



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría en Ciencias Bioquímicas

Identificación de lncRNAs que participan en la respuesta transcripcional a la activación de la vía ER en un modelo celular de cáncer de mama luminal A

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Maestro en Ciencias

PRESENTA:

Ángel David Rodríguez Reyes

TUTOR PRINCIPAL

Dr. Félix Recillas Targa  
[Instituto de Fisiología Celular](#)

COMITÉ TUTOR

Dra. Aliesha Araceli González Arenas  
[Instituto de Investigaciones Biomédicas](#)

Dr. César Augusto Poot Hernández  
[Instituto de Fisiología Celular](#)

Ciudad Universitaria, CDMX. Enero, 2024



Universidad Nacional  
Autónoma de México



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

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

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

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



**PROTESTA UNIVERSITARIA DE INTEGRIDAD Y  
HONESTIDAD ACADÉMICA Y PROFESIONAL  
(Graduación con trabajo escrito)**

De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la Institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado: "Identificación de lncRNAs que participan en la respuesta transcripcional a la activación de la vía ER en un modelo celular de cáncer de mama luminal A", que presenté para obtener el grado de ---Maestría es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi programa de posgrado, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de graduación.

**Atentamente**

Una firma manuscrita en tinta azul que parece decir "Ángel David Rodríguez Reyes".

**Ángel David Rodríguez Reyes  
No. De cuenta: 522008063**

“La epigenética es el estudio de las moléculas y mecanismos con la capacidad de perpetuar diversos estados alternativos en la actividad génica, tratándose de la misma secuencia de DNA.” (Fitz-James & Cavalli, 2022)

*A mi familia que me ha apoyado siempre, compartir  
todo esto con ellos lo hace muy especial.*

## **Agradecimientos Académicos**

La presente tesis se llevó a cabo bajo la orientación del Dr. Félix Recillas Targa, del Departamento de Genética Molecular del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México. Expreso mi profundo agradecimiento a los miembros del comité tutor por sus valiosas observaciones y contribuciones a este trabajo:

Félix Recillas Targa	Instituto de Fisiología Celular
Aliesha Araceli González Arenas	Instituto de Investigaciones Biomédicas
César Augusto Poot Hernández	Instituto de Fisiología Celular

Quisiera destacar y agradecer de manera especial al aspirante a Doctor Hober Nelson Núñez por su excepcional asesoría y respaldo académico, fundamental para la ejecución de todos los análisis computacionales y experimentales detallados en este estudio.

Asimismo, deseo expresar mi reconocimiento por la inestimable guía proporcionada por la Biól. Georgina Guerrero Avendaño y la Dra. Sylvia Patricia Garza Manero. Además, agradezco sinceramente el respaldo brindado por la laboratorista Gianelli Cortés González, y, finalmente, extendiendo mi gratitud a cada uno de los integrantes del Laboratorio 122 NTE del Instituto de Fisiología Celular.

Es relevante señalar que este trabajo se realizó gracias al respaldo financiero proporcionado por la beca CONAHCYT (Consejo Nacional de Humanidades, Ciencias y Tecnologías).

Por último, agradezco sinceramente las significativas observaciones formuladas por los miembros del jurado, las cuales contribuyeron a enriquecer este trabajo de investigación:

MACIAS SILVA MARINA	PMDCBQ	PRESIDENTE
REYNAUD GARZA ENRIQUE ALEJANDRO	PMDCBQ	SECRETARIO
MIRANDA RÍOS JUAN	PMDCBQ	VOCAL
ROBLES FLORES MARTHA	PMDCBQ	VOCAL
ROMERO CORDOBA SANDRA LORENA	PMDCBQ	VOCAL

# ÍNDICE

Listado de Abreviaturas en Inglés .....	3
Resumen .....	6
Abstract .....	7
1. Introducción.....	8
1.0 Receptores Nucleares.....	8
1.1 Cáncer de mama y su relación con el ER $\alpha$ .....	10
1.2 El receptor de estrógenos y su vía de señalización.....	11
1.1.1 Los ligandos de la vía.....	14
1.1 Vías genómicas de señalización.....	16
1.1.4 Vía no genómica.....	20
1.1.5 Genes que responden a la vía de señalización con estradiol (E2) .....	21
1.2 Respuesta transcripcional ante la estimulación con estradiol (E2) .....	22
1.3 RNAs largos no codificantes (lncRNAs).....	25
1.4 Respuesta transcripcional indirecta de la vía p53 mediada por lncRNAs .....	27
2. Antecedentes.....	30
2.1 lncRNAs responsivos a estradiol (E2) .....	30
2.2 Estudio de expresión a escala del genoma completo ( <i>GRO-seq</i> ).....	31
2.3 La Familia de genes <i>FOX</i> .....	33
2.3.1 La proteína FOXF1 y sus funciones.....	34
3. Planteamiento del problema.....	37
4. Hipótesis.....	38
5. Objetivo General .....	38
6. Objetivos particulares .....	39
8. Resultados .....	42
8.1 Objetivo 1: Identificación de lncRNAs responsivos a E2 y regulados por ER $\alpha$ en la línea celular de cáncer de mama MCF-7.....	42
8.1.1 Análisis bioinformáticos .....	42
8.1.2 Búsqueda Exhaustiva en Bases de Datos Públicas.....	42
8.1.3 Fuentes de Datos Utilizadas.....	42
8.2 Objetivo 2: Identificación de lncRNAs con Potencial Regulatorio en Cis en Genes Codificantes Respondientes a E2, Independientes de la Vía de ER. ....	55
8.2.1 Resultados Experimentales .....	59

8.2.2 Distribución del transcrito del <i>LINC01082</i> en la célula y su posible función regulatoria.....	60
8.3 Objetivo 3: Caracterización Funcional de un lncRNA Candidato y su Impacto en el Gen Codificante Asociado. ....	61
9. Discusión .....	71
10. Conclusiones .....	79
11. Perspectivas .....	80
12. Materiales y métodos .....	82
12.1 Cultivo Celular .....	82
12.2 RT-qPCR.....	82
12.3 Análisis de base de datos de CHIP-seq.....	82
12.4 Análisis de base de datos de RNA-seq .....	83
12.5 Extracción de RNA.....	83
12.6 Ensayo de proliferación y viabilidad celular de MTT .....	83
12.7 Prueba estadística t de Student.....	84
13. Bibliografía.....	86

## Listado de Abreviaturas en Inglés

AF: Activating function domain

AF1: Activating function domain 1

AF2: Activating function domain 2

AMP: Adenosine Monophosphate

AP1: Activator protein 1

ATAC-seq: Assay for Transposase-Accessible Chromatin using sequencing

BMP-4: Bone Morphogenetic Protein 4

bFGF: basic Fibroblast Growth Factor

C/EBP $\beta$ : CCAAT/Enhancer-Binding Protein Beta

cAMP: Cyclic Adenosine Monophosphate

cDNA: Complementary DNA

CREB1: cAMP Response Element-Binding Protein 1

CT: Carboxyl Terminal Domain

DBD: DNA Binding Domain

DGE: Differential Gene Expression

DNA: Deoxyribonucleic Acid

E1: Estrona

E2: Estradiol

E3: Estriol

E4: Estetrol

ERE: Estrogen Receptor Element

ER $\alpha$ : Estrogen Receptor Alpha



ER $\beta$ : Estrogen Receptor Beta

ESR1: Estrogen Receptor Alpha Gene

eNOS: Endothelial Nitric Oxide Synthase

FOXF1: Forkhead Box F1

FGF-10: Fibroblast Growth Factor 10

FBS: Fetal Bovine Serum

GPER1: G Protein-Coupled Estrogen Receptor 1

GLI3: Glioma-Associated Oncogene Homolog 3

GRO-seq: Global Run-On Sequencing

Hi-C: High-throughput Chromosome Conformation Capture

HOTAIR: HOX Transcript Antisense RNA

IGF1: Insulin-Like Growth Factor 1

kDA: Kilo Daltons

Kb: Kilobases

LDL: Low-Density Lipoprotein

LINC01082: Long Intergenic Non Coding RNA 01082

LincRNAs: Long Intergenic Non Coding RNA

lncRNA: Long Non Coding RNA

LKLF: Lung Krüppel-Like Factor

MALAT1: Metastasis-Associated Lung Adenocarcinoma Transcript 1

MCF-7: Michigan Cancer Foundation-7 Cell Line

MMP: Matrix Metalloproteinase

nM: Nanomolar

NF- $\kappa$ B: Nuclear Factor-Kappa B

NTD: Nitrogen Terminal Domain

PRC2: Polycomb Repressive Complex 2

qRT-PCR: Quantitative Reverse Transcription Polymerase Chain Reaction

RNA: Ribonucleic Acid

RNA-seq: RNA Sequencing

RPMi: Roswell Park Memorial Institute Culture Medium

sgRNA: Single Guide RNA

SP1: Stimulating Protein 1

STAT: Signal Transducer and Activator of Transcription

STAT5: Signal Transducer and Activator of Transcription 5

TBX: T-Box Genes

TF: Transcription Factor

TSS: Transcriptional Start Site

VEGFR2: Vascular Endothelial Growth Factor Receptor 2

VEGF: Vascular Endothelial Growth Factor

## Resumen

La regulación de la expresión génica es un proceso altamente complejo en el que los RNA largos no codificantes (lncRNA) han emergido como actores clave. Estos lncRNA desempeñan roles fundamentales en procesos biológicos esenciales, desde la diferenciación celular hasta la patogénesis de enfermedades, incluyendo el cáncer. Esto ha convertido a los lncRNA en un campo de investigación vital para comprender las bases moleculares de diversas afecciones.

Este trabajo se centró en la regulación génica a través de la vía del receptor de estrógenos (ER), la cual desempeña un papel crucial en el cáncer de mama. El objetivo principal fue investigar cómo los lncRNA contribuyen a la regulación de genes asociados al ER y los mecanismos subyacentes en esta regulación. Durante el proyecto, se identificaron lncRNA específicos que se activan a través de la vía del receptor de estrógenos. Uno de ellos, el LINC01082, destacó debido a su potencial para regular la expresión génica de manera indirecta a la vía ER $\alpha$ .

El LINC01082 al ser sobre expresado ectópicamente, mostró su capacidad para influir en el aumento transcripcional de los genes: *FOXF1* y *FENDRR*. Estos hallazgos proporcionan una visión más completa de cómo los lncRNA pueden afectar la regulación de la vía del ER en el contexto del cáncer de mama. Además, plantean preguntas interesantes sobre los mecanismos subyacentes que requerirán un análisis más detallado en futuras investigaciones.

En resumen, este estudio resalta la relevancia de los lncRNA en la regulación génica en la vía del ER en el cáncer de mama. Asimismo, esto ofrece una comprensión más profunda de cómo estos lncRNA pueden desempeñar un papel crucial al influir en la expresión de genes clave, lo que potencialmente podría conducir a avances en el entendimiento y tratamiento de esta enfermedad.

## Abstract

The regulation of gene expression is a highly complex process in which long non-coding RNAs (lncRNAs) have emerged as key players. These lncRNAs play fundamental roles in essential biological processes, from cellular differentiation to the pathogenesis of diseases, including cancer. This has made lncRNAs a vital research field to comprehend the molecular foundations of various conditions.

This thesis focused on gene regulation within the context of the estrogen receptor (ER) pathway, which plays a crucial role in breast cancer. The main objective was to investigate how lncRNAs contribute to the regulation of genes associated with ER and the underlying mechanisms in this regulation. During the study, specific lncRNAs responsive to ER activation were identified. Among them, LINC01082 stood out due to its potential to indirectly regulate gene expression through the ER $\alpha$  pathway.

When ectopically overexpressed, LINC01082 demonstrated its ability to influence the transcriptional upregulation of the genes FOXF1 and FENDRR. These findings provide a more comprehensive understanding of how lncRNAs can impact ER pathway regulation in the context of breast cancer. Moreover, they raise interesting questions about the underlying mechanisms that will require more detailed analysis in future research.

In summary, this study highlights the relevance of lncRNAs in gene regulation within the ER pathway in breast cancer. It also offers a deeper insight into how these lncRNAs may play a crucial role in influencing the expression of key genes, potentially leading to advancements in the understanding and treatment of this disease.

# 1. Introducción

## 1.0 Receptores Nucleares

La superfamilia de receptores nucleares comprende a un grupo de proteínas encargadas de regular la expresión génica, en respuesta a señales hormonales y otros estímulos. Los miembros de esta familia comparten una estructura común que consta de un dominio de unión al DNA altamente conservado y un dominio de unión que permite la interacción al ligando que es menos conservado. Se dividen en dos grupos principales: 1) los receptores de esteroides, que se unen a hormonas esteroides como los estrógenos y la testosterona, y 2) los receptores de no esteroides, que se unen a diversas moléculas, como la vitamina D y los ácidos grasos (Perissi & Rosenfeld, 2005). Estos receptores son fundamentales para una amplia gama de procesos biológicos, que incluyen entre otros, el desarrollo embrionario, la homeostasis y la respuesta inmunitaria, entre otros. Además, son blancos importantes en el desarrollo de fármacos, ya que su modulación puede tener aplicaciones terapéuticas en enfermedades como el cáncer y la diabetes (Yu et al., 1992).

La regulación de la transcripción génica por parte de los receptores nucleares se efectúa mediante la unión a secuencias específicas de DNA conocidas como elementos de respuesta, que se localizan en los promotores de los genes, funcionando como factores de transcripción (TFs). Estos receptores tienen la capacidad de activar o reprimir la transcripción génica, y lo pueden llevar a cabo tanto de manera directa, al unirse al DNA, como de manera indirecta, al interaccionar con otros factores de transcripción. Además, los receptores nucleares pueden establecer interacciones con proteínas coactivadoras o corepresoras para modular su actividad transcripcional (Perissi & Rosenfeld, 2005).

Cuando un receptor nuclear se une a estos elementos de respuesta, puede desencadenar la activación o inhibición de la transcripción del gen correspondiente. La especificidad en la regulación génica se logra a través de

interacciones con cofactores y la formación de complejos proteicos que regulan la actividad transcripcional. De igual forma, esta regulación puede ser influenciada por otros factores de transcripción y señales celulares, lo que permite una respuesta coordinada y adaptable a las variaciones en el entorno celular (Yu et al., 1992).

Los miembros de la superfamilia de receptores nucleares para esteroides y para no esteroides pueden emplear estrategias distintas para ejercer un control preciso sobre la regulación génica. El espaciado y la orientación de los sitios de unión desempeñan un papel significativo en la especificidad de la respuesta transcripcional. Por ejemplo, los receptores de glucocorticoides y otros receptores relacionados operan como homodímeros en secuencias palindrómicas de DNA separadas por tres nucleótidos; en contraste, el elemento de respuesta a estrógenos se compone de un motivo palindrómico con motivos AGGTCA separados por tres pares de bases (pb). La identificación de aminoácidos específicos en el receptor, como los del segundo par de cisteínas-cisteínas situados en el primer dedo de zinc (conocido como cuadro proximal), ha demostrado ser fundamental para discernir entre estos diferentes motivos centrales en el DNA (Yu et al., 1992).

Los receptores nucleares tienen la capacidad de interactuar con diversos cofactores, ya sean coactivadores o corepresores, dependiendo de si están unidos o no a su ligando específico. Por lo general, cuando los receptores nucleares no están unidos a su ligando, tienden a interactuar preferentemente con corepresores, lo que conduce a la represión de la expresión génica. Por otro lado, cuando los receptores nucleares están unidos a su ligando específico, funcionan como activadores de la transcripción, ya que tienen la capacidad de reclutar proteínas coactivadoras y remodeladores de la cromatina (Perissi & Rosenfeld, 2005).

Todo lo expuesto anteriormente confirma el papel central que tienen los receptores nucleares en la regulación de una variedad de procesos moleculares en organismos multicelulares, dado que orquestan la expresión génica en respuesta a

una amplia diversidad de señales, tanto internas como externas (Evans & Mangelsdorf, 2014).

### 1.1 Cáncer de mama y su relación con el ER $\alpha$

La relación entre el cáncer de mama y el receptor de estrógeno alfa (ER $\alpha$ ) es un tema de gran relevancia en la investigación oncológica. El estrógeno (una hormona esteroidea) ejerce una influencia significativa en el desarrollo y la progresión del cáncer de mama, y gran parte de sus efectos en este contexto son mediados por el receptor de estrógeno (ER $\alpha$ ). Además, el receptor de progesterona (PR) ha sido también objeto de atención en relación con esta enfermedad debido a que es regulado por ER $\alpha$ , ha sido objeto de atención en relación con esta enfermedad. La elevación de los niveles de ER $\alpha$  y PR en lesiones premalignas y malignas de mama, en comparación con el tejido normal, enfatiza la importancia de estos receptores en el cáncer de mama (Sommer & Fuqua, 2001).

El ER $\alpha$ , es un factor transcripcional clave que desempeña un papel fundamental en el pronóstico del cáncer de mama. Estudios que han analizado a pacientes pre y postmenopáusicas han revelado diferencias significativas en la supervivencia y el intervalo libre de enfermedad en función de la determinación del ER $\alpha$ . Específicamente, se ha observado que las pacientes premenopáusicas con determinación positiva del ER $\alpha$  tienen una supervivencia más larga en comparación con aquellas con determinación negativa. Sin embargo, en el caso de las pacientes postmenopáusicas, no se han encontrado diferencias significativas en la supervivencia en función del estado del receptor de estrógeno (Eni, 1967).

Del mismo modo, se ha establecido que la terapia hormonal es más efectiva en pacientes con tumores que expresan ER $\alpha$ , mostrando un notable 48% de pacientes con tumores ER $\alpha$ + que respondieron a este tratamiento, en contraste con el 6% de pacientes con tumores ER $\alpha$ - (Eni, 1967). También se ha observado una diferencia en los patrones de metástasis, ya que los tumores ER $\alpha$ + tienen una

mayor predilección a metastatizar en la piel y los huesos, mientras que los tumores ER $\alpha$ - tienden a hacerlo en órganos y el cerebro (Eni, 1967).

En conclusión, el estado del receptor de estrógeno es un factor crucial en el pronóstico y la respuesta al tratamiento del cáncer de mama. Los pacientes que tienen tumores que expresan ER $\alpha$  tienen una mejor supervivencia y una mayor probabilidad de responder a la terapia hormonal. La terapia endocrina, que se centra en bloquear los efectos del estrógeno en las células cancerosas, es un enfoque principal en el tratamiento del cáncer de mama ER $\alpha$ + y ha demostrado ser efectiva (Johnston, 2010).

Por tanto, el entendimiento de los intrincados procesos moleculares relacionados con la vía de señalización del estrógeno y la regulación génica a través de ER $\alpha$  puede contribuir a una mayor comprensión de las diferencias fisiológicas entre los cánceres de mama que responden al estradiol y los que no. Esto, a su vez, podría conducir al desarrollo de nuevos enfoques terapéuticos y diagnósticos en el futuro.

## 1.2 El receptor de estrógenos y su vía de señalización

El receptor de estrógenos alfa (ER $\alpha$ ) es una proteína intracelular que se une al estrógeno y actúa como un regulador clave de la expresión génica. Cuando se une al estrógeno, ER $\alpha$  forma complejos con factores de transcripción y se transloca al núcleo, donde regula la expresión de una serie de genes (Ikeda et al., 2015). Algunos de los genes regulados por ER $\alpha$  incluyen:

- Genes relacionados con la proliferación celular: *ciclina D1*, *c-fos*, *c-jun*, *c-myc*
- Genes relacionados con la apoptosis: *BCL2*, *BCL-xL*, *MCL1*
- Genes relacionados con la invasión y la migración celular: *MMP1*, *MMP2*, *MMP9*
- Genes relacionados con la angiogénesis: *VEGF*, *bFGF*
- Genes relacionados con la diferenciación celular: *involucrina*, *laminina*, *queratina*.



La regulación de estos genes por ER $\alpha$  es importante para la respuesta estrógeno-dependiente en tejidos como el mamario y puede tener implicaciones importantes en el desarrollo y el tratamiento del cáncer de mama (Jameera Begam et al., 2017).

Asimismo, la vía de señalización del receptor de estrógenos constituye una parte integral de las cascadas de señalización que impulsan los procesos de proliferación y migración en células de cáncer de mama sensibles a los estrógenos. El estrógeno se une al receptor ER $\alpha$ , el cual forma un dímero y se conecta directamente al genoma en los sitios de unión del ER $\alpha$  (ERBS, por sus siglas en inglés, ER $\alpha$  binding sites) o, de manera indirecta, a través del reclutamiento facilitado por factores transcripcionales como NF- $\kappa$ B o CREB. Estos factores transcripcionales establecen una conexión entre el ER y coactivadores o corepresores (Klinge et al., 2004).

Después de la unión del ER $\alpha$  al genoma, se desencadena el reclutamiento de complejos proteicos corregulatorios, los cuales pueden interactuar ya sea de manera directa o indirecta al formar una estructura de andamiaje con los coreguladores. Es relevante destacar que el receptor de estrógenos (ER) fue el primer receptor caracterizado que demostró su capacidad para unirse a una hormona, evidenciando así su habilidad para regular procesos biológicos en tejidos de órganos reproductivos femeninos (Fuentes & Silveyra, 2019).

En seres humanos, el ER $\alpha$  se encuentra codificado por el gen ESR1, situado en el cromosoma 6, específicamente en el locus 6q25.1. En un principio, se caracterizó el receptor ER $\alpha$ , que comparte con su isoforma ER $\beta$  la naturaleza de ser un receptor nuclear capaz de translocarse al núcleo y regular la transcripción de genes sensibles a los estrógenos. Posteriormente, se describió un tercer receptor conocido como el Receptor de Estrógenos Acoplado a Proteína G (GPER1, por sus siglas en inglés).

Cabe señalar que el ER $\alpha$  es una proteína de 67 kDa, compuesta por 595 aminoácidos, mientras que el ER $\beta$  tiene un peso de 59 kDa. La principal diferencia

entre estas dos isoformas radica en que la parte amino terminal es más corta en el ER $\beta$ . (Figura 1).

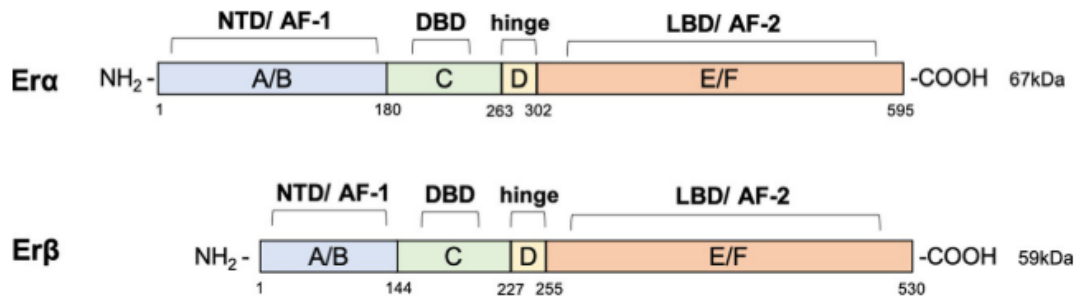


Figura 1: Organización estructural del ER. Representación gráfica de los dominios estructurales y funcionales el receptor de estrógenos  $\alpha$  y  $\beta$ : Imagen tomada de (Fuentes & Silveyra, 2019).

Ambos receptores comparten seis dominios estructurales y funcionales, como se ilustra en la Figura 1:

- La región A/B, también conocida como la región amino terminal (NTD), desempeña un papel crucial en la activación de la transcripción génica. Además, presenta una sección con dedos de zinc que facilita la unión a secuencias específicas en el DNA.
- La región C se refiere al dominio de unión al DNA (DBD por sus siglas en inglés). Este dominio no solo contribuye a la dimerización del ER $\alpha$ , sino que también posibilita su unión a secuencias específicas en la cromatina, identificadas como elementos de respuesta a estrógenos (*estrogen response elements*, EREs)
- El dominio D constituye una región bisagra que conecta los dominios C y E, además de tener la capacidad de interactuar con proteínas chaperonas. Esta sección también alberga la secuencia de señalización para la translocación nuclear, que se vuelve accesible tras la unión del estrógeno, permitiendo que el complejo ligando/receptor se desplace hacia el núcleo.
- La región E/F, también denominada la parte carboxilo terminal (CTD), representa el dominio de unión al ligando. Este dominio tiene la capacidad de unirse a coactivadores y correpresores.

En última instancia, se identifican dos reguladores adicionales de la actividad transcripcional del ER, conocidos como dominios de función activadora (AF). Estos son denominados AF1 y AF2, siendo AF1 notable por no requerir la unión del ligando para su activación. (Fuentes & Silveyra, 2019).

En síntesis, la comprensión detallada de la estructura y función de los receptores de estrógeno (ER $\alpha$  y ER $\beta$ ) resulta esencial tanto en el ámbito de la investigación científica como en la práctica médica. En el contexto de la investigación, este conocimiento puede ampliar nuestra comprensión de los complejos procesos de regulación génica, especialmente en investigaciones vinculadas al cáncer de mama. Esto nos permite adentrarnos en la comprensión de cómo estos receptores interaccionan con el ADN y otros elementos celulares. Por otro lado, en el ámbito médico, esta comprensión facilita el desarrollo y perfeccionamiento de terapias hormonales específicas, el diseño de fármacos precisos y la mejora de los métodos de diagnóstico y pronóstico del cáncer de mama.

A pesar de compartir similitudes en la mayoría de sus dominios estructurales y funcionales, estas dos isoformas exhiben diferencias notables, siendo la longitud de la región amino terminal una de las más destacadas. Estas divergencias pueden influir en su papel como reguladores de la transcripción génica y en los mecanismos específicos que rigen cada isoforma. Este conocimiento resulta fundamental para avanzar en nuestra comprensión de la compleja regulación hormonal.

