAR (CPs: 0.09, 0.1295 and 0.058, respectively).

AC009283.1 is significantly up-regulated in the HER2-enriched subtype in our Mexican cohort, with a fold change of 4.02 (microarray analysis; p-value < 0.0001) (Supplementary data 2.1). Independent validation with TCGA data showed an increased expression in HER2 enriched tumors (RNA-Seq analysis; foldchange 2.46 and p-value < 0.0001) (Fig. 3A). Breast cancer genomic analysis revealed that AC009283.1 maps close to the HER2 gene and is frequently co-amplified within the core region of the HER2-amplicon (17q12) (Fig. 3B). Dedicated copy number alteration analysis on TCGA HER2-amplified tumors have demonstrated that the amplicon can span over other neighboring regions on 12q-21.3, where a number of protein-coding and non-coding genes are encoded and consequently amplified (Fig. 3B). The expression levels of AC009283.1 were significantly correlated with HER2 copy number alterations evaluated through the Genomic Identification of Significant Targets in Cancer (GISTIC) from TCGA data (Fig. 3C). Overall, these data suggest that AC009283.1 is a relevant feature in HER2-enriched breast cancer programs. Thus, to investigate whether the altered expression of AC009283.1 affects HER2-enriched tumor biology, we studied its in-vitro activity in a HER2-enriched cell line model.

Expression of AC009283.1 was evaluated in a panel of BRCA cell lines, identifying SKBR3 cell line as a potential biological model for further analysis. SKBR3 cells over-expressing AC009283.1, carry a 17q12 amplification, and according to PAM50 subtypes represents a HER2-enriched subtype (Supplementary Figure 4). To infer the regulatory role of the lncRNA in cancer programs, its subcellular localization was evaluated. Overall, AC009283.1 was preferentially located within the nucleus, which may suggest its predominant function as a regulator of protein-coding gene expression through transcriptional and epigenetic mechanisms (Supplementary Figure 5).

Knockdown of AC009283.1 alters the transcriptional profile of HER2-enriched breast cancer cells. To explore AC009283.1 function in the HER2-enriched BRCA subtype, we inhibited its expression using two shRNA sequences (shRNA1 and shRNA2). Both sequences target a common exon in 10 of 14 isoforms reported for AC009283.1. 96 h post-transfection, the most relevant reduction of AC009283.1 expression level (70% *vs* negative control) was achieved with shRNA2, thus this shRNA sequence was used for further experiments (Fig. 4A).

Since lncRNAs are capable of regulating genes located nearby (*cis*) or in distant genomic regions (*trans*), we used a high-throughput microarray approach to evaluate transcriptome alterations after AC009283.1 shRNA-mediated knockdown, comparing with shRNA negative control, in SKBR3 cells. 158 genes were differentially expressed among the experimental conditions and distributed throughout the genome (> 1.2, < − 1.2-fold) (ANOVA p-value < 0.05) (Fig. 4B,C). Genes such as *CASCA4*, *NOCTH3*, *TNFa*, *FOSB*, *BLC2A1*, *DNML1,* and *KLF6* were deregulated in AC009283.1 knockdown cells (Supplementary data 4.1). To further to explore the effect of inhibiting AC009283.1 in biological processes, we performed a pathway enrichment analysis of the differentially expressed genes, showing proliferation and apoptosis as significantly affected pathways (Fig. 4D). Other oncogenic pathways were also altered, such as oxidative phosphorylation, mitochondrial dysfunction, p38 MAPK Signaling, PI3K/AKT signaling, among others (Supplementary Figure 5). Interestingly, proliferation, apoptosis and cell cycle were also observed in HER2-enriched tumors from patients in TCGA containing up-regulated AC009283.1 (upper vs lower quartile) (Fig. 4E) (Supplementary data 4.2). These results suggest that AC009283.1 plays a role in biological processes by regulating gene expression of important genes related with carcinogenic pathways.

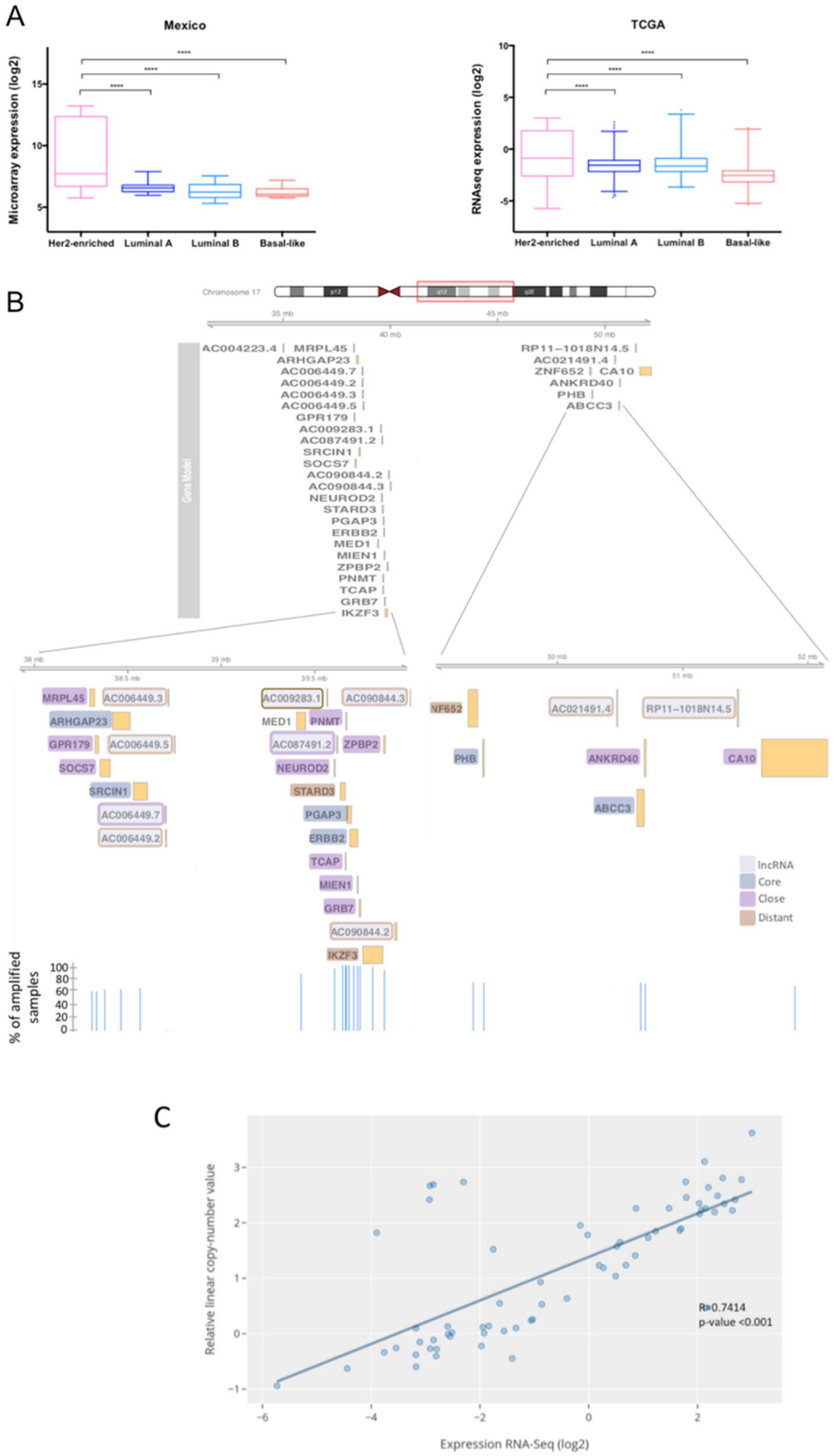

Figure 3. (**A**) AC009283.1 expression across different molecular subtypes in the Mexican cohort (n = 74; microarray data) and TCGA/TANRIC cohort (n = 853; RNA-Seq data). (**B**) Map of amplicon 17q12 with mRNAs and lncRNAs. (**C**) Correlation analysis between expression and copy-number value of AC009283.1 in HER2-enriched from TCGA data.

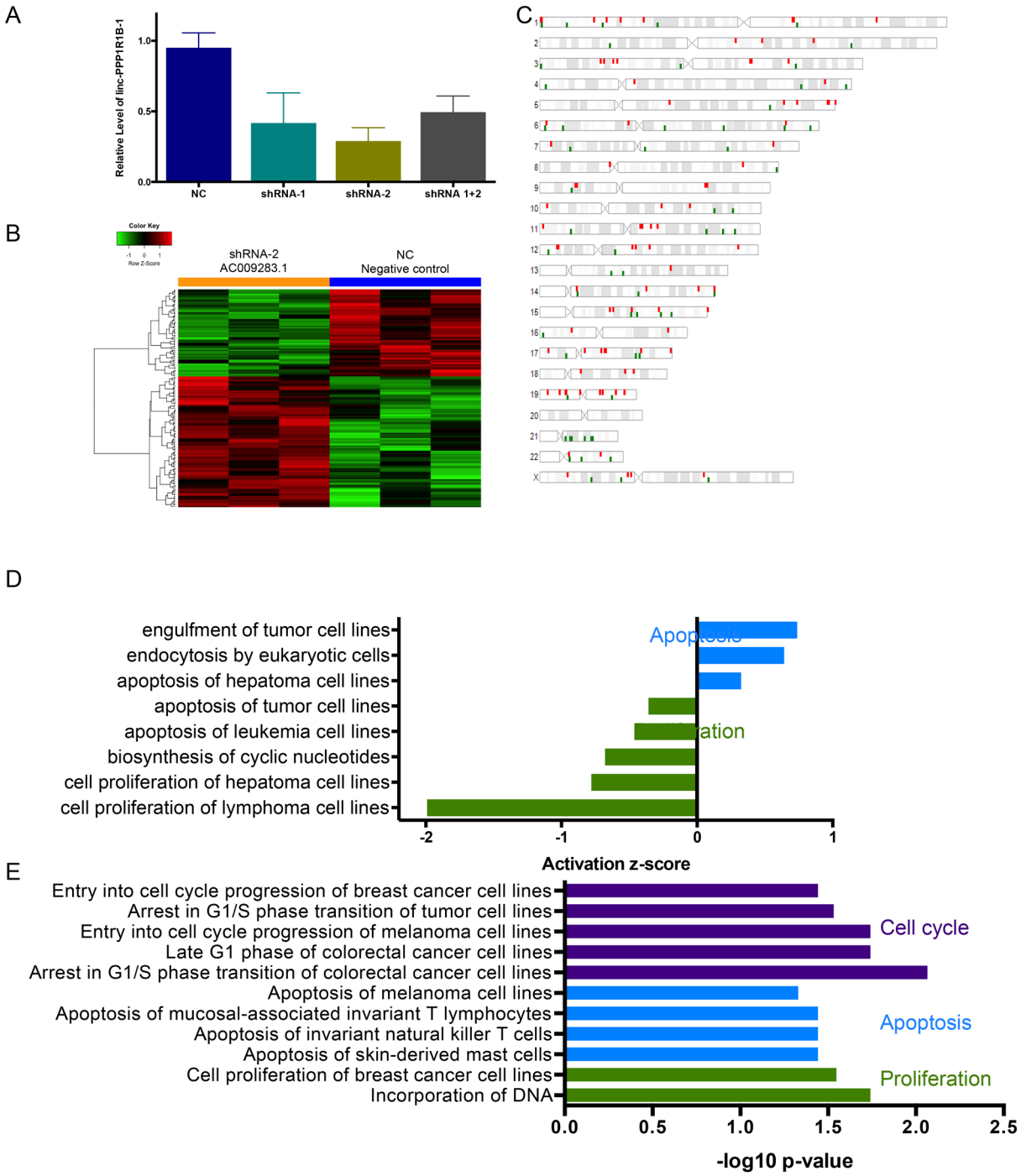

Figure 4. (**A**) Real-time qPCR illustrating gene silencing by shRNA against AC009283.1 in SKBR3, a HER2-enriched cell line model. (**B**) Supervised hierarchical cluster analysis shows 158 differentially expressed genes after AC009283.1 knockdown (Foldchange < −1.2, > 1.2 p-value < 0.05); we observed 94 up-regulated and 64 down-regulated genes in shRNA-2 vs NC and (**C**). Their distribution across the genome. (**D**) Ingenuity Pathway Analysis of 158 genes differentially expressed after knockdown of AC009283.1 in SKBR3 cells. (**E**) Enrichment pathway analysis of genes differently expressed in samples of HER2-enriched tumors from TCGA with high vs low expression of AC009283.1.

AC009283.1 knockdown inhibits proliferation and induces apoptosis in HER2-enriched breast cancer cells. To validate the pathway analysis, we performed cell proliferation and apoptosis assays in vitro. Automated cell counting showed that knockdown of AC009283.1 attenuated cell proliferation compared with the control group (NC) in the SKBR3 cell line. Significantly lower cell numbers are observed from the second day after transfection (Fig. 5A). To confirm this result, proliferation rates were evaluated by flow cytometry. The proliferation index on the fourth day after transfection was significantly decreased (6.01 vs 10.68 in negative

control cells) in SKBR3 cells with reduced AC009283.1 (Fig. 5B). A representative graph of the decreased proliferation rates observed with CFSE assay in AC009283.1 knockdown cells is represented in Fig. 5C. To understand how cell proliferation is controlled by AC009283.1, we performed a cell cycle assay after four days post-transfection with shRNA2 against AC009283.1. A significant arrest in S phase was observed in cells with low levels of AC009283.1 (Fig. 5D). These results suggest that AC009283.1 may regulate cell proliferation by arresting cells in S phase.

Additionally, apoptosis assays using flow cytometry revealed that AC009283.1 knockdown enhances the percentage of cells undergoing early apoptosis (21.3% vs 15.9% in control cells) (Fig. 6A,B). Furthermore, we measured caspase-3 enzymatic activity and found that AC009283.1 knockdown significantly increases the level of caspase-3 activity at day four, compared with negative control cells (shRNA NC) (Fig. 6C). These results suggest that AC009283.1 promotes apoptosis resistance in HER2-enriched breast cancer cells.

AC009283.1 is over-expressed in HER2-amplified tumors and might regulate relevant oncogenic pathways. Based on the above described correlation between HER2-gain and the lncRNA level, we hypothesized that AC009283.1 over-expression is not unique to the HER2-enriched subtype but rather an event associated with HER2-amplification. Since the HER2-enriched PAM50 subtype captures some, but not all of the HER2+ tumors, we expanded our analysis to all HER2-amplified tumors, evaluated through microarray or immunochemistry (IHC) analysis, from public available datasets from TCGA and Gene Expression Omnibus (GEO).

We first explored the expression portraits of HER2-neighboring lncRNA genes encoded in the 17q12-21.3 amplicon. We systematically explored and integrated somatic copy number alteration data of HER2 amplification in breast tumors, independently of the PAM50 molecular subtype (Fig. 3B). A set of lncRNAs contained in or neighboring amplified genes within the 17q12-21.3 amplicon showed a coordinated expression pattern with HER2 mRNA level, suggesting that these correlated transcripts may be needed to sustain oncogenic programs prompted by HER-amplification (Supplementary Figure 7). Importantly, AC009283.1 was among the lncRNAs most correlated with ERBB2 expression (R:0.61 $p < 0.001$) (Supplementary Figure 7). Furthermore, among the overall evaluated lncRNAs, AC009283.1 presented the highest expression in HER-amplified tumors compared with tumors without copy number changes or amplicon loss (Fig. 7A). In accordance with these data, we found a significant up-modulation of AC009283.1 in HER2+ tumors evaluated by IHC, independent of the hormone receptor status (Fig. 7B).

After determining a relevant correlation between HER2-amplification/over-expression and AC009283.1 expression, we performed a gene set enrichment analysis (GSEA) to identify relevant biological processes enriched in HER2+ tumors with diverse AC009283.1 expression patterns. As shown in Fig. 7C, GSEA analysis confirm previous findings in our in vitro assays. We found a significant, exclusive enrichment in G2–M checkpoint, G0–G1 transition, regulation of cell-cycle phase transition, DNA replication, and regulation of intrinsic apoptotic signaling pathways in HER2+ tumors over-expressing AC009283.1.

To validate our observations in a more sophisticated characterization, normalized ssGSEA (single-sample GSEA) scores were computed on HER2+ breast cancer samples retrieved from TCGA (n = 183) and Gene Expression Omnibus (GEO) datasets (n = 226). Once again, tumors over-expressing AC009283.1 (High group) were found to be predominantly enriched in cell proliferation, cell cycle, cell cycle G1–S transition, cell cycle G2–M transition, regulation of cell G2–M transition and negative regulation of cell cycle arrest, compared with low-expressing tumors (Fig. 7D).

Modular expression analysis of AC009283.1 in HER2-amplified tumors. Assessing co-expression patterns may help identify novel functional connections between AC009283.1 and mRNAs under HER2+ amplification. Thus, we performed a gene co-expression network analysis with CEMiTool package using the expression profiles of HER2+ tumors divided in accordance with AC009283.1 levels into a low, medium or high category. Four-gene modules were revealed (Fig. 8A). M1, M2 and M3 were highly enriched in AC009283.1 (Fig. 8A) and gene-set enrichment analysis revealed a significant over-representation of cell metabolism and growth (M1), immune signaling (M2) and oncogenic pathways, such as KRAS (M3), of all them markedly related with aberrant cell proliferation, cell-cycle progression and apoptotic rates, which corroborated our in vitro results (Fig. 8B).

We then integrated co-expression information to identify main regulators and hubs within the modules. Interestingly, as shown in the interaction network, AC009283.1 correlated with genes contained in modules M1, M2 and M3 (Fig. 8C). This analysis showed that AC009283.1 lncRNA may unveil potential transregulatory relationships within the HER2 response network.

Overall, these data suggested that high expression of AC009283.1 is strongly associated with cell proliferation, cell cycle progression and apoptosis, and further reveal that the biological role of AC009283.1 in HER2-enriched breast cancer is probably modulated and associated with HER2-amplification.

