



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

ESTUDIO GENÓMICO DE LA DESMETILASA KDM4A Y SU PARTICIPACIÓN  
EN LA REGULACIÓN DE LA EXPRESIÓN DE GENES INVOLUCRADOS EN  
PROCESOS TUMORALES

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
ITZEL ALEJANDRA HERNÁNDEZ ROMERO

TUTOR PRINCIPAL

DR. ERNESTO SOTO REYES SOLÍS  
Universidad Autónoma Metropolitana

MIEMBROS DEL COMITÉ TUTOR

DRA. LORENA AGUILAR ARNAL  
Instituto de Investigaciones Biomédicas

DR. VÍCTOR JULIÁN VALDÉS RODRÍGUEZ  
Instituto de Fisiología Celular

Ciudad de México, 2020



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.

## **AGRADECIMIENTOS**

Agradezco al Posgrado de Maestría y Doctorado en Ciencias Bioquímicas de la Universidad Nacional Autónoma de México y a los profesores que lo coordinan, por haberme permitido realizar mis estudios de maestría.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo brindado a través del programa de becarios (CVU 886138), el Fondo CB-SEP-CONACyT (284748), el Fondo Sectorial de Investigación en Salud y Seguridad Social (FOSISS, Grant No.0261181), Fondo CB-SEP-CONACyT (284748) y el Fondo Desarrollo Tecnológico e Innovación COVID-19 (312021); así como a la Universidad Autónoma Metropolitana por el apoyo recibido mediante del Fondo UAM-PTC-704.

Agradezco también, al Programa de Apoyo a los Estudios del Posgrado (PAEP) por beneficiar mi asistencia al Congreso XLIV Congreso Nacional de Genética Humana.

A mi Tutor Principal, el Dr. Ernesto Soto Reyes Solís por brindarme su guía y amistad durante mis estudios de maestría en la Universidad Autónoma Metropolitana, Unidad Cuajimalpa.

A los miembros de mi Comité Tutor, la Dra. Lorena Aguilar Arnal y el Dr. Víctor Julián Valdés Rodríguez, por la asesoría y el apoyo recibidos durante el desarrollo del proyecto.

A los miembros del Jurado de Tesis, Dra. Dimitrova Dinkova Tzvetanka, Dra. Elizabeth Langley, Dr. Rodrigo González Barrios de la Parra, Dr. Ángel Alfonso Zarain Herzberg y al Dr. Alejandro Manuel García Carrancá por sus valiosas observaciones al trabajo escrito.

Al Dr. Alejandro García Carrancá y la Dra. Yessenia Sánchez, por brindarme su apoyo durante mi estancia en el Instituto Nacional de Cancerología.

Agradezco al Dr. Nicolás Alcaraz adscrito a la Universidad Técnica de Munich, por su apoyo y tiempo dedicado al análisis bionfornático de los datos de secuenciación masiva.

A mis compañeros de laboratorio, en especial a M. en C. Lissania Guerra por su amistad y consejo.

A mi familia, especialmente, a mi hermana Naomi Hernández y a mis padres Alejandra Romero y David Hernández, que siempre me acompañan de manera incondicional.

A mis amigos, Teresa Alejo, Stella Quiroz, Manuela Rincón y Javier Fregoso por su apoyo a través de los años.

## ABREVIATURAS

<b>ANOVA</b>	Análisis de Varianza
<b>BCSC</b>	Células Troncales de Cáncer de mama
<b>ChIP</b>	Ensayos de Inmunoprecipitación de la Cromatina
<b>ChIP-seq</b>	Secuenciación Masiva de los Ensayos ChIP
<b>CRISPR/Cas9</b>	Agrupación de Repeticiones Palindrómicas Cortas Regularmente Intercaladas y Proteína Asociada Cas9
<b>CTCF</b>	Factor de unión a CCCTC
<b>DAPI</b>	4',6-diamidino-2-fenilindol
<b>DNA</b>	Ácido desoxirribonucleico
<b>ER</b>	Receptor de Estrógenos
<b>HER2</b>	Factor de Crecimiento Epidermal 2
<b>KDM4A</b>	Desmetilasa de Lisinas 4A
<b>KD</b>	Knock Down/Subexpresión
<b>KO</b>	Knock Out/Silenciamiento
<b>MCF7</b>	Línea celular de cáncer de mama. Acrónimo para "Michigan Cancer Foundation-7"
<b>MCF7-KDM4A<sup>KO</sup></b>	Células MCF7 Knock Out para KDM4A
<b>ME</b>	Matriz Extracelular
<b>PR</b>	Receptor de Progesterona
<b>qPCR</b>	Cuantificación de la Reacción en Cadena de la Polimerasa
<b>RNA</b>	Ácido ribonucleico
<b>RFP</b>	Proteína roja florescente
<b>TP53</b>	Proteína tumoral p53
<b>TSS</b>	Sitio del Inicio de la Transcripción
<b>UTR</b>	Regiones no Traducidas
<b><math>\alpha</math>-KG</b>	$\alpha$ -cetoglutarato