### 1.1.1 Los ligandos de la vía

Los estrógenos, hormonas sexuales esteroideas con una amplia gama de funciones fisiológicas, desempeñan un papel crucial en la regulación de diversos procesos biológicos. Estas funciones abarcan la regulación del ciclo menstrual, la densidad ósea, la función cerebral, la movilización del colesterol, el control de la inflamación y el desarrollo del tejido mamario y de los órganos sexuales (Fuentes & Silveyra, 2019).

Cabe destacar que los estrógenos pertenecen a la familia de los esteroides y se dividen en cuatro tipos principales: estrona (E1), estradiol (E2), estriol (E3) y estetrol (E4) (Figura 2). Su estructura básica se compone de un anillo de benceno, un grupo hidroxilo fenólico y un grupo cetona en el caso de la estrona (E1); o bien, un grupo hidroxilo en la posición 17 $\beta$  (17 $\beta$ -estradiol o E2), dos grupos hidroxilo en el caso del estriol (E3) y tres grupos en el estetrol (E4).

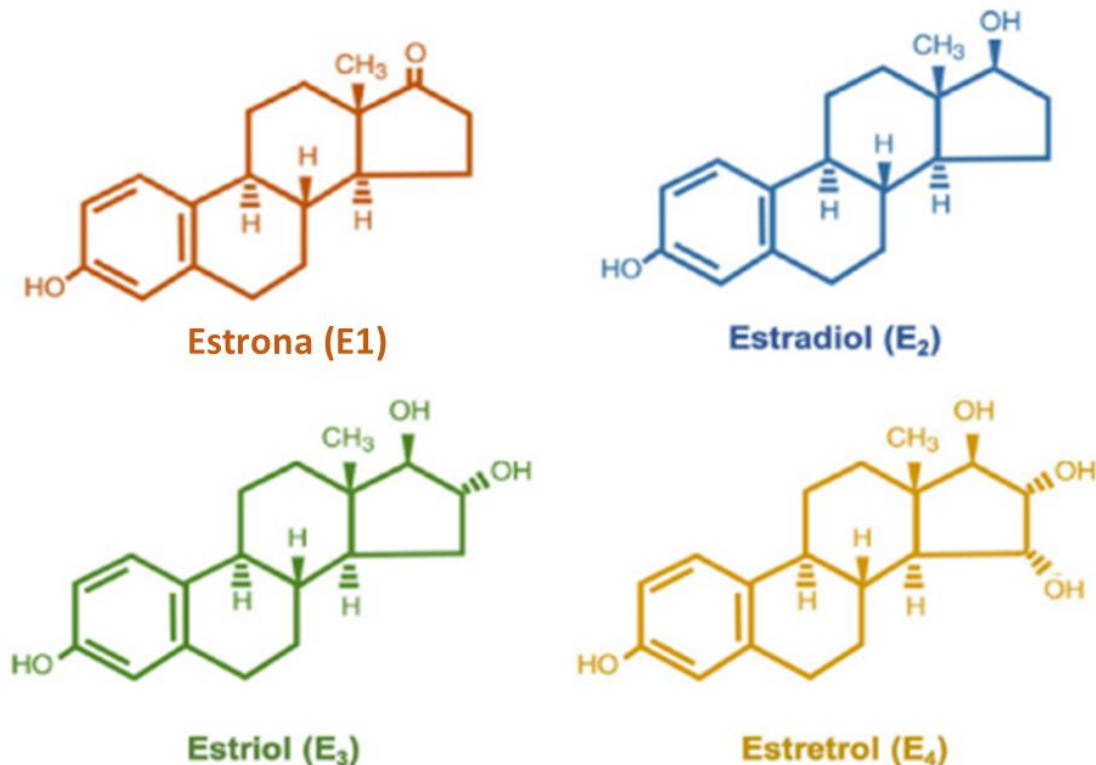


Figura 2: Estructura química de los estrógenos. Estrona (E1; naranja), estradiol (E2; azul), estriol (E3; verde) y estetrol (E4; amarillo). Imagen tomada de: (Fuentes & Silveyra, 2019).

Los estrógenos se sintetizan principalmente en los ovarios, específicamente en los folículos, aunque también se producen en las glándulas suprarrenales y en el tejido adiposo. Los cuatro tipos de estrógenos tienen la capacidad de unirse tanto a receptores de estrógenos nucleares como membranales, aunque con diferentes afinidades y potencias de respuesta.

El estradiol (E2) destaca como el estrógeno circulante más predominante en humanos durante los años reproductivos, siendo, por ende, de mayor relevancia

fisiológica. En particular, el E2 estimula la proliferación de células epiteliales en el endometrio uterino y las glándulas mamarias (Fuentes & Silveyra, 2019).

En resumen, el estradiol (E2) no solo ha demostrado ser fundamental para los procesos celulares normales, sino que también adquiere una notable relevancia en contextos patológicos, como el cáncer. Este hecho se evidencia claramente en la progresión del cáncer de mama luminal A, donde la respuesta a esta hormona conlleva a la inducción de procesos como la proliferación celular.

### 1.1 Vías genómicas de señalización

Las funciones fisiológicas de los estrógenos, mencionadas en la sección anterior, se ven mediadas en parte por su interacción con el receptor de estrógenos (ER) a través de dos vías celulares conocidas como la vía genómica y la no genómica.

En la vía genómica (Figura 3), el  $17\beta$ -estradiol (E2) se une al ER, localizado en el citoplasma, generando la dimerización del receptor. Este proceso provoca la translocación del dímero hacia el núcleo celular. Una vez en el núcleo, el dímero reconoce la secuencia de los elementos de respuesta a estrógenos (ERE, por sus siglas en inglés) en el ADN para regular la transcripción de los genes blanco de esta vía (Słowikowski et al., 2017).

Resulta interesante destacar que, aunque se han identificado EREs en varios promotores de genes y en otras regiones regulatorias, se ha informado que más de un tercio de los genes regulados por el ER carecen de secuencias EREs (Fuentes & Silveyra, 2019). Es por esta razón que la vía genómica se divide, a su vez, en directa e indirecta, como se explicará más adelante.

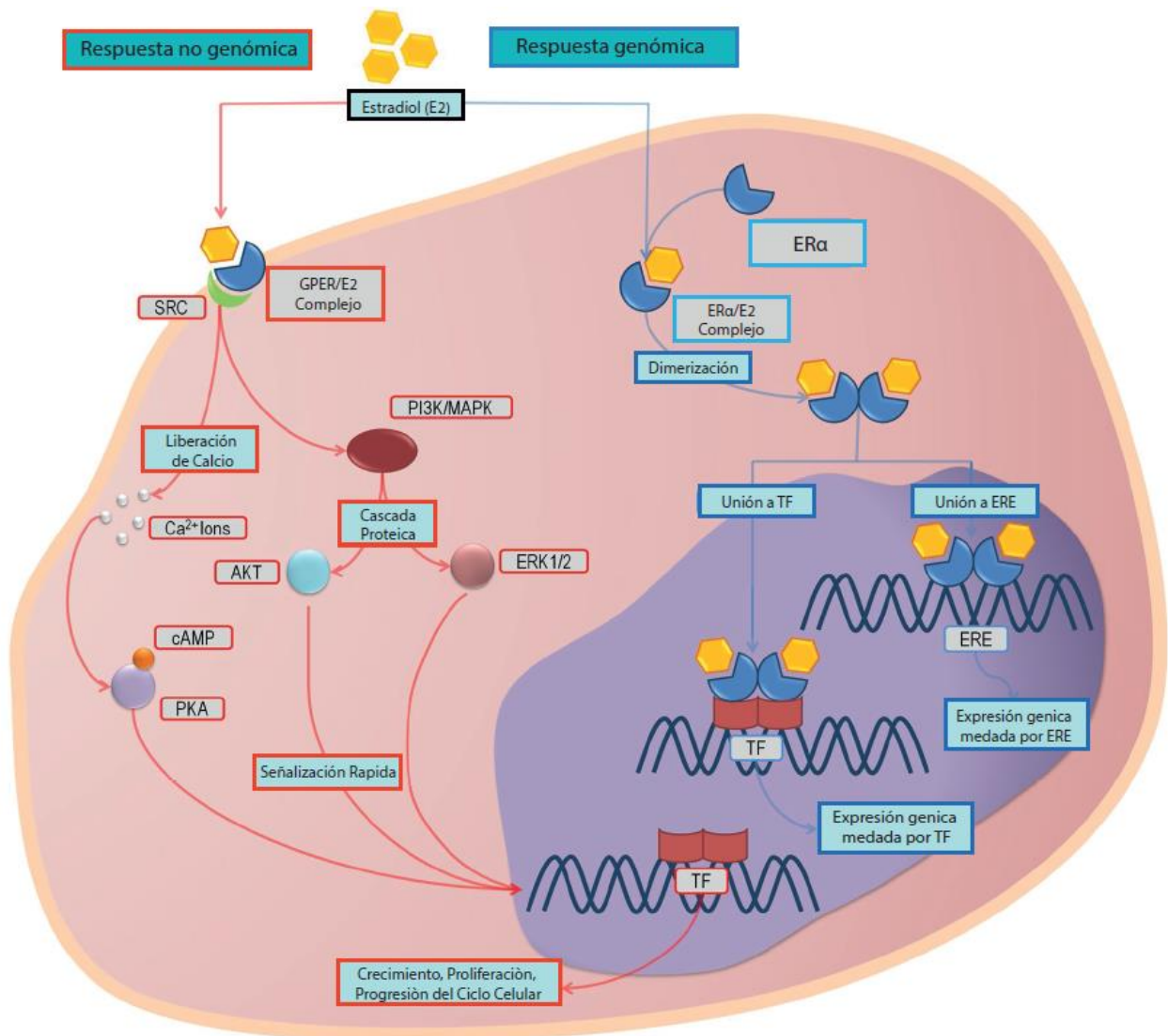


Figura 3: Representación Simplificada de la Vía de Señalización del Receptor de Estrógenos (ER) y su Respuesta Transcripcional. Se presenta un esquema que sintetiza las vías de señalización del estrógeno, diferenciando entre la vía no genómica (líneas rojas) y la vía genómica (líneas azules). Se resaltan las estructuras responsivas (recuadros sombreados en gris) y los procesos celulares (recuadros sombreados en azul cian). Imagen adaptada de: Słowikowski et al., 2017.

### 1.1.2 Vía directa genómica de señalización

En la vía directa, también conocida como la vía clásica, ambas isoformas del ER, ER $\alpha$  y ER $\beta$ , actúan como factores de transcripción activados por ligando. Cuando el ligando E2 se une a cualquiera de estas isoformas del receptor, se desencadena un cambio conformacional que facilita la posterior translocación al núcleo (Fuentes & Silveyra, 2019).

Un estudio bioinformático de escala genómica (genome wide screening) identificó aproximadamente 70,000 secuencias tipo ERE en los genomas humano y de ratón, de las cuales alrededor de 17,000 están ubicadas cerca de los sitios de inicio de la transcripción o TSS (del inglés *transcription start site*), y solo 660 son sitios conservados en ambas especies. La interacción del ER con estos sitios fue validada in vivo, subrayando la importancia del ER en la unión a secuencias promotoras y su papel como factor de transcripción (TF) en la regulación transcripcional de genes que responden a la vía de señalización.

Aunque las secuencias ERE comparten una alta similitud, su composición intrínseca puede influir en la afinidad del ER para unirse al ADN. Por ejemplo, el ER tiene una afinidad elevada por la secuencia ERE ubicada en el gen de la *vitelogenina A2*, pero una menor afinidad por las EREs situadas en el gen de la *oxitocina*.

Es evidente que las diferencias en estas secuencias, debido a la variabilidad genética del individuo o mutaciones, pueden impactar la activación de la expresión génica. Además, secuencias ERE específicas pueden inducir cambios alostéricos en la estructura del receptor, modificando la capacidad del complejo para reclutar coactivadores y factores de transcripción (Fuentes & Silveyra, 2019).

En resumen, estos hallazgos contribuyen a una comprensión más completa del papel regulador específico desempeñado por el ER $\alpha$  en células MCF-7, incluyendo los tipos de interacciones que establece en la cromatina y los efectos resultantes de dichas interacciones.



### 1.1.3 Vía indirecta genómica de señalización (transcriptional cross-talk)

Dentro de la vía del ER, como se mencionó anteriormente, se despliegan diversos mecanismos mediante los cuales este factor de transcripción establece interacciones con la cromatina. Uno de estos procesos es la vía genómica indirecta, donde el ER $\alpha$  requiere interactuar con otros elementos para lograr su unión y, potencialmente, ejercer su efecto regulador en la región diana.

Se estima que aproximadamente el 35% de los genes regulados por E2 carecen de secuencias tipo EREs asociadas a sus elementos regulatorios (Fuentes & Silveyra, 2019). La vía indirecta proporciona una explicación parcial de cómo opera este fenómeno regulador en el 35% de los genes que responden a E2, a pesar de no contar con motivos EREs en sus secuencias promotoras. En este contexto, el ER puede formar complejos multiproteicos con factores transcripcionales como SP1 (Stimulating Protein 1) y AP1 (Activator Protein 1).

SP1 desempeña un papel crucial como mediador de esta respuesta indirecta. Su unión a las regiones promotoras, especialmente en sitios ricos en nucleótidos GC, se ve potenciada por la presencia de ERs. Algunos ejemplos de genes inducidos por esta interacción incluyen el receptor de LDL, el receptor de progesterona, *eNOS*, *GATA1*, *STAT5* y el receptor del ácido retinoico 1- $\alpha$ . Otro escenario posible es que AP1, en colaboración con ER $\alpha$ , induzca la expresión de genes como *IGF1*, *colagenasa*, *receptor IGF1*, *ovalbumina* y *ciclina D1* (Fuentes & Silveyra, 2019).

Así, los complejos ER actúan en esta vía mediante interacciones proteína-proteína con otros factores transcripcionales y sus respectivos elementos responsivos. Esto culmina en un proceso de activación o supresión de genes diana bajo la influencia de E2.

En resumen, aunque esta vía brinda una comprensión parcial de los genes regulados por ER $\alpha$  a través de la unión mediada por otros factores y complejos, hay un considerable número de genes que, sin mostrar una conexión directa o indirecta con ER $\alpha$ , responden transcripcionalmente a esta vía. Esta respuesta podría atribuirse a la influencia de otros elementos reguladores activados por la vía de ER $\alpha$ , tales como los cambios en la estructura y organización de la



cromatina, así como la transcripción de ARN no codificante. Este último aspecto será abordado con mayor detalle próximamente, ya que constituye una de las premisas fundamentales en las que se basa la formulación de este proyecto.

#### 1.1.4 Vía no genómica

Este mecanismo de la vía no genómica del ER $\alpha$  se desarrolló para explicar las respuestas biológicas rápidas desencadenadas por el E2 en las células, donde el GPER1 desempeña un papel protagónico (Figura 3). Esta respuesta implica la activación de mecanismos de transducción de señales, con la consiguiente producción de segundos mensajeros intracelulares, como es el caso del AMP cíclico (cAMP). La regulación del cAMP y la activación de cascadas de señalización por parte de proteínas cinasas resultan en cambios indirectos en la expresión génica.

Cuando GPER1 se une a los estrógenos, estimula la activación de la adenilil ciclasa y del receptor para el factor de crecimiento epidermal (EGFR por sus siglas en inglés). La fosforilación subsiguiente de factores de transcripción por las proteínas cinasas puede alterar su función y capacidad para unirse a las secuencias genómicas, afectando así la expresión de genes. Algunos de estos factores susceptibles a este mecanismo incluyen Elk-1, CREB, C/EBP $\beta$  (del inglés *CCAAT-enhancer-binding protein beta*), el complejo NF- $\kappa$ B y la familia de STAT (del inglés *signal transducer and activator of transcription*).

Cabe destacar que tanto ER $\alpha$  como ER $\beta$  son blanco de la fosforilación por proteínas cinasas, incluyendo las MAPKs, lo que sugiere que los mecanismos no genómicos de los estrógenos pueden implicar la autorregulación de la expresión del receptor (Fuentes & Silveyra, 2019).

En resumen, el mecanismo de la vía de ER $\alpha$ , centrado en la acción de GPER1 y la activación de cascadas de señalización intracelular, ilustra cómo los estrógenos pueden inducir respuestas biológicas rápidas y, al mismo tiempo, influir en la expresión génica mediante la fosforilación de factores de transcripción clave. Este

proceso sugiere una interconexión entre la autorregulación de los receptores de estrógeno y su impacto en la expresión génica, proporcionando así una visión más completa de los mecanismos de acción de los estrógenos en las células.

#### 1.1.5 Genes que responden a la vía de señalización con estradiol (E2)

Una vez descritos los diversos mecanismos de las vías mediante los cuales las células responden a la estimulación con E2, resulta fundamental investigar los genes que desencadenan una respuesta transcripcional y desempeñan un papel crucial en la respuesta fisiológica de las células. Con este fin, se ha seleccionado un modelo celular en el que se ha estudiado ampliamente la respuesta transcripcional a E2, así como la vía del ER $\alpha$ .

La línea celular MCF-7, una línea celular de cáncer de mama humana, ha desempeñado un papel destacado en la investigación biológica y clínica. Se trata de una línea celular epitelial de cáncer de mama que ha sido ampliamente utilizada como modelo para estudiar la biología de esta patología y evaluar posibles tratamientos oncológicos. Estas células, conocidas por su elevada afinidad por el estrógeno, han sido fundamentales en la investigación de la señalización dependiente de estrógenos, abordando aspectos relacionados con la regulación de la proliferación celular y la apoptosis. Además, han contribuido significativamente al estudio de la resistencia a los tratamientos hormonales y a la identificación de nuevas dianas terapéuticas en esta enfermedad.

Los estrógenos, especialmente el estradiol, desempeñan un papel significativo en el crecimiento y desarrollo del cáncer de mama. Investigaciones, como las realizadas por R. Wang et al. (2019), han demostrado que ciertos genes en las células MCF-7 responden a los niveles de estradiol. Algunos ejemplos de estos genes y sus procesos biológicos incluyen:

1. Proliferación celular: *ciclina D1*, *c-fos*, *c-jun*, *c-myc*.
2. Apoptosis: *BCL2*, *BCL-xL*, *MCL1*.
3. Invasión y migración celular: *MMP1*, *MMP2*, *MMP9*.

#### 4. Angiogénesis: *VEGF*, *bFGF*.

Estos genes responden a la influencia del estradiol al unirse a receptores de estrógeno en la superficie celular. Esto desencadena la activación de cascadas de señalización intracelulares, que a su vez regulan la expresión génica.

En resumen, tras describir los mecanismos de respuesta a la estimulación con E2 y de elegir un modelo celular ampliamente estudiado, como la línea celular MCF-7, se comprende la importancia de entender la regulación de genes específicos. Esta célula tumoral, con su alta responsividad al estrógeno, ha sido crucial para la investigación del cáncer de mama, tanto en la evaluación de terapias oncológicas como en la identificación de nuevos blancos terapéuticos. El estradiol, un factor crítico en el crecimiento del cáncer de mama, influye en la expresión de genes clave relacionados con la proliferación, apoptosis, invasión, migración y angiogénesis.

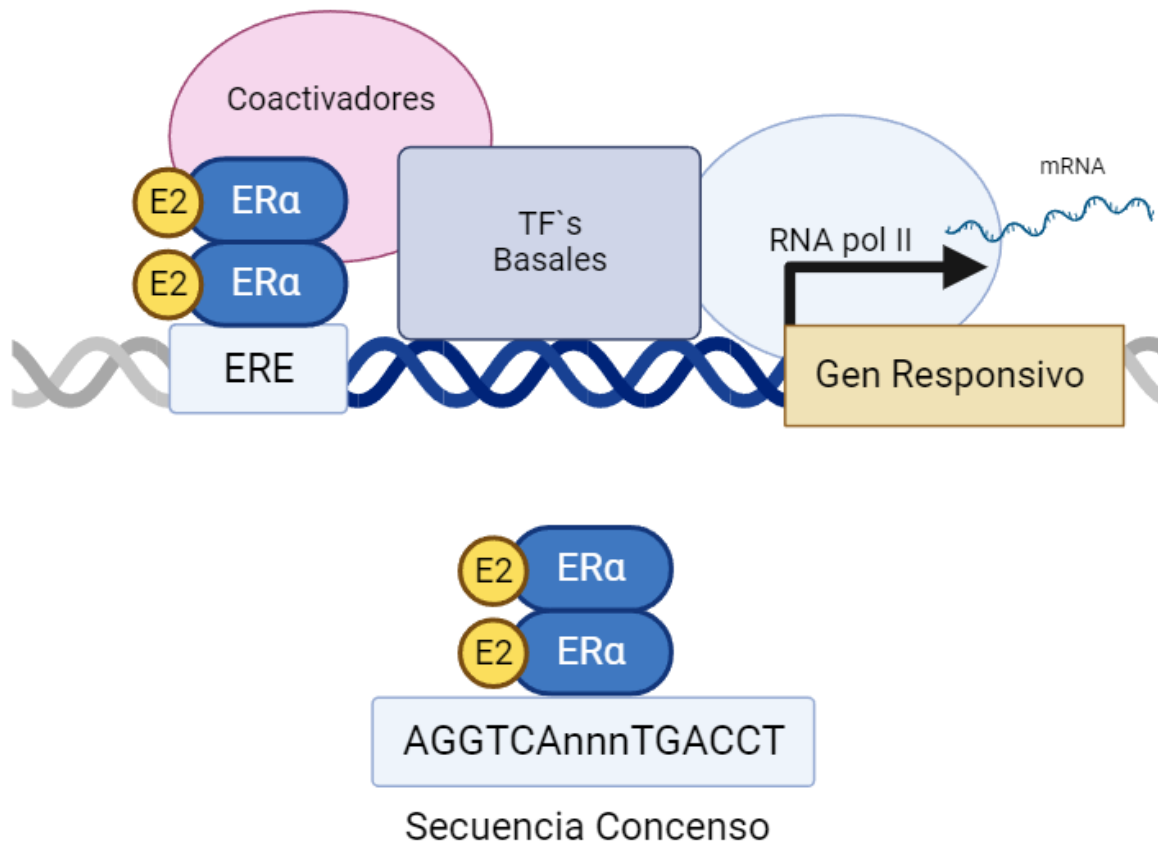
Estos hallazgos subrayan la necesidad de profundizar en la comprensión de cómo el estradiol regula estos genes mediante la interacción con receptores de estrógeno, desencadenando una serie de efectos, que incluyen mecanismos de regulación transcripcional así como efectos epigenéticos. Estos procesos son cruciales para determinar la expresión génica y, por ende, el fenotipo celular en el contexto del cáncer de mama. Cabe destacar que, hasta ahora, solo se ha explorado parcialmente esta compleja red de mecanismos.

#### 1.2 Respuesta transcripcional ante la estimulación con estradiol (E2)

Los motivos EREs se componen de una secuencia palindrómica de motivos 5'-AGGTCA-3' separados por un espaciador de 3 bases (Figura 4). Estos motivos son fundamentales para la respuesta genómica directa, puesto que permiten la unión directa de la región amino terminal del ER $\alpha$  en el genoma.

En estudios llevados a cabo en la línea celular MCF-7 de cáncer de mama luminal A, se observaron más de 1,000 genes responsivos a E2 que presentaron una

alteración en la expresión de al menos dos veces después de la exposición a E2 en diferentes tiempos (Ikeda et al., 2015).



Created in BioRender.com 

Figura 4: Mecanismo de regulación transcripcional por estrógenos. Unión de ER a las secuencias ERE. Imagen creada en BioRender.com con fines educativos.

Un aspecto relevante a tener en cuenta es que, según otro estudio, se identificaron 16,043 factores de transcripción que se unen al ERα en las células MCF-7. Además, se concluyó que la unión de los complejos factor de transcripción-ER al genoma está condicionada por la secuencia de los motivos de unión de estos factores transcripcionales. Esto es evidente en casos como los motivos de las proteínas FoxA1 y AP1, así como en la presencia de la secuencia ERE. Es importante destacar que los sitios de unión para los factores transcripcionales que carecen del motivo ERE representan tan solo el 2% de todas las uniones del receptor en la línea celular MCF-7. (Bojcsuk & Bálint, 2019).

Ahora bien, la secuencia responsiva puede variar respecto a las canónicas en un par de bases. Esto puede resultar en una disminución en la fuerza de unión de la proteína. Sin embargo, aunque el motivo ERE no esté tan conservado y tenga una menor fuerza de unión, el reclutamiento de otros factores puede ayudar a compensarla o incluso incrementarla (Bojcsuk & Bálint, 2019).

En este contexto, los resultados de la inmunoprecipitación de cromatina dirigida al elemento ER $\alpha$  revelaron la presencia de aproximadamente 80,000 sitios de unión del ER $\alpha$  ante la estimulación con E2 en células MCF-7. Estos sitios de unión al ER $\alpha$  se categorizaron según si mostraban cambios en la frecuencia de unión del ER $\alpha$  antes de la estimulación con E2 o si contenían el motivo ERE, dividiéndose en cuatro tipos:

- **Tipo 1:** El ER $\alpha$  se une a secuencias enhancers (potenciadoras de la transcripción) (Bojcsuk et al., 2017).
- **Tipo 2:** El ER $\alpha$  está unido al ADN antes del tratamiento con E2, a través de la secuencia ERE.
- **Tipo 3:** El ER $\alpha$  se une a través del motivo ERE, pero solo con la adición de E2.
- **Tipo 4:** El ER $\alpha$  se une al ADN en respuesta a E2, pero en ausencia de una secuencia ERE.

Un dato relevante es que una proporción significativa de todos los sitios de unión del receptor de estrógeno identificados corresponde al tipo 4 (54,408 de 81,818), donde la unión se establece a través de interacciones del receptor de estrógeno con otros factores transcripcionales u otros mecanismos no completamente elucidados, generalmente después de la estimulación con E2. Sin embargo, es importante señalar que estos sitios exhiben una afinidad de unión reducida, incluso con la adición de E2. Esta observación se confirmó mediante un mapa de calor que reveló la ausencia total de la secuencia de elementos de respuesta a estrógenos (ERE) en dichos sitios de unión. (Bojcsuk & Bálint, 2019).