Discussion

A few decades ago, the non-coding sequences of the genome (98%) were considered non-functional and the search for tumor oncogenes and suppressors focused only on protein coding sequences (2%). Now with the use of high-throughput screening we know that around 70% of the genome is transcribed, including non-coding regions which play an important role in the regulation of gene expression[15]. LncRNAs are deregulated in several cancer types[8]. This study provides a whole transcriptome in-house profile from well-characterized breast cancer tumors, and identifies the deregulated expression of AC009283.1 in the HER2-enriched subtype. We found 307 lncRNAs differentially expressed in tumors compared to normal adjacent tissue and their expression pattern was able to distinguish BRCA tissue from normal adjacent. A small set of deregulated lncRNAs have already

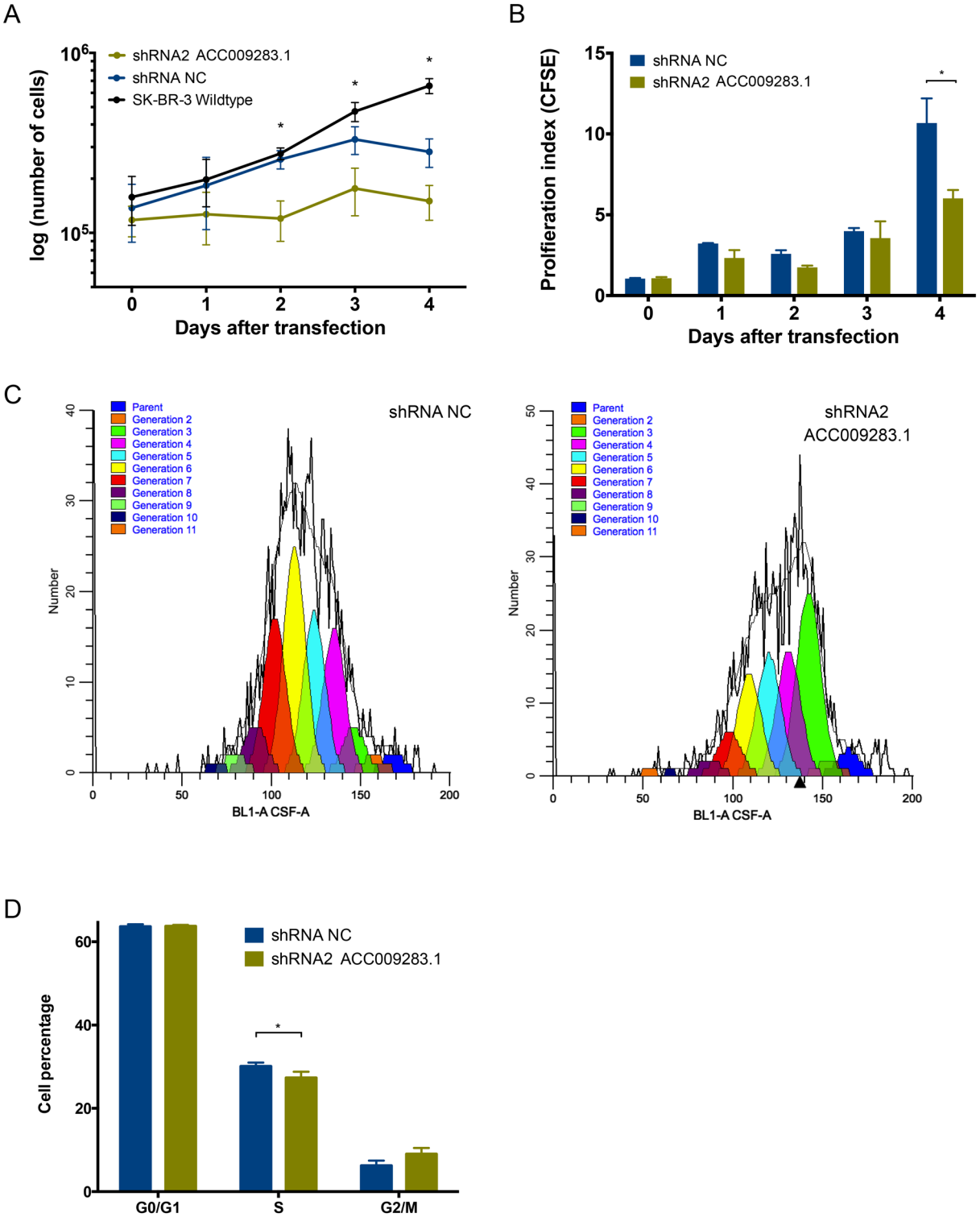

Figure 5. (**A**) Proliferation assay performed following knockdown of AC009283.1 with shRNA2, shRNA NC as negative control, and wildtype cell line. Cells were counted every day for four days. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, comparing with shRNA NC for each condition via Student's *t* test. (**B**) Proliferation index using CFSE assay in SKBR3 cells with shRNA2 AC009283.1 and shRNA NC. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, comparing with shRNA NC for each condition via Student's *t* test. (**C**) Representative CFSE flow-cytometry histograms showing the effect on in vitro SKBR-3 cell proliferation with shRNA NC and shRNA2 (AC009283.1 knock-down). (**D**) Cell cycle assay with flow-cytometry using (propidium iodide) PI on in vitro SKBR-3 cells, AC009283.1 knockdown causes accumulation of cells at S phase. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, comparing to shRNA NC for each condition via Student's *t* test.

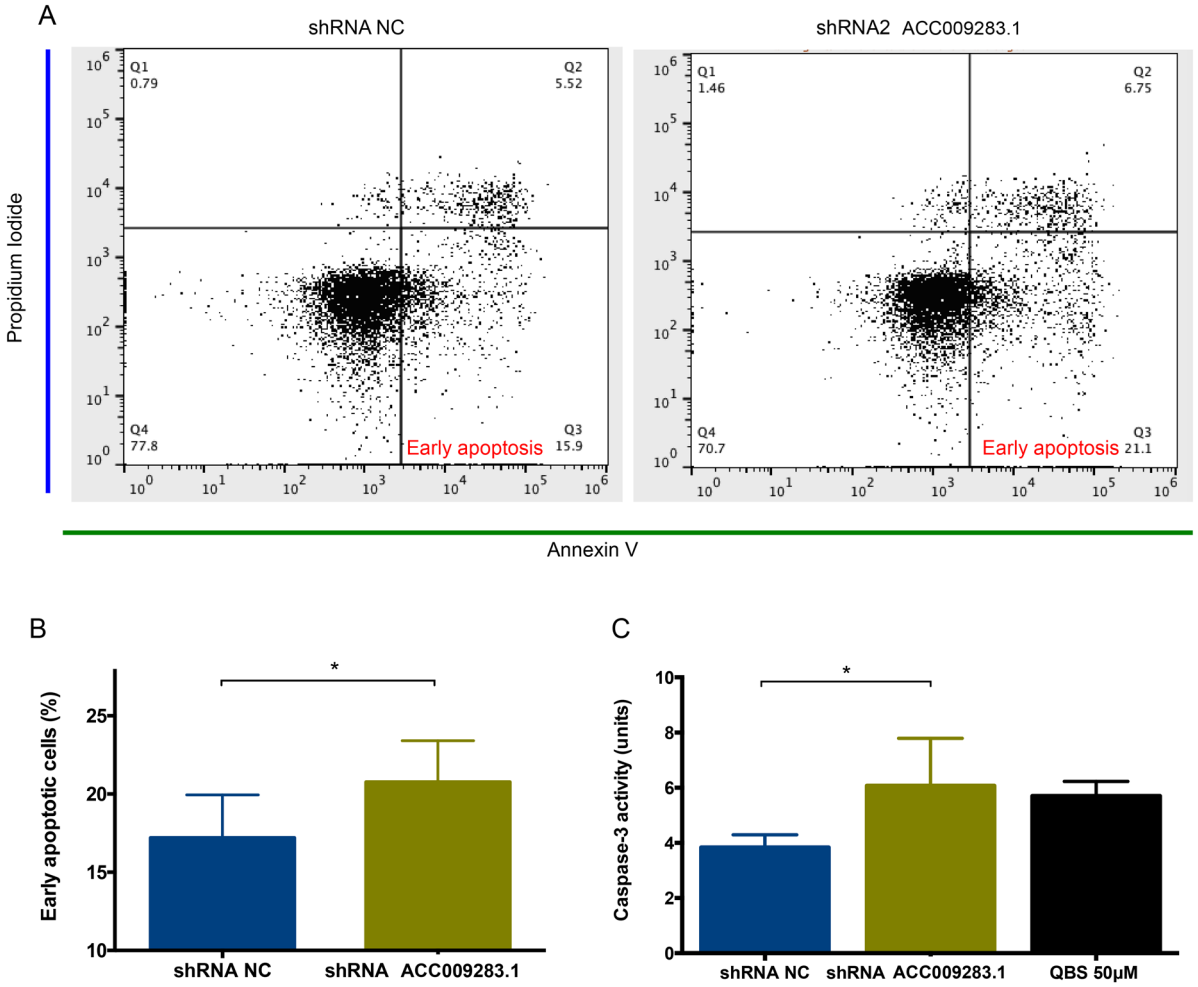

Figure 6. (**A**) Early apoptosis cells were observed to be significantly different between shRNA NC and shRNA2 AC009283.1 using flow cytometry assay on SKBR-3 cells. (**B**) Analysis that represents the percentage of cells in early apoptosis in three biological assays on SKBR-3 cells. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, compared with shRNA NC for each condition via Student's *t* test. (**C**) Quantitative analysis of caspase-3 activity of cells treated with shRNA NC and shRNA2 AC009283.1, units are defined as the amount of enzyme that cleaves 1 nm colorimetric substrate/h. QBS 50uM was used as the positive control for apoptosis. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, compared with shRNA NC for each condition using Student's *t* test.

been reported with a functional role in breast carcinogenesis,however, the biological function of the majority is unknown. Surprisingly, when we validated our gene expression profiles using TCGA data, we found an agreement of 33.8% of the results. The concordance levels between our patients and the TCGA cohort could be influenced by three different variables: (1) sample number disparities (74 for the Mexican cohort and 990 for TCGA), (2) the detection technologies used to obtain gene expression data (microarray for Mexican and RNA-seq for TGCA cohort), and (3) intrinsic biological factors of samples, such as normal/tumor purity percentage. However, the overlapped lncRNAs may have a relevant biological role in breast cancer, as they are present in two independent cohorts. In addition, these data suggest that microarrays and RNA-seq may be complementary methods to identify lncRNAs deregulated in BRCA[16,17].

To acquire insight into the potential biological role of the altered lncRNAs, in our cohort and the TCGA data, among the molecular subtypes, a co-expression analysis of lncRNA–mRNA was performed. It is noteworthy that co-expression analysis has been described as a robust approach for inferring the biological function of lncRNAs[18].

In the Luminal A subtype LINC00504, LINC00993, MIR205HG and PCAT18 were found up-regulated in the analyzed cohorts. Their biological function has not been reported, but according to co-expression analysis (lncRNA–mRNA) they are associated with positive regulation of DNA-templated transcription, a process that regulates the frequency, rate and extent of RNA polymerase II transcription, a target for cancer therapy[19]. This observation might also be related to the fact that in Luminal A BRCA cell line (MCF7), estradiol is a potent stimulator of RNA polymerase II transcription[20]. These four lncRNAs up-regulated in Luminal A tumors also regulate microtubule-based movement, a process that results in the movement of organelles through polymerization or depolymerization of microtubules, this process is also a promising target for breast anticancer drugs[21]. Another biological process enriched was ciliary docking and organization. It has been reported that cilia in breast tumors are more frequent in more differentiated tumors such as those in the luminal A subtype[22]. In

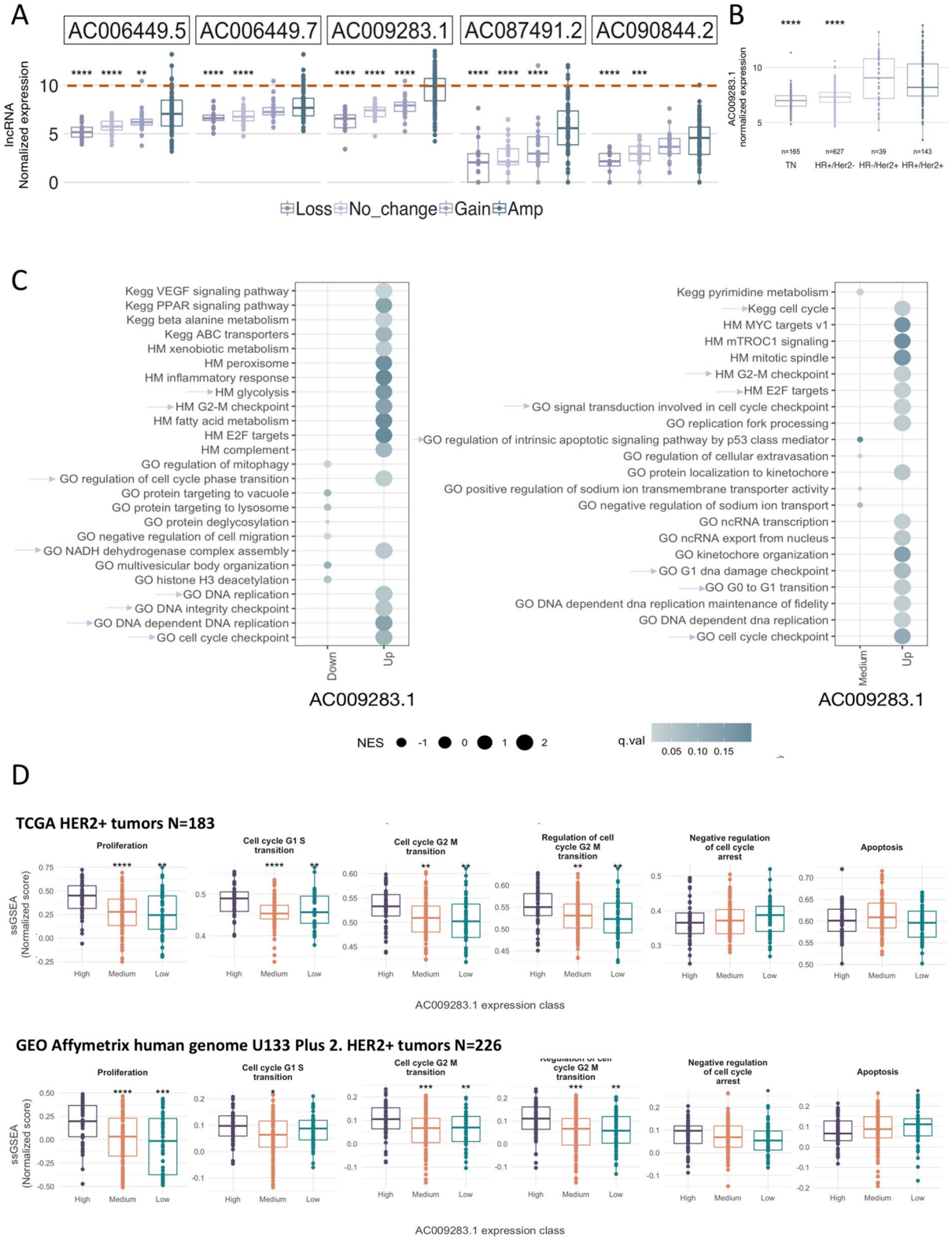

Figure 7. (**A**) Top 5 lncRNAs expressed in 17q12-21.3 vs copy number changes. (**B**) Expression of AC009283.1 in BRCA subtypes including hormone receptor and HER2 receptor status. (**C**) Bubble plot representing GSEA analysis in AC009283.1 high expression breast cancer samples versus AC009283.1 medium and low expression samples. Pink arrows depict biological processes previously shown in in vitro shRNA-mediated knockdown assays. (**D**) ssGSEA analysis for AC009283.1 high, medium and low expression levels, depicting enrichment for cellular proliferation, cell cycle and apoptosis processes in an independent HER2-enriched breast cancer cohort.

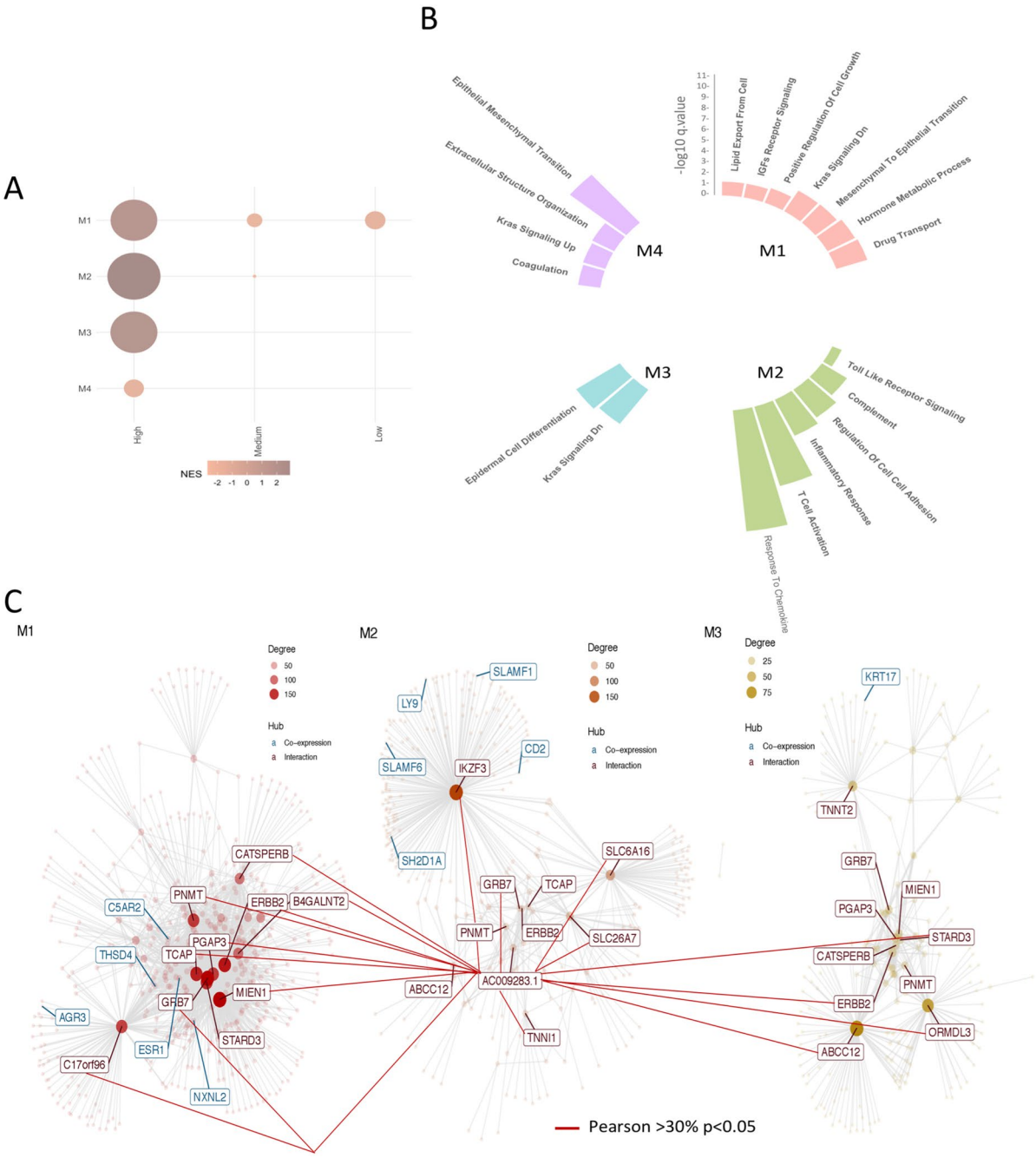

Figure 8. (**A**) Co-expression modules detected by CEMiTool and their enrichment in tumors with high, medium and low expression of AC009283.1. (**B**) Representation of biological processes in each gen module. (**C**) Network graph with the lncRNA in the center and its relationship of co-expression with genes from modules 1, 2 and 3.

addition, it was observed that the promoter of MIR205HG shows a binding site for the estrogen receptor in the Luminal A cell line MCF7, according to Chromatin Immunoprecipitation Sequencing (ChIP-Seq) assays[23], and PCAT18 lncRNA was significantly associated with androgen receptor (AR) signaling in prostate cancer[24]. These data suggest a possible association of these lncRNAs in tumor biology and hormone signaling in this subtype.

In the Luminal B subtype, overexpression of DSCAM-AS1, has been previously reported to mediate tumor progression and tamoxifen resistance. This molecular subtype is the most clinically aggressive ER-positive BRCA, and DSCAM-AS1 is a major discriminator of the luminal subtypes in BRCA[25]. Three other lncRNAs over-expressed in our cohort and TCGA data were LINC0187, RP11-379F12.4 and GATA3-AS1) So far, there are no reports of their biological function in cancer. Co-expression analysis (lncRNA–mRNA) associated them with histone H3K14 acetylation (H3K14ac). Acetylation of specific lysine residues of histones plays a key role in regulating gene expression, in fact, H3K14ac is an active histone mark and correlates with the magnitude of gene expression, probably promoting the expression of oncogenes[26].