## ÍNDICE

<b>RESUMEN</b> .....	<b>1</b>
<b>ABSTRACT</b> .....	<b>2</b>
<b>INTRODUCCIÓN</b> .....	<b>3</b>
1. CÁNCER DE MAMA.....	3
1.1 <i>Características y clasificación</i> .....	3
1.2 <i>Modelos de carcinogénesis</i> .....	4
1.2.1 <b>Modelo Multietapas</b> .....	4
1.2.2 <b>Características distintivas del cáncer</b> .....	6
1.2.3 <b>Plasticidad epigenética en cáncer</b> .....	8
2. EPIGENÉTICA .....	9
2.1 <i>Niveles de compactación de la cromatina</i> .....	9
2.1.1 <b>El factor de transcripción CTCF</b> .....	10
2.2 <i>Modificadores de la cromatina</i> .....	13
3. DESMETILASA DE HISTONAS KDM4A.....	15
3.1 <i>Dominios estructurales de KDM4A</i> .....	16
3.2 <i>La expresión de KDM4A en cáncer</i> .....	18
4. ANTECEDENTES DIRECTOS .....	18
4.1. <i>Genes involucrados en procesos tumorales</i> .....	18
4.1.1 <b>El gen WRAP53</b> .....	18
4.1.2 <b>El gen JAG2</b> .....	20
4.2 <i>Desmetilasas y su interacción con CTCF</i> .....	21
<b>PLANTEAMIENTO DEL PROBLEMA</b> .....	<b>23</b>
<b>HIPÓTESIS</b> .....	<b>24</b>
<b>OBJETIVOS</b> .....	<b>24</b>
<b>OBJETIVO GENERAL</b> .....	<b>24</b>
<b>OBJETIVOS PARTICULARES</b> .....	<b>24</b>
<b>METODOLOGÍA</b> .....	<b>24</b>
<b>RESULTADOS</b> .....	<b>28</b>
1. LOS SITIOS DE UNIÓN DE LA DESMETILASA KDM4A Y EL COMPLEJO CTCF-KDM4A COINCIDEN CON MÚLTIPLES GENES RELACIONADOS CON LA ADHESIÓN DE LAS CÉLULAS MCF7 .....	28
2. VALIDACIÓN DE UN MODELO KNOCK OUT DE KDM4A EN CÉLULAS MCF7 .....	37
3. LA DESMETILASA DE HISTONAS KDM4A REGULA LA CAPACIDAD MIGRATORIA E INVASIVA DE LAS CÉLULAS MCF7.....	40
4. LA DESMETILASA DE HISTONAS KDM4A REGULA LA TRANSCRIPCIÓN DE LOS GENES JAG2 Y WRAP53 .....	43
<b>DISCUSIÓN</b> .....	<b>46</b>
<b>CONCLUSIONES</b> .....	<b>50</b>
<b>PERSPECTIVAS</b> .....	<b>50</b>
<b>REFERENCIAS</b> .....	<b>50</b>
<b>ANEXOS</b> .....	<b>63</b>
1. METODOLOGÍA EN EXTENSO .....	63
2. INFORMACIÓN ACADÉMICA .....	79

## RESUMEN

La desmetilasa de histonas KDM4A es una enzima modificadora de la cromatina, cuya sobreexpresión está relacionada con la formación de tumores colorrectales, pulmonares y mamarios, entre otros. En una investigación previa realizada en nuestro grupo de trabajo, se observó que la sobreexpresión e interacción de KDM4A con el factor de transcripción CTCF regula negativamente la expresión del gen supresor de tumores *CHD5* en una línea celular de cáncer de mama, mediante un mecanismo de regulación transcripcional que tiene lugar en el cuerpo de dicho gen. Dada la posible relevancia de este mecanismo de regulación génica, se decidió evaluar la participación genómica del complejo KDM4A-CTCF en la regulación de la expresión de múltiples genes relacionados con procesos tumorales en el cáncer de mama. Primero, se realizaron estudios de secuenciación masiva a partir de ensayos de inmunoprecipitación de la cromatina (ChIP) y ChIP-ReChIP para identificar los sitios de unión de KDM4A y el complejo KDM4A-CTCF en la línea celular de cáncer de mama luminal MCF7. Un análisis bioinformático reveló que la mayoría de los sitios de unión genómica coinciden con regiones promotoras, intergénicas e intrónicas, en concordancia con el papel regulador de KDM4A en la síntesis y elongación de los transcritos. Entre los oncogenes, supresores de tumores y de reparación enriquecidos que muestran picos de enriquecimiento del complejo KDM4A-CTCF, se encontró una larga lista de genes relacionados con la capacidad migratoria de las células. Para evaluar el efecto de la expresión de KDM4A en el comportamiento de las células neoplásicas, se implementó y validó un modelo de Knock Out de KDM4A en la línea celular MCF7 utilizando un sistema CRISPR/Cas9. Dos ensayos independientes de migración e invasión celular revelaron que el Knock Out de KDM4A disminuye la capacidad de las células para realizar dichas funciones. Además, a través de un análisis de expresión cuantitativa (RT-qPCR) se observó que el Knock Out de esta desmetilasa afecta negativamente la transcripción de dos blancos génicos identificados en los estudios de secuenciación masiva, *JAG2* y *WRAP53*. Los hallazgos del presente trabajo sugieren el papel de KDM4A y el complejo KDM4A-CTCF en la regulación de la expresión de múltiples genes relacionados a la progresión del cáncer de mama.

## ABSTRACT

The KDM4A histone demethylase is a chromatin-modifying enzyme, whose overexpression is related to the formation of colorectal, pulmonary and mammary tumors, among others. In previous studies carried out in our workgroup, it was observed that the overexpression and interaction of KDM4A with the transcription factor CTCF within the body of the tumor suppressor gene *CHD5*, mediates its under-expression in breast cancer. Given the possible relevance of this gene regulation mechanism, we aimed to evaluate the genomic participation of the KDM4A-CTCF complex in the regulation of multiple genes related to tumor processes in breast cancer. First, were performed chromatin immunoprecipitation (ChIP) and ChIP-ReChIP-sequencing assays to identify the binding sites of KDM4A and the KDM4A-CTCF complex in the MCF7 luminal breast cancer cell line. Bioinformatic analysis reveals that most of the genomic binding sites coincided with promotor, intergenic and intronic regions, in accordance with the regulatory role of KDM4A in the synthesis and elongation of transcripts. Among the KDM4A-CTCF enriched oncogenes, tumor suppressors and DNA repair genes, a long list of genes were related to the migratory capacity of cells. To assess the effect of KDM4A expression on the behavior of neoplastic cells, we implemented and validated a KDM4A Knock Out model using a CRISPR/Cas9 system. Two independent cell migration and invasion assays revealed that the KDM4A Knock Out impacts on the ability of cells to perform those functions. Furthermore, through quantitative expression analysis (RT-qPCR) it was observed that Knock Out of this demethylase negatively affects the transcription of two target genes identified in the massive sequencing studies, *JAG2* and *WRAP53*. The findings of this project suggest the role of KDM4A and the KDM4A-CTCF complex in regulating the expression of multiple genes related to breast cancer progression.