Dado que la mayoría de los sitios de unión al receptor de estrógeno alfa (ER $\alpha$ ) carecen de secuencias de elementos de respuesta a estrógenos (ERE), y solo un

2% de la unión de este receptor está mediada exclusivamente por otros motivos de unión de factores transcripcionales en ausencia del motivo ERE, además de que el 35 % de los genes responsivos a la estimulación con E2 carecen de secuencias ERE asociadas, en este estudio se plantea la hipótesis de que parte de la regulación génica mediada por la vía del ER $\alpha$  podría llevarse a cabo mediante un mecanismo de la vía genómica indirecta que involucra la participación de lncRNAs. La literatura científica ha documentado la existencia de lncRNAs que responden a la activación de la vía de ER $\alpha$  y que, a su vez, desencadenan una serie de mecanismos regulatorios y epigenéticos implicados en la modulación transcripcional de los genes blanco de estos lncRNAs. (Sedano et al., 2020). No obstante, es crucial destacar que hasta el momento solo se ha explorado y caracterizado un pequeño subconjunto de este vasto universo de lncRNAs. Esto sugiere la existencia de numerosas funciones y mecanismos aún por descubrir, los cuales podrían poseer una gran relevancia en este contexto.

### 1.3 RNAs largos no codificantes (lncRNAs)

Los lncRNAs, transcritos de 200 nucleótidos o más, se distinguen de los RNAs mensajeros en que no se traducen a proteínas, aunque comparten características como la presencia de una cola poly(A) y un capuchón 5' (Dahariya et al., 2019).

Estos lncRNAs pueden clasificarse en distintos subtipos, según (J. Wang et al., 2020):

- a) **lncRNAs antisentido:** Son lncRNAs endógenos que son parcial o completamente complementarios a otras secuencias transcripcionales.
- b) **lncRNAs enhancers:** Se generan a partir de los enhancers de genes que codifican proteínas.
- c) **lncRNAs (RNAs largos intergénicos no codificantes):** Se codifican de manera independiente en las secuencias ubicadas entre los genes codificantes de proteínas.

- d) **lncRNAs bidireccionales:** Comparten la misma región promotora que los genes codificantes de proteínas, pero la dirección transcripcional del gen está invertida.
- e) **lncRNAs intrónicos:** Se producen a partir de los intrones de los genes.

Los lncRNAs desempeñan un papel clave en la regulación de la expresión génica al participar en procesos como la producción, procesamiento (splicing), degradación o traducción de los mRNA blanco. Esto se logra mediante su capacidad para interactuar con otros componentes clave, ya sea RNA/DNA, RNA/RNA o RNA/proteína (Atianand & Fitzgerald, 2014).

Aunque estos lncRNAs puedan desempeñar papeles regulatorios significativos en el citosol, nos enfocaremos específicamente en aquellos con funciones exclusivas en el núcleo para los propósitos de este proyecto.

Los lncRNAs nucleares tienen la capacidad de interactuar con diversos componentes nucleares, tales como la cromatina, factores de transcripción y proteínas vinculadas a la regulación epigenética. Este tipo de lncRNAs juega un papel crucial en procesos biológicos como la diferenciación celular, la senescencia celular y la apoptosis, al participar en mecanismos epigenéticos de regulación con un impacto directo en la expresión génica (Atianand & Fitzgerald, 2014). Estas interacciones se manifiestan a través de tres modalidades distintas:

**Modelo de señuelo (Decoy):** En este escenario, el lncRNA funciona como señuelo (o decoy en inglés), secuestrando enzimas modificadoras de la cromatina, tales como factores transcripcionales o represores, alejándolos de sus sitios originales de anclaje. Este proceso elimina su función enzimática, resultando en el silenciamiento o activación de genes (Bhan et al., 2017; Zhou et al., 2021).

**Modelo de guía:** En este caso, el lncRNA interacciona con ciertos complejos proteicos reguladores o enzimáticos y los dirige hacia sus correspondientes loci en la cromatina o el promotor del gen (Bhan et al., 2017; Zhou et al., 2021).

**Modelo de formación estructura andamio:** Esta función implica que el lncRNA actúa como un andamio en el cual varios factores o complejos proteicos pueden

anclarse, para luego ser dirigidos a su sitio diana. Por ejemplo, el lncRNA puede interactuar con dos enzimas simultáneamente, funcionando como un andamio que facilita la modificación de la cromatina (Zhou et al., 2021).

En resumen, los lncRNAs nucleares desempeñan un papel crucial en la regulación de la expresión génica y la modulación de procesos biológicos fundamentales. Su capacidad para interactuar con componentes nucleares, como la cromatina, los factores de transcripción y las proteínas epigenéticas, les confiere la capacidad de influir directamente en los mecanismos de regulación epigenética. Estos lncRNAs operan mediante distintos modelos funcionales, cada uno con implicaciones únicas en la modulación de la actividad génica. El modelo de señuelo, el modelo de guía y el modelo de formación de estructura andamio ofrecen diversas formas en las que los lncRNAs pueden dirigir o secuestrar complejos proteicos a ubicaciones específicas en la cromatina, alterando así la expresión génica. En conjunto, la versatilidad y la complejidad de las interacciones de los lncRNAs nucleares con los componentes nucleares subrayan su importancia como actores clave en la regulación epigenética y la función génica.

#### 1.4 Respuesta transcripcional indirecta de la vía p53 mediada por lncRNAs

La idea anterior encuentra respaldo en el ejemplo que se presenta a continuación. Datos publicados por Huarte y colaboradores identificaron lincRNAs (large intergenic non-coding RNAs) que están regulados por p53 (Huarte et al., 2010). Los lincRNAs forman parte de los lncRNAs ubicados entre secuencias de genes, es decir, en regiones intergénicas. Mediante un método de correlación entre la expresión de estos lincRNAs y la expresión de genes codificadores de proteínas, se observó un grupo de lincRNAs fuertemente asociado a la vía transcripcional de p53, un gen supresor tumoral implicado en mantener la integridad del genoma. Aunque se ha estudiado ampliamente cómo p53 activa la transcripción de genes, el proceso por el cual genera la represión génica había permanecido sin aclarar.

Se ha informado que los lincRNAs están físicamente vinculados a los complejos modificadores de cromatina, que constan de proteínas que alteran la estructura de



la cromatina y, por ende, regulan la expresión génica. Por ejemplo, el lncRNA denominado FIRRE se asocia con el complejo modificador de cromatina Polycomb PRC2 y regula la expresión génica en células madre embrionarias de ratón. Otro estudio ha identificado varios lncRNAs con funciones similares a las de los elementos enhancer que se asocian con complejos modificadores de cromatina (Núñez-Martínez & Recillas-Targa, 2022). En términos generales, se cree que los lncRNAs vinculados a complejos modificadores de cromatina pueden regular la expresión génica al modificar la estructura de la cromatina y controlar la accesibilidad de los factores de transcripción al DNA. Por lo tanto, es plausible que p53 ejerza su función represora al activar la transcripción de ciertos lincRNAs.

Se ha observado que los lincRNAs establecen conexiones físicas con los complejos modificadores de cromatina, formados por proteínas que modifican la estructura de la cromatina, ejerciendo así un control sobre la expresión génica. Un ejemplo destacado es el lncRNA conocido como *FIRRE*, el cual se asocia con el complejo modificador de cromatina Polycomb PRC2, desempeñando un papel regulador en la expresión génica de células madre embrionarias de ratón. Investigaciones adicionales han identificado varios lncRNAs que, al igual que los elementos enhancer, se vinculan con complejos modificadores de cromatina (Núñez-Martínez & Recillas-Targa, 2022). En términos generales, se sostiene la creencia de que los lncRNAs asociados a estos complejos pueden modular la expresión génica al alterar la estructura de la cromatina y regular la accesibilidad del DNA a los factores de transcripción. Por lo tanto, es plausible que p53 ejerza su función represora al activar la transcripción de ciertos lincRNAs.

En base a lo anterior, se emprendió una búsqueda computacional de regiones promotoras enriquecidas con motivos conservados de unión a p53. Se descubrió que los promotores de los lincRNAs inducidos por p53 presentan una abundancia significativa de estos motivos. Se seleccionaron dos lincRNAs, *lincRNA-p21* y *lincRNA-Mkl1*, que mostraron los motivos canónicos de unión más conservados.

El *lincRNA-p21*, ante el daño de ADN, se expresa en niveles elevados en fibroblastos humanos gracias a la unión de p53 a su promotor específico,

sugiriendo su participación en la vía de p53. Se llevó a cabo un knockdown por separado de p53 y *lincRNA-p21* mediante interferencia de ARN para explorar genes cuya expresión se viera afectada por ambos knockdowns, midiendo los cambios en los niveles de mRNA mediante microarreglos. De los genes analizados, se identificaron 930 genes comunes en respuesta al knockdown de ambos genes. Aproximadamente el 80% de estos genes comunes mostraron una sobreexpresión, indicando que el *lincRNA-p21* desempeña un papel en la regulación negativa de la represión transcripcional dependiente de p53. Esto sugiere que el *lincRNA-p21* ejerce un efecto represor sobre diversos genes vinculados con la apoptosis y la regulación del ciclo celular, a través de una respuesta transcripcional dependiente de p53 (Huarte et al., 2010).

En síntesis, estos descubrimientos subrayan la relevancia de los lncRNAs en la amplificación de vías celulares, como la vía de p53, para la regulación indirecta de genes asociados a esta vía. En determinadas circunstancias, la expresión de estos lncRNAs se torna crucial para la regulación de genes objetivos. Por lo tanto, es de suma importancia investigar posibles lncRNAs que exhiban una respuesta transcripcional ante la vía del receptor de estrógenos (ER $\alpha$ ) y que, simultáneamente, puedan desempeñar una función reguladora en la transcripción de otros genes implicados en la mencionada vía.

## 2. Antecedentes

Al explorar el panorama de los lncRNAs y su impacto en la regulación génica, resulta esencial dirigir nuestra atención a la comprensión de cómo estos procesos regulatorios podrían estar siendo orquestados por lncRNAs que responden a la vía del receptor de estrógenos.

### 2.1 lncRNAs responsivos a estradiol (E2)

Se ha descubierto que ciertos lncRNAs son regulados por estrógenos, desempeñando funciones cruciales en la interacción con correguladores para potenciar cambios transcripcionales de ER $\alpha$  (Sedano et al., 2020). Además de su papel en la regulación de genes codificantes, se ha demostrado que el ER $\alpha$  puede influir en la expresión de lncRNAs. Algunos ejemplos de lncRNAs regulados por ER $\alpha$  en modelos celulares de cáncer de mama incluyen:

- **HOTAIR:** Implicado en la regulación de la invasión y migración (Bhan et al., 2013).
- **MALAT1:** Involucrado en la regulación de la proliferación celular y la apoptosis (Aiello et al., 2016).
- **lincRNA-p21:** Participa en la regulación de la senescencia celular y la apoptosis (Y. H. He et al., 2021).

Estos lncRNAs regulados por ER $\alpha$  pueden ser importantes para la respuesta dependiente de estrógeno en tejidos como el tejido mamario y pueden tener implicaciones importantes en el desarrollo y el tratamiento del cáncer de mama. Sin embargo, se requieren más investigaciones para entender con más detalle el papel de estos lncRNAs en la respuesta estrógeno-dependiente y su potencial como dianas terapéuticas en el cáncer de mama.

Una técnica ampliamente empleada en la actualidad para la identificación de RNAs no codificantes de baja abundancia en transcritos es la secuenciación del transcriptoma RNA-seq. Experimentos previos con esta y otras técnicas han sido documentados en la literatura, específicamente en células MCF-7, donde se indujo

la respuesta ante el estradiol (E2) con el propósito de identificar el conjunto completo de RNAs no codificantes. Este enfoque será discutido con mayor detalle en la siguiente sección.

## 2.2 Estudio de expresión a escala del genoma completo (*GRO-seq*)

En el marco de esta investigación, se ha empleado la técnica de *GRO-seq* (Global Run-On Sequencing) para obtener una visión detallada de los procesos de transcripción que ocurren en un momento específico. En este enfoque, se aíslan los núcleos de las células y se congelan con el objetivo de capturar la transcripción en curso. Posteriormente, esta transcripción se reanuda *in vitro* a una temperatura de 30°C y se incorporan nucleótidos radioactivos para medir el RNA nascente mediante la elongación a cargo de la RNA polimerasa. Una vez completada esta etapa, el RNA se convierte en cDNA para la preparación de las librerías de secuenciación (Tzerpos et al., 2021). Estas técnicas ofrecen información valiosa sobre la dinámica de la transcripción y permiten desentrañar los eventos moleculares en tiempo real que ocurren en el proceso de regulación génica.

Las ventajas inherentes a esta técnica son notables, especialmente en comparación con el RNA-seq convencional. Una distinción fundamental radica en su capacidad para detectar transcritos en niveles reducidos, una característica particularmente relevante para muchos lncRNAs que suelen presentar estas características en el ámbito nuclear. Además, esta técnica permite la identificación del RNA transcrito en momentos temporales específicos, evitando que los transcritos acumulados de períodos previos interfieran en los resultados. Esta característica confiere una ventaja significativa al proceso de cuantificación de los transcritos generados en respuesta al efecto del estradiol (E2) en intervalos temporales específicos. Asimismo, esta metodología posibilita la detección de un número mayor de lncRNAs, especialmente aquellos presentes en concentraciones bajas, enriqueciendo la capacidad de análisis y comprensión de la regulación génica.

Empleando esta técnica, se llevó a cabo un análisis exhaustivo del transcriptoma de MCF-7 con el propósito de identificar qué transcritos están sujetos a regulación por el estradiol, ya sea incrementando o reduciendo su nivel de transcripción. Para este propósito, se examinó una ventana de 2 kb en la región de terminación 5' de cada transcrito. Esta elección se basó en la intención de capturar cambios en los primeros 10 minutos que aún no se hubieran extendido hacia la terminación 3' de los transcritos más largos. Sorprendentemente, los resultados revelaron que un 26% del transcriptoma presenta alteraciones en respuesta al tratamiento con E2 en al menos un punto de tiempo específico (Hah et al., 2011).

De manera intrigante, se evidenció que la mayoría de los transcritos experimentan regulación a los 40 minutos, con un equilibrio entre los niveles de aumento y disminución de la transcripción. No obstante, a los 160 minutos, predominan los transcritos con niveles reducidos (Hah et al., 2011). Esta observación insinúa que el intervalo de tiempo entre 10 y 40 minutos podría albergar los transcritos de regulación inmediata. Esta pauta temporal sugiere que, durante este lapso se activa la transcripción de lncRNAs que posiblemente desempeñen una función reguladora en otros genes a nivel transcripcional.

El análisis de GRO-seq revela la regulación de diversos transcritos no codificantes por E2, incluyendo aquellos de naturaleza divergente, antisentido e intergénicos, muchos de los cuales aún no tienen funciones totalmente esclarecidas. A través de un análisis integrativo de GRO-seq/RNA-seq, se logró identificar lncRNAs que responden a la estimulación con E2.

El proceso se dividió en tres etapas fundamentales:

1. Mapeo y ensamblaje de transcripciones de RNA mediante RNA-seq poliadenilado (poliA+) (transcripciones en estado estacionario; 0 y 3 horas en presencia de E2).
2. Integración de perfiles de RNA naciente mediante ensayos de GRO-seq (transcripciones primarias o unidades de transcripción; 0, 10, 25, 40 y 160 min de E2).

3. Procesamiento y filtrado de transcritos de lncRNAs considerando: longitud, cobertura, niveles de expresión y potencial de codificación.

En el estudio llevado a cabo por Sun y colaboradores, se identificaron un total de 1,888 lncRNAs, de los cuales el 38% no estaban anotados en ninguna de las cuatro bases de datos analizadas.

En resumen, este estudio representa un esfuerzo significativo que ha conducido a avances sustanciales en la caracterización de los transcritos de lncRNAs que experimentan cambios a nivel transcripcional en respuesta a la estimulación con estradiol a nivel celular. Sin embargo, en muchos casos, aún se desconoce su posible implicación en la regulación transcripcional de otros genes. De confirmarse esta implicación, estos lncRNAs se revelarían como actores cruciales en la regulación de la respuesta génica celular, ejerciendo una influencia sustancial en los procesos fisiológicos desencadenados por el estradiol.

### 2.3 La Familia de genes *FOX*

La familia de genes *FOX* se destaca por sus factores de transcripción, caracterizados por poseer un motivo de unión al ADN conocido como "caja de forkhead" o "motivo de forkhead". Estos factores desempeñan un papel crucial en una amplia gama de procesos biológicos, abarcando la regulación génica, la diferenciación celular, la proliferación, la apoptosis, la senescencia y la reparación del ADN. En humanos, se han identificado un total de 49 genes *FOX*, agrupados en 19 subfamilias según la similitud de secuencia y la estructura del dominio de forkhead, según informes previos (Kato et al., 2013). Cada subfamilia exhibe funciones biológicas distintas y presenta patrones específicos de expresión en diversos tejidos y etapas del desarrollo.

Es crucial resaltar que los genes de la familia *FOX* están estrechamente vinculados a diversas enfermedades humanas, que abarcan desde el cáncer hasta trastornos autoinmunitarios, cardiovasculares y neurológicos. La comprensión de la función de estos genes *FOX* y su contribución a la patogénesis de estas

enfermedades es un área activa de investigación en genética y genómica del cáncer, con implicaciones significativas para el diagnóstico y el desarrollo de terapias específicas (Kato et al., 2013).

### 2.3.1 La proteína FOXF1 y sus funciones

Hasta la fecha, se ha explorado exhaustivamente la función de FOXF1 en diversos contextos, con un enfoque particular en su papel en la progresión del cáncer colorrectal (CRC). La investigación ha revelado una correlación positiva entre la expresión de *FOXF1* y la progresión del CRC, incluida su capacidad para inducir metástasis. Además, se ha confirmado que FOXF1 impulsa la transición epitelial-mesenquimal (EMT) en las células de CRC al activar la expresión del gen *SNA11* (Wang et al., 2018).

Para respaldar estos descubrimientos, se llevaron a cabo diversos experimentos, como ensayos de invasión y migración celular, análisis de expresión génica y ensayos de trasplante ortotópico en modelos murinos. Los resultados demostraron que la inhibición de *FOXF1* tenía el potencial de reducir la capacidad invasiva y metastásica de las células de CRC (Wang et al., 2018).

Además de su papel en el CRC, FOXF1 es un componente clave en la formación de tumores en diversos tipos de cáncer, incluido el rabdomiosarcoma de fusión positiva (FP-RMS). En el contexto del FP-RMS, FOXF1 juega un papel crucial como regulador fundamental de la tumorigénesis. Es importante destacar que en la mayoría de los casos de FP-RMS, se ha identificado una proteína de fusión denominada PAX3-FOXO1, la cual activa la expresión de FOXF1, según lo demostrado en un estudio reciente (Wang et al., 2018).

Un aspecto fundamental de la influencia de FOXF1 en el FP-RMS radica en su capacidad para interactuar con los enhancers de varios genes asociados con la firma génica característica de este tipo de cáncer. Entre estos genes se incluyen *ELMO1*, *TOX3*, *FGFR2*, *MYCN* y *JARID2*, entre otros. Investigaciones recientes respaldan la idea de que la inhibición de *FOXF1* podría representar una estrategia

terapéutica efectiva para el tratamiento del FP-RMS. Se ha documentado que la inhibición de *FOXF1* reduce la proliferación celular y la formación de tumores en modelos murinos de FP-RMS. Además, esta estrategia terapéutica aumenta la sensibilidad de las células de FP-RMS a la quimioterapia, lo que sugiere su potencial aplicación en futuros enfoques terapéuticos (Milewski et al., 2021).

Un estudio de gran relevancia centró su atención en *FOXF1* para identificar células mesenquimales clave en el proceso de fibrogénesis en injertos pulmonares. Este estudio, realizado tanto en ratones como en seres humanos, se enfocó en la identificación de células mesenquimales en diversas ubicaciones anatómicas dentro del pulmón (Braeuer et al., 2021).

Los investigadores distinguieron dos grupos principales de células mesenquimales, caracterizados por su marcada expresión de *Col1a1* y diferenciados por su alta expresión de *Foxf1* o *Itga8*. Estos grupos fueron denominados *Foxf1\_MCs* e *Itga8\_MCs*, respectivamente. El estudio no solo se centró en la identificación de estas células mesenquimales, sino que también exploró el impacto de *FOXF1* en la regulación de la fibrosis pulmonar (Braeuer et al., 2021).

Los resultados revelaron que la sobreexpresión de *FOXF1* en células mesenquimales estaba asociada con un aumento significativo en la producción de colágeno y, en consecuencia, en el desarrollo de fibrosis en el tejido pulmonar de los ratones. Estos hallazgos arrojan luz sobre el papel fundamental de *FOXF1* en la regulación de procesos fibrogénicos en el pulmón y sugieren su potencial implicación en condiciones patológicas relacionadas con la fibrosis pulmonar (Braeuer et al., 2021).

#### 2.3.1.1 Papel de *FOXF1* en cáncer de mama.

La sobreexpresión de *FOXF1* en el cáncer de mama se asocia con la transición epitelial-mesenquimal (EMT), un proceso biológico en el cual las células epiteliales adquieren características de células mesenquimales. Esta transformación les



confiere la capacidad de migrar e invadir tejidos circundantes, y se ha vinculado con la progresión del cáncer y la formación de metástasis. Por consiguiente, la sobreexpresión de *FOXF1* en el cáncer de mama podría contribuir a la progresión de la enfermedad al promover la EMT (Kato et al., 2013).

En cuanto a la regulación a la baja de *FOXF1* en las células de cáncer de mama, esta se debe a la represión epigenética. La restauración de la expresión de *FOXF1* mediante enmascaramiento farmacológico se relaciona con la detención del ciclo celular en la fase G1. Además, la sobreexpresión de *FOXF1* se asocia con un fenotipo mesenquimal y un mayor potencial invasivo en el cáncer de mama. También induce la EMT, manifestada por un fenotipo mesenquimal y la detención del ciclo celular en la fase G1 en el cáncer de mama (Kato et al., 2013).

Por último, es importante destacar que las mutaciones inactivantes de *FOXF1* están relacionadas con la displasia capilar alveolar y otras malformaciones. La displasia capilar alveolar es una enfermedad pulmonar rara y grave que se presenta en recién nacidos y se caracteriza por la falta de desarrollo de los capilares pulmonares. La función normal de *FOXF1* en el desarrollo pulmonar y su papel en la patogénesis de la displasia capilar alveolar son temas de investigación activa en la actualidad (Kato et al., 2013).

### 3. Planteamiento del problema

En el marco de este estudio, nos enfocamos en la identificación de lncRNAs con una rápida respuesta transcripcional frente al estímulo de estradiol (E2), en un intervalo temporal de hasta una hora. Al mismo tiempo, exploramos la posible regulación de estos lncRNAs por parte del receptor de estrógenos ER $\alpha$ , mediada por su interacción con los promotores respectivos. La investigación posterior se centra en un lncRNA específico de esta categoría, con el propósito de examinar su capacidad potencial como regulador transcripcional en uno o varios genes, en el contexto de la línea celular MCF-7.

En el transcriptoma de la línea celular MCF-7, se ha observado que aproximadamente el 26% de los transcritos son objeto de regulación por el estradiol (E2). Dentro de este conjunto de genes, se identifican numerosos lncRNAs que responden a la vía del ER $\alpha$ , tanto aquellos que están previamente anotados en la literatura como nuevos descubrimientos.

Es relevante destacar la presencia de una cantidad significativa de genes codificantes de proteínas que, aunque no presentan enriquecimiento de ER $\alpha$  en la región promotora, muestran cambios notables en su expresión en respuesta a la estimulación con E2. A partir de estos hallazgos, el presente proyecto plantea la búsqueda de lncRNAs candidatos que podrían participar en la regulación transcripcional de genes responsivos a la vía del ER $\alpha$  durante la estimulación con E2, incluso si carecen de la interacción directa del ER $\alpha$  en su promotor. Se hipotetiza que estos lncRNAs podrían desempeñar un papel fundamental en el ajuste de la expresión de este conjunto de genes, añadiendo así un nivel adicional de regulación mediante la influencia de lncRNAs en la respuesta de la vía del ER $\alpha$ .

## 4. Hipótesis

Los lncRNAs, cuya expresión es provocada por la estimulación con estrógeno/estradiol, juegan un papel crucial en la compleja red de regulación transcripcional de la vía del receptor de estrógeno  $\alpha$  (ER $\alpha$ ). Se postula que estos lncRNAs activan la transcripción de genes que no exhiben una unión directa al ER en sus promotores, sugiriendo así su participación en la modulación de la expresión génica más allá de las interacciones tradicionales de este receptor con los elementos promotores de los genes.

## 5. Objetivo General

Identificar lncRNAs a través de análisis bioinformáticos y validación experimental, cuya transcripción se vea activada en respuesta a la estimulación con estrógeno/estradiol, con el propósito de elucidar su contribución al establecimiento de la red de regulación transcripcional de la vía del receptor de estrógeno  $\alpha$  (ER $\alpha$ ) en un modelo celular representativo de cáncer de mama luminal A.