In the Basal-like subtype we observed 280 deregulated lncRNAs, importantly, the function of most of them is still unknown. This makes evident the lack of knowledge about the function of the lncRNAs deregulated in the Basal-like tumor phenotype, one of the most aggressive subtypes lacking targeted therapy. For instance, the up-regulation of lncRNAs U62317.1 and U62317.2 (in our cohort and TCGA) is correlated with genes involved in the positive regulation of the JNK cascade, an intracellular protein kinase cascade. Its known that persistent activation of JNKs is involved in cancer development and progression[27]. Another pathway enriched with these two lncRNAs is the activation of cysteine-type endopeptidases involved in the apoptotic process, such as caspases, enzymes known for their role in the initiation and execution of apoptosis[28]. Additionally, these two lncRNAs are associated with positive regulation of NF-κB, deregulated NF-κB pathway leads to the disruption of the balance between cell proliferation and death through the positive regulation of anti-apoptotic proteins[29].

Another four lncRNAs (LINC02095, LINC01436, VIM-AS1, LINC01819 and AFAP1-AS1) were also up-regulated in Basal-like subtype in our cohort and the TCGA cohort. They have an expression correlation with 103 mRNAs in both cohorts, associated with activation of MAPKKK (MAP3K1), which functions in cell survival, apoptosis, and cell migration in multiple tumor cell types[30].

In the HER2-enriched subtype, 8 lncRNAs are up-regulated, of which, only the function of LINC00152 is known, since it regulates cell proliferation, promotes cell cycle arrest at G1 phase, triggers late apoptosis, reduces the epithelial mesenchymal transition program, and suppresses cell migration and invasion in gastric cancer[31]. MIR4435-1HG is up-regulated in HER2-enriched tumors and has been associated with reduced OS and disease-free survival in head and neck cancer[32]. It is noteworthy that the functions of LINC00152 and MIR4435-1HG, mentioned above, have been described in gastric and head and neck cancer models, but not breast cancer.

There are no reports of biological function in cancer of lncRNAs AC116021 and LINC00672, that were found co-expressed with 90 mRNAs in HER2-enriched tumors in our cohort and TCGA, and they are associated with positive regulation of VEGF production. It has been documented that in tumors with overexpressed HER2 the expression and secretion of VEGF is induced, and expression increases when tumors become resistant to treatment[33]. Additionally, negative regulation of anoikis is also an enriched pathway. Anoikis is a hallmark of metastatic cells, and it has been reported that HER2 cells are resistant to this process[34]. None of the up-regulated lncRNAs have been described in HER2-enriched BRCA.

Co-expression analysis of lncRNA–mRNA highlights important functions of specific lncRNAs for each molecular subtype. Some groups of subtype-specific altered lncRNAs are involved in common pathways (RNA polymerase transcription, histone acetylation), while others are more specialized (in the HER2 subtype, those related to VEGF production), explaining the high heterogeneity of the molecular biology of breast tumors. This suggests that the up-regulation of lncRNAs may regulate processes associated with carcinogenesis, which makes them interesting targets for further studies to understand the biology of the molecular subtypes of BRCA.

The HER2-enriched molecular subtype is characterized by the amplification of genes contained in the 17q12 amplification, as well as, the high expression rate of ERBB2 gen and other five cancer promoting genes (*GRB7*, *PNMT*, *STARD3*, *PGAP3*, and *MGC1483*)[35]. HER2-enriched tumors are also one of the subtypes exhibiting the lowest progression free survival[36] and the highest rate of metastasis and lower OS in the Mexican cohort. Therefore, biological characterization of HER2-enriched tumors is a relevant task, especially for deciphering the role of genes located within the HER2 amplification zone.

Few reports have focused on the study of lncRNAs in the HER2-enriched molecular subtype[37]. To date, no one has investigated the functional role of the AC009283.1, a lncRNA up-regulated in HER2-enriched subtype and localized in an amplified region that has been documented to be necessary for the carcinogenic process in this subtype. To better understand the function of the AC009283.1 in the regulation of global gene expression, we examined the effect of its knockdown on the SKBR3 cell transcriptome. To the best of our knowledge, AC009283.1 alterations and biological roles have not previously been reported in breast cancer, thus this would be the first study to describe them.

Microarray analysis revealed that AC009283.1 knockdown regulates 158 genes, some associated with proliferation, cell cycle and apoptosis, as shown by enrichment analysis pathway. This analysis found that AC009283.1 knockdown altered the expression of *NOTCH3*, *TNFa* and *FOSB*, genes that have been previously suggested to be drivers for proliferation and cell cycle in cancer. Up-regulation of *NOTCH3* induces cell cycle arrest at the G0/G1–S phase, and inhibits proliferation and colony-formation in BRCA cell lines[38]. *NOTCH3* was up-regulated by the AC009283.1 knockdown and we observed cycle cell arrest in S phase. *TNFa* gene expression was down-regulated by the knockdown of AC009283.1 and *TNFa* up-regulation has also been shown to increase cellular proliferation[39]. The oncogene FOSB was up-regulated after AC009283.1 knockdown, and *FOSB* has been associated with suppressed cell proliferation in gastric cancer cell lines[40]. In vitro assays (automated cell counting, proliferation index and cycle cell assay by flow cytometry) confirmed that knockdown of AC009283.1 reduced proliferation and decreased S phase in SKBR3 cells, potentially through the modulation of *NOCTH3*, *TNFa* and *FOSB* genes.

Our experimental analysis showed that the expression of *BCL2A1*, *DNML1* and *KLF6*, previously reported to be associated with apoptosis, was affected by knockdown of AC009283.1. BCL2A1, an anti-apoptotic gene, is overexpressed in a variety of cancer cells, including hematological malignancies and solid tumors[41] and *DNM1L* mediates mitochondrial and peroxisomal division and is involved in the regulation of apoptosis. In a previous work, the knockdown of DNM1L expression caused significant increase in apoptosis[42]. Both of these genes were down-regulated after AC009283.1 knockdown. Finally KLF6 is a tumor suppressor that is down-regulated or mutated in several types of cancers, suppressing tumor growth through the activation of p21 and inducing apoptosis in prostate cancer[43]. Knockdown of AC009283.1 causes up-regulation of *KLF6*. Thus, increased apoptosis in SKBR3 cells with knockdown of AC009283.1, may be through the modulation of BCL2A1, DNML1 and KLF6 genes.

Pathway enrichment analysis from microarray analyses and our in vitro assays demonstrated that AC009283.1 knockdown has an anti-proliferative effect and enhances apoptosis in the SKBR3 cell line. Thus, overexpression of AC009283.1 in breast tumors induces apoptosis evasion and sustained proliferation, both processes being hallmarks of cancer[44]. The dual regulation of proliferation and apoptosis that we observe in this work seems to be common for cancer lncRNAs. One example is LSINCT5 that regulates 36 protein-coding genes in breast and ovarian cancer[45], GAS5 in lung cancer[46], and HOXA11-AS in gastric cancer through regulation of miR-1297[47] and UCA1 in colorectal cancer[48]. In this context, our results suggest that AC009283.1 contributes to the malignant phenotype in the HER2-enriched subtype through the interaction with key genes, leading to increased proliferation and resistance to death by apoptosis.

To further validate the functional biological role of AC009283.1 in HER2-enriched tumors driven by HER2 amplification, we performed a GSEA and ssGSEA on two independent breast cancer cohorts. GSEA analysis shows that over-expression of AC009283.1 is associated with over-representation of cellular proliferation and cell cycle progression enrichment pathways. This data was confirmed by ssGSEA. Moreover, cellular processes are driven by multiple molecules that interact with each other. Knowing that genes that participate in the same signaling and that share similar functions will tend to co-express[49], Cemitool co-expression analysis was performed. We found modules significantly enriched in HER2+ tumors overexpressing AC009283.1 related to cell proliferation and metabolism. Some immunological process were also over-represented. Since the HER2 subtype has been described as an immunogenic tumor, a detailed study of this phenotype and the establishment of its relation with AC009283.1 should be included in future studies[50].

Data resulting from diverse computational strategies suggest that AC009283.1 is a potential driver for the malignant phenotype in HER2 enriched/amplified tumors, confirming our results from the in-vitro knockdown assays. The molecular mechanism by which AC009283.1 regulates the expression of target genes involved in proliferation, cell cycle and apoptosis remains unknown. Previous studies have reported that a large part of lncRNAs regulate transcription through chromatin interaction and modulation[9]. We found the AC009283.1 is enriched in the cell nucleus so it is likely that its functions take place in the chromatin. Therefore, further studies are needed in order to understand the molecular mechanism of AC009283.1 in HER2-enriched BRCA.

Conclusions

We found that the expression of a significant number of lncRNAs is deregulated in BRCA. High expression of AC009283.1 may be causally associated with carcinogenesis and we suggest that it plays a potential role in HER2-enriched breast cancer. Additional studies to clarify the regulatory mechanisms of lncRNAs in BRCA will improve our understanding of their contribution to the tumor phenotype.

Methods

Breast tissue collection. Human BRCA specimens (n = 74) (Luminal A = 24, Luminal B = 23, HER2 = 14, Basal = 13) and adjacent breast tissue (n = 12) were obtained from the Institute of Breast Diseases, FUCAM, between 2008–2015. None of the patients received neoadjuvant therapy. The adjacent tissues were collected 2 cm from the tumor margin. All tissues were frozen in liquid nitrogen and stored at − 80 °C until further use. Histological evaluation by two trained pathologists confirmed diagnosis and only samples with more than 80% of tumor cells were included in the analysis. This study was approved by the Research and Ethics Committee of National Institute of Genomic Medicine and the Institute of Breast Diseases, FUCAM (CE2009/11). Research was performed in accordance with relevant guidelines and regulations. Written informed consent was obtained from each patient before any procedure.

Overall survival (OS) of the BRCA TCGA and METABRIC patients was analyzed with the multivariable Cox regression model. Relative Risk was also calculated. This analysis was performed using the PASW statistics software (SPSS, IBM). Disease-free survival (DFS) of our Mexican cohort was analyzed with the Kaplan–Meier model. For all statistical tests, the level of significance was < 0.05.

Cell lines. Human BRCA cell lines included in the analyses (MCF7, ZR-75-1, MDA-MB-361, SKBR3, MDA-MB-468, HCC1187, HS578T, MDA-MB-231, MDA-MB-453) and non-tumorigenic epithelial cell line MCF10A, were purchased from American Type Culture Collection (ATCC, Manassas, VA, USA). Cell lines were grown according to ATCC guidelines, supplemented with 10% fetal bovine serum (FBS) (ATCC) in a humidified incubator at 37 °C with 5% CO_2.

RNA extraction. Frozen tissues were disrupted using a TissueRuptor (Qiagen Inc, Valencia, CA) and RNA extraction was performed with the DNA-RNA AllPrep System (Qiagen Inc, Valencia, CA). RNA from cell lines was extracted by Trizol (Invitrogen). RNA concentration was evaluated by spectrophotometry (NanoDrop Technologies, Wilmington, Delaware), and RNA integrity was analyzed using the BioAnalyzer 2100 (Agilent Technologies, Palo Alto, CA). Only samples with a RNA integrity number (RIN) above 8 were used for microarray analysis. Total RNA was stored at − 80 °C until processing.

Quantitative reverse transcription PCR (RT-qPCR). cDNA was synthetized using SuperScript III RT-PCR (Invitrogen) following the manufacturer's recommendations. Briefly, 100 nanograms (ng) of total RNA from cell lines were used to synthesize cDNA in a final reaction volume of 20 µL. The PCR mixture contained 1 µL of cDNA, 5 µL 2× TaqMan Universal Master Mix (Applied Biosystems, Cat 4304437), 0.5 µL TaqMan probes (Hs_01590788 custom from AC009283.1) and 3.5 µL of nuclease-free water. GAPDH (Hs99999905) and SCARNA (Hs03391742_cn) were used as endogenous controls.

Microarray expression. Human Transcriptome Array 2.0 (Affymetrix, Inc, Santa Clara, CA) was used to measure the mRNA and lncRNA expression following the manufacturer's recommendations. 200 ng of total RNA were processed with the WT Plus Reagent Kit protocol and Affymetrix hybridization kit.

Gene expression profiles. Transcriptome data from Affymetrix HTA2 data was analyzed using Affymetrix's Expression and Transcriptome Analysis Consoles. According to the manufacturer, the array platform contains 44,699 protein-coding genes and 22,829 non-protein coding genes, also called Transcript Cluster ID (TCID), however, only 62% of all TCID (41,572) is annotated, the remaining 38% of TCID (25,956) is unannotated. In order to have a high-quality reference gene annotation, the TCIDs in the Affymetrix array were re-annotated with the BioMart-Ensembl tool. In this manner, we identified 10,153 lncRNAs that were included in the microarray HTA 2.0 (Supplementary data 1.4).

Gene expression data from our Mexican cohort was processed to determine BRCA molecular subtypes according to the PAM50 expression profile[3]. Initially, tumor samples were compared to normal adjacent tissues, and secondly each molecular subtype was compared to the other subtypes, using an unpaired one-way ANOVA test. Transcripts were considered to be differentially expressed when fold change was > 2.0, ANOVA p-value < 0.05, and false discovery rate (FDR) < 0.05.

To validate our findings regarding the highly expressed lncRNAs in breast tumors from the Mexican cases, we used TCGA level 3 data and selected lncRNAs with potential clinical significance in breast tumors through TANRIC. In the TCGA cohort, for the validation of lncRNA expression in tumors relative to normal adjective tissue, we used RNAseq data from The Cancer lncRNome Atlas (TCLA)[51]. Expression data from The Cancer Genome Atlas in the Atlas of Non-Coding RNA (TANRIC) (https://ibl.mdanderson.org/tanric/_design/basic/download.html) was used as an independent validation dataset for the differential expression of lncRNAs across molecular subtypes.

Co-expression analysis. To identify co-expressed lncRNA–mRNA pairs, Pearson correlation coefficients were calculated based on the expression value between every differentially expressed lncRNA and mRNA pair using LncSubpathway for the Mexico data and the TANRIC tool for the TCGA data. The threshold of Pearson correlation coefficients was set to > 0.3 from TCGA data and > 0.4 from Mexico data, and the corresponding p-value was set to < 0.05. The Gene ontology (GO) analyses for co-expression sets were completed using the ClueGO app in Cytoscape, using mRNA that correlates with lncRNAs in both cohorts.

Breast cancer TCGA data sets for HER2 analysis. The HT-Seq raw counts from gene expression RNAseq were downloaded from Xena data base (https://xenabrowser.net/datapages/). Data was annotated with biomaRt[52], Gencode v33 (https://www.gencodegenes.org/human/) and LNCipedia[53]. All transcripts annotated as processed transcript, lncRNA, lincRNA, antisense, nc RNA, ncRNA intronic, sense intronic, sense overlapping and prime overlapping ncrna were considered as lncRNAs. Segmentation file from Affymetrix Genome-Wide Human SNP Array 6.0 was also retrieved from the Xena data base (copy number segments—after removing germline cnv). Additionally, clinical information was downloaded. HER2+ tumors, based on immunochemistry (IHC) evaluations, validated when available with CISH HER2 information, were selected for further analysis. GEO data sets processed with the Affymetrix Human Genome U133 Plus 2.0 Array included in this study: GSE20711, GSE29431 ,GSE5460, GSE17907, GSE29044, and GSE66305 only HER2+ tumors by IHC.

RNA-seq expression profiles. Low expressed (< 10 counts) raw count data were filtered. Then, normalization and differential expression profiles were computed using DESeq2 on DESeq2 package on R[53], independently analyzing coding and lncRNA transcripts.

Generation of somatic copy number alteration profiles. To determine significantly recurrent regions on segmented copy number data from HER2+ (IHC) tumors, GISTIC (Genomic Identification of Significant Targets in Cancer) 2.0 algorithm[54,55] was computed on a Genepattern server (https://cloud.genepattern.org/gp/pages/index.jsf) with the following parameters: deletion threshold = 0.1, cap values = 1.5, broad length cutoff = 0.7, remove X-Chromosome = 0, confidence Level = 0.90, Join Segment Size = 4, arm level Peel Off = 1, Maximum sample segments = 2000 and Gene GISTIC = 1.

Neighboring coding and noncoding transcripts within the 17q-21 amplicon. Genes contained (core genes) in each significant identified wide peak (q < 0.05) on 17q-21 amplicon by GISTIC analysis, were determined by Genome browser table tool, using Gencode v32. Then, a set of lncRNA–mRNA and mRNA–mRNA pair neighboring core genes were assembled using the GenomicRanges package[56]. Two categories were defined based on the proximity to the core genes: (1) close gene: the nearest neighbor genes and (2) distant genes: the nearest neighbor gene of the defined close genes. Genes were considered if they were expressed in at least 50% of the evaluated samples.

Co-expression module identification and building of consensus network. HER2+ tumors with different AC009283.1 expression levels were classified as: low expressing tumors (AC009283.1 levels < 1st quantile), medium (> 1st quantile and < 3rd quantile) and high (> 3rd quantile). DeSeq2 normalized counts of coding genes and AC009283.1, belonging to the above described classes were submitted to CEMiTool[57] on R environment, using default parameters (Pearson correlation coefficient and unbiased selection of genes using a variance-based filter). The Pearson correlation coefficient between each pair of genes in the evaluated tumor expression

profiles was computed. Genes from each module were connected through a network. Genes significantly correlated (Pearson > 30%, $p < 0.05$) with AC009283.1 were annotated in each network. CEMiTool GSEA was carried out by using the curated gene sets[58] (C1-Hallmark gene sets, C2-Kegg curated genes, and C5GO-BP gene sets) of the Molecular Signature Database (MSigDB) (https://www.broad.mit.edu/gsea/).

Implementation of GSEA and ssGSEA. A non-preranked GSEA[59] was performed by applying the C1-Hallmark gene sets[58,] C2-Kegg curated genes and C5GO-BP gene sets of the MSigDB using the R package GSVA[60]. FDR < 0.25 is considered to be statistically significant for GSEA. Individual scores for each tumor sample were computed by the ssGSEA method[61] implemented in GSVA R package using the following parameters: min gene set size of 5 and normalization. The evaluated gene sets were: proliferation UNC337 intrinsic clustering, hallmark apoptosis, GO regulation of cell cycle g2 m phase transition, GO negative regulation of cell cycle arrest and GO cell cycle g2 m phase transition. ssGSEA scores were then plotted by ggplot library. DeSeq2 normalized RNA-Seq data from coding genes were employed to compute GSEA and ssGSEA scores.

Subcellular fractionation. Cellular fractionation (cell nucleus and cytoplasm) was performed using Ambion PARIS Kit (Thermo Fisher, Cat AM1921) according to manufacturer's instructions. RNA was isolated and qRT-PCR was performed as described above.

AC009283.1 knockdown. Short hairping RNAs (shRNAs) specifically targeting AC009283.1 were designed using BLOCK-iT U6 RNAi Entry Vector Kit (Invitrogen, Cat K494500) following the manufacturer's instructions. We generated the double-stranded oligo (dsoligo) and subsequently performed the ligation reaction of ds oligo into the pENTR /U6 vector, which was used to transform competent *E. coli* One Shot TOP10. We used Sanger sequencing to corroborate the presence and correct orientation of the dsoligo on the vector.

Two designed shRNA were used to knockdown AC009283.1 (shRNA1 and shRNA2). shRNA1: top 5′CACCGAGGAAGTAGGTTAGGGAATGCGAACATTCCCTAACCTACTTCCTC3′, bottom 5′AAAAGAGGAAGTAGGTTAGGGAATGTTCGCATTCCCTAACCTACTTCCTC3′ and shRNA2: top 5′CACCGCAAATAGGTGTCTCATAGCTCGAAAGCTATGAGACACCTATTTGC3′ and bottom 5′AAAAGCAAATAGGTGTCTCATAGCTTTCGAGCTATGAGACACCTATTTGC3′. negative control (NC): top: 5′CACCGGAATTACGGAGTCTTCTTCGCGAACGAAGAAGACTCCGTAATTCC3′ and bottom: 5′ AAAAGGAATTACGGAGTCTTCTTCGTTCGCGAAGAAGACTCCGTAATTCC3′ with a sequence that does not target mRNAs in the human genome (Green Fluorescence Protein, GFP).