# INTRODUCCIÓN

## 1. Cáncer de mama

### 1.1 Características y clasificación

El cáncer es un grupo de enfermedades caracterizadas por el crecimiento descontrolado y la propagación de células que dejan de ser sensibles a las señales tisulares que regulan su diferenciación, sobrevivencia, proliferación y muerte; causando lesiones locales y diversas respuestas sistémicas («Cancer - Latest research and news | Nature» s. f.). De acuerdo con el Instituto Nacional de Cáncer de los Estados Unidos (NCI), existen más de 200 tipos de cáncer responsables de miles de decesos alrededor del mundo (Song, Merajver, y Li 2015). De éstos, el cáncer de mama es el tipo de cáncer con la mayor incidencia en mujeres y el segundo con la más alta tasa de mortalidad a nivel mundial (Siegel, Miller, y Jemal 2018).

La Organización Mundial de la Salud (OMS) ha clasificado al cáncer de mama como una enfermedad heterogénea que comprende 18 subtipos histológicos y cuatro subtipos moleculares (Oh, Wang, y Muscat 2017). La clasificación histológica de los tumores es compleja y ha cambiado considerablemente a lo largo de los años. La distinción más general de los tipos histológicos de cáncer de mama considera la existencia de carcinomas ductales y lobulares, con base en sus patrones estructurales, características citológicas y perfiles inmunohistoquímicos (Weigelt y Reis-Filho 2009). En la década pasada, los estudios basados en microrreglos de expresión sentaron las bases para identificar las características moleculares del cáncer de mama (Perou et al. 2000). Hoy en día, los subtipos moleculares son identificados a través de biomarcadores que reflejan la respuesta hormonal de los tumores, dada por la presencia o ausencia del receptor de estrógenos (ER), de progesterona (PR) y del factor de crecimiento epidérmico 2 (HER2). Dichas hormonas regulan de manera fisiológica el crecimiento y la proliferación de las células que componen al tejido mamario; y su expresión en cáncer de mama constituye un importante factor predictivo de la respuesta terapéutica (Beckmann et al. 1997). Los cuatro subtipos moleculares de cáncer de mama son: luminal A (ER+ y/o PR+; HER2-), luminal B (ER+ y/o PR+; HER2+), tipo basal (ER-, PR- y HER2-) y HER2 enriquecido (ER-, PR- y HER2+) (Byler et al. 2014). En la clínica, los tipos de cáncer Luminal A y B se asocian a un mejor pronóstico, siendo los Luminal A los de mejor respuesta. El fenotipo Luminal A representa alrededor del 71% de los casos de cáncer y los Luminal B el 12%. El tipo HER2 enriquecido representa al 5% de los casos de cáncer. Mientras que el tipo basal tiene un mal pronóstico y prevalece en el 12% de los tumores invasivos (DeSantis et al. 2019; Tang et al. 2016).

La respuesta hormonal de los tumores puede estar condicionada a múltiples factores ambientales y fisiológicos. A pesar de que el cáncer es típicamente considerado una enfermedad genética, de acuerdo con las Estadísticas Globales de Cáncer, solo entre el 5-10% de los casos de cáncer de mama están asociados a la herencia de mutaciones genéticas y más



del 80% de las neoplasias ocurren en mujeres sin evidencia familiar previa (Bray et al. 2018). Lo anterior sugiere que más de un factor etiológico interviene en la carcinogénesis; es decir, el proceso por el cual una célula normal se transforma en cancerosa.

## **1.2 Modelos de carcinogénesis**

Existen varios modelos acerca de la adquisición celular de características malignas; algunos de ellos muestran coincidencias conceptuales y podrían considerarse complementarios, pero no existe a la fecha un modelo de carcinogénesis unificado.