## 6. Objetivos particulares

### 1. Identificación de lncRNAs responsivos a E2 regulados por ER $\alpha$ en la línea celular de cáncer de mama MCF-7.

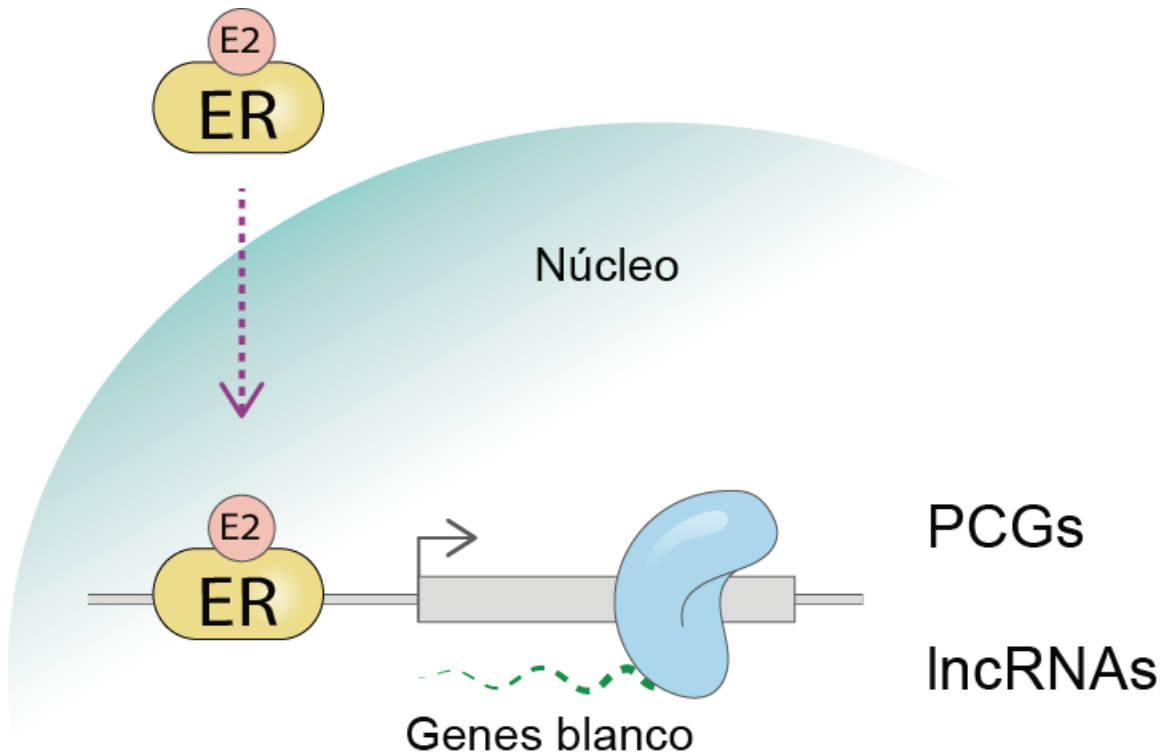


Figura 5: Mecanismo por el cual el ER $\alpha$  al unirse con E2, se une a secuencias promotoras promoviendo la transcripción de genes codificantes para proteínas (PCGs) y RNAs largos no codificantes (lncRNAs).

- 1.1. Identificar los sitios regulatorios de genes de lncRNAs en respuesta a la unión del ER $\alpha$ , mediante el análisis de datos de CHIP-seq antes y después de 40 minutos de tratamiento con E2.
- 1.2. Identificar los transcritos de lncRNAs con mayores variaciones transcripcionales en respuesta a E2, específicamente después de 1 hora de tratamiento.
- 1.3. Realizar una integración de las dos bases de datos obtenidas para identificar aquellos lncRNAs que presenten una unión al ER $\alpha$  en una proximidad de +/- 5 kbs con respecto al TSS. Esta selección se

complementará con el requisito de una diferencia significativa en el aumento de la expresión en respuesta a la estimulación con E2.

## 2. Identificación de lncRNAs con Potencial Regulatorio en Cis en Genes Codificantes Respondientes a E2, Independientes de la Vía de ER.

- 2.1. 2.1. Identificar lncRNAs candidatos que presenten una correlación positiva con la expresión de genes codificantes en respuesta a E2, sin estar directamente asociados a la vía de ER.
- 2.2. 2.2. Evaluar la interacción espacial entre los lncRNAs candidatos y los promotores de los genes blanco, verificando que ambos componentes muestren un aumento transcripcional en presencia de E2.

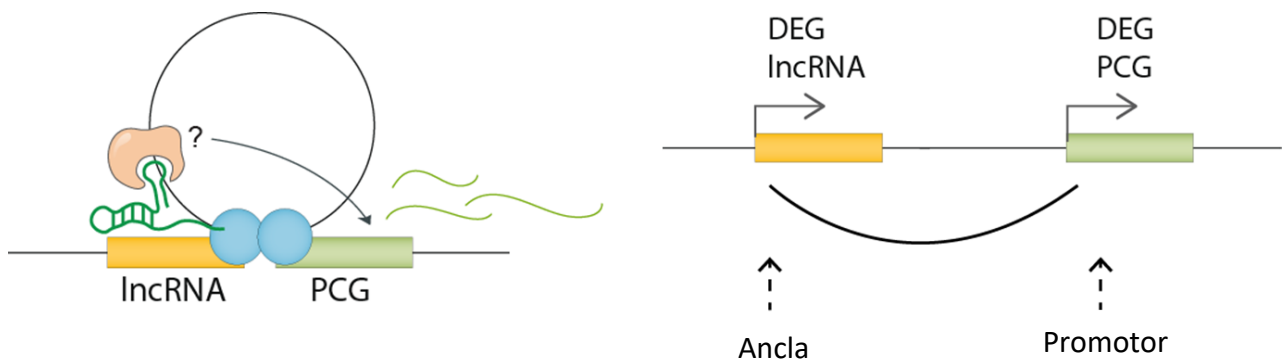


Figura 6: Representación del mecanismo por el cual los genes de los lncRNAs establecen contacto con los promotores de genes codificantes para proteínas (PCG) en la línea celular MCF-7. Este hallazgo sugiere la posibilidad de que el transcrito del lncRNA desempeñe una función regulatoria sobre el gen con el que entra en contacto.

### 3. Caracterización Funcional de un lncRNA Candidato y su Impacto en el Gen Codificante Asociado.

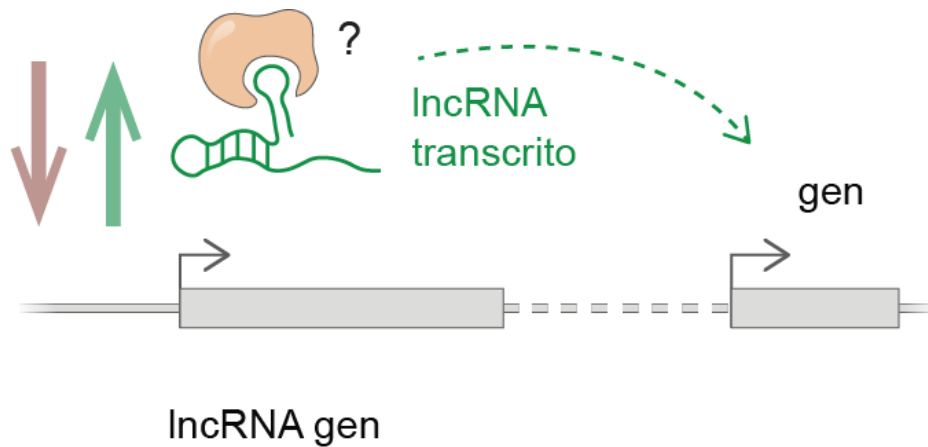


Figura 7: Ilustración de la cuestión planteada sobre las posibles implicaciones al modificar los niveles de transcripción del lncRNA de manera ectópica, en relación a los niveles de transcripción del gen al que está vinculado.

3.1. Seleccionar un lncRNA candidato con base en los criterios establecidos y explorar sus efectos sobre el gen codificante vinculado.

3.2. Implementar técnicas de sobreexpresión y represión del lncRNA seleccionado utilizando el sistema CRISPRi para la inhibición y CRISPRa para la activación, respectivamente.

3.3. Evaluar las implicaciones funcionales en el gen codificador, investigando posibles alteraciones en sus niveles de expresión como respuesta a las manipulaciones en los niveles de expresión del lncRNA.

## 8. Resultados

8.1 Objetivo 1: Identificación de lncRNAs responsivos a E2 y regulados por ER $\alpha$  en la línea celular de cáncer de mama MCF-7.

### 8.1.1 Análisis bioinformáticos

Se llevaron a cabo una serie de análisis bioinformáticos para cumplir con el **objetivo 1**, que consiste en la **identificación de lncRNAs responsivos a E2 y regulados por ER $\alpha$  en la línea celular de cáncer de mama MCF-7**.

En términos simples, los datos bioinformáticos son información especial que obtenemos cuando estudiamos células y genes. Sirviendo como pistas o huellas que nos ayudan a entender cómo funcionan las células

### 8.1.2 Búsqueda Exhaustiva en Bases de Datos Públicas

Se llevó a cabo una búsqueda intensiva en bases de datos públicas para identificar lncRNAs que responden a la estimulación con E2 y presentan al ER $\alpha$  como factor de transcripción en su promotor. Además, se buscó aquellos lncRNAs con cambios en su expresión ante la estimulación con E2.

### 8.1.3 Fuentes de Datos Utilizadas

Se utilizaron conjuntos de datos provenientes de experimentos de ChIP-seq y RNA-seq en células de cáncer de mama MCF-7, realizados en condiciones de ausencia y presencia de E2 durante lapsos específicos. También se analizaron datos de ATAC-seq en células estimuladas con E2.

A continuación, explicaremos la utilidad de los datos bioinformáticos utilizados, así como el fundamento de las técnicas experimentales de donde se obtuvieron.

#### 8.1.3.1 ChIP-seq (Secuenciación de Inmunoprecipitación de la Cromatina)

La técnica ChIP-seq (Chromatin Immunoprecipitation followed by Sequencing) es una herramienta poderosa en la investigación genética. Su objetivo principal es entender qué regiones del ADN interactúan con proteínas específicas dentro de las células.

La técnica consiste en los siguientes pasos:

1. **Inmunoprecipitación (IP):** Se añaden anticuerpos específicos a las proteínas de interés (por ejemplo, el receptor de estrógeno alfa, ER $\alpha$ ) a las células. Estos anticuerpos actúan como 'imanes' que se adhieren a las proteínas específicas.
2. **Fijación Cruzada:** Las células se tratan para conservar la interacción entre las proteínas y el ADN.
3. **Fragmentación del ADN:** Se corta el ADN en pequeños fragmentos.
4. **Purificación:** Se purifica el ADN unido a las proteínas de interés.
5. **Secuenciación:** Se secuencia el ADN para identificar las regiones específicas que estaban vinculadas a las proteínas.

En el caso del ER $\alpha$ , el ChIP-seq es especialmente útil para entender cómo este receptor se une al ADN y qué genes pueden estar siendo regulados en respuesta a la presencia del estrógeno. Puede revelar las ubicaciones precisas en el genoma donde el ER $\alpha$  está interactuando, proporcionando información clave sobre los procesos de regulación génica relacionados con la respuesta al estrógeno.

#### 8.1.3.2 RNA-seq (Secuenciación del ARN)

La técnica RNA-seq (Secuenciación de ARN) es una herramienta avanzada que permite analizar de manera global la expresión génica en una muestra biológica. A continuación, se describe brevemente en qué consiste y su utilidad para estudiar el receptor de estrógeno alfa (ER $\alpha$ ).



### **Principio básico de RNA-seq:**

1. **Extracción de ARN:** Se extrae el ARN total de las células o tejidos de interés. El ARN representa las moléculas que transmiten la información genética desde el ADN y se utiliza como un reflejo de la actividad génica.
2. **Conversión a cDNA:** El ARN se convierte en ADN complementario (cDNA). Este paso es crucial para su posterior secuenciación.
3. **Secuenciación:** El cDNA se secuencia, generando millones de fragmentos de secuencias cortas que representan los diferentes transcritos de ARN presentes en la muestra.
4. **Análisis bioinformático:** Las secuencias se ensamblan y alinean con el genoma de referencia para determinar la abundancia relativa de cada transcripción, permitiendo cuantificar la expresión de cada gen.

### **Utilidad para estudiar el ER $\alpha$ :**

- Identificación de genes regulados: RNA-seq permite identificar qué genes están siendo activados o desactivados en respuesta a la presencia del receptor de estrógeno alfa (ER $\alpha$ ). Esto proporciona información crucial sobre los efectos de la señalización hormonal en el nivel de expresión génica.
- Perfil completo de expresión génica: Permite obtener un perfil completo de cómo la expresión de todos los genes en el genoma está siendo modulada por la actividad del ER $\alpha$ . Esto ayuda a comprender los mecanismos moleculares y los procesos biológicos que están siendo regulados.
- Descubrimiento de nuevos genes relacionados: RNA-seq no solo confirma la expresión de genes conocidos, sino que también puede descubrir nuevos genes que están regulados por el ER $\alpha$ , lo que contribuye a una comprensión más completa de la red de regulación génica.

### 8.1.3.3 ATAC-seq (Ensayo para Cromatina Accesible mediante Transposasa y secuenciación)

ATAC-seq es una técnica de secuenciación que se utiliza para mapear las regiones accesibles de la cromatina en el genoma. La sigla ATAC se refiere a "Assay for Transposase-Accessible Chromatin," lo que significa "Ensayo para Cromatina Accesible por Transposasa" en español.

#### **Principio Básico de ATAC-seq:**

1. Preparación de Cromatina: Se inicia con células que contienen cromatina, una estructura formada por ADN y proteínas. La cromatina se extrae y se trata con una enzima llamada transposasa.
2. Transposición: La transposasa corta y agrega adaptadores a regiones accesibles de la cromatina. Las regiones más accesibles son aquellas donde la cromatina está menos compactada, lo que generalmente coincide con regiones reguladoras y promotoras de genes.
3. Secuenciación: La cromatina modificada se somete a secuenciación de nueva generación, generando fragmentos cortos que representan las regiones accesibles de la cromatina.
4. Análisis Bioinformático: Las secuencias obtenidas se mapean al genoma de referencia para identificar las regiones donde la cromatina es más accesible. Esto proporciona un perfil de accesibilidad cromatínica.

Utilidad para estudiar el ER $\alpha$  y Cambios de Accesibilidad por Estimulación con E2:

- Mapeo de Sitios de Unión de ER $\alpha$ : ATAC-seq se utiliza para mapear sitios de unión del receptor de estrógeno alfa (ER $\alpha$ ) en el genoma. Estos sitios corresponden a regiones donde el ER $\alpha$  se une para modular la expresión génica.
- Cambios de Accesibilidad por E2: La estimulación con estradiol (E2), una forma de estrógeno, induce cambios en la accesibilidad cromatínica. ATAC-seq permite identificar las regiones que experimentan cambios en su accesibilidad después de la estimulación con E2, lo que indica las regiones reguladoras activadas por el ER $\alpha$  en respuesta a las señales hormonales.

- Identificación de Regiones Reguladoras: ATAC-seq facilita la identificación de elementos reguladores clave, como enhancers y promotores, que son fundamentales para controlar la expresión génica mediada por el ER $\alpha$ .
- Caracterización de la Cromatina: Proporciona una visión global de la estructura de la cromatina, destacando las regiones que pueden ser importantes en la regulación de genes asociados con la respuesta al estrógeno.

En resumen, ATAC-seq es una técnica valiosa para estudiar la accesibilidad cromatínica y, en el contexto del ER $\alpha$ , ayuda a comprender cómo las señales hormonales, como la estimulación con E2, afectan la estructura de la cromatina y la regulación génica asociada con el receptor de estrógeno alfa.

La primera parte del análisis bioinformático se llevó a cabo utilizando conjuntos de datos obtenidos de experimentos de ChIP-seq (Baran-Gale et al., 2016) y RNA-seq (Franco et al., 2015) realizados en células de cáncer de mama MCF-7. Estos experimentos se realizaron tanto en condiciones de ausencia como de presencia de estimulación con E2 durante períodos específicos: 40 minutos para ChIP-seq y 1 hora para RNA-seq. Las concentraciones de E2 utilizadas fueron de 100 nM para ChIP-seq y 10 nM para RNA-seq, como se muestra en la Figura 8A. Además, se realizaron análisis de datos de ATAC-seq (Guan et al., 2019) en células estimuladas con E2 durante 45 minutos a una concentración de 1 nM.

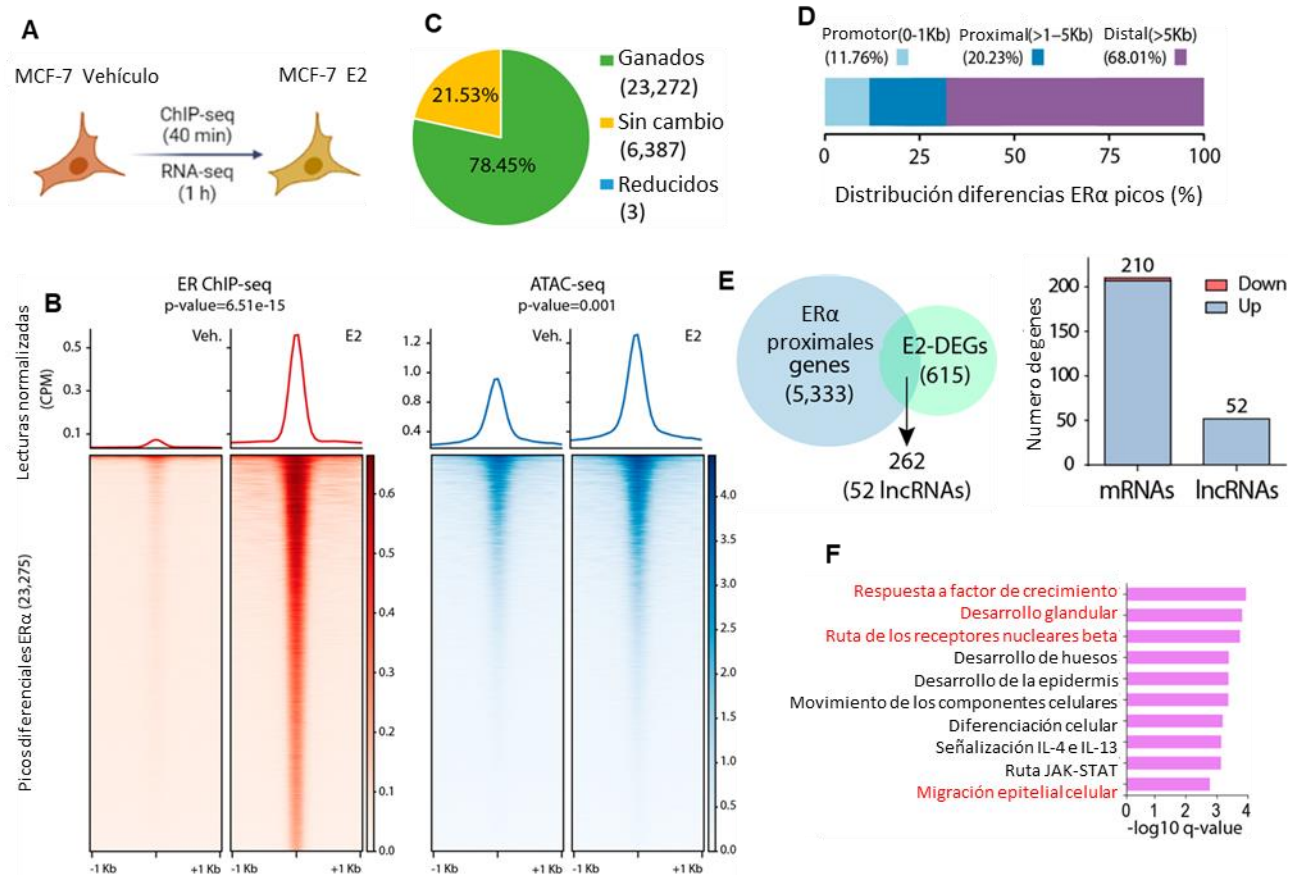


Figura 8. Análisis bioinformático para la identificación de lncRNAs regulados por el ER en respuesta al estradiol en células MCF-7. Los datos utilizados para los análisis bioinformáticos se obtuvieron de fuentes previas, incluyendo ChIP-seq de (Franco et al., 2015), RNA-seq de (Baran-Gale et al., 2016), y ATAC-seq de (Guan et al., 2019). Panel A: Representación esquemática del experimento en células MCF-7 bajo diferentes condiciones de estimulación con E2, tanto en términos de concentración como de tiempo, para la generación de los conjuntos de datos de RNA-seq y ChIP-seq. Panel B: Mapas de calor que ilustran el aumento en la señal de unión del ER $\alpha$  a la cromatina en células MCF-7 después de la estimulación con E2 (en rojo) y el incremento en la accesibilidad de la cromatina posterior al tratamiento con E2. Panel C: Gráfico de pastel que presenta los resultados de ChIP-seq después del tratamiento con E2, mostrando las regiones con enriquecimiento de unión del ER $\alpha$  (en verde), las regiones sin cambios significativos en la unión del ER $\alpha$  (en amarillo) y las regiones que perdieron la unión del ER $\alpha$  (en azul). Panel D: Análisis de las regiones enriquecidas en relación con la ubicación de los TSS, que se dividen en regiones promotoras (0 a 1 kb), regiones proximales (0 a 5 kb) y regiones distales (más de 5 kb). Panel E: En la parte izquierda, un diagrama de Venn que muestra la intersección de las bases de datos entre las regiones promotoras/proximales con unión del ER $\alpha$  (en azul) y los genes con expresión diferencial (DEGs). En la parte derecha, un diagrama de barras que indica mRNAs y lncRNAs que aumentaron su nivel de transcripción y que presentan unión del ER $\alpha$  (en azul) frente a los

que no presentaron unión del ER $\alpha$  (en rojo). Panel F: Representación de Ontología de Genes de los 262 genes con unión del ER $\alpha$  y cambio en la expresión transcripcional.

Iniciamos nuestra evaluación utilizando datos procedentes de experimentos de ChIP-seq (40 minutos) y RNA-seq (1 hora) realizados en condiciones de tratamiento con vehículo y E2 (Figura 8A). Estos conjuntos de datos fueron recopilados de recursos públicos proporcionados por los grupos de investigación (Franco et al., 2015) y (Baran-Gale et al., 2016).

A continuación, presentamos un mapa de calor que ilustra el incremento general en la unión del ER $\alpha$  a la cromatina después del tratamiento con E2 (Figura 8B, mapa de calor en rojo). En este mapa, se observa un enriquecimiento en la señal de unión de ER $\alpha$  a la cromatina, representado por un aumento en la intensidad del color rojo. Este aumento en la señal de unión sugiere que la estimulación con E2 desempeña un papel crucial en la interacción del ER $\alpha$  con la cromatina.

Adicionalmente, exploramos los impactos en la accesibilidad utilizando datos de ATAC-seq recopilados a los 45 minutos de la estimulación con E2, y observamos un aumento generalizado en la accesibilidad cromatínica posterior a la estimulación (Figura 8B, mapa de calor en azul). En este mapa, se evidencia un enriquecimiento en la señal que indica un aumento en la accesibilidad cromatínica, representado por una intensificación del color azul. Este hallazgo sugiere que la estimulación con E2 desempeña un papel clave en la generación de un incremento en la accesibilidad cromatínica a nivel global.

Es importante destacar que el aumento en la accesibilidad cromatínica, representado por un mayor color azul en el mapa de calor, indica por lo general una mayor actividad transcripcional en esas regiones. Este fenómeno se asocia a una ganancia de accesibilidad, convirtiéndose así en un factor de interés fundamental para nuestro estudio. Estos datos fueron obtenidos de fuentes públicas proporcionadas por el grupo de investigación (Guan et al., 2019).

La intensificación de la señal de unión del ER $\alpha$  en la cromatina se manifestó en una proximidad de 1 kb (+/-) con respecto al centro del pico del receptor de

estrógeno, simultáneamente con un aumento en la accesibilidad cromatínica (Figura 8B).

Al analizar los sitios de unión del ER $\alpha$ , se observó que aproximadamente el 80% de estos sitios experimentaron un enriquecimiento en la unión a la cromatina, el 22% no mostró cambios significativos (ni enriquecimiento ni pérdida de unión de ER $\alpha$ ), y solo el 3% presentó una pérdida de la unión de ER $\alpha$ , todo ello posterior a la estimulación con E2 (Figura 8C).

Dado nuestro interés en comprender las regiones del genoma donde el ER $\alpha$  podría tener un impacto en el aumento transcripcional debido a su función como factor de transcripción, seleccionamos el primer conjunto que ganó unión (80% de los sitios). Clasificamos estos sitios según su ubicación en el genoma: Promotor (0-1 kb), Proximal (>1-5 kb) y Distal (>5 kb). Nos centramos especialmente en las regiones Promotor y Proximal, que abarcan aproximadamente el 30% de los sitios totales de unión del ER $\alpha$  y podrían estar asociadas a regiones promotoras de genes. Esto nos proporciona información valiosa para identificar genes que podrían estar siendo regulados transcripcionalmente por la unión del ER $\alpha$  a sus regiones promotoras (Figura 8D). Sin embargo, para validar esta hipótesis, se requerirán análisis adicionales de cambios en la expresión génica en respuesta al estímulo con E2, lo que ayudará a delimitar nuestro conjunto de genes de interés.

De los 23,272 sitios que ganan unión de ER $\alpha$  después del estímulo con E2, identificamos que 5,333 están asociados con regiones Promotor/Proximal con respecto a los genes. Esta información nos permitió la identificación de los genes vinculados a estas regiones Promotor/Proximal basándonos en su proximidad física. Luego, llevamos a cabo un cruce de esta lista con los datos de RNA-seq (secuenciación del transcriptoma) obtenidos después de 1 hora de tratamiento con E2. Estos datos proporcionan una visión integral de los genes en las células MCF-7 que experimentan un cambio transcripcional significativo tras la estimulación con E2 durante 1 hora. De la intersección entre estas dos bases de datos, identificamos 262 genes que presentaron un aumento en la unión de ER $\alpha$  en sus regiones promotoras y, simultáneamente, experimentaron un cambio

transcripcional significativo en respuesta a la estimulación con E2. Dentro de este conjunto, resaltamos la identificación de 52 lncRNAs de respuesta rápida (Figura 8E). Es notable que el 100% de estos lncRNAs exhibieron un incremento en su expresión en respuesta al estímulo con E2 durante 1 hora, sugiriendo una posible regulación transcripcional positiva mediada por la unión de ER $\alpha$ .