Knockdown of AC009283.1 in HER2-enriched cell line (SKBR3 ATCC HTB-30) was accomplished using two shRNAs (shRNA-1 and shRNA-2), which were transfected using Xfect Transfection Reagent Protocol-At-A-Glance (Clontech, Palo Alto, CA, Cat 631318). Briefly, 3.5×10^5 SKBR3 cells were seeded in 6-well plates and 3 µg of plasmid were transfected during 48 h. Nanoparticle complexes were removed and replaced with complete growth medium McCoy5A (ATCC). 96 h post-transfection, expression of AC009283.1 was evaluated by RT-qPCR. The RNA was extracted by TRIzol (Thermo Scientific, Cat 15596026) and stored at − 80 °C. Independent experiments were performed in triplicate.

The effect of AC009283.1 knockdown on the transcriptional landscape of SKBR3 cell was analyzed with the HTA 2.0 microarray (Affymetrix, Santa Clara, USA) as described above, we performed three microarrays for shRNA NC and three for shRNA 2 (AC009283.1 knockdown). Genes with a fold change of > 1.2 and < − 1.2, p-value < 0.05 were considered as significant and selected for biological pathway analysis using Ingenuity Pathway Analysis (IPA). To explore the effect of the AC009283.1 in tumor samples, we performed a differential expression analysis of HER2-enriched tumor samples with high and low expression of AC009283.1 (first and fourth quantile). Count tables of RNASeq were obtained from TCGA data portal and subsequently differential expression was performed with DESeq2. We obtained a list of differentially expressed genes and used it for biological pathway analysis with IPA.

Cellular proliferation assay. Cell proliferation of AC009283.1 knockdown cells was analyzed using two methods: (1) counting the total number of positive cells stained with trypan blue in a TC20 Automated Cell Counter (BIO-RAD), and (2) detecting proliferation rate using Cell Trace CFSE Cell Proliferation Kit (Invitrogen, Cat C34554), according to the manufacturer's protocol, using flow cytometry (Attune, Applied Biosystems). 7.7×10^4 SKBR3 cells were seeded in 24-well plates and were transfected with shRNA or NC shRNA. Cell proliferation assays were carried out at 24, 48, 72 and 96 h. At least three independent experiments were performed for each assay.

Apoptosis assay. 7.7×10^4 SKBR3 cells were seeded in 24-well plates, and incubated with shRNA for 96 h. Cells were harvested by trypsinization and washed with PBS, cells were treated with QBS 20 µM for 24 h to perform the apoptosis assay. Then, the cells were re-suspended in binding buffer and stained with Annexin V and PI (FITC Annexin V/Dead cell Apoptosis kit, Invitrogen, Carlsbad, CA, USA) for 15 min in the dark at room temperature. The stained cells were examined by flow cytometry (Attune, Applied Biosystems). The cells were categorized into early and late apoptotic cells. At least three independent experiments were performed. We conducted a second validation test of apoptosis using a caspase-3 activity assay, performed following the manufacturer's instructions (Merck #235419). We used QBS 50 µM as positive control of apoptosis and FlowJo T v10.0 software for analysis.

Statistical analysis. Statistical significance was estimated using GraphPadPrism (version 6, San Diego, USA). ANOVA and Student's *t* test were performed for all comparisons involving categorical variables. Correlation between variables was determined by Pearson's correlation coefficient. The Kaplan–Meier method was assessed by log-rank test and COX. p value < 0.05 was considered as significant ($^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.0005$).

Ethical approval and informed consent. This study was approved by the Research and Ethics Committee of the National Institute of Genomic Medicine and the Institute of Breast Diseases, FUCAM (CE2009/11). Written informed consent was obtained from each patient before any procedure.

Data availability

To review GEO accession GSE134254: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134254. Enter token enopkmqerzmvzut into the box. To review GEO accession GSE134359: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134359 Enter token ehofieiuhpkbdkp into the box.

Received: 26 July 2019; Accepted: 19 July 2020
Published online: 04 August 2020

References

1. Bray, F. G. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*. https://doi.org/10.3322/caac.21492 (2018).
2. Allred, D. C. Issues and updates: evaluating estrogen receptor-alpha, progesterone receptor, and HER2 in breast cancer. *Mod. Pathol.* **23**(Suppl 2), S52–S59. https://doi.org/10.1038/modpathol.2010.55 (2010).
3. Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167. https://doi.org/10.1200/JCO.2008.18.1370 (2009).
4. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752. https://doi.org/10.1038/35021093 (2000).
5. Haque, R. *et al.* Impact of breast cancer subtypes and treatment on survival: an analysis spanning two decades. *Cancer Epidemiol. Biomark. Prev.* **21**, 1848–1855. https://doi.org/10.1158/1055-9965.EPI-12-0474 (2012).
6. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70. https://doi.org/10.1038/nature11412 (2012).
7. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773. https://doi.org/10.1093/nar/gky955 (2019).
8. Gibb, E. A., Brown, C. J. & Lam, W. L. The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* **10**, 38. https://doi.org/10.1186/1476-4598-10-38 (2011).
9. Liu, Y., Sharma, S. & Watabe, K. Roles of lncRNA in breast cancer. *Front. Biosci. (Schol Ed)* **7**, 94–108 (2015).
10. Ye, N. *et al.* Functional roles of long non-coding RNA in human breast cancer. *Asian Pac. J. Cancer Prev.* **15**, 5993–5997 (2014).
11. Zhong, B. L. *et al.* Identification of key genes involved in HER2-positive breast cancer. *Eur. Rev. Med. Pharmacol. Sci.* **20**, 664–672 (2016).
12. Liu, Y. R. *et al.* Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. *Breast Cancer Res.* **18**, 33. https://doi.org/10.1186/s13058-016-0690-8 (2016).
13. Zhao, W., Luo, J. & Jiao, S. Comprehensive characterization of cancer subtype associated long non-coding RNAs and their clinical implications. *Sci. Rep.* **4**, 6591. https://doi.org/10.1038/srep06591 (2014).
14. Su, X. *et al.* Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. *Oncotarget* **5**, 9864–9876. https://doi.org/10.18632/oncotarget.2454 (2014).
15. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789. https://doi.org/10.1101/gr.132159.111 (2012).
16. Guo, Y. *et al.* Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. *PLoS ONE* **8**, e71462. https://doi.org/10.1371/journal.pone.0071462 (2013).
17. Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509–1517. https://doi.org/10.1101/gr.079558.108 (2008).
18. Guo, Q. *et al.* Comprehensive analysis of lncRNA–mRNA co-expression patterns identifies immune-associated lncRNA biomarkers in ovarian cancer malignant progression. *Sci. Rep.* **5**, 17683. https://doi.org/10.1038/srep17683 (2015).
19. Villicana, C., Cruz, G. & Zurita, M. The basal transcription machinery as a target for cancer therapy. *Cancer Cell Int.* **14**, 18. https://doi.org/10.1186/1475-2867-14-18 (2014).
20. Wu, J., Cai, Y. & Zhao, G. Bioinformatic analysis of changes in RNA polymerase II transcription stimulated by estradiol in MCF7 cells. *Neoplasma* **65**, 14–20. https://doi.org/10.4149/neo_2018_161214N637 (2018).
21. Higa, G. M. The microtubule as a breast cancer target. *Breast Cancer* **18**, 103–119. https://doi.org/10.1007/s12282-010-0224-7 (2011).
22. Menzl, I. *et al.* Loss of primary cilia occurs early in breast cancer development. *Cilia* **3**, 7. https://doi.org/10.1186/2046-2530-3-7 (2014).
23. Niknafs, Y. S. *et al.* The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression. *Nat. Commun.* **7**, 12791. https://doi.org/10.1038/ncomms12791 (2016).
24. Crea, F. *et al.* Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer. *Oncotarget* **5**, 764–774. https://doi.org/10.18632/oncotarget.1769 (2014).
25. Ma, Y., Bu, D., Long, J., Chai, W. & Dong, J. LncRNA DSCAM-AS1 acts as a sponge of miR-137 to enhance Tamoxifen resistance in breast cancer. *J. Cell. Physiol.* https://doi.org/10.1002/jcp.27105 (2018).
26. Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H. & Tora, L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**, 424. https://doi.org/10.1186/1471-2164-13-424 (2012).
27. Bubici, C. & Papa, S. JNK signalling in cancer: in need of new, smarter therapeutic targets. *Br. J. Pharmacol.* **171**, 24–37. https://doi.org/10.1111/bph.12432 (2014).
28. Krepela, E. Cysteine proteinases in tumor cell growth and apoptosis. *Neoplasma* **48**, 332–349 (2001).
29. Biswas, D. K. *et al.* NF-kappa B activation in human breast cancer specimens and its role in cell proliferation and apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10137–10142. https://doi.org/10.1073/pnas.0403621101 (2004).
30. Pham, T. T., Angus, S. P. & Johnson, G. L. MAP3K1: genomic alterations in cancer and function in promoting cell survival or apoptosis. *Genes Cancer* **4**, 419–426. https://doi.org/10.1177/1947601913513950 (2013).

31. Zhao, J. *et al.* Long non-coding RNA Linc00152 is involved in cell cycle arrest, apoptosis, epithelial to mesenchymal transition, cell migration and invasion in gastric cancer. *Cell Cycle* **14**, 3112–3123. https://doi.org/10.1080/15384101.2015.1078034 (2015).
32. Nohata, N., Abba, M. C. & Gutkind, J. S. Unraveling the oral cancer lncRNAome: identification of novel lncRNAs associated with malignant progression and HPV infection. *Oral Oncol.* **59**, 58–66. https://doi.org/10.1016/j.oraloncology.2016.05.014 (2016).
33. Schoppmann, S. F. *et al.* HER2/neu expression correlates with vascular endothelial growth factor-C and lymphangiogenesis in lymph node-positive breast cancer. *Ann. Oncol.* **21**, 955–960. https://doi.org/10.1093/annonc/mdp532 (2010).
34. Hou, J. *et al.* HER2 reduces breast cancer radiosensitivity by activating focal adhesion kinase in vitro and in vivo. *Oncotarget* **7**, 45186–45198. https://doi.org/10.18632/oncotarget.9870 (2016).
35. Hongisto, V., Aure, M. R., Makela, R. & Sahlberg, K. K. The HER2 amplicon includes several genes required for the growth and survival of HER2 positive breast cancer cells—a data description. *Genom Data* **2**, 249–253. https://doi.org/10.1016/j.gdata.2014.06.025 (2014).
36. Sontrop, H. M. J., Reinders, M. J. T. & Moerland, P. D. Breast cancer subtype predictors revisited: from consensus to concordance?. *BMC Med. Genomics* **9**, 26. https://doi.org/10.1186/s12920-016-0185-6 (2016).
37. Yang, F. *et al.* Expression profile analysis of long noncoding RNA in HER-2-enriched subtype breast cancer by next-generation sequencing and bioinformatics. *Onco Targets Ther* **9**, 761–772. https://doi.org/10.2147/OTT.S97664 (2016).
38. Chen, C. F. *et al.* Notch3 overexpression causes arrest of cell cycle progression by inducing Cdh1 expression in human breast cancer cells. *Cell Cycle* **15**, 432–440. https://doi.org/10.1080/15384101.2015.1127474 (2016).
39. Wang, X. H. *et al.* Tumor necrosis factor alpha promotes the proliferation of human nucleus pulposus cells via nuclear factor-kappaB, c-Jun N-terminal kinase, and p38 mitogen-activated protein kinase. *Exp. Biol. Med. (Maywood)* **240**, 411–417. https://doi.org/10.1177/1535370214554533 (2015).
40. Tang, C. *et al.* Abnormal expression of FOSB correlates with tumor progression and poor survival in patients with gastric cancer. *Int. J. Oncol.* **49**, 1489–1496. https://doi.org/10.3892/ijo.2016.3661 (2016).
41. Vogler, M. BCL2A1: the underdog in the BCL2 family. *Cell Death Differ.* **19**, 67–74. https://doi.org/10.1038/cdd.2011.158 (2012).
42. Xu, X. W. *et al.* DNM1L, a key prognostic predictor for gastric adenocarcinoma, is involved in cell proliferation, invasion, and apoptosis. *Oncol Lett.* **16**, 3635–3641. https://doi.org/10.3892/ol.2018.9138 (2018).
43. Huang, X., Li, X. & Guo, B. KLF6 induces apoptosis in prostate cancer cells through up-regulation of ATF3. *J. Biol. Chem.* **283**, 29795–29801. https://doi.org/10.1074/jbc.M802515200 (2008).
44. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674. https://doi.org/10.1016/j.cell.2011.02.013 (2011).
45. Silva, J. M., Boczek, N. J., Berres, M. W., Ma, X. & Smith, D. I. LSINCT5 is over expressed in breast and ovarian cancer and affects cellular proliferation. *RNA Biol.* **8**, 496–505. https://doi.org/10.4161/rna.8.3.14800 (2011).
46. Shi, X. *et al.* A critical role for the long non-coding RNA GAS5 in proliferation and apoptosis in non-small-cell lung cancer. *Mol. Carcinog.* **54**(Suppl 1), E1–E12. https://doi.org/10.1002/mc.22120 (2015).
47. Yu, W., Peng, W., Jiang, H., Sha, H. & Li, J. LncRNA HOXA11-AS promotes proliferation and invasion by targeting miR-124 in human non-small cell lung cancer cells. *Tumour Biol.* **39**, 1010428317721440. https://doi.org/10.1177/1010428317721440 (2017).
48. Han, Y. *et al.* UCA1, a long non-coding RNA up-regulated in colorectal cancer influences cell proliferation, apoptosis and cell cycle distribution. *Pathology* **46**, 396–401. https://doi.org/10.1097/PAT.0000000000000125 (2014).
49. Wang, J. *et al.* Single-cell co-expression analysis reveals distinct functional modules, co-regulation mechanisms and clinical outcomes. *PLoS Comput. Biol.* **12**, e1004892. https://doi.org/10.1371/journal.pcbi.1004892 (2016).
50. Griguolo, G., Pascual, T., Dieci, M. V., Guarneri, V. & Prat, A. Interaction of host immunity with HER2-targeted treatment and tumor heterogeneity in HER2-positive breast cancer. *J. Immunother. Cancer* **7**, 90. https://doi.org/10.1186/s40425-019-0548-6 (2019).
51. Yan, X. *et al.* Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **28**, 529–540. https://doi.org/10.1016/j.ccell.2015.09.006 (2015).
52. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191. https://doi.org/10.1038/nprot.2009.97 (2009).
53. Volders, P. J. *et al.* LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* **47**, D135–D139. https://doi.org/10.1093/nar/gky1031 (2019).
54. Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 20007–20012. https://doi.org/10.1073/pnas.0710052104 (2007).
55. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905. https://doi.org/10.1038/nature08822 (2010).
56. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118. https://doi.org/10.1371/journal.pcbi.1003118 (2013).
57. Russo, P. S. T. *et al.* CEMiTool: a bioconductor package for performing comprehensive modular co-expression analyses. *BMC Bioinform.* **19**, 56. https://doi.org/10.1186/s12859-018-2053-1 (2018).
58. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425. https://doi.org/10.1016/j.cels.2015.12.004 (2015).
59. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550. https://doi.org/10.1073/pnas.0506580102 (2005).
60. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7. https://doi.org/10.1186/1471-2105-14-7 (2013).
61. Barbie, D. A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112. https://doi.org/10.1038/nature08460 (2009).

Acknowledgements

Alberto Cedro Tanda is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), and received fellowship (CVU 440707) 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 and M.Sc. Nelly Patiño of the cytometry unit and PhD Elizabeth Langley McCarron for her English proofread. 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

C.T.A. carried out most of the experiments, data analysis and wrote the manuscript. B.P.V.: histopathological review of the cases. D.R.C., V.C.F., T.T.A.: patient identification and clinical follow up. R.V.R., A.R.L.: sample collection and processing. C.T.A., R.R.M., C.V.M. discussed the design of the experiments and the results. R.C.S.:

provided transcriptomic analysis. R.C.S., R.R.M., C.V.M., J.M.S., B.F.O., A.H.M.: discussion and edit the paper. H.M.A.: project coordination and leadership. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-69905-z.

Correspondence and requests for materials should be addressed to A.H.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5. DISCUSIÓN

Los RNAs largos no codificantes ejercen numerosas funciones en los cánceres humanos, ya que sus actividades biológicas implican la regulación de la proliferación, la muerte, la diferenciación, la migración, la invasión y la metástasis de las células. La desregulación de la expresión de los lncRNAs también se ha asociado con el resultado clínico de los pacientes con cáncer. Los lncRNAs pueden afectar a la expresión de miles de genes, por lo que se consideran reguladores celulares maestros.

En este trabajo, el objetivo fue examinar si la expresión de los lncRNAs FAM83H- AS1 y LINC00460 se encuentra desregulada en diferentes tumores y si está asociada con las características clínicas y patológicas de estos cánceres. Como objetivos particularesse planteó identificar el papel biológico de lncRNAs FAM83H-AS1 y LINC00460 en CaMa, así como sus potenciales genes blanco y la caracterización funcional de éstos. En la siguiente sección se discutirán los hallazgos relacionados a FAM83H-AS1 y posteriormente, se discutirán los resultados asociados a LINC00460.

5.1. FAM83H-AS1

FAM83H-AS1 está sobreexpresado en nueve tipos de tumores diferentes en la base de datos TCGA. En particular, FAM83H-AS1 está sobreexpresado y significativamente correlacionado con un peor resultado clínico en los subtipos CaMa positivos (detectados por IHC), en la cohorte de CaMa del TCGA. Un estudio anterior (Yang *et al*, 2016) muestra que la alta expresión de FAM83H-AS1 se asocia a un pronóstico desfavorable en el CaMa luminal y es un marcador pronóstico independiente. Nuestro estudio es el primero que sugiere una interacción de variables entre FAM83H-AS1, PR y ER (detectados por IHC) en el contexto del pronóstico clínico. Además, este trabajo demostró que los niveles de expresión de RE y RP pueden actuar como potenciadores de mal pronóstico de OS de FAM83H-AS1. En particular, los niveles de expresión elevados de RE y RP, junto con la sobreexpresión de FAM83H-AS1, confieren un riesgo muy elevado (HR=57) de fallecimiento en las pacientes con CaMa. Estos datos sugieren un importante papel clínico de FAM83H-AS1 en el CaMa ER/PR positivo. Sin embargo, en la actualidad se desconoce

si estas interacciones estadísticas y clínicas se reflejan a nivel biológico o molecular, por lo que futuros estudios deberán abordar esta materia.

También encontramos una correlación significativa entre la alta expresión de FAM83H-AS1 y la mala respuesta al tamoxifeno en las pacientes con CaMa. Esta asociación podría explicar parcialmente la menor respuesta clínica en el grupo de alta expresión de FAM83H-AS1 en comparación con los casos que lo subexpresaron .