### **1.2.1 Modelo Multietapas**

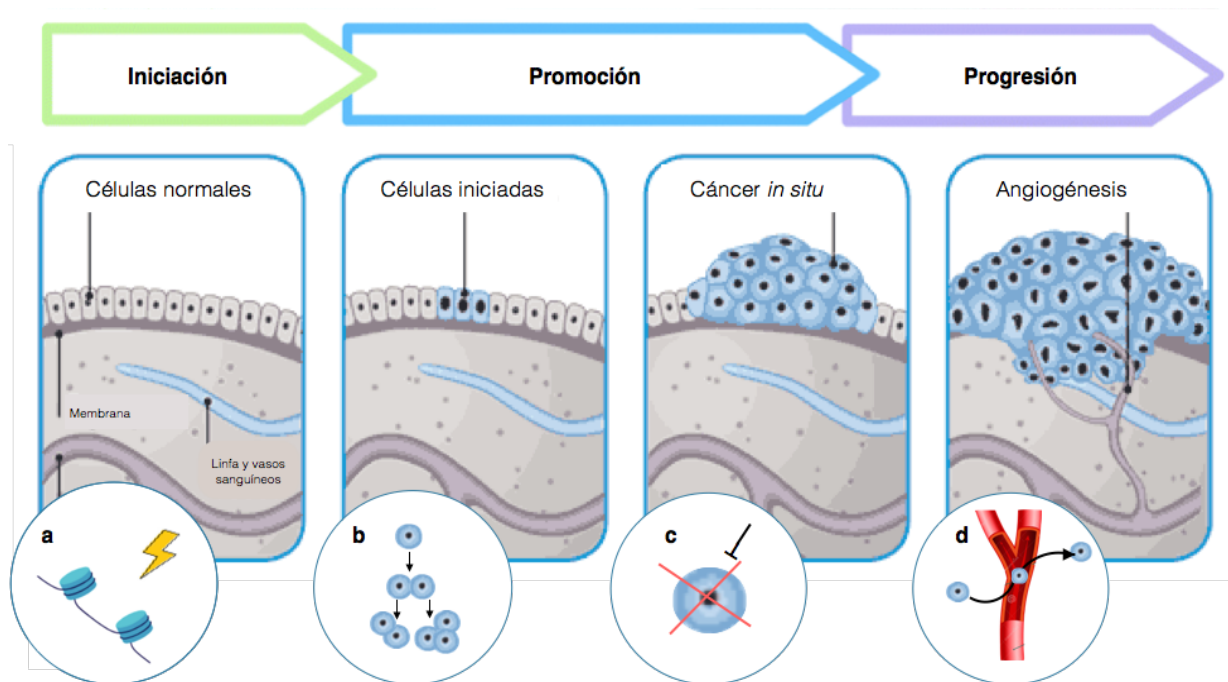
El modelo “Multietapas” de la carcinogénesis propone que la transformación de las células neoplásicas ocurre a través de múltiples eventos moleculares, donde las células sufren una acumulación sucesiva de alteraciones genéticas que cambian su comportamiento (Foulds 1954; Sharma, Kelly, y Jones 2010; Virani et al. 2012). El modelo Multietapas distingue por lo menos tres etapas en el proceso carcinogénico: la iniciación, la promoción y la progresión (Figura 1).

La iniciación es la etapa en la que una célula normal presenta un daño genético que no es reparado y la “inicia” en el proceso carcinogénico (Figura 1a). El estímulo iniciador o carcinógeno puede provenir de agentes químicos, físicos y biológicos; y causar daño al ácido desoxirribonucleico (DNA), guiando a la adquisición de mutaciones y al encendido o apagado de los genes (Tanaka et al. 2013). Este modelo no descarta que la acción de un carcinógeno se sume al efecto de otros factores de riesgo hereditarios o relacionados al estilo de vida; como el sobrepeso, el padecimiento de diabetes y la exposición de hormonas esteroides durante tratamientos de menopausia y la anticoncepción, entre otros (Bray et al. 2018; DeSantis et al. 2019).

Las alteraciones adquiridas durante la iniciación pueden seguirse acumulando en la etapa de promoción y afectar la expresión de genes que promueven la sobrevivencia y proliferación celular (oncogenes) y/o genes que inhiben el desarrollo de los tumores (genes supresores de tumores). La etapa de promoción involucra la expansión clonal de las células iniciadas en una lesión focalizada (Figura 1b). Estas células capaces de evadir las señales de muerte celular, son susceptibles al estímulo de “promotores”, sustancias dependientes de la dosis y el tiempo de exposición que estimulan la proliferación (Tanaka et al. 2013). Los promotores suelen ser fármacos u hormonas que no interactúan directamente con el DNA, pero de alguna manera influyen la transcripción génica. A medida que el tiempo pasa, los tumores crecen hasta alcanzar un tamaño que es detectable por medio de las herramientas de detección y diagnóstico que son aplicadas en la clínica (Tanaka et al. 2013).

La heterogeneidad de los tumores es una de las características más notables en cáncer que refleja el desarrollo sucesivo de varias subpoblaciones celulares que coexisten dentro de un tumor primario. Las sub-poblaciones difieren en su comportamiento, morfología y marcadores de superficie (Beckmann et al. 1997; Esparza-López et al. 2017). La adquisición de estas características se atribuye a factores genéticos y a la presión que ejerce un proceso llamado selección clonal. La hipótesis de la selección clonal propone que las clonas con el mejor desempeño para un ambiente en particular serán seleccionadas durante la promoción del tumor (Almendro y Fuster 2011).

En cáncer de mama, la heterogeneidad intra-tumoral se adquiere durante la selección clonal y tiene como resultado la amplificación diferencial de los genes que codifican a la proteína tumoral p53 (*TP53*), *HER2*, *ER* y *PR* en diversas regiones del tumor primario; sin embargo, la heterogeneidad parece perderse a medida que se acerca la etapa de progresión de la carcinogénesis. Un estudio donde se analizó la evolución de 21 tumores de cáncer de mama, encontró que un solo linaje sub-clonal puede llegar a representar hasta al 50% de las células tumorales (Esparza-López et al. 2017). Este fenómeno se ha observado en otros modelos, donde después de un primer ciclo de expansión clonal, la muerte celular disminuye al igual que el número de divisiones celulares, dando como resultado una población celular menos variable (Labi y Erlacher 2015).



**Figura 1.** Modelo de carcinogénesis Multietapas en cáncer de mama. En la iniciación, **a)** las células normales reciben un estímulo carcinogénico que desencadena la aparición de mutaciones y alteraciones epigenéticas en oncogenes y genes supresores de tumores. El daño provocado por un carcinógeno puede sumarse al efecto de otros factores de riesgo hereditarios y relacionados con el estilo de vida. En la promoción, **b)** Las células que han sido iniciadas son susceptibles a la acción de sustancias promotoras que favorecen la expansión clonal de diferentes subpoblaciones tumorales *in situ*. **c)** Las subpoblaciones son seleccionadas de acuerdo con su desempeño en el microambiente tumoral y su capacidad de resistir a las señales de muerte celular. En la progresión, **d)** las células tumorales adquieren un fenotipo maligno, la habilidad de inducir angiogénesis y hacer metástasis. Modificado de referencia (Mothoneos y Pelosi 2017). BioRender.