Concluimos nuestro estudio llevando a cabo un análisis de ontología génica centrado en los 262 genes que logramos identificar. Este minucioso análisis nos brindó insights significativos sobre los procesos moleculares y celulares relacionados con la regulación inducida por la unión del receptor de estrógeno alfa (ER $\alpha$ ) y la estimulación con estradiol (E2) durante el periodo de una hora. De manera notable, se destacó la proliferación celular como uno de los procesos más prominentes en este contexto.

La proliferación celular, conocida por estar estrechamente vinculada a la vía de ER, emergió como un indicador clave de la efectividad de nuestros análisis bioinformáticos en la identificación de lncRNAs de particular interés. Este proceso biológico, esencial en el crecimiento y desarrollo celular, ha sido consistentemente asociado a la activación de la vía de ER. La representación gráfica de estos hallazgos se muestra de manera detallada en la Figura 8F.

En síntesis, al destacar la proliferación celular como uno de los resultados sobresalientes de nuestro análisis de ontología génica, reforzamos la relevancia de los lncRNAs identificados en la regulación de procesos biológicos críticos, consolidando así nuestra comprensión de la intrincada red de regulación génica que opera en el contexto específico del cáncer de mama luminal A, en respuesta a la acción de ER $\alpha$  y E2. Estos resultados proporcionan una valiosa perspectiva para futuras investigaciones y comprensión más profunda de los mecanismos subyacentes en esta enfermedad.

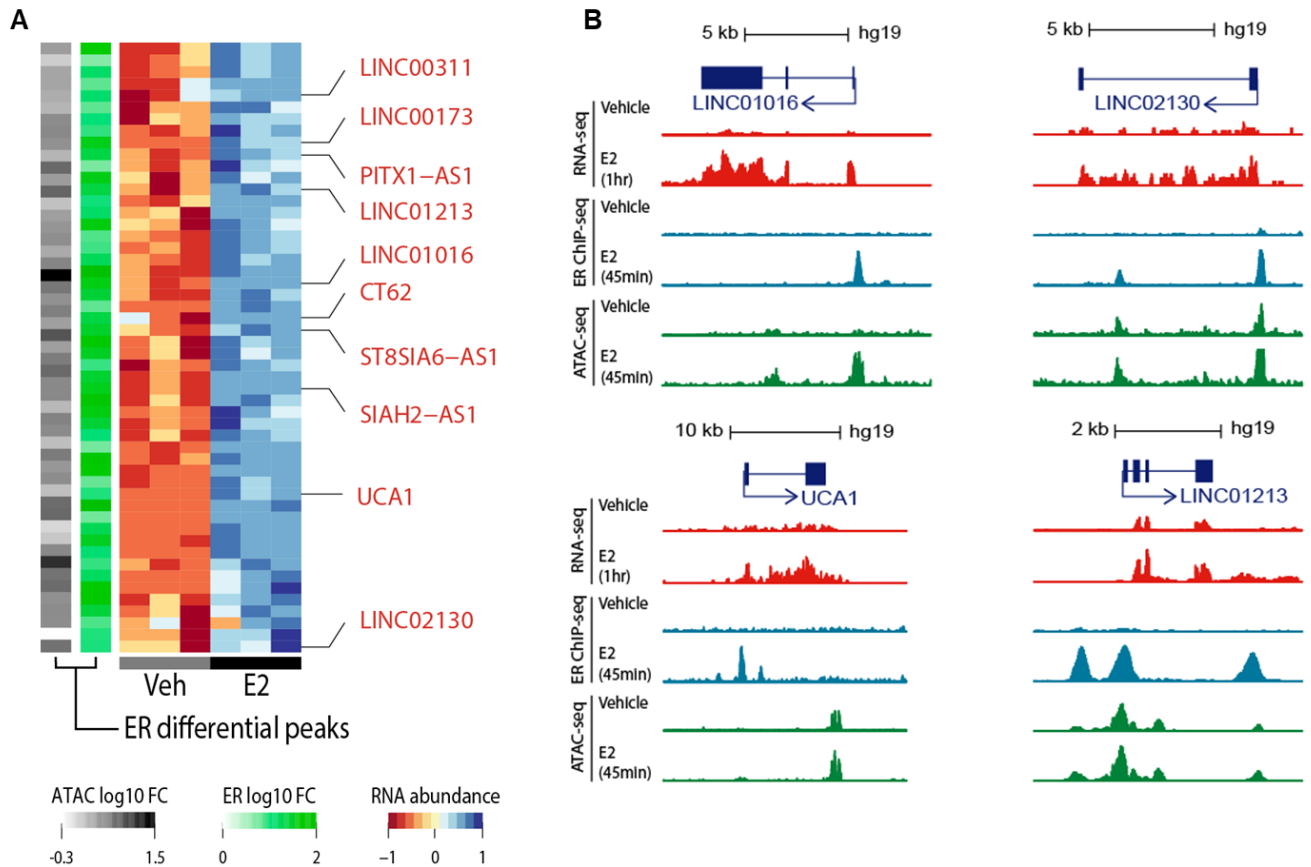


Figura 9. Identificación de lncRNAs con unión del ER $\alpha$  y aumento en la transcripción en respuesta a la estimulación con estradiol (1 h) en células MCF-7. En el panel izquierdo (A), se resaltan los lncRNAs identificados mediante el análisis bioinformático, algunos de los cuales han sido previamente mencionados en la literatura científica. En el panel derecho (B) se presentan representaciones gráficas de los criterios utilizados para su identificación tras la estimulación con E2, que incluyen la unión cercana del ER $\alpha$  al TSS, el incremento en la expresión transcripcional y la ganancia de accesibilidad de la cromatina.

Llevamos a cabo un análisis minucioso de los 52 genes que corresponden a lncRNAs en términos de Cambio Diferencial de Expresión de Genes (DGE) (Figura 9A). En el mapa de calor generado con estos 52 lncRNAs, se pueden identificar diversas características, como el aumento en la accesibilidad cromatínica en la región de unión del ER $\alpha$ , representado por la intensidad de color gris.

En la intensidad de color verde, se aprecia el enriquecimiento de la señal de unión de ER $\alpha$  después de la estimulación con E2 en cada promotor de los 52 lncRNAs. Los lncRNAs transcripcionalmente reprimidos se muestran en rojo, mientras que



aquellos que experimentaron ganancia transcripcional se destacan en azul. Esta representación abarca dos columnas: la primera corresponde a las células MCF-7 tratadas con vehículo, y la segunda muestra las células después de la estimulación con E2 durante una hora. Un hallazgo interesante es que, tras el tratamiento con E2, todos los lncRNAs muestran ganancia transcripcional. Este diagrama resalta todos estos elementos clave para la selección de los mejores candidatos a ser evaluados o estudiados en detalle.

En los cuatro ejemplos de la parte derecha, se evidencian los criterios fundamentales para su selección. Se busca que los genes presenten una unión de ER $\alpha$  muy próxima al Sitio de Inicio de la Transcripción (TSS), con un cambio de señal de unión notable y un aumento de accesibilidad en esta zona, en concordancia con la señal de ER $\alpha$ . En términos de la señal de transcripción, se espera que esté localizada en los exones del gen, además de experimentar un cambio sustancial tras el estímulo con E2.

Es importante destacar que, de los 52 genes identificados, 10 de ellos se han reportado en la bibliografía y solo algunos de estos se ha validado experimentalmente que responden al tratamiento con E2, según la literatura. A continuación, se presentan cuatro ejemplos con una distribución genómica de especial interés (Figura 9B). Estas representaciones muestran cómo la unión de ER $\alpha$  se ubica en las proximidades del Sitio de Inicio de la Transcripción (TSS), seguida por un aumento en la transcripción y, finalmente, una ganancia en la accesibilidad cromatínica.

El LINC01016 (ver Figura 9, Panel B, superior izquierda) es un ARN largo no codificante (lncRNA) que ha sido objeto de estudio en diversos tipos de cáncer. En el cáncer de endometrio, se ha observado una sustancial sobreexpresión de LINC01016, y su silenciamiento ha demostrado inhibir el comportamiento maligno de las células de cáncer de endometrio (Pan et al., 2018).

En relación con el estradiol, se ha evidenciado que LINC01016 muestra una sobreexpresión en el subtipo de cáncer de mama positivo para el receptor de estrógeno (ER), lo que podría estar asociado con un mal pronóstico (Sun et al.,

2023). No obstante, los detalles precisos sobre cómo LINC01016 interactúa con el estradiol o el receptor de estrógeno aún no están completamente esclarecidos.

LINC01213 (ver Figura 9, Panel B, inferior derecha) ha sido vinculado al cáncer de mama y al cáncer de próstata. En el contexto del cáncer de próstata, se ha identificado que LINC01213 desempeña un papel crucial en la transición de las células de cáncer de próstata, inicialmente dependientes de andrógenos, hacia formas independientes de andrógenos (Guo et al., 2022). Se ha evidenciado que LINC01213 presenta una sobreexpresión en las células de cáncer de próstata, y dicha sobreexpresión activa la vía de señalización Wnt/ $\beta$ -catenina (Guo et al., 2022). A pesar de estos hallazgos, aún no se ha explorado su papel específico en la respuesta al estradiol y en la vía del receptor de estrógeno (ER).

UCA1 (Urothelial Cancer Associated 1), un lncRNA que figura en el Panel B, inferior izquierda de la Figura 9, ha sido identificado como sobreexpresado en el cáncer de vejiga. Se presume que UCA1 desempeña un papel crucial en la regulación del desarrollo embrionario, así como en la invasión y progresión del cáncer de vejiga, al modular la expresión de diversos genes implicados en la tumorigénesis y/o el desarrollo embrionario (R. Z. He et al., 2021). Se ha sugerido que la sobreexpresión de UCA1, mediada por la modificación m6A, promueve la proliferación y supervivencia de las células cancerosas mediante UCA12 (R. Z. He et al., 2021).

A pesar de su asociación conocida con el cáncer de vejiga, la relación de UCA1 con el cáncer de mama, así como su respuesta hormonal al estradiol y su participación en la vía del receptor de estrógeno (ER), aún no han sido exploradas.

En cuanto al lncRNA LINC02130, su asociación con el cáncer de mama se menciona únicamente en la información disponible en [genecards.org](https://genecards.org). Sin embargo, los detalles específicos sobre su función en la enfermedad no están detallados.

En resumen, de los cuatro ejemplos estudiados, se ha investigado extensamente el papel de LINC01016 en la vía del receptor de estrógeno (ER) ante la

estimulación con estradiol (E2). Aunque se conoce su sobreexpresión en el subtipo de cáncer de mama positivo para el receptor de estrógeno, aún persisten incógnitas, como los genes blanco que podría estar regulando, los procesos moleculares en los que participa y la posible caracterización de su mecanismo epigenético.

En cambio, los otros tres lncRNAs (LINC01213, UCA1 y LINC02130) cuentan con información limitada en cuanto a su relación con el cáncer de mama y prácticamente no se dispone de datos sobre su participación en la vía del receptor de estrógeno. Por lo tanto, esta lista de lncRNAs identificados emerge como un conjunto intrigante que podría beneficiarse de estudios más detallados. Explorar cada uno de estos lncRNAs a fondo podría revelar información valiosa sobre su papel en la progresión del cáncer de mama a través de la activación de la vía del receptor de estrógeno. Estos esfuerzos podrían arrojar luz sobre mecanismos de regulación génica actualmente desconocidos, lo cual podría ser fundamental para avanzar en la comprensión de la enfermedad y el desarrollo de tratamientos más efectivos en el futuro. En última instancia, nuestros análisis bioinformáticos ofrecen una sólida base para comprender la compleja red de regulación génica en el contexto específico del cáncer de mama luminal A.

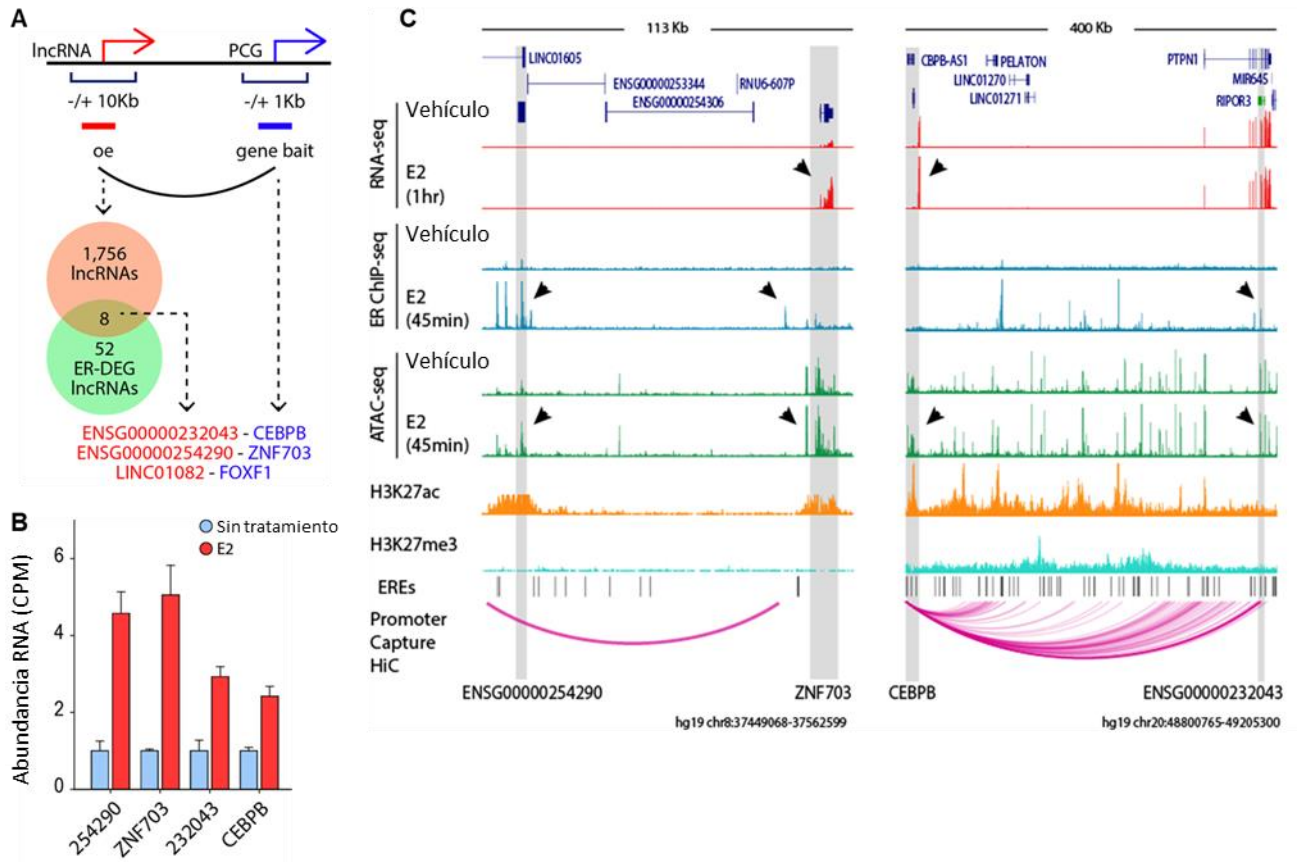


Figura 10. Identificación de lncRNAs regulados por ER en respuesta a estradiol en células MCF-7, que se encuentran asociados con promotores de genes codificantes de proteínas (PCGs). A través de la técnica de promoter capture Hi-C se obtuvo el conjunto de lncRNAs vinculados a PCGs. Posteriormente, se realizó una intersección con el grupo de 52 lncRNAs previamente identificados para obtener los posibles candidatos a ser estudiados. Los datos para los análisis bioinformáticos fueron obtenidos de (Beesley et al., 2020) panel A y de (Franco et al., 2015) panel B. Panel C: Mapas genómicos de los lncRNAs que contactan a un gen codificante.

## 8.2 Objetivo 2: Identificación de lncRNAs con Potencial Regulatorio en Cis en Genes Codificantes Respondientes a E2, Independientes de la Vía de ER.

Se realizaron análisis bioinformáticos con el propósito de cumplir con el **objetivo 2**, centrado en la **Identificación de lncRNAs con Potencial Regulatorio en Cis en Genes Codificantes Respondientes a E2, Independientes de la Vía de ER**. Estos análisis se basaron en datos de Promoter Capture Hi-C, con el objetivo de anticipar las interacciones entre los promotores de genes y lncRNAs.

Promoter Capture Hi-C es una técnica avanzada que combina la tecnología de captura de la cromatina con la secuenciación Hi-C. Hi-C es una técnica que mapea las interacciones tridimensionales en el genoma, revelando cómo las distintas regiones cromosómicas se pliegan y se relacionan entre sí en el núcleo celular.

**Principio Básico:**

1. **Captura de Cromatina:** En el caso de Promoter Capture Hi-C, se utiliza una sonda de captura que se dirige específicamente a los promotores de genes conocidos. Estas sondas tienen la capacidad de "capturar" y aislar la cromatina que está en proximidad física a los promotores seleccionados.
2. **Marcado y Secuenciación:** Una vez que la cromatina ha sido capturada, se marca y se secuencia para identificar las regiones específicas del genoma que interactúan espacialmente con los promotores seleccionados.

**Utilidad en el Proyecto:**

- **Identificación de Interacciones:** Promoter Capture Hi-C es especialmente útil para identificar interacciones físicas entre los promotores de genes y regiones específicas del genoma, incluidos los promotores de lncRNAs.
- **Genómica Funcional:** Permite comprender cómo los elementos reguladores, como los promotores de lncRNAs, interactúan con los promotores de genes codificantes. Esto proporciona información clave sobre cómo la regulación génica se lleva a cabo en el contexto tridimensional de la cromatina.
- **Conexiones Específicas:** Ayuda a establecer conexiones específicas entre lncRNAs y genes codificantes, lo que es esencial para comprender su coordinación en respuestas celulares, como las inducidas por tratamientos hormonales.
- **Validación de Interacciones:** Permite validar experimentalmente las interacciones previamente predichas o identificadas mediante otras técnicas, brindando una capa adicional de confianza en los resultados.
- **Construcción de Redes de Regulación:** Facilita la construcción de redes de regulación génica, destacando las relaciones entre los promotores de

lncRNAs y los genes codificantes, lo cual es esencial para comprender la complejidad de la regulación génica.

En resumen, Promoter Capture Hi-C es una herramienta avanzada que aporta información detallada sobre las interacciones físicas en el genoma, siendo crucial para comprender la regulación génica, especialmente en el contexto de la participación de lncRNAs en respuestas celulares específicas

El enfoque para optar por el uso de estos datos, se sustenta en la observación de que algunos lncRNAs, con funciones en cis, operan mediante la formación de bucles en la cromatina (Gil & Ulitsky, 2020).

A partir de este análisis, identificamos un conjunto de 1,756 lncRNAs que presentan interacciones con los promotores de genes en células MCF-7. Luego, al cruzar esta información con los 52 lncRNAs previamente identificados en fases anteriores de nuestro estudio, se reveló un hallazgo notable: solo 3 lncRNAs presentaron intersección (Figura 10A).

Es importante señalar que los tres lncRNAs identificados mostraron un aumento significativo en sus niveles de transcripción, al igual que los genes con los que interactúan, en respuesta al estímulo con E2 (Figura 10B). Además, se proporciona un esquema visual que detalla las interacciones o contactos entre cada lncRNA y su gen codificante correspondiente, representados en tono rosa (Figura 10C). Este descubrimiento sugiere un potencial papel crucial de estos lncRNAs en la respuesta transcripcional mediada por E2 y respalda la hipótesis de su participación en la regulación de genes cercanos mediante interacciones espaciales específicas en la cromatina.

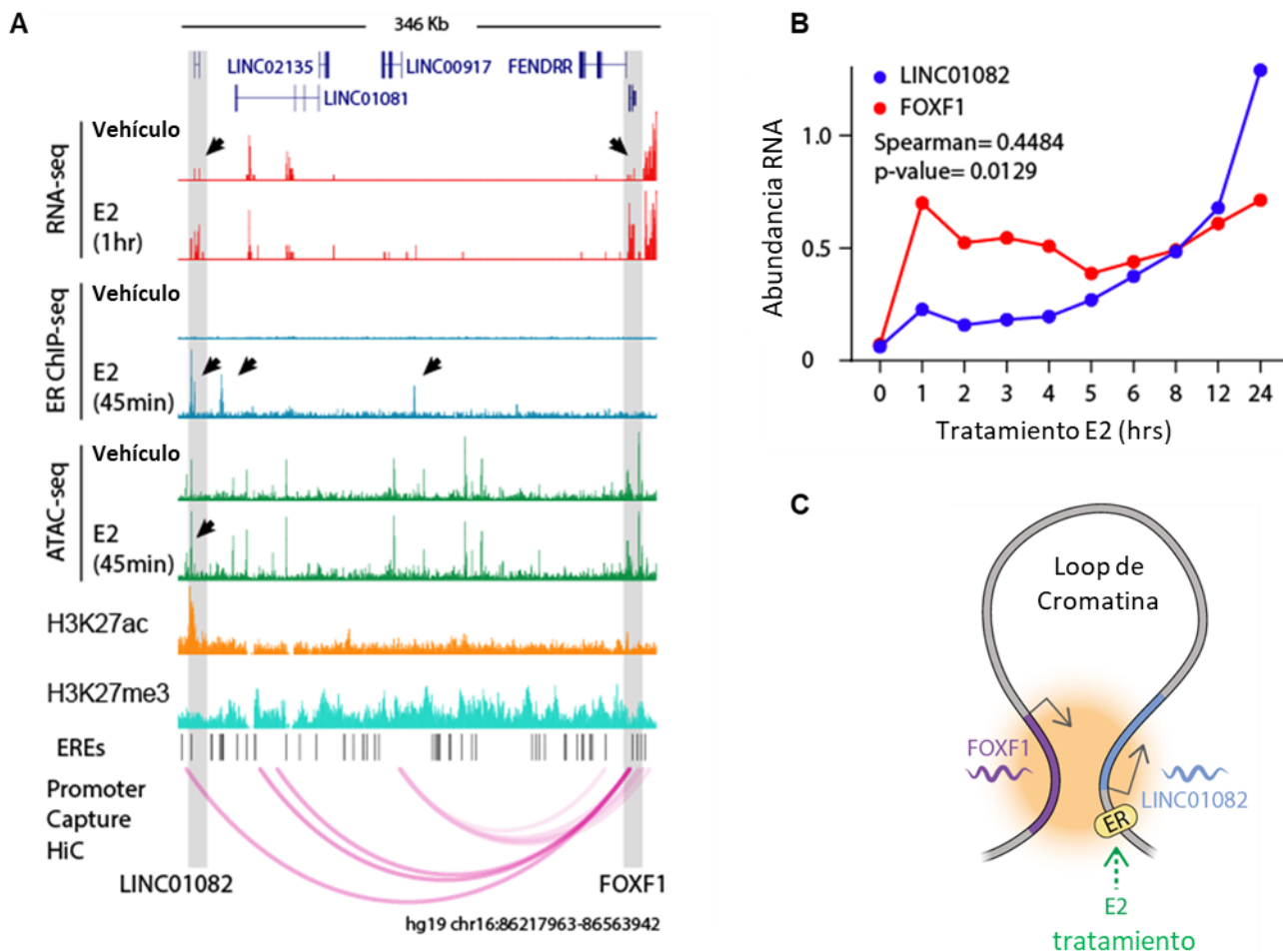


Figura 11. Análisis del *LINC01082* y su relación con *FOXF1* en respuesta al estradiol en células MCF-7. *LINC01082* muestra enriquecimiento de ER $\alpha$  ante E2, mientras *FOXF1* carece de unión ER $\alpha$  (A). Ambos genes incrementan su transcripción progresivamente durante 24 hrs. con estradiol (B), indicando posible co-regulación (C). Los datos para los análisis bioinformáticos fueron obtenidos de (Beesley et al., 2020) panel A y de (Franco et al., 2015) panel B. Panel C: Modelo propuesto en base a los datos de interacción entre el *LINC01082* y *FOXF1*, y su respuesta transcripcional ante E2.

En el análisis de los tres casos previamente mencionados, se destaca especialmente la investigación del *LINC01082*, que establece una interacción con el gen *FOXF1*. Un rasgo distintivo del *LINC01082* es su enriquecimiento de ER $\alpha$  en respuesta a E2, mientras que su gen asociado, *FOXF1*, no muestra una unión a este factor de transcripción. Este descubrimiento sugiere la posibilidad de que la respuesta transcripcional del *LINC01082* a E2 esté mediada por un mecanismo diferenciado (Figura 11A). De manera intrigante, ambos genes revelan un

aumento progresivo en su transcripción a lo largo de 24 horas en presencia de estradiol (Figura 11B). Esta observación señala hacia una posible relación co-regulatoria, indicando que el aumento en la transcripción de *FOXF1* podría estar relacionado con el incremento en la transcripción del *LINC01082* (Figura 11C). Este fenómeno podría sugerir una interdependencia funcional entre ambos elementos en la respuesta celular al estradiol.