La sobreexpresión de FAM83H-AS1 en las muestras de CaMa depositadas en el TCGA está asociada a la regulación a la baja de transcritos relacionados con la migración y la muerte celular, como FGF4, FGF21, LEP, CLDN17, TP53, BAX y TNFRSF11B (Ríos-Romero *et al*., 2020. Ver tabla 2). En consecuencia, también se logró descubrir que el silenciamiento de FAM83H-AS1 desregula significativamente la migración y los genes relacionados con la apoptosis, como TP63 y CLND1 (Ríos-Romero *et al*., 2020. Ver tabla 3). Los ensayos de migración mostraron que, efectivamente, la migración celular aumenta tras el silenciamiento de FAM83H-AS1 (Ríos-Romero *et al*., 2020. Ver figura 6). Se ha demostrado que tanto LEP como CLDN1 inducen la migración celular y la transición epitelial a mesenquimal (EMT) (que constituye el primer paso de la migración) en las células de CaMa (Zhao *et al.*, 2015; Wei *et al.*, 2016; Strong *et al.*, 2015; Li *et al.*, 2016; Huang *et al.*, 2017), lo que sugiere además un papel de FAM83H-AS1 en las pri- meras fases de la migración. En conjunto, tanto los datos publicados anteriormente como nuestros resultados podrían explicar los mecanismos subyacentes relacionados con la migración celular de FAM83H-AS1 en las células de CaMa.

FAM83H-AS1 podría desempeñar un papel doble, probablemente debido al contexto celular. FAM83H-AS1 se encuentra implicado en la regulación de los procesos de proliferación, migración e invasión celular, que disminuyeron tras el silenciamiento de FAM83H-AS1 en las células de cáncer de pulmón AS1 (Zhang *et al.*, 2017). Otros análisis del mismo estudio indicaron que el ciclo celular se detiene en la fase G2 tras la eliminación de FAM83H-AS1 (Zhang *et al.*, 2017). Estos resultados contradictorios pueden deberse al contexto celular o a mecanismos de regulación específicos y, por tanto, a blancos moleculares específicos.

También se identificó que la sobreexpresión de FAM83H-AS1 está asociada a la

regulación a la baja de transcritos relacionados con la muerte celular, como BAX, TNFRSF11B y P53 (Ríos-Romero *et al*., 2020. Ver figura 4). Los ensayos *in vitro* también muestran que el silenciamiento de FAM83H-AS1 aumenta la muerte celular, posiblemente mediante la regulación al alza de genes como *p63* (Ríos-Romero *et al*., 2020. Ver figura 5 D y E y tabla 3). Un estudio previo (Lu *et al.*, 2018) demostró que la muerte celular aumenta notablemente tras la supresión de FAM83H-AS1 en líneas celulares de cáncer colorrectal. FAM83H-AS1, Notch1 y Hes1 aumentaron significativamente en las muestras y líneas celulares de cáncer colorrectal. La proliferación celular se inhibió con el silenciamiento de FAM83H-AS1 y este efecto mediado por FAM83H-AS1 pudo ser revertido por los reguladores de Notch1 (Lu *et al.*, 2018).

Sin embargo, actualmente no está claro si FAM83H-AS1 tiene un efecto directo o indirecto en la regulación de los genes. A este respecto, se ha demostrado que FAM83H-AS1 silencia epigenéticamente al gen *CDKN1A* mediante la unión a EZH2 en células de glioma (Bi *et al.*, 2018). En nuestro análisis de expresión diferencial, demostramos que FAM83H-AS1 regula mayoritariamente a la baja la expresión génica. Estos datos podrían sugerir que FAM83H-AS1 desempeña una función de regulación inhibidora. Los estudios futuros deben abordar estos mecanismos, ya que no se ha descartado que FAM83H-AS1 pueda regular la expresión de genes maestros a través del reclutamiento de complejos epigenéticos (por ejemplo, EZH2). Adicionalmente, el papel exacto de FAM83H-AS1 en los genes regulados al alza sigue siendo poco claro. Sigue latente la posibilidad de que este lncRNA juegue un papel sutil y alternativo de en la activación de genes, y futuros estudios deberán abordar esta cuestión. Nuestros resultados también muestran que FAM83H-AS1 está presente tanto en el núcleo como en el citoplasma de las células de CaMa. Es posible que este lncRNA desempeñe un papel diferente en el citoplasma y los estudios futuros deben centrarse en esta cuestión.

5.2. LINC00460

En el caso de LINC00460, el estudio se centró en su papel clínico en el CaMa agresivo del tipo parecido a basal y en la identificación de posibles blancos de LINC00460 en este modelo. El objetivo era saber si LINC00460 puede dirigirse a miRNAs *in-silico*, que a su vez pueden unirse a mRNAs relevantes. Por último, se examinó si algunos de los posibles

genes candidatos tendrían un papel combinatorio para la OS y la respuesta a la terapia en CaMa parecido a basal.

LINC00460 se encuentra desregulado en siete tipos de tumores diferentes en la base de datos TCGA, incluyendo dos tumores no reportados previamente: Glioma de bajo grado y Glioblastoma Multiforme (LGG y GBM) (Cisneros-Villanueva *et al.,* 2021. Ver figuras 1A y S1). También se confirmó la desregulación de LINC00460 en CaMa utilizando dos cohortes independientes (Cisneros-Villanueva *et al.,* 2021. Ver figuras 1B y S2). La sobreexpresión de LINC00460 está asociada a tumores clínicamente agresivos, como adenocarcinoma de pulmón, carcinoma escamoso de pulmón y carcinoma de riñón de células claras (LUAD, LUSC y KIRC, respectivamente, por sus siglas en inglés) en etapas avanzadas, alto grado histológico en carcinomas de cabeza y cuello (KIRC, HNSC) () negativo al Virus de Papiloma Humano (VPH) y CaMa parecido a basal (Cisneros-Villanueva *et al.,* 2021. Ver figuras S3 y S4 del material suplementario), lo cual sugiere un papel importante de LINC00460 en la progresión o agresividad intrínseca de estos tumores. Se ha demostrado previamente que la expresión de LINC00460 puede promover la progresión del cáncer (Liu *et al.*, 2018; Li y Kong, 2019a; Lian *et al.*, 2019; Xie *et al.*, 2019), la metástasis (Zhang *et al.*, 2019) e influye en la respuesta a la terapia (Ma *et al.*, 2019).

Algunos reportes muestran que LINC00460 es un marcador de mal pronóstico para OS en diferentes tumores, como cáncer cérvico uterino de células escamosas y endocervicales (CESC) (Li, Zhu y Wang, 2020), HNSC (Zhang *et al.*, 2020), KIRC (Zhang, Zeng y Hu, 2020), LUAD (Nakano *et al.*, 2019) y adenocarcinoma pancreático (PAAD) (Sun *et al.*, 2019). En este trabajo, se describió que la alta expresión de LINC00460 se asocia significativamente con una mala supervivencia en tres tumores diferentes (GBM, LGG y sarcoma, SARC) pero se relaciona con una tasa de supervivencia favorable en CaMa, es decir, su asociación con el resultado clínico varía entre los tumores (Cisneros-Villanueva *et al.,* 2021). Estos datos sugieren que los niveles de expresión del LINC00460 y su impacto en la OS, RFS y sobrevida libre de metástasis (DMFS) pueden ser específicos de cada tejido. En estudios previos, se ha demostrado que un mismo lncRNA puede ejercer un doble papel pronóstico en distintos tumores. Por ejemplo, se ha reportado de que la expresión elevada de MALAT-1 es un marcador de mal pronóstico en diversos tumores,

como el adenocarcinoma de colon (COAD), adenocarcinoma de estómago (STAD), PAAD y carcinoma de esófago (ESCA), entre otros (Zhu *et al.*, 2015), pero también es un factor de buen pronóstico para el CaMa, ya que actúa como supresor de metástasis (Kim *et al.*, 2018). Otro ejemplo de este fenómeno dual es la expresión de XIST. Se ha demostrado que la alta expresión de este lncRNA está relacionada con un mal pronóstico clínico en diferentes tipos de cáncer (Zhou *et al.*, 2018), pero en otro estudio, los autores demuestran que la alta expresión de XIST está relacionada con una mayor DMFS cerebral en pacientes CaMa (Xing *et al.*, 2018). Tomando en cuenta los resultados de la presente investigación y los antecedentes previamente publicados, sugerimos que LINC00460 puede funcionar como un marcador pronóstico dual específico de tejido, similar a MALAT-1 y XIST; aunque no hay suficiente evidencia para descartar mecanismos alternativos, como por ejemplo las diferencias en las variantes de *splicing* analizadas entre los estudios. Estos patrones de expresión de las transcripciones variantes deben tenerse en cuenta para evaluar los biomarcadores predictivos basados en lncRNAs (Meseure *et al.*, 2016). Esta última posibilidad debe abordarse en futuros estudios.

En el modelo CaMa derivado de la cohorte de TCGA, la expresión de LINC00460 se asocia con la composición fenotípica del tumor, donde la sobre-expresión de LINC00460 se asocia con un resultado negativo a los receptores hormonales RE y RP cuando se evalúa por IHQ. También se observó que, en el fenotipo más agresivo, el CaMa parecido a basal, la expresión de LINC00460 favorece el resultado clínico (Cisneros-Villanueva *et al.,* 2021. Ver figura 6) incluso al ser sometido a análisis en los diferentes subtipos de TNBC. Su comportamiento similar a través de los diferentes subtipos de CaMa y de las diferentes cohortes de CaMa, nos llevó a proponer a LINC00460 como un biomarcador potencial para mejorar la predicción de la OS, la RFS y posiblemente la DMFS en el CaMa parecido a basal (Cisneros-Villanueva *et al.,* 2021. Ver figuras 4, 5 y 6). Este comportamiento puede explicarse potencialmente a través de los genes candidatos coexpresados encontrados en los análisis de expresión diferencial y análisis de sobrerrepresentación (ORA, por sus siglas en inglés) (Cisneros-Villanueva *et al.,* 2021. Ver figura 7). Proponemos cinco mecanismos principales que potencialmente explican el papel de LINC00460 en el buen pronóstico clínico de las pacientes CaMa: (1) regulación/coexpresión de genes relacionados con el buen pronóstico clínico (Cisneros-Villanueva *et al.,* 2021. Ver figura 7 y tablas 2 y 3); (2) coexpresión/modulación de genes que promueven

un nicho inmunogénico (Cisneros-Villanueva *et al.,* 2021. Ver figura 7C y tablas1,2 y 3); (3) disminución de la proliferación de células tumorales (Cisneros-Villanueva *et al.,* 2021. Ver figura 7B); (4) regulación de WNT7A a través de la captura de miR-103-a-1 (Cisneros-Villanueva *et al.,* 2021. Ver figuras 8A, 8B y tablas 4,5,6) y (5) promoción de la respuesta patológica completa a la quimioterapia (Antraciclinas) (Cisneros-Villanueva *et al.,* 2021. Ver figura 8C). Dado que la mayoría de estas interacciones fueron predichas *in-silico*, o se asocian con características clínicas en cohortes de pacientes, se necesita una validación experimental futura para apoyar estas hipótesis.

En este sentido, observamos que la expresión de LINC00460 está significativamente enriquecida en el subtipo BL2 de TNBC. Este hallazgo es de especial interés, ya que el subtipo BL2 muestra una variedad de ontologías génicas enriquecidas en componentes y vías implicadas en la proliferación celular, el crecimiento, la supervivencia y las vías de diferenciación celular, como la vía WNT (Lehmann *et al.*, 2011). De acuerdo con este hallazgo, identificamos miembros importantes de la vía WNT, como WNT7A, como potencialmente regulados por LINC00460 (Cisneros-Villanueva *et al.,* 2021.Ver la figura 7C). Estos datos refuerzan el papel de LINC00460 en BL2 TNBC y aclaran aún más su mecanismo de acción en estos tumores.

Detectamos que algunos de los genes co-expresados junto a LINC00460, como *SFRP5* y *HOXD13*, también están relacionados con una mayor tasa de supervivencia en diferentes tumores. La alta expresión de SFRP5 se asocia significativamente con un mejor pronóstico en PAAD (Zhou *et al.,* 2018) *y* CaMa (Wu *et al.*, 2020). Los niveles de la proteína HOXD13 están relacionados con una mayor OS en CaMa (Zhong *et al.*, 2015). Por lo tanto, estos genes blanco candidatos de LINC00460 contribuyen potencialmente al efecto de aumento de la supervivencia observado en los pacientes de CaMa.

Por otra parte, se observó que varios genes coexpresados con LINC00460 (*TRIML2, SFRP5, FOSL1, IFNK, CSF2, DUSP7, DEFB103A* e *IL1AA*) (Cisneros-Villanueva *et al.,* 2021. Ver tabla 4) están relacionados con la inmunidad. Las vías de respuesta inmunitaria son clínicamente relevantes, ya que se ha descrito previamente que un nicho altamente inmunogénico en un tumor puede mejorar el resultado de la enfermedad (Giraldo *et al.*, 2017). Estos genes y vías están frecuentemente enriquecidos en TNBC (Lehmann *et al.*, 2011; Rody *et al.*, 2011). Además, el TNBC también muestra un enriquecimiento de

linfocitos infiltrantes de tumores (TILs) (Stanton and Disis, 2016). Sugerimos que LINC00460 se relaciona con un buen pronóstico en el TNBC/parecido a basal debido a su posible relación con estos genes y con la presencia de linfocitos infiltrantes de tumores (TILs), como se muestra en nuestra cohorte (Cisneros-Villanueva *et al.,* 2021. Ver la tabla 1). Se ha demostrado previamente que la presencia de TILs mejora el pronóstico ya que modula la progresión del cáncer y mejora la respuesta a la quimioterapia en el TNBC, confiriendo una inmunidad protectora en estas pacientes (Denkert *et al.*, 2018).

Por lo tanto, sugerimos que la correlación entre los niveles de expresión de estos genes inmunogénicos y LINC00460 puede explicar parcialmente el comportamiento clínico (buen pronóstico) en las cohortes estudiadas de CaMa (Cisneros-Villanueva *et al.,* 2021. Ver figuras 4, 5 y 6), ya que la sobreexpresión de LINC00460 se asocia con la regulación al alza de factores inmunogénicos que, a su vez, permiten la migración de componentes del sistema inmune. Estos mecanismos pueden promover un entorno tumoral inmunogénico y favorecer así la muerte de las células tumorales. Se necesitan más estudios para validar experimentalmente estos datos, ya que son relevantes para el resultado clínico en el TNBC agresivo.

Además de promover el nicho inmunogénico, LINC00460 podría disminuir la proliferación de las células tumorales, ya que observamos que la vía de proliferación está enriquecida negativamente en el grupo con alta expresión de LINC00460 en las muestras de CaMa. Esto también podría explicar el aumento de las tasas de supervivencia de los pacientes CaMa. Ciertamente, múltiples líneas de evidencia sugieren que LINC00460 puede modular la proliferación celular, la muerte celular, la migración e invasión y la EMT, a través de su actividad de esponja y dirigiendo varios transcritos claves en diversos tipos de cáncer (F. Wang *et al.*, 2018; K. Li *et al.*, 2018; Dong y Quan, 2019; Wen *et al.*, 2019; Yuan *et al.*, 2020; Zou *et al.*, 2020).

LINC00460 se une potencialmente a miR-103-a-1 *in-silico*, que, a su vez, puede unirse a WNT7A. A este respecto, aunque ha habido informes que sugieren que WNT7A es una oncoproteína (Avgustinova *et al.*, 2016), también se ha demostrado que la pérdida de expresión de WNT7A se asocia significativamente con una mala RFS en CaMa (Yi *et al.*, 2017) y también está implicada en la diferenciación de las células tumorales (Liu *et al.*, 2013). Por lo tanto, el papel exacto de WNT7A en el CaMa no está claro actualmente.

Nuestros resultados podrían sugerir que WNT7A desempeña un importante papel clínico en el CaMa, ya que LINC00460 podría actuar como esponja de miR-103-a-1 y, por tanto, liberar WNT7A. Otras investigaciones posteriores deberán validar estas predicciones experimentalmente.

Hemos observado que la relación LINC00460:WNT7A es un marcador compuesto que puede predecir una OS y una DMFS favorables en el TNBC (Cisneros-Villanueva *et al.,* 2021. Ver figura 8). Estos resultados ponen de manifiesto el papel clínico y biológico de los transcritos LINC00460 y WNT7A en el TNBC y constituyen datos valiosos, ya que las simples proporciones de los niveles de expresión génica pueden utilizarse para diagnosticar con precisión (Gordon *et al.*, 2002) y predecir los resultados clínicos cáncer (Gordon *et al.*, 2003; Ma *et al.*, 2004), al tiempo que evitan muchas de las limitaciones que impiden el uso de las técnicas de microarreglos en aplicaciones clínicas extensas (Golub *et al.*, 1999; Shipp *et al.*, 2002). De acuerdo con nuestras observaciones relativas a la posible regulación al alza de los factores inmunogénicos y con los hallazgos previos en este campo (Denkert *et al.*, 2010, 2018; Stanton and Disis, 2016), también se identificó una asociación significativa entre el enriquecimiento de la expresión de LINC00460 y WNT7A con las pacientes con RE que responden a las antraciclinas. Este hallazgo refuerza aún más el papel benéfico de ambos transcritos en la predicción y el pronóstico de las pacientes, ya que se ha demostrado que los cánceres de mama ER-negativos con altos niveles de TILs tienen una mayor sensibilidad a la quimioterapia basada en antraciclinas (West *et al.*, 2011), y que los TILs son un predictor independiente de buena respuesta a la quimioterapia neoadyuvante con antraciclinas/taxanos (Denkert *et al.*, 2010). Todas estas observaciones, sin embargo, están limitadas por el tamaño de las cohortes de pacientes del repositorio Gene Expression Omnibus (GEO) y requerirán validación posterior.

En cuanto al papel de miR-103-a-1, identificamos una asociación marginal con una mala OS en CaMa (Cisneros-Villanueva *et al.,* 2021. Ver figura S7 del material suplementario). Este resultado concuerda con hallazgos anteriores, ya que se ha demostrado que miR-103-a-1 actúa como oncogén para promover la migración y la invasión de las células en TNBC (Xiong *et al.*, 2017). En otro estudio, los autores muestran que la sobreexpresión

de miR-103 en suero se correlacionó significativamente con factores clínicos de mal pronóstico, así como con una peor OS y RFS en el cáncer colorrectal (Wang *et al.*, 2018). La expresión de miR-103/107 muestra funciones promotoras de la capacidad de crecimiento y un perfil de expresión alta de miR-103/107 y baja de Axin2 se correlaciona con un mal pronóstico en pacientes con cáncer colorrectal (Chen *et al.*, 2019). En pacientes con cáncer gástrico, la expresión elevada de miR-103 se asoció significativamente con una mala supervivencia global y libre de enfermedad, y es un factor clave que contribuye a la progresión del tumor (Zheng *et al.*, 2017). En conjunto, estos datos sugieren que miR-103-a-1 es un marcador de mal pronóstico en varios tumores, incluido CaMa. Éste es el primer estudio que sugiere una posible conexión entre LINC00460, WNT7A y miR-103-a-1. Las investigaciones futuras deberán dilucidar los mecanismos exactos implica-dos en esta red potencial de 3 genes, y su impacto en la biología del CaMa parecido a basal.

6. CONCLUSIONES

6.1. FAM83H-AS1

- La sobreexpresión de FAM83H-AS1 se asocia a un mal pronóstico clínico en CaMa. Los resultados sugieren que este efecto se potencia cuando la paciente presenta una alta expresión de RE y RP. Adicionalmente, la alta expresión de FAM83H-AS1 se asocia a una menor respuesta a tamoxifeno en pacientes con CaMa.
- La sobreexpresión de FAM83H-AS1 en muestras de CaMa del TCGA está asociada a la regulación a la baja de transcritos relacionados con la migración y la muerte celular, como *FGF4, FGF21, LEP, CLDN17, TP53, BAX* y *TNFRSF11B*.
- El silenciamiento de FAM83H-AS1 está asociado a la sobreexpresión de transcritos relacionados con la muerte celular, como p63. Los ensayos in vitro también muestran que el silenciamiento de FAM83H-AS1 aumenta la muerte celular.
- Los resultados muestran que FAM83H-AS1 juega un papel importante en la biología de CaMa y, a su vez, influye en el curso clínico de las pacientes.