La resistencia a la muerte celular puede verse influenciada por la pérdida de función de los genes supresores de tumores y la acción sinérgica de los oncogenes (Figura 1c). Por ejemplo, las mutaciones que causan la pérdida de función de *TP53* resultan en procesos de resistencia a la apoptosis y son frecuentemente asociados con un mal pronóstico (Labi y Erlacher 2015). Durante la selección clonal, la muerte de las células iniciadas genera nichos vacantes dentro del ambiente tumoral, que son repoblados por clonas más agresivas. De manera que, la apoptosis incrementa la presión proliferativa y promueve la selección clonal (Labi y Erlacher 2015).

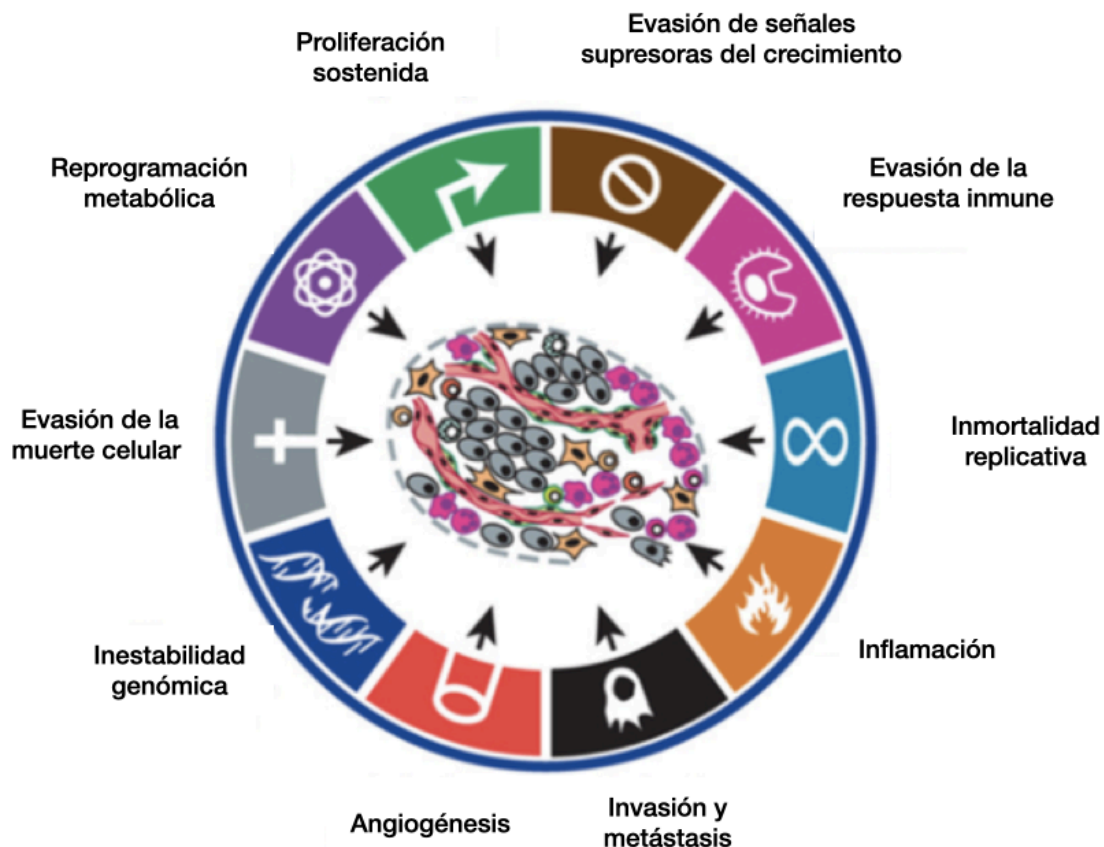
La “conversión” de las células tumorales durante la etapa de progresión es irreversible e implica la ganancia de fenotipos malignos con la capacidad de migrar e invadir otros tejidos (Tanaka et al. 2013) (Figura 1d). Las células convertidas proliferan independientemente del estímulo y tienen la capacidad de transformar su ambiente mediante la liberación de factores de crecimiento. La metástasis inicia con la invasión del tejido adyacente al tumor primario, la incursión de las células tumorales en los vasos sanguíneos o linfáticos y finaliza con la diseminación a órganos distantes (Scully et al. 2012). Si las células tumorales tienen éxito en su nuevo ambiente, el proceso de migración e invasión se puede repetir y producir una metástasis secundaria. Se ha observado que el cáncer de mama hace metástasis con más frecuencia a los huesos y los pulmones que a órganos como el hígado y el cerebro (Gandhi y Das 2019). En esta etapa de la carcinogénesis las lesiones o neoplasmas son altamente distinguibles del resto del parénquima y las células experimentan cambios metabólicos (Scully et al. 2012).

### **1.2.2 Características distintivas del cáncer**

Las “características distintivas del cáncer” fueron descritas en un inicio como “las reglas que gobiernan la transformación de las células normales en cancerosas” (D. Hanahan y Weinberg 2000); aunque constituyen en realidad una recapitulación de las habilidades más importantes que adquieren las células durante el desarrollo tumoral: 1) la capacidad proliferativa e 2) inmortalidad replicativa, 3) la habilidad para evadir señales supresoras del crecimiento y 4) de muerte celular; así como, 5) la competencia para inducir angiogénesis y 6) activar procesos de invasión y metástasis (D. Hanahan y Weinberg 2000).

Las características distintivas del cáncer conservan vigencia hasta nuestros días y no se contraponen de manera directa con la visión adaptativa del modelo Multietapas. De hecho, este modelo recapitula muchos de los elementos del modelo Multietapas, como la relevancia de las mutaciones en oncogenes y/o genes supresores de tumores que se pueden heredar y favorecer la adquisición de las marcas distintivas del cáncer. Este modelo también propone que la acumulación de alteraciones conduce a un fenotipo maligno, donde la sinergia de varios oncogenes puede “reprogramar los circuitos” que regulan las vías de señalización celular (D. Hanahan y Weinberg 2000).