### 8.2.1 Resultados Experimentales

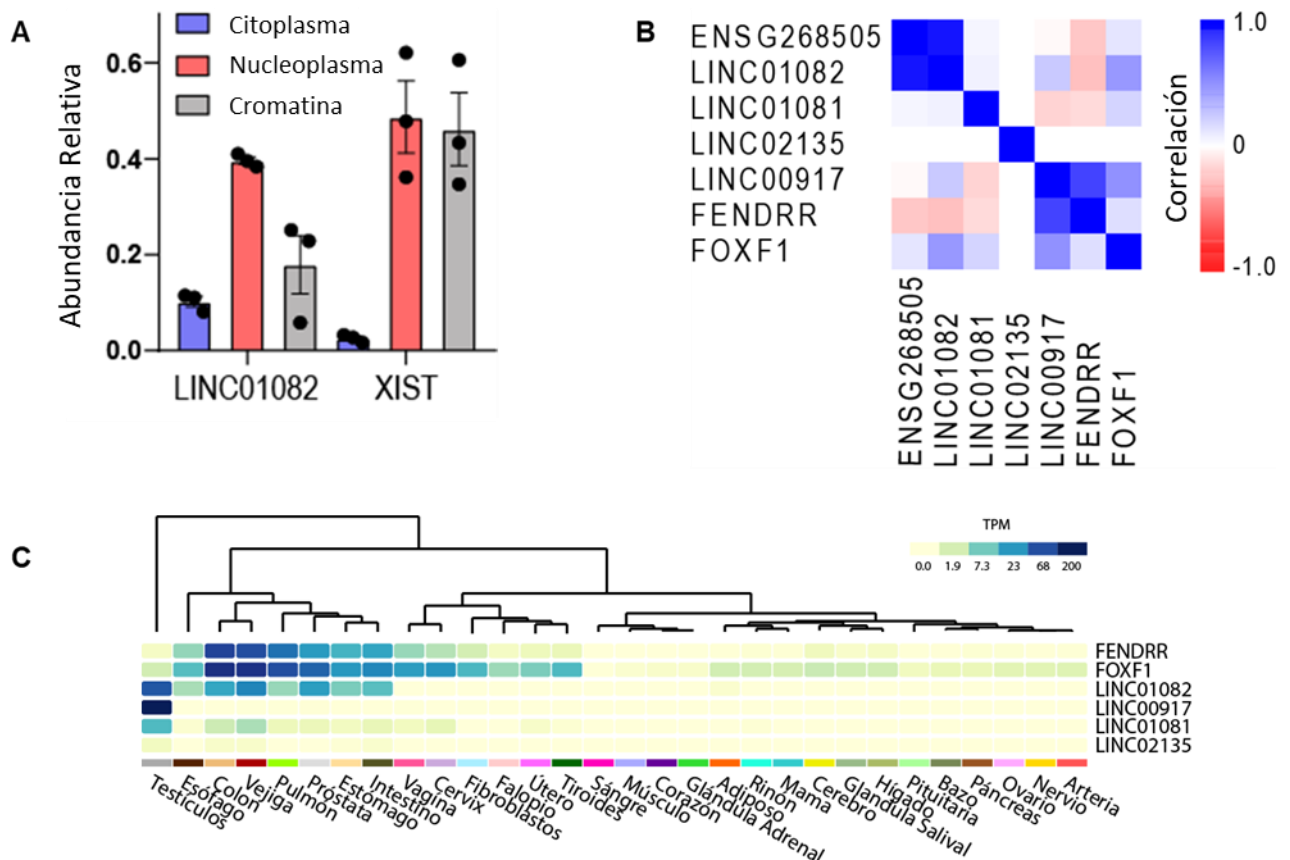


Figura 12. Asociación de *LINC01082* con *FOXF1* y su localización subcelular. En A, se utilizó la técnica de extracción de RNA por fraccionamiento para determinar la distribución subcelular predominante del *LINC01082*. En B, se presenta un mapa de calor basado en la correlación de Pearson, que revela la coexpresión entre los transcritos. Finalmente, en C, se exhiben datos transcriptómicos de la base de datos GTEx para los genes de interés en distintos tejidos.



### 8.2.2 Distribución del transcrito del *LINC01082* en la célula y su posible función regulatoria.

Nuestro objetivo principal consistió en esclarecer la ubicación predominante de nuestro lncRNA de interés en el contexto celular. Dada la sugerencia de la literatura de que los lncRNAs no codificantes con funciones nucleares tienden a concentrarse en el núcleo y la cromatina, nos propusimos abordar su posible papel como reguladores transcripcionales. Realizamos un ensayo de RT-qPCR por fraccionamiento, que permitió la separación y análisis del RNA en fracciones de citosol, nucleoplasma y cromatina. Esta técnica nos proporcionó la oportunidad de identificar la fracción enriquecida con el transcrito de interés. Inspirados por observaciones similares en lncRNAs como *XIST*, elegimos llevar a cabo este análisis.

Los resultados de la RT-qPCR por fraccionamiento revelaron un marcado enriquecimiento del *LINC01082* en el núcleo (Figura 12A), sugiriendo su participación en diversos procesos reguladores nucleares, como la transcripción y la estructura de la cromatina. A continuación, realizamos un análisis de correlación lineal entre la transcripción de todos los elementos presentes en el locus (*LINC01082* y el gen *FOXF1*). Identificamos una correlación positiva entre *LINC01082* y *FOXF1*, así como una correlación inversa entre el *LINC01082* y el gen *FENDRR* (un lncRNA antisentido al gen *FOXF1*). Además, observamos una correlación positiva entre *LINC00917* con *FOXF1* y *FENDRR*. Esto sugiere que la expresión del gen *FOXF1* está condicionada por la expresión de *LINC01082*, y lo mismo se aplica a *LINC00917* y *FENDRR*. Aunque una correlación negativa en el caso de *FENDRR* sugiere que su expresión disminuye a medida que *LINC01082* se incrementa. Este hallazgo intensifica el interés en explorar a fondo este locus (Figura 12B).

Dando un paso más, evaluamos la expresión de estos genes en diversos tejidos (Figura 12C) y notamos que los genes *FENDR*, *FOXF1* y *LINC01082* se encuentran co-expresados en tejidos epiteliales de varios órganos (excepto el tejido mamario en estado sano) (Figura 12C). Este análisis robustece la

implicación de estos elementos en procesos celulares específicos, subrayando su relevancia potencial en la regulación de la expresión génica en contextos tisulares.

Este cuestionamiento nos condujo a indagar: ¿Cómo se verían afectados los niveles de expresión del gen *FOXF1* y otros elementos del locus si manipulamos de manera exógena los niveles del *LINC01082*? ¿Persistiríamos en la observación de los mismos patrones previamente identificados?

En resumen, para cumplir el objetivo 2 de nuestra investigación, empleamos la técnica avanzada de Promoter Capture Hi-C para explorar las interacciones genómicas en células MCF-7 bajo la estimulación con E2. Identificamos 1,756 lncRNAs que interactúan con promotores de genes, y al cruzar estos datos con lncRNAs previamente identificados, destacamos 3, incluido el *LINC01082*. Estos lncRNAs, enriquecidos con ER $\alpha$  en respuesta a E2, mostraron un aumento significativo en la transcripción, sugiriendo su papel en la respuesta transcripcional mediada por E2. Especialmente, el *LINC01082* reveló interacciones específicas con el gen *FOXF1* y una distribución enriquecida en el núcleo, indicando una posible participación en procesos reguladores nucleares. La correlación entre estos elementos y su co-expresión en tejidos epiteliales subrayan su relevancia en la regulación génica. Este estudio proporciona una base sólida para futuras investigaciones sobre la manipulación de *LINC01082* y su impacto en la expresión génica en el contexto del cáncer de mama.

### 8.3 Objetivo 3: Caracterización Funcional de un lncRNA Candidato y su Impacto en el Gen Codificante Asociado.

Las consideraciones anteriores nos llevaron a diseñar experimentos para cumplir con el objetivo 3: Caracterización Funcional de un lncRNA Candidato y su Impacto en el Gen Codificante Asociado. Para lograr esto, se idearon RNAs guías (Figura 13A) con el propósito de llevar a cabo la sobreexpresión y represión del *LINC01082* mediante el sistema CRISPR-dCas, específicamente utilizando los sistemas de activación CRISPRa e inhibición CRISPRi para modular su estructura

de la cromatina. El sistema CRISPRi emplea la dCas junto con remodeladores de la cromatina que, localmente (a través de sgRNAs específicos), promueven la formación de una cromatina compacta, ejerciendo así represión sobre la actividad transcripcional del gen adyacente (Figura 13B). En cambio, en el sistema CRISPRa, la dCas se asocia con remodeladores de la cromatina que favorecen una cromatina propensa a la actividad transcripcional (Figura 13B).

Para llevar a cabo esta serie de experimentos, se diseñaron dos guías cercanas al sitio de inicio de la transcripción del gen *LINC01082* (Figura 13A). En los experimentos subsiguientes, se empleó una combinación de estas guías.

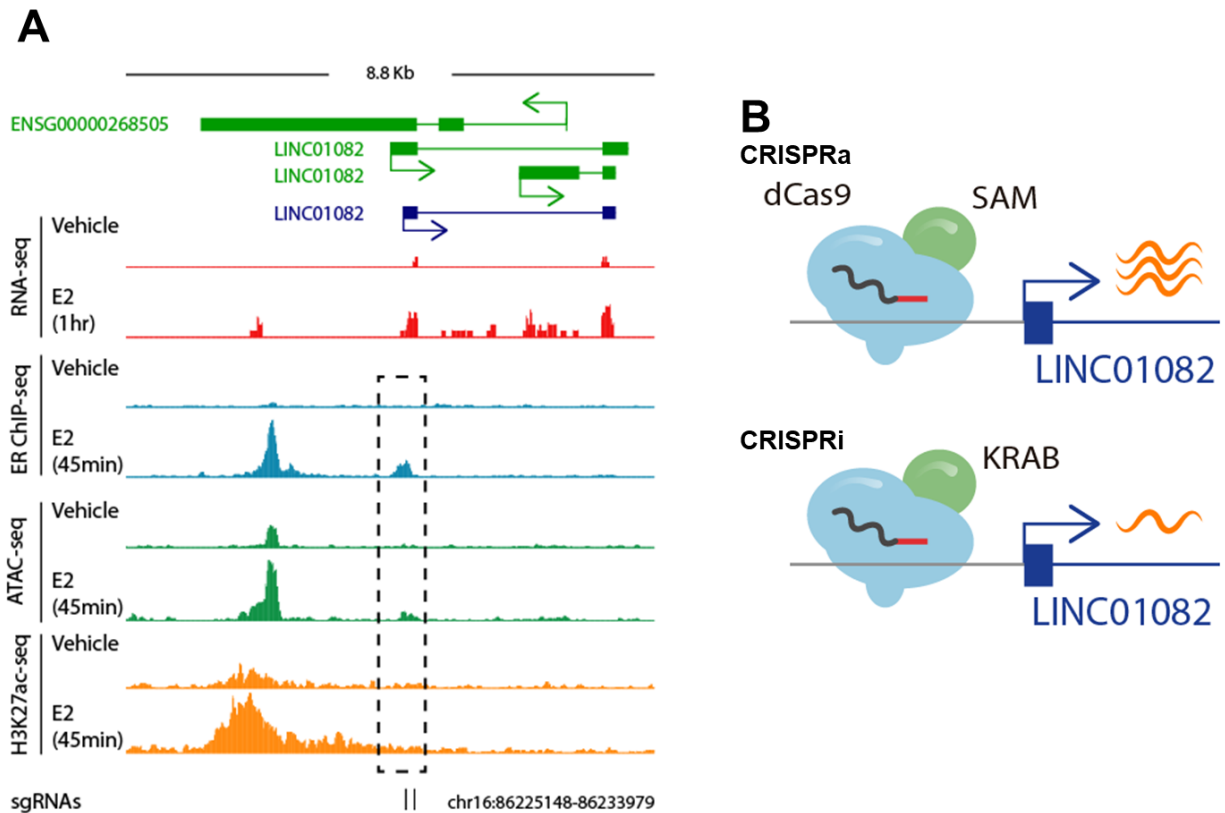


Figura 13: Enfoques experimentales para analizar la función regulatoria del *LINC01082*. Se llevo a cabo el diseño de RNAs guías (sgRNAs) para llevar a cabo el reclutamiento de los sistemas de sobre expresión y represión ectópica, CRISPRa y CRISPRi, respectivamente. Panel A: Región blanco de los sgRNAs para los sistemas de sobre expresión CRISPRa y de represión CRISPRi del transcrito del *LINC01082*. Panel B: Modelo gráfico de los sistemas de sobre expresión CRISPRa y de represión CRISPRi del transcrito del *LINC01082* empleados.

La técnica de transfección de los vectores que codifican para los sistemas CRISPRi/a se realizó de manera independiente en las células MCF-7 mediante la infección con lentivirus generados a partir de las células HEK293. Las células se infectaron de forma individual con cuatro lentivirus distintos:

KRAB vector vacío (EV)

KRAB sg1+2

SAM vector vacío (EV)

SAM sg1+2

Posteriormente, se llevó a cabo la selección con puromicina para permitir la supervivencia solo de las células que incorporaron el sistema CRISPR. Seguidamente, se procedió a la extracción de RNA y se realizó un ensayo de expresión mediante RT-qPCR para evaluar los niveles de expresión del *LINC01082* en estas condiciones.

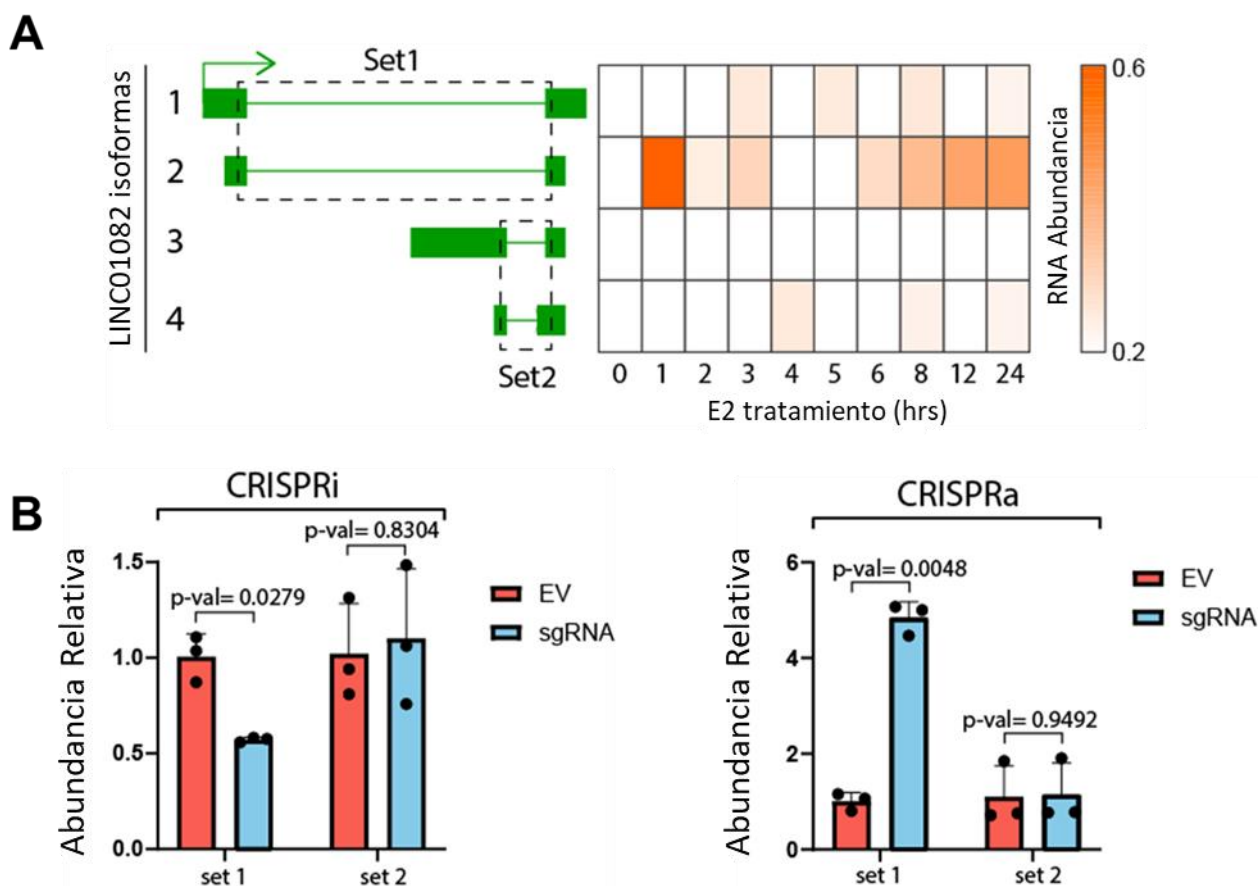


Figura 14: A) Heat map que muestra la señal de expresión obtenida de RNA-seq por cada isoforma existente del LINC01082. B) Datos de qRT-PCR con la expresión de las isoformas ante el sistema de represión CRISPRi (izquierda) y el sistema de sobre expresión CRISPRa (derecha).

Debido a la presencia de múltiples isoformas del gen *LINC01082* (transcritos de diferentes tamaños), se diseñaron primers específicos para identificar las isoformas largas 1 y 2 (Set 1, Figura 14A) y las isoformas cortas 3 y 4 (Set 2, Figura 14A). El objetivo de este diseño fue emplear la técnica de qRT-PCR para determinar cuáles de estas isoformas experimentaban sobreexpresión debido al sistema CRISPRa y verificar si esto coincidía con lo observado en el análisis bioinformático.

El análisis bioinformático de la expresión durante un periodo de 0 a 24 horas (RNA-seq) durante el tratamiento con E2 reveló que, tras la implementación del experimento de sobreexpresión, la isoforma 2 del Set 1 se manifestó como la más

predominante en términos de niveles transcripcionales. En este análisis, observamos que, de las cuatro isoformas asociadas con LINC01082, solo la isoforma 2 del Set 1 experimenta un aumento transcripcional significativo en la primera hora, mostrando así una respuesta rápida al tratamiento con E2 (Figura 14A). Esta isoforma podría desempeñar una función de regulación transcripcional en otros genes que también muestran una respuesta temprana a la estimulación con E2 y, particularmente, podría estar vinculada al aumento transcripcional simultáneo del gen FOXF1 en la primera hora.

En relación con los resultados del sistema de represión CRISPRi/KRAB, se observó una disminución significativa en la transcripción del LINC01082 en el contexto de las isoformas largas (Set 1, Figura 14B). No obstante, no se registró una disminución significativa en la transcripción de las isoformas cortas (Set 2) (Figura 14B). En términos estadísticos, el valor p de 0.0279 (Figura 14B, izquierda) indica que hay un 2.79% de probabilidad de obtener los resultados observados si la hipótesis nula es verdadera. Con un nivel de significancia predefinido del 5%, el valor p de 0.0279 es inferior a  $\alpha$ , lo que indica evidencia suficiente para rechazar la hipótesis nula. Por lo tanto, se concluye que existen diferencias significativas en la expresión del transcrito del LINC01082 entre los grupos EV y sg, específicamente en el caso de las isoformas largas (Set 1, Figura 14B). Para las isoformas cortas, al tener un valor p de 0.9492, no hay evidencia suficiente para rechazar la hipótesis nula, sugiriendo que no hay diferencias significativas entre los grupos EV y sgRNA en el sistema de represión del LINC01082 mediante CRISPRi.

En cuanto al sistema CRISPRa/SAM, se observó una sobreexpresión altamente enriquecida del LINC01082 en el contexto de la isoforma larga (Set 1). Sin embargo, esta sobreexpresión no se manifestó en las isoformas del Set 2. Esto respalda la conclusión de que al amplificar solo las isoformas largas, el efecto transcripcional observado en FOXF1 podría explicarse por el aumento en la transcripción de las isoformas del Set 1, especialmente la isoforma 2, lo que concuerda con los datos de RNA-seq (Figura 14A). Con un valor p de 0.0048 (considerando un nivel de significancia del 5%), se concluye que existen

diferencias significativas en la expresión del transcrito del LINC01082 entre los grupos EV y sg, específicamente en el caso de las isoformas largas (Set 1, Figura 14B). Para las isoformas cortas, con un valor p de 0.9492, no hay evidencia suficiente para rechazar la hipótesis nula, sugiriendo que no hay diferencias significativas entre los grupos EV y sgRNA en el sistema de sobreexpresión del LINC01082 mediante CRISPRi.

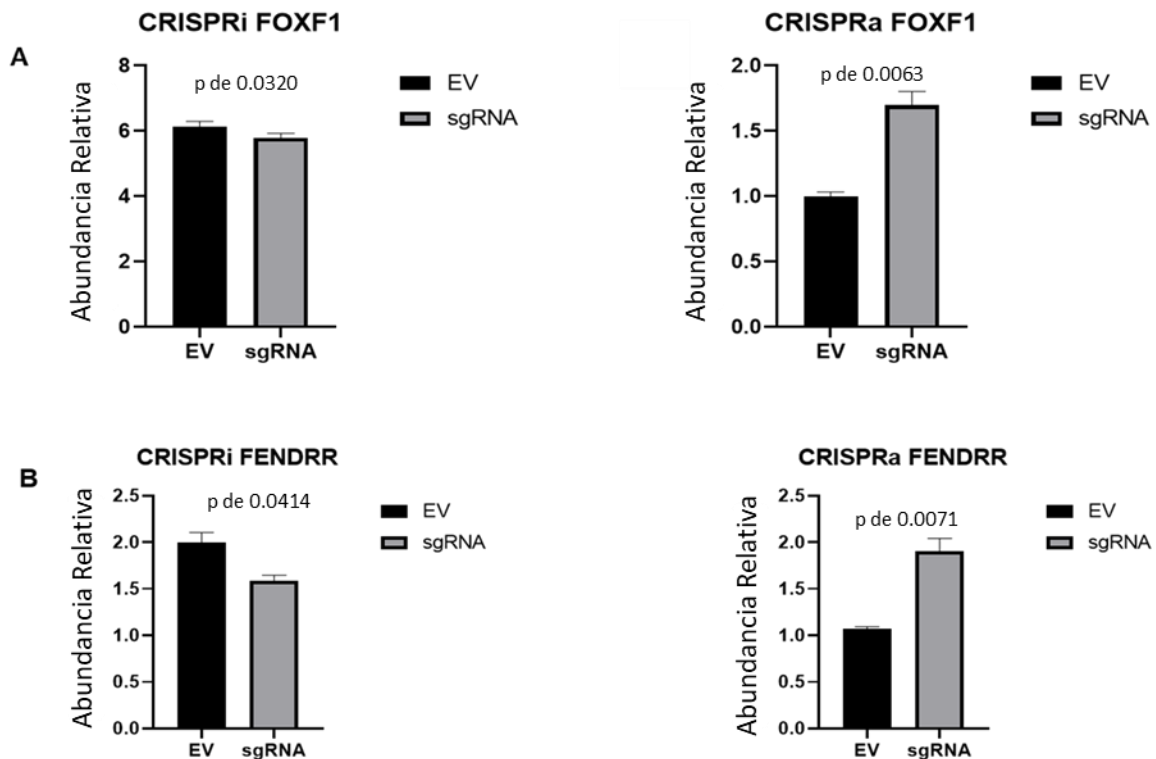


Figura 15: Análisis del efecto de la sobre expresión del *LINC01082* en los niveles transcripcionales de los genes *FOXF1* y *FENDRR*. A) Resultados de RT-qPCR que muestran la expresión de *FOXF1* ante la represión del *LINC01082* mediante el sistema CRISPRi (izquierda) y la expresión de *FOXF1* ante la sobreexpresión del *LINC01082* mediante el sistema CRISPRa (derecha). B) Resultados de RT-qPCR que ilustran la expresión de *FENDRR* ante la represión del *LINC01082* con el sistema CRISPRi (izquierda) y la expresión de *FENDRR* ante la sobreexpresión del *LINC01082* con el sistema CRISPRa (derecha).

Después de analizar los patrones de expresión de las isoformas Set1 y Set2 del LINC01082, nuestro objetivo fue evaluar su impacto en la expresión de los genes FOXF1 y FENDRR. En el sistema CRISPRi (Figura 15A), se observa una

disminución significativa en la expresión de los genes FOXF1 y FENDRR. Este resultado respalda la idea de que la intervención específica del LINC01082 mediante CRISPRi tiene un efecto negativo en la expresión de estos genes, sugiriendo un papel regulador del LINC01082 sobre FOXF1 y FENDRR.

En términos prácticos, con un nivel de significancia ( $\alpha$ ) establecido en el 5%, el valor p de 0.0320 se encuentra por debajo de  $\alpha$ , indicando evidencia suficiente para rechazar la hipótesis nula. Esto implica que existen diferencias significativas en la expresión del transcrito de FOXF1 entre los grupos EV y sg, particularmente en el sistema de represión de la transcripción del LINC01082 (Figura 15A, izquierda). De manera similar, para el transcrito de FENDRR, con un valor p de 0.0414, se evidencian diferencias significativas en los niveles de expresión entre los grupos EV y sg, específicamente en el sistema de represión de la transcripción del LINC01082.

En el análisis del sistema CRISPRa, se destaca una sobreexpresión significativa tanto del gen FOXF1 como del gen FENDRR (Figura 15A). En términos prácticos, con un nivel de significancia ( $\alpha$ ) establecido en el 5%, el valor p de 0.0063 se sitúa por debajo de  $\alpha$ , indicando evidencia suficiente para rechazar la hipótesis nula. Por lo tanto, existen diferencias significativas en la expresión del transcrito de FOXF1 entre los grupos EV y sg, focalizadas especialmente en el sistema de sobreexpresión de la transcripción del LINC01082 (Figura 15A, izquierda). De manera similar, para el transcrito de FENDRR, con un valor p de 0.0071, se evidencian diferencias significativas en los niveles de expresión entre los grupos EV y sg, específicamente en el sistema de sobreexpresión de la transcripción del LINC01082.

Este hallazgo refuerza nuestra hipótesis inicial de que el transcrito del LINC01082 desempeña un papel crucial en la regulación transcripcional de FOXF1 e incluso de FENDRR. La activación de la expresión de ambos genes tras la sobreexpresión del LINC01082 respalda la idea de que este lncRNA tiene un impacto positivo en la transcripción de los genes asociados en el locus, destacando su papel como regulador clave en este contexto específico.