6.2. LINC00460

- La expresión de LINC00460 es un doble marcador potencial de fenotipos agresivos y de mal resultado clínico en distintos tumores, incluyendo HNSC, KIRC LUSC y LUAD, que también se asocia con un mayor pronóstico en CaMa de tipo basal. LINC00460 está enriquecido en el TNBC BL2 y regula potencialmente la vía de diferenciación WNT.
- LINC00460 también puede modular una plétora de genes relacionados con la inmunidad en CaMa, como *SFRP5, FOSL1, IFNK, CSF2, DUSP7 e IL1A*, e interactúa con miR-103-a-1, *in-silico*, que, a su vez, no puede unirse a WNT7A.
- La tasa LINC00460:WNT7A constituye un marcador compuesto que puede predecir una OS y una DLMS favorables en TNBC, y la combinación de la sobreexpresión de LINC00460 y WNT7A se asocia con la respuesta patológica completa tras el tratamiento con antraciclinas en pacientes con ER- CaMa. Estos datos confirman que LINC00460 es un regulador maestro en los circuitos moleculares de CaMa y que influye en el resultado clínico.

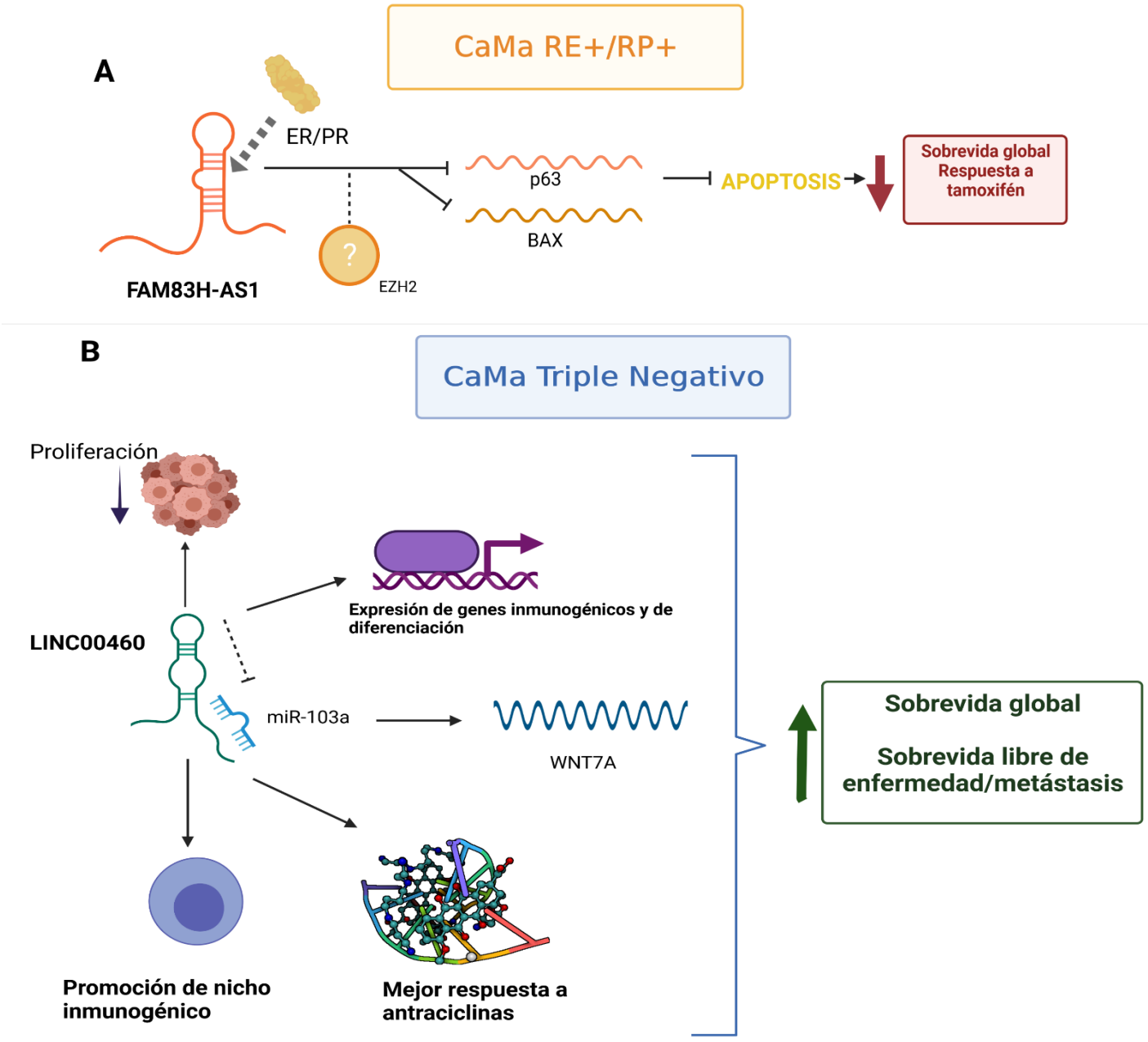

Figura 7. Modelo de los mecanismos de acción propuestos para FAM83H-AS1 y LINC00460.
A)F AM83H-AS1 es un lncRNA que está desregulado en múltiples tumores, y es una molécula que puede actuar como factor pronóstico independiente en el CaMa RE/ RP positivo. FAM83H-AS1 puede regular a genes relacionados con la apoptosis, como *p63 y BAX.* No está claro si esta regulación es directa o indirecta, por medio de factores epigenéticos como EZH2. La desregulación de FAM83H-AS1 está asociada a una menor sobrevida global y respuesta a tamoxifén. B) LINC00460 potencialmente funciona como esponja de miR-103A, lo cual a su vez podría promover la disponibilidad del transcrito WNT7A. La sobre-expresión de LINC00460 se asocia a una mayor expresión de genes relacionados a la diferenciación, genes inmunogénicos y a una disminución de la proliferación celular. Adicionalmente, podría asociarse a la promoción de un nicho inmunogénico y buena respuesta a antraciclinas. Todos estos factores, en su conjunto, potencialmente contribuyen a una mejor OS, RFS y DMFS en CaMa Triple negativo. Imagen creada con BioRender.

4. REFERENCIAS BIBLIOGRÁFICAS

Albanell, J., et al. (2012). «Prospective TransGEICAM Study of the Impact of the 21-Gene Recurrence Score Assay and Traditional Clinicopathological Factors on Adjuvant Clinical Decision Making in Women with Estrogen Receptor-Positive (ER+) Node-Negative Breast Cancer». Annals of Oncology, vol. 23, n.o 3, marzo, pp. 625-31. DOI: https://doi.org/10.1093/annonc/mdr278.

Allred DC, Harvey JM, Berardo M, Clark GM. (1998). Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Modern Pathology: an Official Journal of the United States and Canadian Academy of Pathology, Inc. Feb;11 (2):155-168. PMID: 9504686.

Andrews, S. J., & Rothnagel, J. A. (2014). Emerging evidence for functional peptides encoded by short open reading frames. *Nature reviews. Genetics*, *15*(3), 193–204. https://doi.org/10.1038/nrg3520

Aprile, M. *et al.* (2020) 'LncRNAs in Cancer: From garbage to Junk', *Cancers*, 12(11), p. 3220. doi: 10.3390/cancers12113220.

Avgustinova, A. *et al.* (2016) 'Tumour cell-derived Wnt7a recruits and activates fibroblasts to promote tumour aggressiveness', *Nature Communications*, 7(1), p. 10305. doi: 10.1038/ncomms10305.

Bamodu, O. A. *et al.* (2016) 'Aberrant KDM5B expression promotes aggressive breast cancer through MALAT1 overexpression and downregulation of hsa-miR-448', *BMC Cancer*, 16(1), p. 160. doi: 10.1186/s12885-016-2108-5.

Bargallo, J. E. *et al.* (2015) 'A study of the impact of the 21-gene breast cancer assay on the use of adjuvant chemotherapy in women with breast cancer in a Mexican public hospital', *Journal of Surgical Oncology*, 111(2), pp. 203–207. doi: 10.1002/jso.23794.

Barr, J. A. *et al.* (2019) 'Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells', *Scientific Reports*, 9(1), p. 3662. doi: 10.1038/s41598-019-40094-8.

Basak, P. *et al.* (2018) 'Long Non-Coding RNA H19 Acts as an Estrogen Receptor Modulator that is Required for Endocrine Therapy Resistance in ER+ Breast Cancer Cells', *Cellular Physi-ology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry,and Pharmacology*, 51(4), pp. 1518–1532. doi: 10.1159/000495643.

Bhan, A. *et al.* (2013) 'Antisense Transcript Long Noncoding RNA (lncRNA) HOTAIR is Transcriptionally Induced by Estradiol', *Journal of Molecular Biology*, 425(19), pp. 3707–3722. doi: 10.1016/j.jmb.2013.01.022.

Bhan, A. *et al.* (2014) 'Bisphenol-A and diethylstilbestrol exposure induces the expression of breast cancer associated long noncoding RNA HOTAIR in vitro and in vivo', *The Journal of Steroid Biochemistry and Molecular Biology*, 141, pp. 160–170. doi: 10.1016/j.jsbmb.2014.02.002.

Bi, Y. et al. (2018). Long noncoding RNA FAM83H-AS1 exerts an oncogenic role in glioma through epigenetically silencing CDKN1A (p21). *Journal of cellular physiology*, *233*(11), 8896–8907. https://doi.org/10.1002/jcp.26813.

Borsani, G. *et al.* (1991) 'Characterization of a murine gene expressed from the inactive X chromosome', *Nature*, 351(6324), pp. 325–329. doi: 10.1038/351325a0.

Bray, F. *et al.* (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: a cancer journal for clinicians*, 68(6), pp. 394–424. doi: 10.3322/caac.21492.

Brown, C. J. *et al.* (1991) 'A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome', *Nature*, 349(6304), pp. 38–44. doi: 10.1038/349038a0.

Cai, X. and Cullen, B. R. (2007) 'The imprinted H19 noncoding RNA is a primary microRNA precursor', *RNA (New York, N.Y.)*, 13(3), pp. 313–316. doi: 10.1261/rna.351707.

Cao, W. *et al.* (2017) 'A three-lncRNA signature derived from the Atlas of ncRNA in cancer (TANRIC) database predicts the survival of patients with head and neck squamous cell carcinoma', *Oral Oncology*, 65, pp. 94–101. doi: 10.1016/j.oraloncology.2016.12.017.

Cárdenas-Sánchez, J. *et al.* (2019) 'Consenso Mexicano sobre diagnóstico y tratamiento del cáncer mamario. Octava revisión. Colima 2019', *Gaceta Mexicana de Oncolog*[illegible]*a*, 18(3), p. 2600. doi: 10.24875/j.gamo.M19000180.

Cedro-Tanda, A. *et al.* (2020) 'A lncRNA landscape in breast cancer reveals a potential role for AC009283.1 in proliferation and apoptosis in HER2-enriched subtype', *Scientific Reports*, 10(1), p. 13146. doi: 10.1038/s41598-020-69905-z.

Chaumeil, J. *et al.* (2006) 'A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced', *Genes & Development*, 20(16), pp. 2223–2237. doi: 10.1101/gad.380906.

Cheang, M. C. U., et al. (2012). «Responsiveness of Intrinsic Subtypes to Adjuvant Anthracycline Substitution in the NCIC.CTG MA.5 Randomized Trial». Clinical Cancer Research, vol. 18, n.o 8, abril de 2012, pp. 2402-12. DOI: https://doi.org/10.1158/1078-0432.CCR-11-2956

Cheang, M. C. U. *et al.* (2009) 'Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer', *JNCI: Journal of the National Cancer Institute*, 101(10), pp. 736–750. doi: 10.1093/jnci/djp082.

Chen, C.-K. *et al.* (2016) 'Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing', *Science (New York, N.Y.)*, 354(6311), pp. 468–472. doi: 10.1126/science.aae0047.

Chen, H.-Y. *et al.* (2019) 'miR-103/107 prolong Wnt/β-catenin signaling and colorectal cancer stemness by targeting Axin2', *Scientific Reports*, 9(1), p. 9687. doi: 10.1038/s41598-019-41053-z.

Chen, L.-L. (2016) 'Linking Long Noncoding RNA Localization and Function', *Trends in Biochemical Sciences*, 41(9), pp. 761–772. doi: 10.1016/j.tibs.2016.07.003.

Chen, Q., Zhu, C. and Jin, Y. (2020) 'The Oncogenic and Tumor Suppressive Functions of the Long Noncoding RNA MALAT1: An Emerging Controversy', *Frontiers in Genetics*, 11, p. 93. doi: 10.3389/fgene.2020.00093.

Chen, T. *et al.* (2019) 'Down-regulation of long non-coding RNA HOTAIR sensitizes breast cancer to trastuzumab', *Scientific Reports*, 9(1), p. 19881. doi: 10.1038/s41598-019-53699-w.

Chia, Stephen K., et al. «A 50-Gene Intrinsic Subtype Classifier for Prognosis and Prediction of Benefit from Adjuvant Tamoxifen». Clinical Cancer Research, vol. 18, n.o 16, agosto de 2012, pp. 4465-72. DOI: https://doi.org/10.1158/1078-0432.CCR-12-0286.

Chou, J. *et al.* (2016) 'MALAT1 induced migration and invasion of human breast cancer cells by competitively binding miR-1 with cdc42', *Biochemical and Biophysical Research Communications*, 472(1), pp. 262–269. doi: 10.1016/j.bbrc.2016.02.102.

Chu, C. *et al.* (2015) 'Systematic Discovery of Xist RNA Binding Proteins', *Cell*, 161(2), pp. 404–416. doi: 10.1016/j.cell.2015.03.025.

Chu, C. and Rana, T. M. (2006) 'Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54', *PLoS biology*, 4(7), p. e210. doi: 10.1371/journal.pbio.0040210.

Clemson, C. M. *et al.* (2009) 'An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles', *Molecular Cell*, 33(6), pp. 717–726. doi: 10.1016/j.molcel.2009.01.026.

Cordero, F. *et al.* (2015) 'Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the European Prospective Investigation into Nutrition and Cancer (EPIC-Italy) cohort', *Carcinogenesis*, 36(10), pp. 1144–1153. doi: 10.1093/carcin/bgv102.

Creighton, C. (2012) 'The molecular profile of luminal B breast cancer', *Biologics: Targets and Therapy*, p. 289. doi: 10.2147/BTT.S29923.

Cui, Y. *et al.* (2020) 'LncRNA linc00460 sponges miR-1224-5p to promote esophageal cancer metastatic potential and epithelial-mesenchymal transition', *Pathology, Research and Practice*, 216(7), p. 153026. doi: 10.1016/j.prp.2020.153026.

Da, J. *et al.* (2019) 'Upregulation of the long non-coding RNA FAM83H-AS1 in gastric cancer and its clinical significance', *Pathology, Research and Practice*, 215(10), p. 152616. doi: 10.1016/j.prp.2019.152616.

Deng, L.-L. *et al.* (2016) 'LINC00978 predicts poor prognosis in breast cancer patients', *Scientific Reports*, 6, p. 37936. doi: 10.1038/srep37936.

Denkert, C. *et al.* (2010) 'Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer', *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 28(1), pp. 105–113. doi: 10.1200/JCO.2009.23.7370.

Denkert, C. *et al.* (2018) 'Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy', *The Lancet. Oncology*, 19(1), pp. 40–50. doi: 10.1016/S1470-2045(17)30904-X.

Derrien, T. *et al.* (2012) 'The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression', *Genome Research*, 22(9), pp. 1775–1789. doi:

10.1101/gr.132159.111.

Derrien, T., Guigó, R. and Johnson, R. (2012) 'The Long Non-Coding RNAs: A New (P)layer in the "Dark Matter"', *Frontiers in Genetics*, 2. doi: 10.3389/fgene.2011.00107.

Deva Magendhra Rao, A. K. *et al.* (2019) 'Identification of lncRNAs associated with early-stage breast cancer and their prognostic implications', *Molecular Oncology*, 13(6), pp. 1342–1355. doi: 10.1002/1878-0261.12489.

Dimitrova, N. *et al.* (2014) 'LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint', *Molecular Cell*, 54(5), pp. 777–790. doi: 10.1016/j.molcel.2014.04.025.

Dong, Y. and Quan, H.-Y. (2019) 'Downregulated LINC00460 inhibits cell proliferation and promotes cell apoptosis in prostate cancer', *European Review for Medical and Pharmacological Sciences*, 23(14), pp. 6070–6078. doi: 10.26355/eurrev_201907_18420.

Dou, Q. *et al.* (2019) 'LncRNA FAM83H-AS1 contributes to the radioresistance, proliferation, and metastasis in ovarian cancer through stabilizing HuR protein', *European Journal of Pharmacology*, 852, pp. 134–141. doi: 10.1016/j.ejphar.2019.03.002.

Dressman, M. A. *et al.* (2001) 'Genes that co-cluster with estrogen receptor alpha in microarray analysis of breast biopsies', *The Pharmacogenomics Journal*, 1(2), pp. 135–141. doi: 10.1038/sj.tpj.6500022.

Engreitz, J. M. *et al.* (2013) 'The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome', *Science*, 341(6147), p. 1237973. doi: 10.1126/science.1237973.

Feng, L. *et al.* (2019) 'Long noncoding RNA 00460 (LINC00460) promotes glioma progression by negatively regulating miR-320a', *Journal of Cellular Biochemistry*, 120(6), pp. 9556–9563. doi: 10.1002/jcb.28232.

Feng, T. *et al.* (2016) 'miR-124 downregulation leads to breast cancer progression via LncRNA-MALAT1 regulation and CDK4/E2F1 signal activation', *Oncotarget*, 7(13), pp. 16205–16216. doi: 10.18632/oncotarget.7578.

Foulkes, W. D. *et al.* (2004) 'The Prognostic Implication of the Basal-Like (Cyclin E ^{high} /p27 ^{low} /p53 $^{+}$ /Glomeruloid-Microvascular-Proliferation $^{+}$) Phenotype of *CaMa1* -Related Breast Cancer', *Cancer Research*, 64(3), pp. 830–835. doi: 10.1158/0008-5472.CAN-03-2970.

Fox, A. H. and Lamond, A. I. (2010) 'Paraspeckles', *Cold Spring Harbor Perspectives in Biology*, 2(7), p. a000687. doi: 10.1101/cshperspect.a000687.

GenBank (no date) *Homo sapiens long intergenic non-protein coding RNA 460 (LINC00460), transcript variant 1, long non-coding RNA*. Available at: https://www.ncbi.nlm.nih.gov/nuccore/1636409762/ (Accessed: 10 November 2020).

Ghoncheh, M., Pournamdar, Z. and Salehiniya, H. (2016) 'Incidence and Mortality and Epidemiology of Breast Cancer in the World', *Asian Pacific journal of cancer prevention: APJCP*, 17(S3), pp. 43–46. doi: 10.7314/apjcp.2016.17.s3.43.

Giraldo, N. A. *et al.* (2017) 'Tumor-Infiltrating and Peripheral Blood T-cell Immunophenotypes Predict Early Relapse in Localized Clear Cell Renal Cell Carcinoma', *Clinical Cancer Research*, 23(15), pp. 4416–4428. doi: 10.1158/1078-0432.CCR-16-2848.