La diferencia más importante entre la visión del modelo Multietapas y las características distintivas del cáncer es que, si bien las células cancerosas son capaces de acumular diferentes habilidades, la adquisición de éstas no sigue un único orden cronológico. El modelo sugiere que eventualmente todas las células cancerosas adquieren más de una característica distintiva del cáncer, pero que el orden de adquisición puede variar dependiendo del tipo de cáncer y que incluso, pueden adquirirse dos o más habilidades al mismo tiempo (D. Hanahan y Weinberg 2000).



**Figura 2.** Las características distintivas del cáncer. La ilustración reúne las diez habilidades que adquieren las células cancerosas durante su desarrollo. Modificado de referencia (Douglas Hanahan y Weinberg 2011).

Después de 10 años de su publicación, el modelo de las características distintivas del cáncer incorporó dos elementos más: las características facilitadoras y emergentes del cáncer. Como su nombre lo indica, las características facilitadoras del cáncer promueven la adquisición de otras habilidades a través de 7) la inflamación e 8) inestabilidad genómica. Las características emergentes del cáncer están involucradas en la patogénesis, guiando 9) la reprogramación metabólica y 10) la evasión de la respuesta inmune (Douglas Hanahan y Weinberg 2011) (Figura 2).

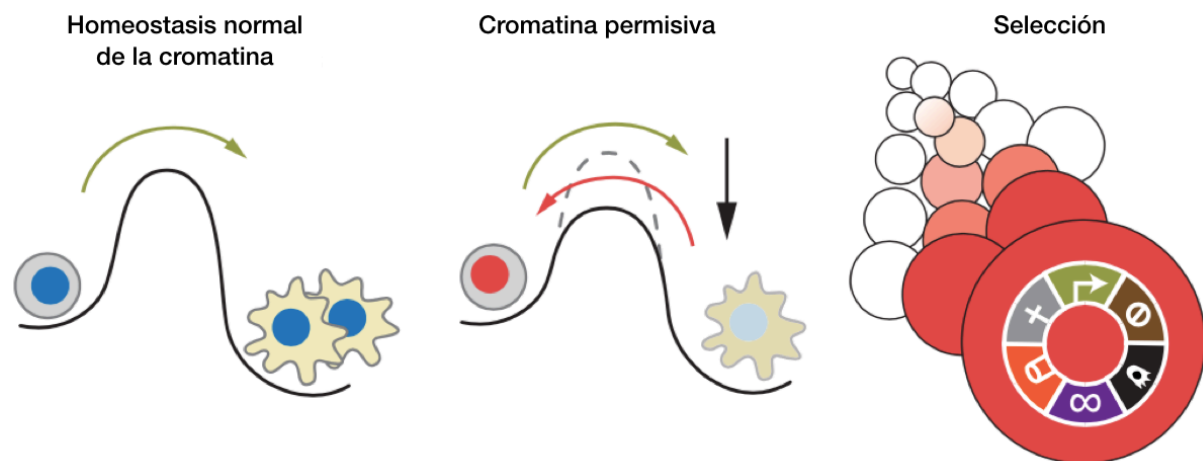
La nueva propuesta de las características distintivas del cáncer destaca la importancia de la evolución del “microambiente tumoral” y las interacciones entre las células cancerosas y las

células adyacentes que conforman el tejido. Además, sitúa a los mecanismos epigenéticos como elementos heredables que influyen la expresión de oncogenes y genes supresores de tumores (Douglas Hanahan y Weinberg 2011).

### 1.2.3 Plasticidad epigenética en cáncer

La epigenética se define como el estudio de los cambios heredables en la expresión génica que no dependen de la secuencia desnuda del DNA. Entre los mecanismos de regulación epigenética que coordinan la organización del núcleo y la transcripción de los genes, se encuentran la metilación del DNA, las modificaciones de histonas, el posicionamiento de las histonas y la regulación de las moléculas no codificantes de RNA (Kanwal, Gupta, y Gupta 2015).

Desde que se estableció el término en 1942, Conrad Waddington sugirió que la epigenética era un puente entre el genotipo y el fenotipo, capaz de dar explicación a los procesos de diferenciación celular que ocurren en un mismo organismo y otorgan a las células diversidad en forma y función. Bajo esta premisa se construiría más adelante el concepto de “paisaje epigenético”, como el conjunto de variables que intervienen en la toma de decisiones de una célula indiferenciada y la guían durante el desarrollo hacia diferentes destinos celulares (Goldberg, Allis, y Bernstein 2007).



**Figura 3.** Modelo de Plasticidad epigenética y cáncer. La homeostasis normal de la cromatina estabiliza la expresión génica y la diferenciación celular (azul). Las alteraciones ambientales, genéticas y metabólicas cambian la conformación de la cromatina haciéndola permisiva a cambios en la expresión génica. La plasticidad epigenética puede conferir habilidades que seleccionan a las células cancerosas (rojo). Modificado de referencia (Flavahan, Gaskell, y Bernstein 2017).

El modelo de la “plasticidad epigenética” retoma los principios básicos de las características distintivas del cáncer y sustenta el papel de la epigenética en la adquisición de fenotipos malignos. Este modelo propone que las alteraciones genéticas y epigenéticas comparten una relación estrecha durante el desarrollo de los tumores, donde el estado de la cromatina confiere

plasticidad para activar o reprimir la expresión de oncogenes y genes supresores de tumores (Flavahan, Gaskell, y Bernstein 2017) (Figura 3).