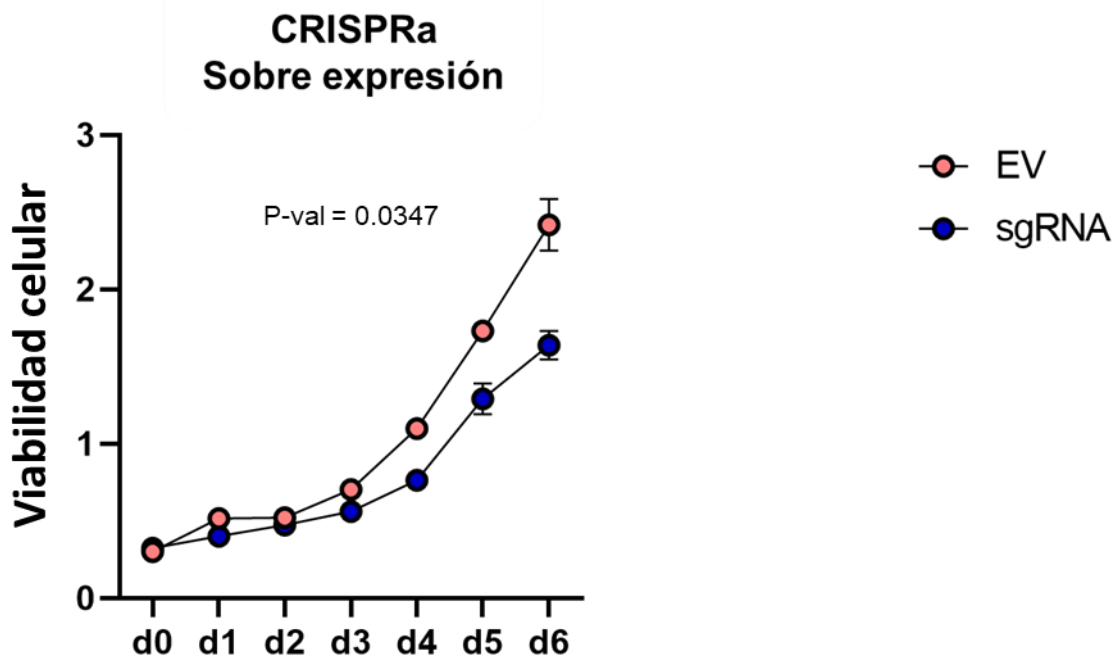


Figura 16. Ensayo de MTT para evaluar el efecto de la sobre expresión del LINC01082 con el sistema CRISPRa, en la proliferación y viabilidad celular de la línea MCF-7. En azul podemos observar el sistema de sobre expresión funcional con los guías sgRNAs. Mientras que en naranja tenemos el sistema CRISPRa *Empty Vector* (EV) como control.

Al llevar a cabo la infección mediante los sistemas lentivirales CRISPRa y CRISPRi, se observó una aparente variación en la tasa de proliferación, especialmente en las células tratadas con los sgRNAs en el caso de CRISPRa. Este fenómeno concuerda con los resultados de expresión obtenidos (Figura 14B, lado derecho), donde se evidencia una sobreexpresión ectópica del transcrito del *LINC01082* en estas células. Esta observación sugiere que la sobreexpresión de este lncRNA podría influir en los procesos celulares relacionados con la proliferación y viabilidad celular. Es plausible que esta sobreexpresión ejerza una función regulatoria en los genes presentes en la línea celular MCF-7, lo cual podría estar vinculado con las diferencias en la tasa de proliferación observadas.

Para evaluar cómo la sobreexpresión del LINC01082 afecta la proliferación celular, llevamos a cabo un experimento de viabilidad celular utilizando el ensayo MTT. Los resultados obtenidos revelan de manera clara que la sobreexpresión del

LINC01082 está asociada con una disminución en la tasa proliferativa en comparación con el grupo de control, como se muestra en la Figura 16.

Para determinar la significancia estadística de esta observación, aplicamos la prueba t de Student. El valor de p obtenido fue 0.0347. En términos prácticos, este valor de p representa la probabilidad de obtener un resultado tan extremo o más extremo que el observado, suponiendo que no hay diferencia real en la proliferación celular entre los grupos y que cualquier diferencia es debida al azar. Dado que 0.0347 es menor que el nivel de significancia convencional de 0.05 (5%), se considera que hay evidencia estadística suficiente para rechazar la hipótesis nula.

En consecuencia, podemos concluir que la diferencia observada en los niveles de proliferación entre los grupos que sobreexpresan el LINC01082 (EV) y el grupo de control (sgRNA) no es simplemente el resultado del azar. En lugar de eso, sugiere de manera significativa que la sobreexpresión del LINC01082 tiene un efecto real y negativo en la tasa de proliferación celular.

Este hallazgo sugiere que el *LINC01082* podría influir en la regulación transcripcional de otros genes, además del gen *FOXF1*, con consecuencias en la proliferación celular.

Adicionalmente, sospechamos que el aumento transcripcional del gen *FOXF1* podría ser en gran parte responsable de los cambios celulares que afectan la proliferación y viabilidad celular. *FOXF1* es un factor transcripcional crucial en el desarrollo del mesénquima pulmonar, el tejido embrionario precursor de los vasos sanguíneos del pulmón y el tracto gastrointestinal (Ustiyana et al., 2018a). En este sentido, la sobreexpresión del gen *FOXF1* podría estar regulando otros genes relacionados con procesos fisiológicos en la línea celular MCF-7. Estos resultados sugieren que tanto el *LINC01082* como *FOXF1* pueden desempeñar un papel vital en la regulación de la expresión génica y la modulación de procesos celulares cruciales en esta línea celular.

En la consecución del objetivo 3 de caracterización funcional del lncRNA *LINC01082*, empleamos estrategias innovadoras de ingeniería genética, utilizando el sistema CRISPR-dCas para modular la estructura de la cromatina y evaluar sus efectos en la expresión génica. La sobreexpresión y represión dirigidas por CRISPR-dCas revelaron patrones distintivos en las isoformas del *LINC01082*, especialmente la isoforma 2 del Set 1, que mostró una respuesta rápida a la estimulación con E2. Este patrón coincidió con el análisis bioinformático, validando la efectividad de la manipulación. Además, la intervención CRISPRi demostró la regulación negativa del *LINC01082* sobre los genes *FOXF1* y *FENDRR*, mientras que CRISPRa reveló un impacto positivo, respaldando nuestra hipótesis inicial. La variación en la tasa de proliferación celular asociada con la sobreexpresión del *LINC01082* indica su función regulatoria en procesos celulares más amplios, destacando su influencia en la expresión génica y la viabilidad celular. Estos hallazgos no solo aportan a la comprensión de los mecanismos de acción de los lncRNAs, sino que también señalan al *LINC01082* y su interacción con *FOXF1* como componentes clave en la regulación de la expresión génica y procesos celulares en la línea celular MCF-7.

En resumen, estos hallazgos ofrecen nuevas perspectivas para futuras investigaciones y sugieren la existencia de mecanismos de regulación transcripcional complejos y entrelazados en la línea celular MCF-7. Investigaciones posteriores deberán centrarse en identificar los genes específicos regulados por el *LINC01082* y *FOXF1*, así como comprender los mecanismos moleculares involucrados en estos procesos. Esta información contribuirá significativamente a nuestro conocimiento sobre la función de estos elementos regulatorios y su relevancia en el contexto del cáncer de mama luminal A.

## 9. Discusión

Este estudio representa un avance significativo al lograr demostrar de manera concluyente el impacto funcional del lncRNA en un sistema de regulación que se revela excepcionalmente complejo (Figura 17). La manipulación de los niveles de expresión del linc01082 reveló un sorprendente y directo efecto en los niveles transcripcionales de FOXF1 y FENDRR, dos sistemas cuidadosamente regulados por este lncRNA específico. Este hallazgo no solo resalta la importancia funcional del linc01082 en la modulación de la expresión génica, sino que también destaca la complejidad intrínseca de este sistema de regulación.

Comparativamente, este sistema de regulación presenta similitudes asombrosas con otros sistemas bien conocidos, como el del lincRNA-p21 regulado por p53. La capacidad de alterar los niveles de expresión del linc01082 y observar un impacto directo en la transcripción de genes específicos subraya la sofisticación y eficacia de este sistema de regulación particular. A diferencia de sistemas más convencionales, donde la intervención puede resultar en respuestas predecibles, la capacidad de nuestro lncRNA para modular la expresión de genes críticos en la vía de regulación del cáncer de mama presenta una complejidad igualmente asombrosa y reveladora. Estos resultados no solo resaltan la singularidad de nuestro sistema de estudio, sino que también subrayan la importancia de profundizar en la comprensión de estos intrincados mecanismos reguladores en la biología celular.

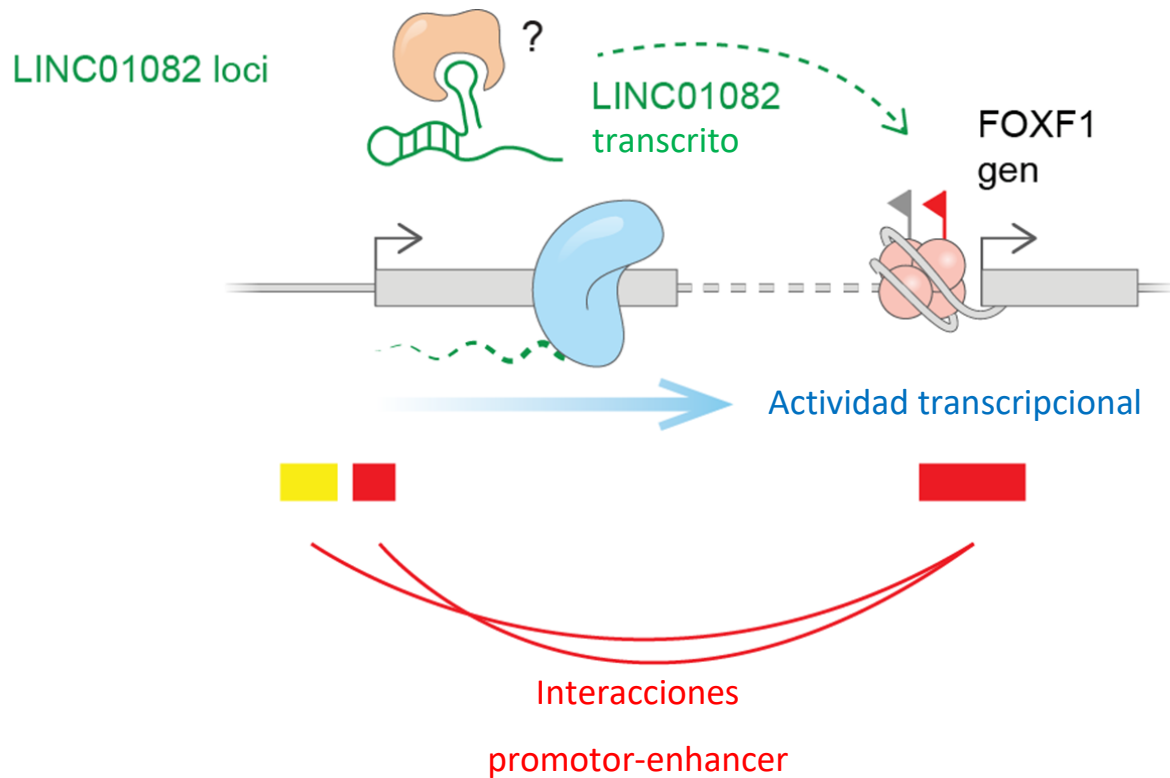


Fig. 17. Modelo propuesto de los posibles mecanismos regulatorios o epigenéticos que podría estar ejerciendo el transcrito del LINC01082 en el gen FOXF1 para llevar a cabo su aumento transcripcional en respuesta a E2 a 1 h de manera indirecta a la vía de ER $\alpha$ .

El modelo propuesto en este proyecto (Figura 17) constituye un sistema de regulación génica altamente complejo, desencadenado por la estimulación con E2, el ligando del receptor de estrógeno alfa (ER $\alpha$ ). Cuando ER $\alpha$  se dimeriza y transloca al núcleo de la célula, se une a regiones promotoras de genes, incluyendo lncRNAs que se han reconocido por su potencial papel epigenético y su función como reguladores transcripcionales. En este estudio, se logró demostrar de manera convincente que la alteración de los niveles del lncRNA LINC01082, que cumple con las características mencionadas, tiene un impacto discernible en el cambio transcripcional de los genes FOXF1 y FENDRR. La complejidad de este proceso regulador se destaca especialmente en el caso de FOXF1, un factor de transcripción con la capacidad de influir en múltiples genes blanco, intensificando así la sofisticación de este sistema.

Por otro lado, FENDRR, al ser un lncRNA, abre la posibilidad de influir en la regulación de otros genes diana, ampliando aún más el espectro de su función reguladora. Este modelo sienta las bases para explorar un sistema regulador fascinante, ofreciendo una oportunidad única para profundizar en los intrincados procesos de regulación génica que las células emplean en respuesta a su entorno, específicamente en el contexto de la patología del cáncer de mama y su progresión. El estudio de este sistema puede aportar conocimientos valiosos que no solo ampliarán nuestra comprensión de la biología molecular, sino que también podrían tener implicaciones significativas en el desarrollo de enfoques terapéuticos más precisos y efectivos para el cáncer de mama.

La superfamilia de receptores nucleares se revela como un componente esencial en la maquinaria reguladora de la expresión génica en organismos multicelulares, mostrando una precisa respuesta a diversas señales hormonales y estímulos. Su estructura compartida, que consta de dominios de unión al ADN altamente conservados y dominios de unión al ligando menos conservados, constituye la base de su versatilidad funcional. Clasificados en receptores de esteroides y no esteroides, estos se vinculan a distintos tipos de moléculas señaladoras. Su influencia abarca procesos biológicos cruciales, como el desarrollo embrionario, el mantenimiento de la homeostasis y la regulación de la respuesta inmunológica. Destaca su potencial como blancos terapéuticos en enfermedades como el cáncer y la diabetes, resaltando su importancia médica.

La regulación precisa de la transcripción génica por parte de estos receptores se logra mediante la interacción con elementos de respuesta en el ADN y la formación de complejos proteicos que coordinan la actividad transcripcional. Esta regulación puede ser influenciada por otros factores de transcripción y señales celulares, permitiendo una respuesta adaptativa a las cambiantes condiciones del entorno. En particular, la estrategia de unión a secuencias específicas de ADN y el papel de cofactores, como coactivadores y corepresores, resultan cruciales para modular su actividad transcripcional y determinar su función precisa en la célula. Comprender plenamente estos mecanismos es esencial para avanzar en la

investigación biomédica y el desarrollo de terapias dirigidas con un alto potencial terapéutico.

La investigación científica enfocada en entender los intrincados mecanismos de regulación génica, nos llevó a hacer un análisis meticuloso para identificar y validar lncRNAs que en respuesta a la estimulación con E2 muestran una activación transcripcional. El objetivo es entender su papel en la red de regulación transcripcional, específicamente en la vía del receptor de estrógeno alfa ( $ER\alpha$ ), en un modelo de cáncer de mama luminal A. La importancia de entender los mecanismos subyacentes a la respuesta celular al estrógeno radica en su relevancia clínica y su potencial para arrojar luz sobre los procesos de proliferación y diferenciación celular, factores esenciales en la patogénesis del cáncer de mama.

Con el propósito general como marco de referencia, esta investigación adopta un enfoque integral que fusiona análisis bioinformáticos avanzados con validaciones experimentales. La utilización de técnicas de secuenciación de última generación ha posibilitado la exploración exhaustiva del transcriptoma en un modelo de cáncer de mama luminal A en respuesta a la exposición al estradiol. La información resultante, obtenida de bases de datos públicas, se ha sometido a procedimientos de filtrado y análisis muy rigurosos con el objetivo de identificar lncRNAs cuya transcripción se activa de manera específica y significativa en respuesta al estímulo estrogénico.

Los lncRNAs surgen como elementos clave en la regulación génica, y su participación en la red de regulación transcripcional impulsada por la vía  $ER\alpha$  añade un nivel de complejidad a este proceso. Esta discusión se enfoca en analizar detalladamente los resultados obtenidos mediante la integración de enfoques bioinformáticos y experimentales. Además, se explorarán las implicaciones de estos hallazgos en el contexto del cáncer de mama luminal A, así como su potencial para revelar nuevos mecanismos reguladores y abrir perspectivas en la investigación oncológica y terapéutica.

En este trabajo se han identificado posibles lncRNAs que responden a la vía ER $\alpha$  y que podrían desempeñar un papel potencial como reguladores transcripcionales o desencadenantes de mecanismos epigenéticos.

En un análisis inicial que involucra la combinación de datos provenientes de ChIP-seq de ER $\alpha$  y RNA-seq, ambos recopilados después de la exposición al estradiol durante aproximadamente una hora, se identificaron 52 lncRNAs con un aumento en la unión del factor transcripcional ER $\alpha$  y un incremento en su expresión. Posteriormente, al examinar los datos del interactoma mediante el ensayo de captura de Hi-C (promoter capture Hi-C) con el propósito de identificar el conjunto completo de 1,756 lncRNAs en la línea celular MCF-7 que están vinculados o interactúan con promotores génicos, se llevó a cabo una comparación cruzada de estos dos conjuntos de datos. Este cruce resultó en la identificación de solo 8 lncRNAs en común. Entre estos 8, solo 3 lncRNAs establecían una conexión con genes que también experimentaban un aumento en su transcripción en respuesta al tratamiento con E2.

Dentro de este conjunto, el *LINC01082* parecía el lncRNA de mayor interés y se seleccionó para un análisis más detallado. Esta elección se fundamentó en su capacidad para cumplir con los criterios mencionados anteriormente, así como en su particularidad de estar asociado a un gen codificante, *FOXF1* (mediante interacciones 3D de la cromatina "Hi-C"), que carece de unión del ER $\alpha$  en su región promotora o cercana al TSS. Esto contrasta con los otros dos casos, donde los genes codificantes en cuestión presentaban una unión del factor transcripcional ER $\alpha$ .

Este resultado es muy importante ya que sugiere la posibilidad de una regulación en cis del gen *FOXF1* por parte del transcrito del *LINC01082*. La secuencia de eventos propuesta para esta respuesta transcripcional inicia con la entrada del estradiol a la célula. Posteriormente, el ER $\alpha$  se une al estradiol y forma un dímero, permitiendo su translocación al núcleo celular. En este lugar, el ER $\alpha$  se une a elementos reguladores en la cromatina, incluyendo los promotores de genes. En el caso específico del *LINC01082*, la unión del ER $\alpha$  en su región promotora conduce



a un aumento en su expresión transcripcional, lo cual a su vez establece una interacción con el promotor del gen *FOXF1* antes de la estimulación con E2. Esto plantea la posibilidad de que el transcrito del *LINC01082* ejerza su efecto en cis sobre el gen *FOXF1*, regulándolo transcripcionalmente a través de los mecanismos epigenéticos característicos de los lncRNAs, como "scaffold", "guide" y "decoy" (Figura 17). Este conjunto de eventos culmina en una respuesta transcripcional ante el E2 y, de manera indirecta, en la activación de la vía del ER $\alpha$ .

En el contexto de estudios en ratones, la eliminación del gen *FOXF1* resulta en la aparición de una malformación congénita en la cual el esófago no se desarrolla de manera apropiada, según ha sido documentado en una investigación anterior (Shaw-Smith, 2010). En pacientes humanos, se han identificado mutaciones heterocigotas en el gen *FOXF1*, asociadas con la displasia capilar alveolar y malformaciones pulmonares. Estas mutaciones también ejercen un impacto en la expresión de otros genes esenciales para el desarrollo pulmonar, como *VEGFA*, *VEGFR2*, *BMP-4*, *TBX*, *LKLF*, *FGF-10*, *GLI3* y *Notch-2*, afectando además la formación adecuada de los vasos sanguíneos y linfáticos en los pulmones (Stankiewicz et al., 2009).

En investigaciones adicionales, se ha resaltado la función crucial de *FOXF1* en la promoción de la morfogénesis pulmonar al inducir la proliferación celular en el tejido mesenquimal pulmonar fetal. Este factor transcripcional se expresa en diversos linajes celulares mesenquimales y desempeña un papel esencial en el desarrollo de varios sistemas de órganos, abarcando los sistemas respiratorio, cardiovascular y gastrointestinal (Ustiyani et al., 2018). Específicamente, *FOXF1* ejerce una función clave en la regulación de la señalización mesenquimal-epitelial y estimula la proliferación celular en el tejido mesenquimal pulmonar durante la etapa fetal (Ustiyani et al., 2018).

A partir de la información expuesta, es evidente que el gen *FOXF1* emerge como un factor transcripcional de relevancia en los procesos de desarrollo de los tejidos pulmonar y gastrointestinal, a pesar de que su estudio hasta la fecha ha sido

relativamente limitado. Sin embargo, resulta importante considerar su posible función en la modulación de la respuesta transcripcional frente a la estimulación con estradiol, específicamente en el contexto de la vía del receptor de estrógeno (ER) en la línea celular MCF-7.

Este planteamiento cobra una importancia significativa dado que FOXF1, como se ha mencionado previamente, desempeña un papel modulador en la proliferación celular y en la transición de la morfología celular desde el estado epitelial al mesenquimal. Estos procesos son de suma relevancia en la progresión del cáncer de mama luminal A, al cual pertenece la línea celular MCF7. Además, esta línea celular exhibe una respuesta genética y fisiológica específica a la vía de señalización del ER. Por lo tanto, explorar la potencial interacción entre FOXF1 y la respuesta al estradiol en esta vía podría abrir nuevas perspectivas en la comprensión de los mecanismos subyacentes a esta forma de cáncer de mama.

En experimentos realizados en ratones, se observó que, aunque no existen homólogos directos de los lncRNAs *LINC01082* y *LINC01081*, existe un lncRNA llamado Gm26878 ubicado en la región del ratón que es sinténica con la secuencia enhancer upstream del gen humano *FOXF1*. Sin embargo, no se observaron cambios en la expresión de *FOXF1* tras la delección de la secuencia del gen mediante el sistema CRISPR/Cas9. Esto sugiere que la regulación transcripcional de este gen difiere entre humanos y ratones (Jiska Cohen-Mansfield, Maha Dakheel-Ali, MD, Marcia S. Marx, PhD, Khin Thein, MD, and Natalie G. Regier & Waage et al., 2017). Este hallazgo resalta aún más la importancia de estudiar los lncRNAs para comprender los mecanismos regulatorios complejos que distinguen a los seres humanos de otros organismos biológicos.

Además, es crucial explorar las implicaciones del transcrito del lncRNA en los mecanismos moleculares que desencadenan la activación de *FOXF1*. Existe la posibilidad de que el lncRNA reclute proteínas activadoras, secuestre correceptores u desempeñe otras funciones aún desconocidas, con potenciales efectos no solo en el gen *FOXF1*, sino también en otros genes. Por lo tanto, es

esencial profundizar en este aspecto para obtener una comprensión más completa de los mecanismos subyacentes y el impacto global de este lncRNA en la regulación transcripcional.

Con el propósito de abordar este aspecto, propongo llevar a cabo un estudio adicional mediante el uso de RNA-seq en células MCF-7 que sobreexpresen tanto el gen *LINC01082* como *FOXF1* de manera individual. Este enfoque nos permitirá identificar el conjunto de genes que experimentan cambios significativos en su expresión y que podrían estar bajo la regulación de estos dos genes. Asimismo, sugiero realizar un análisis de ChIP-seq de FOXF1 en células MCF-7 para determinar si este grupo de genes compartidos también presenta la unión del factor transcripcional FOXF1 en regiones promotoras o enhancers.

Estos experimentos proporcionarán información crucial para dilucidar si los efectos observados en la proliferación y viabilidad de las células MCF-7 son el resultado de una regulación transcripcional directa del *LINC01082* sobre *FOXF1*. Además, nos permitirán explorar si el *LINC01082* podría estar regulando de manera directa otros genes o si su influencia es principalmente a través de la modulación de FOXF1. Cabe destacar que llevar a cabo la sobreexpresión del *LINC01082* mediante un plásmido, en lugar del sistema de activación CRISPRa, nos permitirá diferenciar si es el transcrito del lncRNA o la activación transcripcional del gen lo que conduce al aumento en la expresión de *FOXF1*. Estas investigaciones adicionales serán fundamentales para esclarecer la compleja red de regulación génica en juego.

## 10. Conclusiones

El presente estudio ofrece pruebas que respaldan la implicación del *LINC01082* y el gen *FOXF1* en la respuesta transcripcional mediada por el estradiol en la línea celular MCF-7. Aunque no se pudo realizar una validación experimental completa de los hallazgos computacionales en este contexto específico (unión de ER $\alpha$  y aumento transcripcional ante E2), observamos un aumento en la transcripción de ambos genes en presencia de estradiol. Se ha planteado la hipótesis de que el *LINC01082* actúa como un regulador transcripcional, siendo activado por el factor de transcripción ER $\alpha$  y, a su vez, promoviendo la sobreexpresión del gen *FOXF1*.

Este estudio subraya la importancia crítica del *LINC01082* y el gen *FOXF1* en los procesos fisiológicos celulares que se ven perturbados en el contexto del cáncer. Estos actores contribuyen a la progresión del cáncer al estimular la proliferación celular y otros procesos a través de una respuesta transcripcional ante la estimulación hormonal. Además, se plantea la posibilidad de que el lncRNA tenga funciones adicionales, como la capacidad de reclutar proteínas activadoras o modular la expresión de otros genes en respuesta a E2 a través de la vía de ER $\alpha$ , lo que establece un mecanismo de regulación génica altamente complejo e intrigante (Figura 17).