Giuliano, A. E. *et al.* (2017) 'Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual', *CA: a cancer journal for clinicians*, 67(4), pp. 290–303. doi: 10.3322/caac.21393.

Gökmen-Polar, Y. and Badve, S. (2012) 'Molecular profiling assays in breast cancer: are we ready for prime time?', *Oncology (Williston Park, N.Y.)*, 26(4), pp. 350–357, 361.

Golub, T. R. *et al.* (1999) 'Molecular classification of cancer: class discovery and class prediction by gene expression monitoring', *Science (New York, N.Y.)*, 286(5439), pp. 531–537. doi: 10.1126/science.286.5439.531.

Gong, C. and Maquat, L. E. (2011) 'lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements', *Nature*, 470(7333), pp. 284–288. doi: 10.1038/nature09701.

Gong, Y.-B. and Zou, Y.-F. (2019) 'Clinical significance of lncRNA FAM83H-AS1 in ovarian cancer', *European Review for Medical and Pharmacological Sciences*, 23(11), pp. 4656–4662. doi: 10.26355/eurrev_201906_18045.

Gordon, G. J. *et al.* (2002) 'Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma', *Cancer Research*, 62(17), pp. 4963–4967. doi: Published September 2002. PMID: 12208747

Gordon, G. J. *et al.* (2003) 'Using Gene Expression Ratios to Predict Outcome Among Patients With Mesothelioma', *JNCI Journal of the National Cancer Institute*, 95(8), pp. 598–605. doi: 10.1093/jnci/95.8.598.

Gupta, R. A. *et al.* (2010) 'Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis', *Nature*, 464(7291), pp. 1071–1076. doi: 10.1038/nature08975.

Han, J. *et al.* (2018) 'Knockdown of lncRNA H19 restores chemo-sensitivity in paclitaxel- resistant triple-negative breast cancer through triggering apoptosis and regulating Akt signaling pathway', *Toxicology and Applied Pharmacology*, 359, pp. 55–61. doi: 10.1016/j.taap.2018.09.018.

Han, L. *et al.* (2016) 'Circulating long noncoding RNA GAS5 as a potential biomarker in breast cancer for assessing the surgical effects', *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, 37(5), pp. 6847–6854. doi: 10.1007/s13277-015-4568-7.

Han, Y. *et al.* (2013) 'Hsa-miR-125b suppresses bladder cancer development by down- regulating oncogene SIRT7 and oncogenic long non-coding RNA MALAT1', *FEBS letters*, 587(23), pp. 3875–3882. http://dx.doi.org/10.1016/j.febslet.2013.10.023

Hirose, T. *et al.* (2014) 'NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies', *Molecular Biology of the Cell*, 25(1), pp. 169–183. doi: 10.1091/mbc.E13-09-0558.

Hong, W. *et al.* (2020) 'lncRNA LINC00460 Silencing Represses EMT in Colon Cancer through

Downregulation of ANXA2 via Upregulating miR-433-3p', *Molecular Therapy. Nucleic Acids*, 19, pp. 1209–1218. doi: 10.1016/j.omtn.2019.12.006.

Housman, G. and Ulitsky, I. (2016) 'Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs', *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1859(1), pp. 31–40. doi: niazi.

Howlader, N. *et al.* (2018) 'Differences in Breast Cancer Survival by Molecular Subtypes in the United States', *Cancer Epidemiology Biomarkers & Prevention*, 27(6), pp. 619–626. doi: 10.1158/1055-9965.EPI-17-0627.

Hu, G., Lou, Z. and Gupta, M. (2014) 'The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation', *PloS One*, 9(9), p. e107016. doi: 10.1371/journal.pone.0107016.

Hu, Xuguang *et al.* (2019) 'Knockdown of lncRNA HOTAIR sensitizes breast cancer cells to ionizing radiation through activating miR-218', *Bioscience Reports*, 39(4). doi: 10.1042/BSR20181038.

Hu, Xiaodong *et al.* (2019) 'Long noncoding RNA LINC00460 aggravates invasion and metastasis by targeting miR-30a-3p/Rap1A in nasopharyngeal carcinoma', *Human Cell*, 32(4), pp. 465–476. doi: 10.1007/s13577-019-00262-4.

Huang, G.-W. *et al.* (2018) 'A three-lncRNA signature predicts overall survival and disease-free survival in patients with esophageal squamous cell carcinoma', *BMC Cancer*, 18(1), p. 147. doi: 10.1186/s12885-018-4058-6.

Huang, Y. *et al.* (2017) 'Leptin promotes the migration and invasion of breast cancer cells by upregulating ACAT2', *Cellular Oncology*, 40(6), pp. 537–547. doi: 10.1007/s13402-017-0342-8.

Jarroux, J., Morillon, A. and Pinskaya, M. (2017) 'History, Discovery, and Classification of lncRNAs', *Advances in Experimental Medicine and Biology*, 1008, pp. 1–46. doi: 10.1007/978-981-10-5203-3_1.

Jiang, X. and Chen, D. (2020) 'LncRNA FAM83H-AS1 maintains intervertebral disc tissue homeostasis and attenuates inflammation-related pain via promoting nucleus pulposus cell growth through miR-22-3p inhibition', *Annals of Translational Medicine*, 8(22), p. 1518. doi: 10.21037/atm-20-7056.

Jiang, Y. *et al.* (2019) 'LncRNA LINC00460 promotes EMT in head and neck squamous cell carcinoma by facilitating peroxiredoxin-1 into the nucleus', *Journal of Experimental & Clinical Cancer Research*, 38(1), p. 365. doi: 10.1186/s13046-019-1364-z.

Jørgensen, C. L. T., et al. (2014). «PAM50 Breast Cancer Intrinsic Subtypes and Effect of Gemcitabine in Advanced Breast Cancer Patients». Acta Oncologica, vol. 53, n.o 6, junio de 2014, pp. 776-87. DOI: https://doi.org/10.3109/0284186X.2013.865076.

Kallen, A. N. *et al.* (2013) 'The imprinted H19 lncRNA antagonizes let-7 microRNAs', *Molecular Cell*, 52(1), pp. 101–112. doi: 10.1016/j.molcel.2013.08.027.

Kennecke, H. *et al.* (2010) 'Metastatic behavior of breast cancer subtypes', *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 28(20), pp. 3271–3277.

doi: 10.1200/JCO.2009.25.9820.

Kim, J. *et al.* (2018) 'Long noncoding RNA MALAT1 suppresses breast cancer metastasis', *Nature Genetics*, 50(12), pp. 1705–1715. doi: 10.1038/s41588-018-0252-3.

de Kruijf, E. M. *et al.* (2014) 'Comparison of frequencies and prognostic effect of molecular subtypes between young and elderly breast cancer patients', *Molecular Oncology*, 8(5), pp. 1014–1025. doi: 10.1016/j.molonc.2014.03.022.

Lal, S. *et al.* (2017) 'Molecular signatures in breast cancer', *Methods*, 131, pp. 135–146. doi: 10.1016/j.ymeth.2017.06.032.

Lehmann, B. D. *et al.* (2011) 'Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies', *Journal of Clinical Investigation*, 121(7), pp. 2750–2767. doi: 10.1172/JCI45014.

Li, F., Zhu, W. and Wang, Z. (2020) 'Long noncoding RNA LINC00460 promotes the progression of cervical cancer via regulation of the miR-361-3p/Gli1 axis', *Human Cell*, 34, pp. 229–237. doi: 10.1007/s13577-020-00447-2.

Li, G. and Kong, Q. (2019a) 'LncRNA LINC00460 promotes the papillary thyroid cancer progression by regulating the LINC00460/miR-485-5p/Raf1 axis', *Biological Research*, 52(1), pp. 52–61. doi: 10.1186/s40659-019-0269-9.

Li, G. and Kong, Q. (2019b) 'LncRNA LINC00460 promotes the papillary thyroid cancer progression by regulating the LINC00460/miR-485-5p/Raf1 axis', *Biological Research*, 52(1), p. 61. doi: 10.1186/s40659-019-0269-9.

Li, J. *et al.* (2011) 'Triple-Negative Subtype Predicts Poor Overall Survival and High Locoregional Relapse in Inflammatory Breast Cancer', *The Oncologist*, 16(12), pp. 1675–1683. doi: 10.1634/theoncologist.2011-0196.

Li, K. *et al.* (2016) 'Leptin promotes breast cancer cell migration and invasion via IL-18 expression and secretion', *International Journal of Oncology*, 48(6), pp. 2479–2487. doi: 10.3892/ijo.2016.3483.

Li, K. *et al.* (2018) 'Long non-coding RNA linc00460 promotes epithelial-mesenchymal transition and cell migration in lung cancer cells', *Cancer Letters*, 420, pp. 80–90. doi: 10.1016/j.canlet.2018.01.060.

Li, Y. *et al.* (2019) 'A Compound AC1Q3QWB Selectively Disrupts HOTAIR-Mediated Recruitment of PRC2 and Enhances Cancer Therapy of DZNep', *Theranostics*, 9(16), pp. 4608– 4623. doi: 10.7150/thno.35188.

Li, Z. *et al.* (2018) 'LncRNA MALAT1 promotes relapse of breast cancer patients with postoperative fever', *American Journal of Translational Research*, 10(10), pp. 3186–3197.

Lian, H. *et al.* (2019) 'Linc00460 promotes osteosarcoma progression via miR-1224-5p/FADS1 axis', *Life Sciences*, 233, p. 116757. doi: 10.1016/j.lfs.2019.116757.

Lian, Y. *et al.* (2018) 'A Novel lncRNA, LINC00460, Affects Cell Proliferation and Apoptosis by Regulating KLF2 and CUL4A Expression in Colorectal Cancer', *Molecular Therapy - Nucleic*

Acids, 12, pp. 684–697. doi: 10.1016/j.omtn.2018.06.012.

Liang, W.-H. *et al.* (2019) 'DSCAM-AS1 promotes tumor growth of breast cancer by reducing miR-204-5p and up-regulating RRM2', *Molecular Carcinogenesis*, 58(4), pp. 461–473. doi: 10.1002/mc.22941.

Liang, Y. *et al.* (2017) 'A novel long noncoding RNA linc00460 up-regulated by CBP/P300 promotes carcinogenesis in esophageal squamous cell carcinoma', *Bioscience Reports*, 37(5), p. BSR20171019. doi: 10.1042/BSR20171019.

Lin, A. *et al.* (2016) 'The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer', *Nature Cell Biology*, 18(2), pp. 213–224. doi: 10.1038/ncb3295.

Lips, E. H. *et al.* (2013) 'Breast cancer subtyping by immunohistochemistry and histological grade outperforms breast cancer intrinsic subtypes in predicting neoadjuvant chemotherapy response', *Breast Cancer Research and Treatment*, 140(1), pp. 63–71. doi: 10.1007/s10549- 013-2620-0.

Liu, B. *et al.* (2020) 'A Novel Androgen-Induced lncRNA FAM83H-AS1 Promotes Prostate Cancer Progression via the miR-15a/CCNE2 Axis', *Frontiers in Oncology*, 10, p. 620306. doi: 10.3389/fonc.2020.620306.

Liu, X. *et al.* (2018) 'Long non-coding RNA LINC00460 promotes epithelial ovarian cancer progression by regulating microRNA-338-3p', *Biomedicine & Pharmacotherapy*, 108, pp. 1022–1028. doi: 10.1016/j.biopha.2018.09.103.

Liu, Y. *et al.* (2013) 'Overexpression of Wnt7a Is Associated With Tumor Progression and Unfavorable Prognosis in Endometrial Cancer':, *International Journal of Gynecological Cancer*, 23(2), pp. 304–311. doi: 10.1097/IGC.0b013e31827c7708.

Lu, L. *et al.* (2012) 'Association of large noncoding RNA HOTAIR expression and its downstream intergenic CpG island methylation with survival in breast cancer', *Breast Cancer Research and Treatment*, 136(3), pp. 875–883. doi: 10.1007/s10549-012-2314-z.

Lu, R. *et al.* (2018) 'Circulating HOTAIR expression predicts the clinical response to neoadjuvant chemotherapy in patients with breast cancer', *Cancer Biomarkers: Section A of Disease Markers*, 22(2), pp. 249–256. doi: 10.3233/CBM-170874.

Lu, S. *et al.* (2018) 'lncRNA FAM83H-AS1 is associated with the prognosis of colorectal carcinoma and promotes cell proliferation by targeting the Notch signaling pathway', *Oncology Letters*, 15(2), pp. 1861–1868. doi: 10.3892/ol.2017.7520.

Ma, G. *et al.* (2019) 'Long Noncoding RNA LINC00460 Promotes the Gefitinib Resistance of Nonsmall Cell Lung Cancer Through Epidermal Growth Factor Receptor by Sponging miR-769-5p', *DNA and Cell Biology*, 38(2), pp. 176–183. doi: 10.1089/dna.2018.4462.

Ma, X.-J. *et al.* (2004) 'A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen', *Cancer Cell*, 5(6), pp. 607–616. doi: 10.1016/j.ccr.2004.05.015.

Ma, Y. *et al.* (2019) 'LncRNA DSCAM-AS1 acts as a sponge of miR-137 to enhance Tamoxifen resistance in breast cancer', *Journal of Cellular Physiology*, 234(3), pp. 2880–2894. doi: 10.1002/jcp.27105.

Maisonneuve, P. *et al.* (2014) 'Proposed new clinicopathological surrogate definitions of luminal A and luminal B (HER2-negative) intrinsic breast cancer subtypes', *Breast Cancer Research*, 16(3), p. R65. doi: 10.1186/bcr3679.

Martín, M., et al. «PAM50 Proliferation Score as a Predictor of Weekly Paclitaxel Benefit in Breast Cancer». Breast Cancer Research and Treatment, vol. 138, n.o 2, abril de 2013, pp. 457- 66. DOI: https://doi.org/10.1007/s10549-013-2416-2.

Matouk, I. J. *et al.* (2014) 'Oncofetal H19 RNA promotes tumor metastasis', *Biochimica Et Biophysica Acta*, 1843(7), pp. 1414–1426. doi: 10.1016/j.bbamcr.2014.03.023.

McHugh, C. A. *et al.* (2015) 'The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3', *Nature*, 521(7551), pp. 232–236. doi: 10.1038/nature14443.

Meseure, D. *et al.* (2016) 'Prognostic value of a newly identified MALAT1 alternatively spliced transcript in breast cancer', *British Journal of Cancer*, 114(12), pp. 1395–1404. doi: 10.1038/bjc.2016.123.

Mourtada-Maarabouni, M. *et al.* (2009) 'GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer', *Oncogene*, 28(2), pp. 195–208. doi: 10.1038/onc.2008.373.

Murthy, U. M. S. and Rangarajan, P. N. (2010) 'Identification of protein interaction regions of VINC/NEAT1/Men epsilon RNA', *FEBS letters*, 584(8), pp. 1531–1535. doi: 10.1016/j.febslet.2010.03.003.

Nakano, Y. *et al.* (2019) 'Clinical importance of long non-coding RNA LINC00460 expression in EGFR-mutant lung adenocarcinoma', *International Journal of Oncology*, 56, pp. 243–257. doi: 10.3892/ijo.2019.4919.

Nelson, B. R. *et al.* (2016) 'A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle', *Science*, 351(6270), pp. 271–275. doi: 10.1126/science.aad4076.

Niazi, F. and Valadkhan, S. (2012) 'Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3' UTRs', *RNA*, 18(4), pp. 825–843. doi: nelson.

Nielsen, T. O., et al. «A Comparison of PAM50 Intrinsic Subtyping with Immunohistochemistry and Clinical Prognostic Factors in Tamoxifen-Treated Estrogen Receptor–Positive Breast Cancer». Clinical Cancer Research, vol. 16, n.o 21, noviembre de 2010, pp. 5222-32. DOI: https://doi.org/10.1158/1078-0432.CCR-10-1282

Niknafs, Y. S. *et al.* (2016) 'The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression', *Nature Communications*, 7, p. 12791. doi: 10.1038/ncomms12791.

Nora, E. P. *et al.* (2012) 'Spatial partitioning of the regulatory landscape of the X-inactivation centre', *Nature*, 485(7398), pp. 381–385. doi: 10.1038/nature11049.

Nozawa, R.-S. *et al.* (2013) 'Human inactive X chromosome is compacted through a PRC2-

independent SMCHD1-HBiX1 pathway', *Nature Structural & Molecular Biology*, 20(5), pp. 566–573. doi: 10.1038/nsmb.2532.

Parker, Joel S., et al. (2009)«Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes». Journal of Clinical Oncology, vol. 27, n.o 8, marzo de 2009, pp. 1160-67. DOI: https://doi.org/10.1200/JCO.2008.18.1370.

Peng, J. *et al.* (2017) 'Expression profile analysis of long noncoding RNA in ER-positive subtype breast cancer using microarray technique and bioinformatics', *Cancer Management and Research*, Volume 9, pp. 891–901. doi: 10.2147/CMAR.S151120.

Prat, A. *et al.* (2013) 'Prognostic Significance of Progesterone Receptor–Positive Tumor Cells Within Immunohistochemically Defined Luminal A Breast Cancer', *Journal of Clinical Oncology*, 31(2), pp. 203–209. doi: 10.1200/JCO.2012.43.4134.

Rakha, E. A. *et al.* (2009) 'Triple-Negative Breast Cancer: Distinguishing between Basal and Nonbasal Subtypes', *Clinical Cancer Research*, 15(7), pp. 2302–2310. doi: 10.1158/1078-0432.CCR-08-2132.

Reis-Filho, J. S. *et al.* (2010) 'Molecular Profiling: Moving Away from Tumor Philately', *Science Translational Medicine*, 2(47), pp. 47ps43-47ps43. doi: 10.1126/scitranslmed.3001329.

Ren, Y. *et al.* (2019) 'Targeted design and identification of AC1NOD4Q to block activity of HOTAIR by abrogating the scaffold interaction with EZH2', *Clinical Epigenetics*, 11(1), p. 29. doi: 10.1186/s13148-019-0624-2.

Riaz, M. *et al.* (2012) 'Correlation of breast cancer susceptibility loci with patient characteristics, metastasis-free survival, and mRNA expression of the nearest genes', *Breast Cancer Research and Treatment*, 133(3), pp. 843–851. doi: 10.1007/s10549-011-1663-3.

Rinn, J. L. *et al.* (2007) 'Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs', *Cell*, 129(7), pp. 1311–1323. doi: 10.1016/j.cell.2007.05.022.

Rody, A. *et al.* (2011) 'A clinically relevant gene signature in triple negative and basal-like breast cancer', *Breast cancer research: BCR*, 13(5), p. R97. doi: 10.1186/bcr3035.

Russnes, H. G. *et al.* (2017) 'Breast Cancer Molecular Stratification', *The American Journal of Pathology*, 187(10), pp. 2152–2162. doi: 10.1016/j.ajpath.2017.04.022.

Sawe, R. T. *et al.* (2016) 'Aggressive breast cancer in western Kenya has early onset, high proliferation, and immune cell infiltration', *BMC Cancer*, 16(1), p. 204. doi: 10.1186/s12885-016-2204-6.

Sestak, Ivana, et al. (2015) «Prediction of Late Distant Recurrence After 5 Years of Endocrine Treatment: A Combined Analysis of Patients From the Austrian Breast and Colorectal Cancer Study Group 8 and Arimidex, Tamoxifen Alone or in Combination Randomized Trials Using the PAM50 Risk of Recurrence Score». Journal of Clinical Oncology, vol. 33, n.o 8, marzo de 2015, pp. 916-22. DOI: https://doi.org/10.1200/JCO.2014.55.6894.

Shah, S. P. *et al.* (2012) 'The clonal and mutational evolution spectrum of primary triple- negative breast cancers', *Nature*, 486(7403), pp. 395–399. doi: 10.1038/nature10933.