La plasticidad de la cromatina puede verse afectada por estímulos genéticos y no-genéticos que afectan la expresión y actividad de los reguladores epigenéticos que están involucrados en la carcinogénesis (Chatterjee, Rodger, y Eccles 2018). Algunos estímulos no-genéticos son la inflamación, el envejecimiento, la hipoxia, el estrés celular y el metabolismo (Flavahan, Gaskell, y Bernstein 2017). La desregulación epigenética suele tener consecuencias globales en la transcripción génica, por lo que en los últimos años se ha hecho un gran esfuerzo para identificar y validar nuevos blancos terapéuticos (Kelly y Issa 2017; Baylin y Jones 2011).

## **2. Epigenética**

### **2.1 Niveles de compactación de la cromatina**

Un genoma es el compendio completo de información genética de una especie. El genoma humano contiene alrededor de 2,85 billones de nucleótidos contenidos en el núcleo celular, en estructuras super-enrolladas que denominamos cromosomas (International Human Genome Sequencing Consortium 2004). Los cromosomas se componen de cromatina; es decir, DNA y proteínas, que en conjunto pueden adoptar diferentes niveles de compactación. El primer nivel de compactación involucra la formación de estructuras llamadas nucleosomas, donde 146 pares de bases (pb) de DNA se enrollan alrededor de un núcleo de ocho proteínas, denominadas histonas. Las proteínas que conforman de manera canónica a este núcleo son dos dímeros de histonas H2A/H2B y un tetrámero H3/H4 (Cutter y Hayes 2015). La unión entre el DNA y las histonas es posible gracias a interacciones electrostáticas; la carga positiva que las histonas adquieren por su elevado contenido de aminoácidos básicos, como la lisina y arginina, favorece la atracción de las cargas negativas de los grupos fosfato del DNA en condiciones fisiológicas de pH. Las histonas H1 y H5 permanecen al exterior del núcleo de ocho histonas, pero constituye un también un elemento clave en la condensación de la cromatina y la formación de estructuras de orden superior (Happel y Doenecke 2009).

Las dimensiones de un nucleosoma rondan entre los 10 nm de diámetro y 5 nm de altura. A la estructura cromatínica que forma una colección de nucleosomas se le conoce como “collar de perlas”. Esta estructura permite un acomodo laxo de la cromatina (eucromatina) y la unión de la maquinaria de transcripción génica (Felsenfeld y Groudine 2003). Las unidades transcripcionales o genes se componen normalmente de a) un promotor en el extremo 5' de la cadena codificante de DNA, b) dos regiones 5' y 3' no traducidas (UTR) que flanquean al cuerpo del gen y contienen señales para la maduración de los RNA mensajeros (mRNA) y la regulación de la traducción y c) el cuerpo del gen que cuenta con secuencias codificantes (exones) y no codificantes (intrones) (Hegde y Crowley 2019). Las secuencias de nucleótidos que flanquean al 5'-UTR y al 3'-UTR proveen las señales moleculares para el Inicio de la

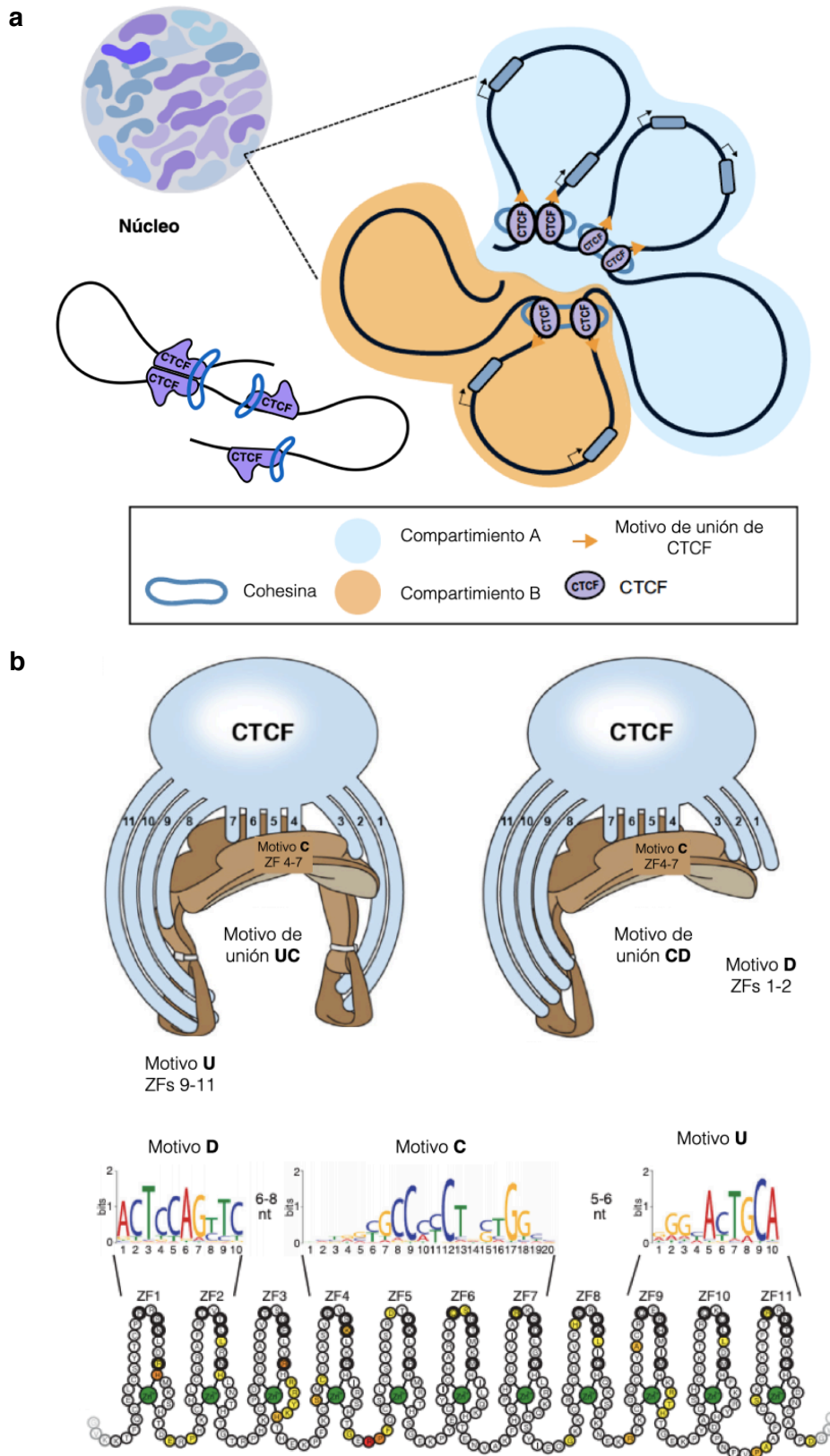
Transcripción (TSS) y el Término de la Transcripción (TTS), respectivamente. A las secuencias de DNA que se encuentran entre unidades transcripcionales se les conoce como regiones intergénicas (Hegde y Crowley 2019).