Es importante destacar que, si bien hemos arrojado luz sobre estos hallazgos, aún se requieren investigaciones adicionales para comprender completamente los mecanismos subyacentes y las interacciones precisas entre el *LINC01082* y *FOXF1*. No obstante, los resultados presentados aquí abren nuevas vías de estudio que permiten explorar el papel funcional y la relevancia clínica de estas moléculas en diversos contextos fisiológicos y patológicos. Estudios futuros podrían contribuir significativamente a nuestro conocimiento de la regulación transcripcional y a la identificación de posibles estrategias terapéuticas dirigidas específicamente a estas moléculas clave. En última instancia, estos avances podrían tener un impacto positivo en la comprensión y el tratamiento de condiciones médicas relacionadas con la regulación hormonal y la respuesta transcripcional.

## 11. Perspectivas

Las perspectivas que surgen a partir de este proyecto son realmente prometedoras y abren diversas direcciones para futuras investigaciones en el campo de los lncRNAs y su relación con la expresión génica mediante la vía de receptores nucleares. Entre las perspectivas más relevantes se encuentran las siguientes:

1. **Validación Experimental:** Aunque este estudio no pudo llevar a cabo una validación experimental completa, es esencial realizar experimentos adicionales para confirmar los hallazgos obtenidos a través de análisis computacionales. Estos experimentos pueden incluir técnicas como ensayos de unión de proteínas, análisis de cromatina y estudios de interacción entre ARN y proteínas para evaluar de manera específica las interacciones entre el *LINC01082*, el gen *FOXF1* y las proteínas reguladoras involucradas en estos procesos.
2. **Mecanismos de Regulación:** Se requiere una investigación más profunda para comprender en detalle los mecanismos precisos mediante los cuales el *LINC01082* regula la expresión del gen *FOXF1*. Esto implica identificar las proteínas y complejos de transcripción involucrados en esta regulación, así como caracterizar las secuencias regulatorias en el *LINC01082* que permiten su interacción con estos factores. Un enfoque más detallado en estos mecanismos puede proporcionar información valiosa sobre la función de estos elementos en la regulación génica.
3. **Papel Funcional del linc01082:** Además de su relación con el gen *FOXF1*, es fundamental investigar si el *LINC01082* tiene un impacto funcional en otros procesos biológicos y si regula la expresión de otros genes. Estudios de pérdida y ganancia de función pueden ser esenciales para comprender los efectos del silenciamiento o la sobreexpresión del *LINC01082* en la expresión génica global y en fenotipos celulares relevantes. Ya que como ha sido reportado en la literatura se ha visto asociado en a procesos proliferativos en células de tipo epitelial de tejido de origen pulmonar y

gástrico (Li et al., 2018; Stankiewicz et al., 2009; Ustiyan et al., 2018b; S. Wang et al., 2018).

4. **Relevancia Clínica:** Explorar la relevancia clínica del *LINC01082* y el gen *FOXF1* en enfermedades humanas, especialmente en trastornos del desarrollo pulmonar y gastrointestinal, podría ser una perspectiva fascinante. Investigaciones que analicen la expresión y la función de estas moléculas en muestras clínicas, así como su relación con la respuesta a terapias específicas, pueden proporcionar información crucial para el diseño de enfoques terapéuticos personalizados. Por ejemplo, el diseño de fármacos que sobre expresen o inhiban al *LINC01082* en pacientes con cáncer de mamá luminal A y les permita tener un mejor pronóstico.

En resumen, las perspectivas futuras derivadas de este proyecto abarcan desde la validación experimental hasta la exploración de los mecanismos de regulación, el papel funcional y la relevancia clínica del *LINC01082* y su relación con el gen *FOXF1*. Estas investigaciones adicionales tienen el potencial de expandir significativamente nuestro conocimiento sobre los lncRNAs y su contribución a la regulación transcripcional. Además, podrían tener implicaciones clínicas significativas en el tratamiento de enfermedades humanas, ofreciendo nuevas oportunidades para la medicina personalizada y la terapia dirigida.

## 12. Materiales y métodos

### 12.1 Cultivo Celular

La línea celular MCF-7, que corresponde a un tipo de cáncer luminal A de tipo epitelial, fue cultivada en medio RPMI 1640. (Thermo Fisher Scientific, Waltham, MA, EE. UU.) suplementado con un 10% de suero fetal bovino (FBS) y un 1% de penicilina/estreptomicina a 37°C en una atmósfera humidificada que contenía un 5% de CO<sub>2</sub>.

### 12.2 RT-qPCR

El RNA total se extrajo de las células MCF-7 utilizando los sistemas de sobreexpresión CRISPRa SAM, que incluyeron sgRNA (para la sobreexpresión del LINC01082) y EV (vector vacío de control), en triplicado, mediante el uso de Trizol siguiendo las instrucciones del fabricante. Posteriormente, el RNA se utilizó directamente para llevar a cabo la qRT-PCR, empleando el kit KAPA SYBR FAST One Step de KAPA Byosystems, con al menos dos réplicas técnicas por muestra y utilizando un sistema de PCR en tiempo real StepOne. En la Tabla 1 se proporciona una lista con todos oligonucleótidos utilizados en este estudio. Los datos obtenidos se analizaron mediante el método  $\Delta\Delta C_t$  y las diferencias estadísticamente significativas en la expresión génica entre las condiciones se determinaron mediante una prueba t de Student (valor  $p < 0,05$ ) utilizando el software Graphpad Prisma 8.3.0.

### 12.3 Análisis de base de datos de ChIP-seq

Enlace de la base de datos: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59530>

Se llevaron a cabo las siguientes etapas de análisis de datos: primero, las lecturas crudas en formato fastq se alinearon al genoma de referencia humano (hg19) utilizando el software de alineación Bowtie2 (v.2.5.0). Posteriormente, los archivos resultantes del proceso de mapeo se importaron al programa deepTools (v3.3.1) para generar archivos de señal bamCoverage (v3.5.2). Finalmente, se procedió al llamado de picos utilizando el software MACS2 (v.2.2.7).

## 12.4 Análisis de base de datos de RNA-seq

Enlace de la base de datos: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78169>

El análisis de datos se llevó a cabo en varias etapas. En primer lugar, se alinearon las lecturas crudas en formato fastq al genoma de referencia humano (hg19) utilizando el software de alineación STAR RNA-seq (v2.7.8). A continuación, se cuantificaron los niveles de abundancia de los transcritos utilizando FeatureCounts (v.2.0.1) y se empleó la anotación de Gencode v39lift37. Se procedió al análisis de expresión diferencial mediante EdgeR, identificando los genes diferencialmente expresados (DEGs) con un valor de FDR < 0.01 y un fold change > 2 como criterios de selección. Por último, se obtuvieron los conteos normalizados (FPKM) mediante el uso de StringTie (v.2.2.1), y se consideraron como genes transcripcionalmente activos aquellos con un FPKM > 5.

## 12.5 Extracción de RNA

La obtención de RNA se realizó mediante la técnica y el reactivo TRIzol. Se procedió a la resuspensión del pellet de células en 1 ml de TRIzol, asegurándose de homogeneizar la muestra completamente. A continuación, se añadieron 200 µl de cloroformo para separar el RNA por centrifugación durante 15 min a 4°C a 10,000 rpm. Posteriormente, se incorporaron 500 µl de isopropanol y se dejó incubar en hielo durante 10 min. Se centrifugó la muestra durante 10 min a 4°C a 10,000 rpm. Finalmente, el pellet resultante se lavó con etanol al 70%, se eliminó el sobrenadante y se dejó secar durante 10 min a temperatura ambiente. Para concluir, el pellet se resuspendió en 10-50 µl de agua inyectable y se llevó a cabo la cuantificación mediante NanoDrop-2000.

## 12.6 Ensayo de proliferación y viabilidad celular de MTT

El ensayo de proliferación de MTT (del inglés Dimethylthiazol Diphenyltetrazolium Bromide) y se utiliza en investigaciones biológicas y médicas para evaluar la viabilidad y la proliferación celular, se basa en la conversión de este compuesto químico en un producto de formazán de color violeta por parte de las células vivas. Este cambio de color se emplea como indicador indirecto de la cantidad de células viables en una muestra.



Para llevar a cabo el ensayo, se sembraron 2,500 células MCF-7 en cada pozo de una placa de 96 pocillos. Se dispuso de múltiples pozos para obtener triplicados de cada muestra y se evaluó el proceso proliferativo de las células durante un período de 6 días. Para realizar la lectura, se agregaron inicialmente 10  $\mu$ l del reactivo MTT en cada pozo, seguido de la adición de 200  $\mu$ l de un buffer de solubilización después de 4 h, con el fin de disolver el formazán y liberar la característica coloración violeta. Cuanto más intensa sea la coloración, mayor será el número de células viables. La cuantificación de este último aspecto se llevó a cabo midiendo la absorbancia a 595 nm utilizando un lector de placas.

### 12.7 Prueba estadística t de Student

Para evaluar las diferencias significativas en los resultados obtenidos, se empleó la prueba t de Student, una herramienta estadística ampliamente utilizada para comparar medias entre dos grupos. Este análisis se llevó a cabo con el objetivo de verificar la presencia de efectos significativos en las variables estudiadas.

Procedimiento:

- **Recopilación de Datos:** Se recopilaron datos experimentales relevantes para la investigación, considerando las condiciones específicas de cada grupo de estudio.
- **Organización de Datos:** Se organizó y estructuró la información recopilada en conjuntos de datos correspondientes a los grupos de interés, lo que permitió una comparación sistemática.
- **Selección de Grupos Comparativos:** Se identificaron los grupos de interés para la comparación, como los sujetos de estudio expuestos a diferentes tratamientos o condiciones.
- **Definición de Hipótesis Nula y Alternativa:** Se formuló la hipótesis nula ( $H_0$ ) que sostiene que no hay diferencias significativas entre los grupos, y la hipótesis alternativa ( $H_1$ ) que propone la presencia de diferencias significativas.

- Elección del Nivel de Significancia ( $\alpha$ ): Se estableció un nivel de significancia predeterminado (comúnmente  $\alpha = 0.05$ ) para determinar la probabilidad de rechazar la hipótesis nula.
- Realización de la Prueba t de Student: Se aplicó la prueba t de Student a los conjuntos de datos, calculando la diferencia de medias y considerando la variabilidad interna de cada grupo.
- Interpretación de Resultados: Los resultados se interpretaron en función del valor p obtenido. Un valor p menor que el nivel de significancia establecido indicaría la presencia de diferencias significativas.
- Consideración de Limitaciones y Supuestos: Se tuvieron en cuenta las limitaciones y supuestos asociados con la prueba t de Student, como la asunción de distribución normal de los datos y homogeneidad de varianzas.

Software Utilizado: El análisis estadístico se llevó a cabo utilizando software especializado, concretamente la versión 8.3.0 del programa GraphPad Prism. Este enfoque metodológico permitió una evaluación rigurosa de las diferencias entre grupos, proporcionando una base sólida para las conclusiones y hallazgos presentados en la investigación.

## 13. Bibliografia

- Aiello, A., Bacci, L., Re, A., Ripoli, C., Pierconti, F., Pinto, F., Masetti, R., Grassi, C., Gaetano, C., Bassi, P. F., Pontecorvi, A., Nanni, S., & Farsetti, A. (2016). MALAT1 and HOTAIR Long Non-Coding RNAs Play Opposite Role in Estrogen-Mediated Transcriptional Regulation in Prostate Cancer Cells. *Scientific Reports*, 6(August), 1–11. <https://doi.org/10.1038/srep38414>
- Atianand, M. K., & Fitzgerald, K. A. (2014). Long non-coding rnas and control of gene expression in the immune system. *Trends in Molecular Medicine*, 20(11), 623–631. <https://doi.org/10.1016/j.molmed.2014.09.002>
- Baran-Gale, J., Purvis, J. E., & Sethupathy, P. (2016). An integrative transcriptomics approach identifies miR-503 as a candidate master regulator of the estrogen response in MCF-7 breast cancer cells. *Rna*, 22(10), 1592–1603. <https://doi.org/10.1261/rna.056895.116>
- Beesley, J., Sivakumaran, H., Moradi Marjaneh, M., Lima, L. G., Hillman, K. M., Kaufmann, S., Tuano, N., Hussein, N., Ham, S., Mukhopadhyay, P., Kazakoff, S., Lee, J. S., Michailidou, K., Barnes, D. R., Antoniou, A. C., Fachal, L., Dunning, A. M., Easton, D. F., Waddell, N., ... Edwards, S. L. (2020). Chromatin interactome mapping at 139 independent breast cancer risk signals. *Genome Biology*, 21(1). <https://doi.org/10.1186/s13059-019-1877-y>
- Bhan, A., Hussain, I., Ansari, K. I., Kasiri, S., Bashyal, A., & Mandal, S. S. (2013). Antisense transcript long noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by estradiol. *Journal of Molecular Biology*, 425(19), 3707–3722. <https://doi.org/10.1016/j.jmb.2013.01.022>
- Bhan, A., Soleimani, M., & Mandal, S. S. (2017). Long noncoding RNA and cancer: A new paradigm. *Cancer Research*, 77(15), 3965–3981. <https://doi.org/10.1158/0008-5472.CAN-16-2634>
- Bojcsuk, D., & Bálint, B. L. (2019). Classification of different types of estrogen receptor alpha binding sites in MCF-7 cells. *Journal of Biotechnology*, 299(April), 13–20. <https://doi.org/10.1016/j.jbiotec.2019.04.016>
- Bojcsuk, D., Nagy, G., & Balint, B. L. (2017). Inducible super-enhancers are organized based on canonical signal-specific transcription factor binding elements. *Nucleic Acids Research*, 45(7), 3693–3706. <https://doi.org/10.1093/nar/gkw1283>
- Braeuer, R. R., Walker, N. M., Misumi, K., Mazzoni-Putman, S., Aoki, Y., Liao, R., Vittal, R., Kleer, G. G., Wheeler, D. S., Sexton, J. Z., Farver, C. F., Welch, J. D., & Lama, V. N. (2021). Transcription factor FOXF1 identifies compartmentally distinct mesenchymal cells with a role in lung allograft fibrogenesis. *Journal of Clinical Investigation*, 131(21), 1–8. <https://doi.org/10.1172/JCI147343>
- Dahariya, S., Paddibhatla, I., Kumar, S., Raghuvanshi, S., Palapati, A., & Gutti, R. K. (2019). Long non-coding RNA : Classification , biogenesis and functions in blood cells. 112(April), 82–92. <https://doi.org/10.1016/j.molimm.2019.04.011>
- Eni. (1967). 濟無 No Title No Title No Title. *Angewandte Chemie International Edition*, 6(11), 951–952., *Mj*, 5–24.

- Evans, R. M., & Mangelsdorf, D. J. (2014). Nuclear receptors, RXR, and the big bang. *Cell*, *157*(1), 255–266. <https://doi.org/10.1016/j.cell.2014.03.012>
- Fitz-James, M. H., & Cavalli, G. (2022). Molecular mechanisms of transgenerational epigenetic inheritance. *Nature Reviews. Genetics*, *0123456789*. <https://doi.org/10.1038/s41576-021-00438-5>
- Franco, H. L., Nagari, A., & Kraus, W. L. (2015). TNF $\alpha$  signaling exposes latent estrogen receptor binding sites to alter the breast cancer cell transcriptome. *Molecular Cell*, *58*(1), 21–34. <https://doi.org/10.1016/j.molcel.2015.02.001>
- Fuentes, N., & Silveyra, P. (2019). Estrogen receptor signaling mechanisms. In *Advances in Protein Chemistry and Structural Biology* (1st ed., Vol. 116). Elsevier Inc. <https://doi.org/10.1016/bs.apcsb.2019.01.001>
- Guan, J., Zhou, W., Hafner, M., Blake, R. A., Chalouni, C., Chen, I. P., De Bruyn, T., Giltnane, J. M., Hartman, S. J., Heidersbach, A., Houtman, R., Ingalla, E., Kategaya, L., Kleinheinz, T., Li, J., Martin, S. E., Modrusan, Z., Nannini, M., Oeh, J., ... Metcalfe, C. (2019). Therapeutic Ligands Antagonize Estrogen Receptor Function by Impairing Its Mobility. *Cell*, *178*(4), 949-963.e18. <https://doi.org/10.1016/j.cell.2019.06.026>
- Guo, Z., Lu, X., Yang, F., He, C., Qin, L., Yang, N., Han, C., & Wu, J. (2022). Exosomal LINC01213 Plays a Role in the Transition of Androgen-Dependent Prostate Cancer Cells into Androgen-Independent Manners. *Journal of Oncology*, *2022*. <https://doi.org/10.1155/2022/8058770>
- Hah, N., Danko, C. G., Core, L., Waterfall, J. J., Siepel, A., Lis, J. T., & Kraus, W. L. (2011). A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell*, *145*(4), 622–634. <https://doi.org/10.1016/j.cell.2011.03.042>
- He, R. Z., Jiang, J., Hu, X., Lei, M., Li, J., Luo, W., Duan, L., Hu, Z., Mo, Y. Y., Luo, D. X., & Peng, W. X. (2021). Stabilization of UCA1 by N6-methyladenosine RNA methylation modification promotes colorectal cancer progression. *Cancer Cell International*, *21*(1), 1–12. <https://doi.org/10.1186/s12935-021-02288-x>
- He, Y. H., Yeh, M. H., Chen, H. F., Wang, T. S., Wong, R. H., Wei, Y. L., Huynh, T. K., Hu, D. W., Cheng, F. J., Chen, J. Y., Hu, S. W., Huang, C. C., Chen, Y., Yu, J., Cheng, W. C., Shen, P. C., Liu, L. C., Huang, C. H., Chang, Y. J., & Huang, W. C. (2021). ER $\alpha$  determines the chemo-resistant function of mutant p53 involving the switch between lincRNA-p21 and DDB2 expressions. *Molecular Therapy - Nucleic Acids*, *25*, 536–553. <https://doi.org/10.1016/j.omtn.2021.07.022>
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., Khalil, A. M., Zuk, O., Amit, I., Rabani, M., Attardi, L. D., Regev, A., Lander, E. S., Jacks, T., & Rinn, J. L. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, *142*(3), 409–419. <https://doi.org/10.1016/j.cell.2010.06.040>
- Ikeda, K., Horie-Inoue, K., & Inoue, S. (2015). Identification of estrogen-responsive genes based on the DNA binding properties of estrogen receptors using high-throughput sequencing technology. *Acta Pharmacologica Sinica*, *36*(1), 24–31. <https://doi.org/10.1038/aps.2014.123>

- Jameera Begam, A., Jubie, S., & Nanjan, M. J. (2017). Estrogen receptor agonists/antagonists in breast cancer therapy: A critical review. In *Bioorganic Chemistry* (Vol. 71, pp. 257–274). Academic Press Inc. <https://doi.org/10.1016/j.bioorg.2017.02.011>
- Jiska Cohen-Mansfield, Maha Dakheel-Ali, MDb, Marcia S. Marx, PhD, Khin Thein, MDb, and Natalie G. Regier, P., & Waage et al. (2017). 乳鼠心肌提取 HHS Public Access. *Physiology & Behavior*, *176*(1), 139–148. <https://doi.org/10.1007/s00335-017-9686-7>. CRISPR/Cas9-mediated
- Johnston, S. R. D. (2010). New strategies in estrogen receptor-positive breast cancer. *Clinical Cancer Research*, *16*(7), 1979–1987. <https://doi.org/10.1158/1078-0432.CCR-09-1823>
- Katoh, M., Igarashi, M., Fukuda, H., Nakagama, H., & Katoh, M. (2013). Cancer genetics and genomics of human FOX family genes. *Cancer Letters*, *328*(2), 198–206. <https://doi.org/10.1016/j.canlet.2012.09.017>
- Klinge, C. M., Jernigan, S. C., Mattingly, K. A., Risinger, K. E., & Zhang, J. (2004). Estrogen response element-dependent regulation of transcriptional activation of estrogen receptors  $\alpha$  and  $\beta$  by coactivators and corepressors. *Journal of Molecular Endocrinology*, *33*(2), 387–410. <https://doi.org/10.1677/jme.1.01541>
- Li, Y., Zhang, W., Liu, P., Xu, Y., Tang, L., Chen, W., & Guan, X. (2018). Long non-coding RNA FENRR inhibits cell proliferation and is associated with good prognosis in breast cancer. *Oncotargets and Therapy*, *11*, 1403–1412. <https://doi.org/10.2147/OTT.S149511>
- Milewski, D., Shukla, S., Gryder, B. E., Pradhan, A., Donovan, J., Sudha, P., Vallabh, S., Pyros, A., Xu, Y., Barski, A., Szabo, S., Turpin, B., Pressey, J. G., Millay, D. P., Khan, J., Kalinichenko, V. V., & Kalin, T. V. (2021). FOXF1 is required for the oncogenic properties of PAX3-FOXO1 in rhabdomyosarcoma. *Oncogene*, *40*(12), 2182–2199. <https://doi.org/10.1038/s41388-021-01694-9>
- Núñez-Martínez, H. N., & Recillas-Targa, F. (2022). Emerging Functions of lncRNA Loci beyond the Transcript Itself. *International Journal of Molecular Sciences*, *23*(11). <https://doi.org/10.3390/ijms23116258>
- Pan, X., Li, D., Huo, J., Kong, F., Yang, H., & Ma, X. (2018). LINC01016 promotes the malignant phenotype of endometrial cancer cells by regulating the miR-302a-3p/miR-3130-3p/NFYA/SATB1 axis. *Cell Death and Disease*, *9*(3). <https://doi.org/10.1038/s41419-018-0291-9>
- Perissi, V., & Rosenfeld, M. G. (2005). Controlling nuclear receptors: The circular logic of cofactor cycles. *Nature Reviews Molecular Cell Biology*, *6*(7), 542–554. <https://doi.org/10.1038/nrm1680>
- Sedano, M. J., Harrison, A. L., Zilaie, M., Das, C., Choudhari, R., Ramos, E., & Gadad, S. S. (2020). Emerging roles of estrogen-regulated enhancer and long non-coding RNAs. *International Journal of Molecular Sciences*, *21*(10). <https://doi.org/10.3390/ijms21103711>

- Shaw-Smith, C. (2010). Genetic factors in esophageal atresia, tracheo-esophageal fistula and the VACTERL association: Roles for FOXF1 and the 16q24.1 FOX transcription factor gene cluster, and review of the literature. *European Journal of Medical Genetics*, 53(1), 6–13. <https://doi.org/10.1016/j.ejmg.2009.10.001>
- Słowikowski, B. K., Lianeri, M., & Jagodziński, P. P. (2017). Exploring estrogenic activity in lung cancer. *Molecular Biology Reports*, 44(1), 35–50. <https://doi.org/10.1007/s11033-016-4086-8>
- Sommer, S., & Fuqua, S. A. W. (2001). Estrogen receptor and breast cancer. *Seminars in Cancer Biology*, 11(5), 339–352. <https://doi.org/10.1006/scbi.2001.0389>
- Stankiewicz, P., Sen, P., Bhatt, S. S., Storer, M., Xia, Z., Bejjani, B. A., Ou, Z., Wiszniewska, J., Driscoll, D. J., Bolivar, J., Bauer, M., Zackai, E. H., McDonald-McGinn, D., Nowaczyk, M. M. J., Murray, M., Shaikh, T. H., Martin, V., Tyreman, M., Simonic, I., ... Shaw-Smith, C. (2009). Genomic and Genic Deletions of the FOX Gene Cluster on 16q24.1 and Inactivating Mutations of FOXF1 Cause Alveolar Capillary Dysplasia and Other Malformations. *American Journal of Human Genetics*, 84(6), 780–791. <https://doi.org/10.1016/j.ajhg.2009.05.005>
- Sun, Y., Zhang, H., Ma, R., Guo, X., Zhang, G., Liu, S., Zhu, W., Liu, H., & Gao, P. (2023). ETS-1-activated LINC01016 over-expression promotes tumor progression via suppression of RFFL-mediated DHX9 ubiquitination degradation in breast cancers. *Cell Death and Disease*, 14(8). <https://doi.org/10.1038/s41419-023-06016-3>
- Tzerpos, P., Daniel, B., & Nagy, L. (2021). Global Run-on Sequencing (GRO-Seq). *Methods in Molecular Biology*, 2351, 25–39. [https://doi.org/10.1007/978-1-0716-1597-3\\_2](https://doi.org/10.1007/978-1-0716-1597-3_2)
- Ustiyani, V., Bolte, C., Zhang, Y., Han, L., Xu, Y., Yutzey, K. E., Zorn, A. M., Kalin, T. V., Shannon, J. M., & Kalinichenko, V. V. (2018). FOXF1 transcription factor promotes lung morphogenesis by inducing cellular proliferation in fetal lung mesenchyme. *Developmental Biology*, 443(1), 50–63. <https://doi.org/10.1016/j.ydbio.2018.08.011>
- Wang, J., Sun, J., & Yang, F. (2020). The role of long non-coding RNA H19 in breast cancer (Review). *Oncology Letters*, 19(1), 7–16. <https://doi.org/10.3892/ol.2019.11093>
- Wang, S., Yan, S., Zhu, S., Zhao, Y., Yan, J., Xiao, Z., Bi, J., Qiu, J., Zhang, D., Hong, Z., Zhang, L., Huang, C., Li, T., Liang, L., Liao, W., Jiao, H., Ding, Y., & Ye, Y. (2018). FOXF1 Induces Epithelial-Mesenchymal Transition in Colorectal Cancer Metastasis by Transcriptionally Activating SNAI1. *Neoplasia (United States)*, 20(10), 996–1007. <https://doi.org/10.1016/j.neo.2018.08.004>
- Yu, V. C., Näär, A. M., & Rosenfeld, M. G. (1992). Transcriptional regulation by the nuclear receptor superfamily. *Current Opinion in Biotechnology*, 3(6), 597–602. [https://doi.org/10.1016/0958-1669\(92\)90002-Z](https://doi.org/10.1016/0958-1669(92)90002-Z)
- Zhou, Y., Sun, W., Qin, Z., Guo, S., Kang, Y., Zeng, S., & Yu, L. (2021). LncRNA regulation: New frontiers in epigenetic solutions to drug chemoresistance. *Biochemical Pharmacology*, 189, 114228. <https://doi.org/10.1016/j.bcp.2020.114228>