Shan, H. *et al.* (2019) 'FAM83H-AS1 is associated with clinical progression and modulates cell proliferation, migration, and invasion in bladder cancer', *Journal of Cellular Biochemistry*, 120(3), pp. 4687–4693. doi: 10.1002/jcb.27758.

Shim, H. J. *et al.* (2014) 'Breast cancer recurrence according to molecular subtype', *Asian Pacific journal of cancer prevention: APJCP*, 15(14), pp. 5539–5544. doi: 10.7314/apjcp.2014.15.14.5539.

Shipp, M. A. *et al.* (2002) 'Diffuse large B-cell lymphoma outcome prediction by gene- expression profiling and supervised machine leRNAing', *Nature Medicine*, 8(1), pp. 68–74. doi: 10.1038/nm0102-68.

Si, X. *et al.* (2016a) 'LncRNA H19 confers chemoresistance in ERα-positive breast cancer through epigenetic silencing of the pro-apoptotic gene BIK', *Oncotarget*, 7(49), pp. 81452–81462. doi: 10.18632/oncotarget.13263.

Si, X. *et al.* (2016b) 'LncRNA H19 confers chemoresistance in ERα-positive breast cancer through epigenetic silencing of the pro-apoptotic gene BIK', *Oncotarget*, 7(49), pp. 81452–81462. doi: 10.18632/oncotarget.13263.

Simon, M. D. *et al.* (2013) 'High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation', *Nature*, 504(7480), pp. 465–469. doi: 10.1038/nature12719.

Sinn, P. *et al.* (2013) 'Multigene Assays for Classification, Prognosis, and Prediction in Breast Cancer: a Critical Review on the Background and Clinical Utility', *Geburtshilfe und Frauenheilkunde*, 73(09), pp. 932–940. doi: 10.1055/s-0033-1350831.

Sørensen, K. P. *et al.* (2013) 'Long non-coding RNA HOTAIR is an independent prognostic marker of metastasis in estrogen receptor-positive primary breast cancer', *Breast Cancer Research and Treatment*, 142(3), pp. 529–536. doi: 10.1007/s10549-013-2776-7.

Sorlie, T., et al. (2001). «Gene Expression Patterns of Breast Carcinomas Distinguish Tumor Subclasses with Clinical Implications». Proceedings of the National Academy of Sciences, vol. 98, n.o 19, septiembre de 2001, pp.10869-74. DOI:https://doi.org/10.1073/pnas.191367098.

Sparano, J. A. *et al.* (2015) 'Prospective Validation of a 21-Gene Expression Assay in Breast Cancer', *New England Journal of Medicine*, 373(21), pp. 2005–2014. doi: 10.1056/NEJMoa1510764.

Stanton, S. E. and Disis, M. L. (2016) 'Clinical significance of tumor-infiltrating lymphocytes in breast cancer', *Journal for Immunotherapy of Cancer*, 4, p. 59. doi: 10.1186/s40425-016-0165-6.

Steuten, L. *et al.* (2019) 'Cost Effectiveness of Multigene Panel Sequencing for Patients With Advanced Non–Small-Cell Lung Cancer', *JCO Clinical Cancer Informatics*, (3), pp. 1–10. doi: 10.1200/CCI.19.00002.

Strong, A. L. *et al.* (2015) 'Leptin produced by obese adipose stromal/stem cells enhances proliferation and metastasis of estrogen receptor positive breast cancers', *Breast Cancer Research*, 17(1), p. 112. doi: 10.1186/s13058-015-0622-z.

Sun, B. *et al.* (2019) 'Research progress on the interactions between long non-coding RNAs and

microRNAs in human cancer (Review)', *Oncology Letters*, 19(1), pp. 595–605. doi: 10.3892/ol.2019.11182.

Tang, S. *et al.* (2019) 'Overexpression of serum exosomal HOTAIR is correlated with poor survival and poor response to chemotherapy in breast cancer patients', *Journal of Biosciences*, 44(2).

Tarighi, M. *et al.* (2021) 'Long non-coding RNA (lncRNA) DSCAM-AS1 is upregulated in breast cancer', *Breast Disease*. doi: 10.3233/BD-201010.

Tran, B. and Bedard, P. L. (2011) 'Luminal-B breast cancer and novel therapeutic targets', *Breast Cancer Research*, 13(6), p. 221. doi: 10.1186/bcr2904.

Tripathi, V. *et al.* (2010) 'The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation', *Molecular Cell*, 39(6), pp. 925–938. doi: 10.1016/j.molcel.2010.08.011.

Tsai, M.-C. *et al.* (2010) 'Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes', *Science*, 329(5992), pp. 689–693. doi: 10.1126/science.1192002.

Tsutsui, S. *et al.* (2003) 'Prognostic significance of the coexpression of p53 protein and c-erbB2 in breast cancer', *American Journal of Surgery*, 185(2), pp. 165–167. doi: 10.1016/s0002-9610(02)01203-5.

van 't Veer, L. J. *et al.* (2002) 'Gene expression profiling predicts clinical outcome of breast cancer', *Nature*, 415(6871), enero 2002, pp. 530–536. doi: 10.1038/415530a.

Vennin, C. *et al.* (2015) 'H19 non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b', *Oncotarget*, 6(30), pp. 29209–29223. doi: 10.18632/oncotarget.4976.

Wang, D. *et al.* (2015) 'LncRNA MALAT1 enhances oncogenic activities of EZH2 in castration-resistant prostate cancer', *Oncotarget*, 6(38), pp. 41045–41055. doi: 10.18632/oncotarget.5728.

Wang, D.-S. *et al.* (2018) 'Upregulation of serum miR-103 predicts unfavorable prognosis in patients with colorectal cancer', *European Review for Medical and Pharmacological Sciences*, 22(14), pp. 4518–4523. doi: 10.26355/eurrev_201807_15506.

Wang, F. *et al.* (2018) 'LINC00460 modulates KDM2A to promote cell proliferation and migration by targeting miR-342-3p in gastric cancer', *OncoTargets and Therapy*, Volume 11, pp. 6383–6394. doi: 10.2147/OTT.S169307.

Wang, H.-X. *et al.* (2020) 'LINC00460 promotes proliferation and inhibits apoptosis of non-small cell lung cancer cells through targeted regulation of miR-539', *European Review for Medical and Pharmacological Sciences*, 24(12), pp. 6752–6758. doi: 10.26355/eurrev_202006_21663.

Wang, J. *et al.* (2010) 'CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer', *Nucleic Acids Research*, 38(16), pp. 5366–5383. doi: 10.1093/nar/gkq285.

Wang, J. *et al.* (2017) 'Huaier Extract Inhibits Breast Cancer Progression Through a LncRNA-H19/MiR-675-5p Pathway', *Cellular Physiology and Biochemistry: International Journal of*

Experimental Cellular Physiology, Biochemistry, and Pharmacology, 44(2), pp. 581–593. doi: 10.1159/000485093.

Wang, K. C. *et al.* (2011) 'A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression', *Nature*, 472(7341), pp. 120–124. doi: 10.1038/nature09819.

Wang, Lan *et al.* (2013) 'CaMa1 is a negative modulator of the PRC2 complex', *The EMBO Journal*, 32(11), pp. 1584–1597. doi: 10.1038/emboj.2013.95.

Wang, R.-X. *et al.* (2016) 'Value of Ki-67 expression in triple-negative breast cancer before and after neoadjuvant chemotherapy with weekly paclitaxel plus carboplatin', *Scientific Reports*, 6(1), p. 30091. doi: 10.1038/srep30091.

Wang, X. *et al.* (2018) 'Upregulated Expression of Long Non-Coding RNA, LINC00460, Suppresses Proliferation of Colorectal Cancer.', *Journal of Cancer*, 9(16), pp. 2834–2843. doi: 10.7150/jca.26046.

Wang, Z.-L. *et al.* (2016) 'Integrative analysis reveals clinical phenotypes and oncogenic potentials of long non-coding RNAs across 15 cancer types', *Oncotarget*, 7(23), pp. 35044– 35055. doi: 10.18632/oncotarget.9037.

Watson, C. N., Belli, A. and Di Pietro, V. (2019) 'Small Non-coding RNAs: New Class of Biomarkers and Potential Therapeutic Targets in Neurodegenerative Disease', *Frontiers in Genetics*, 10, p. 364. doi: 10.3389/fgene.2019.00364.

Wei, L. *et al.* (2016) 'Leptin promotes epithelial-mesenchymal transition of breast cancer via the upregulation of pyruvate kinase M2', *Journal of Experimental & Clinical Cancer Research*, 35(1), p. 166. doi: 10.1186/s13046-016-0446-4.

Wen, L. *et al.* (2019) 'The long non-coding RNA LINC00460 predicts the prognosis and promotes the proliferation and migration of cells in bladder urothelial carcinoma', *Oncology Letters*, 17, pp. 3874–3880. doi: 10.3892/ol.2019.10023.

West, N. R. *et al.* (2011) 'Tumor-infiltrating lymphocytes predict response to anthracycline- based chemotherapy in estrogen receptor-negative breast cancer', *Breast cancer research: BCR*, 13(6), p. R126. doi: 10.1186/bcr3072.

Wu, Q. *et al.* (2017) 'Breast cancer subtypes predict the preferential site of distant metastases: a SEER based study', *Oncotarget*, 8(17), pp. 27990–27996. doi: 10.18632/oncotarget.15856.

Wu, Z.-H. *et al.* (2020) 'Comprehensive Analysis of the Expression and Prognosis for *SFRPs* in Breast Carcinoma', *Cell Transplantation*, 29, p. 096368972096247. doi: 10.1177/0963689720962479.

Xie, X. *et al.* (2019) 'Long non-coding RNA LINC00460 promotes head and neck squamous cell carcinoma cell progression by sponging miR-612 to up-regulate AKT2', *American Journal of Translational Research*, 11(10), pp. 6326–6340. doi: eCollection 2019.

Xing, F. *et al.* (2018) 'Loss of XIST in Breast Cancer Activates MSN-c-Met and Reprograms Microglia via Exosomal miRNA to Promote Brain Metastasis', *Cancer Research*, 78(15), pp. 4316–4330. doi: 10.1158/0008-5472.CAN-18-1102.

Xing, H. *et al.* (2018) 'Long noncoding RNA LINC00460 targets miR-539/MMP-9 to promote meningioma progression and metastasis', *Biomedicine & Pharmacotherapy*, 105, pp. 677–682. doi: 10.1016/j.biopha.2018.06.005.

Xiong, B. *et al.* (2017) 'miR-103 regulates triple negative breast cancer cells migration and invasion through targeting olfactomedin 4', *Biomedicine & Pharmacotherapy*, 89, pp. 1401– 1408. doi: 10.1016/j.biopha.2017.02.028.

Xu, Z. *et al.* (2019) 'Characterization of mRNA Expression and Endogenous RNA Profiles in Bladder Cancer Based on The Cancer Genome Atlas (TCGA) Database', *Medical Science Monitor*, 25, pp. 3041–3060. doi: 10.12659/MSM.915487.

Yan, L. *et al.* (2015) 'Regulation of tumor cell migration and invasion by the H19/let-7 axis is antagonized by metformin-induced DNA methylation', *Oncogene*, 34(23), pp. 3076–3084. doi: 10.1038/onc.2014.236.

Yan, X. *et al.* (2015) 'Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers', *Cancer Cell*, 28(4), pp. 529–540. doi: 10.1016/j.ccell.2015.09.006.

Yang, F., Liu, Y.-H., *et al.* (2016) 'A novel long non-coding RNA FGF14-AS2 is correlated with progression and prognosis in breast cancer', *Biochemical and Biophysical Research Communications*, 470(3), pp. 479–483. doi: 10.1016/j.bbrc.2016.01.147.

Yang, F., Lv, S.-X., *et al.* (2016) 'Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer', *OncoTargets and Therapy*, 9, pp. 7039–7045. doi: 10.2147/OTT.S110055.

Yang, J. *et al.* (2020) 'Long Noncoding RNA LINC00460 Promotes Hepatocellular Carcinoma Progression via Regulation of miR-342-3p/AGR2 Axis', *OncoTargets and Therapy*, Volume 13, pp. 1979–1991. doi: 10.2147/OTT.S239258.

Yang, L. *et al.* (2011) 'ncRNA- and Pc2 Methylation-Dependent Gene Relocation between Nuclear Structures Mediates Gene Activation Programs', *Cell*, 147(4), pp. 773–788. doi: 10.1016/j.cell.2011.08.054.

Yang, L. *et al.* (2019) 'FAM83H-AS1 is upregulated and predicts poor prognosis in colon cancer', *Biomedicine & Pharmacotherapy*, 118, p. 109342. doi: 10.1016/j.biopha.2019.109342.

Yates, A. D. *et al.* (2020) 'Ensembl 2020', *Nucleic Acids Research*, 48(D1), pp. D682–D688. doi: Long non-coding RNA LINC00460 promotes head and neck squamous cell carcinoma cell progression by sponging miR-612.

Yersal, O. and Barutca, S. (2014) 'Biological subtypes of breast cancer: Prognostic and therapeutic implications', *World Journal of Clinical Oncology*, 5(3), pp. 412–424. doi: 10.5306/wjco.v5.i3.412.

Yi, K. *et al.* (2017) 'Wnt7a Deficiency Could Predict Worse Disease-Free and Overall Survival in Estrogen Receptor-Positive Breast Cancer', *Journal of Breast Cancer*, 20(4), p. 361. doi: 10.4048/jbc.2017.20.4.361.

Yoon, J.-H. *et al.* (2012) 'LincRNA-p21 suppresses target mRNA translation', *Molecular Cell*, 47(4), pp. 648–655. doi: 10.1016/j.molcel.2012.06.027.

Yoon, J.-H. *et al.* (2013) 'Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination', *Nature Communications*, 4, p. 2939. doi: 10.1038/ncomms3939.

Yu, F. *et al.* (2015) 'MALAT1 functions as a competing endogenous RNA to mediate Rac1 expression by sequestering miR-101b in liver fibrosis', *Cell Cycle*, 14(24), pp. 3885–3896. doi: 10.1080/15384101.2015.1120917.

Yuan, B. *et al.* (2020) 'Down-Regulation of LINC00460 Represses Metastasis of Colorectal Cancer via WWC2', *Digestive Diseases and Sciences*, 65(2), pp. 442–456. doi: 10.1007/s10620-019-05801-5.

Zhai, L.-L. *et al.* (2015) 'Over-expression of miR-675 in formalin-fixed paraffin-embedded (FFPE) tissues of breast cancer patients', *International Journal of Clinical and Experimental Medicine*, 8(7), pp. 11195–11201. PMID: 26379923. PMCID: PMC4565306

Zhang, A. *et al.* (2015) 'LncRNA HOTAIR Enhances the Androgen-Receptor-Mediated Transcriptional Program and Drives Castration-Resistant Prostate Cancer', *Cell Reports*, 13(1), pp. 209–221. doi: 10.1016/j.celrep.2015.08.069.

Zhang, C. *et al.* (2020) 'A prognostic long non-coding RNA-associated competing endogenous RNA network in head and neck squamous cell carcinoma', *PeerJ*, 8, p. e9701. doi: 10.7717/peerj.9701.

Zhang, D., Zeng, S. and Hu, X. (2020) 'Identification of a three-long noncoding RNA prognostic model involved competitive endogenous RNA in kidney renal clear cell carcinoma', *Cancer Cell International*, 20(1), p. 319. doi: 10.1186/s12935-020-01423-4.

Zhang, J. *et al.* (2017) 'Overexpression of FAM83H-AS1 indicates poor patient survival and knockdown impairs cell proliferation and invasion via MET/EGFR signaling in lung cancer', *Scientific Reports*, 7, p. 42819. doi: 10.1038/srep42819.

Zhang, K. *et al.* (2016) 'Circulating lncRNA H19 in plasma as a novel biomarker for breast cancer', *Cancer Biomarkers: Section A of Disease Markers*, 17(2), pp. 187–194. doi: 10.3233/CBM-160630.

Zhang, Y. *et al.* (2020) 'LncRNA DSCAM-AS1 interacts with YBX1 to promote cancer progression by forming a positive feedback loop that activates FOXA1 transcription network', *Theranostics*, 10(23), pp. 10823–10837. doi: 10.7150/thno.47830.

Zhang, Yueyan *et al.* (2019) 'lncRNA LINC00460 promoted colorectal cancer cells metastasis via miR-939-5p sponging', *Cancer Management and Research*, Volume 11, pp. 1779–1789. doi: 10.2147/CMAR.S192452.

Zhang, Z. and Carmichael, G. G. (2001) 'The Fate of dsRNA in the Nucleus', *Cell*, 106(4), pp. 465–476. doi: 10.1016/S0092-8674(01)00466-4.

Zhao, X. *et al.* (2015) 'Lentiviral Vector Mediated Claudin1 Silencing Inhibits Epithelial to Mesenchymal Transition in Breast Cancer Cells', *Viruses*, 7(6), pp. 2965–2979. doi: 10.3390/v7062755.

Zheng, J. *et al.* (2017) 'miR-103 Promotes Proliferation and Metastasis by Targeting KLF4 in Gastric Cancer', *International Journal of Molecular Sciences*, 18(5). doi: 10.3390/ijms18050910.

Zhong, Z.-B. *et al.* (2015) 'Prognostic significance of HOXD13 expression in human breast cancer', *International Journal of Clinical and Experimental Pathology*, 8(9), pp. 11407–11413. PMID: 26617867. PMCID: PMC4637683

Zhou, M. *et al.* (2016) 'Discovery of potential prognostic long non-coding RNA biomarkers for predicting the risk of tumor recurrence of breast cancer patients', *Scientific Reports*, 6(1), p. 31038. doi: 10.1038/srep31038.

Zhou, Q. *et al.* (2018) 'Long non coding RNA XIST as a prognostic cancer marker – A meta-analysis', *Clinica Chimica Acta*, 482, pp. 1–7. doi: 10.1016/j.cca.2018.03.016.

Zhou, W. & Tian, Ming & Hu, J. & Li, L. & He, Y.. (no date) 'SFRP5 as a prognostic biomarker for patients with pancreatic ductal adenocarcinoma.', 9(3), pp. 3442–3447. doi: IJCEP0020124.

Zhou, X. *et al.* (2015) 'The Interaction Between MiR-141 and lncRNA-H19 in Regulating Cell Proliferation and Migration in Gastric Cancer', *Cellular Physiology and Biochemistry*, 36(4), pp. 1440–1452. doi: 10.1159/000430309.

Zhou, Y. *et al.* (2017) 'Long non-coding RNA HOTAIR enhances radioresistance in MDA-MB231 breast cancer cells', *Oncology Letters*, 13(3), pp. 1143–1148. doi: 10.3892/ol.2017.5587.

Zhu, L. *et al.* (2015) 'Long Noncoding RNA MALAT-1 Can Predict Metastasis and a Poor Prognosis: a Meta-Analysis', *Pathology & Oncology Research*, 21(4), pp. 1259–1264. doi: 10.1007/s12253-015-9960-5.

Zhu, Y. *et al.* (2019) 'Long noncoding RNA Linc00460 promotes breast cancer progression by regulating the miR-489-5p/FGF7/AKT axis', *Cancer Management and Research*, Volume 11, pp. 5983–6001. doi: 10.2147/CMAR.S207084.

Zou, X. *et al.* (2020) 'Long Noncoding RNA LINC00460 Modulates MMP-9 to Promote Cell Proliferation, Invasion and Apoptosis by Targeting miR-539 in Papillary Thyroid Cancer', *Cancer Management and Research*, 12, pp. 199–207. doi: 10.2147/CMAR.S222085.