Al segundo nivel de compactación se conoce como solenoide o fibra de 30 nm y consta de 6 o 7 nucleosomas por giro, dispuestos con sus caras planas paralelas al eje de la fibra (Felsenfeld y Groudine 2003). En este nivel la transcripción génica se dificulta. A la cromatina que puede cambiar de un estado transcripcional a otro se le conoce como heterocromatina facultativa, mientras que la cromatina altamente condensada y transcripcionalmente inactiva se le denomina heterocromatina constitutiva (Huisinga, Brower-Toland, y Elgin 2006). La fibra de 30 nm debe compactarse otras 100 veces para formar asas de cromatina. La sucesiva compactación de las fibras forma dominios de cromatina de 700 nm. Estos dominios se disponen en diferentes regiones del núcleo de las células en interfase y se compactan a 1400 nm durante la división celular para formar el clásico cromosoma metafásico (Felsenfeld y Groudine 2003). Se ha visto que la organización territorial de los cromosomas dentro del núcleo interfásico está altamente regulada y tiene un efecto directo sobre la transcripción de los genes (Cremer y Cremer 2010). La cercanía entre diferentes regiones inter e intra-cromosomales favorece el contacto de múltiples secuencias génicas y elementos regulatorios que cooperan con la maquinaria transcripcional (Fraser y Bickmore 2007). De manera, que la compactación y la disposición de la cromatina en el núcleo juegan un papel crucial en la regulación de la expresión génica.

### **2.1.1 El factor de transcripción CTCF**

La organización de la cromatina al interior del núcleo se concibe gracias a la participación de proteínas arquitectónicas, entre las que destaca el factor de unión a la secuencia CCCTC (CTCF). Esta proteína, conocida en un inicio como NeP1, se sitúa entre los límites de los dominios topológicamente asociados (TAD) y junto con el complejo proteico de la cohesina forma asas de cromatina que facilitan la interacción entre diferentes regiones del genoma (Jabbari, Chakraborty, y Wiehe 2019; Burcin et al. 1997). La formación de asas establece la compartimentalización del núcleo y mantiene una barrera entre la señalización de secuencias intensificadoras de la transcripción llamadas enhancers y los promotores de un dominio y otro (Merkenschlager y Nora 2016) (Figura 4a).

CTCF es una proteína evolutivamente conservada en vertebrados que se compone de un dominio N-terminal, un dominio central con 11 dedos de zinc y un dominio C-terminal (Heger et al. 2012). Los dedos de zinc son un tipo de motivo estructural de las proteínas que les permiten unirse al surco mayor del DNA. Los dedos de zinc consisten en dos hojas antiparalelas  $\beta$  plegada y una  $\alpha$  hélice, que son estabilizadas por la presencia de un ion zinc (Bulyk et al. 2001).





Dentro de un mismo compartimiento e incluso dentro de un mismo gen, pueden encontrarse varios sitios de unión de CTCF. Se estima que CTCF se une en alrededor de 40,000-80,000 sitios en el genoma de los mamíferos, que corresponden predominantemente a regiones intergénicas (53%), intrónicas (30%) y promotoras (12%) que se sobrelapan con secuencias regulatorias como los enhancers (H. Chen et al. 2012). En el genoma humano, CTCF cuenta con más de 77,811 sitios de unión, de los cuales el 50,149 son variables entre un tipo celular y otro (H. Wang et al. 2012).

El reclutamiento de CTCF a sus motivos de unión en el DNA está mediado en el 80% de los casos por los dedos de zinc 3-7 de su dominio central (Hashimoto et al. 2017). A la secuencia de 15-20 pb que reconocen los dedos de zinc 4-7 se le conoce como “motivo C” o “motivo nuclear” (Nakahashi et al. 2013). Se cree que la función de los otros dedos de zinc, 1-2 y 8-11, es estabilizar la unión de CTCF a las secuencias no conservadas que flanquean al motivo de unión (Nakahashi et al. 2013).

El modelo de la “silla de caballo” sugiere que “el motivo U” reconocido por los dedos 9-11 estabiliza la unión de CTCF. Al formarse “el motivo UC” los dedos 1-2 contribuyen a la unión de CTCF al asociarse con secuencias de DNA que no cuentan con ningún motivo consenso. En ausencia del motivo U, se forma “el motivo CD” en el que los dedos 1-2 pierden afinidad por el DNA y el dedo 8 se agrupa con los dedos 9-11 para estabilizar la unión de CTCF. De acuerdo con este modelo, la asociación del dedo 3 es esencial en los sitios sin motivos de unión U (Nakahashi et al. 2013) (Figura 4b).

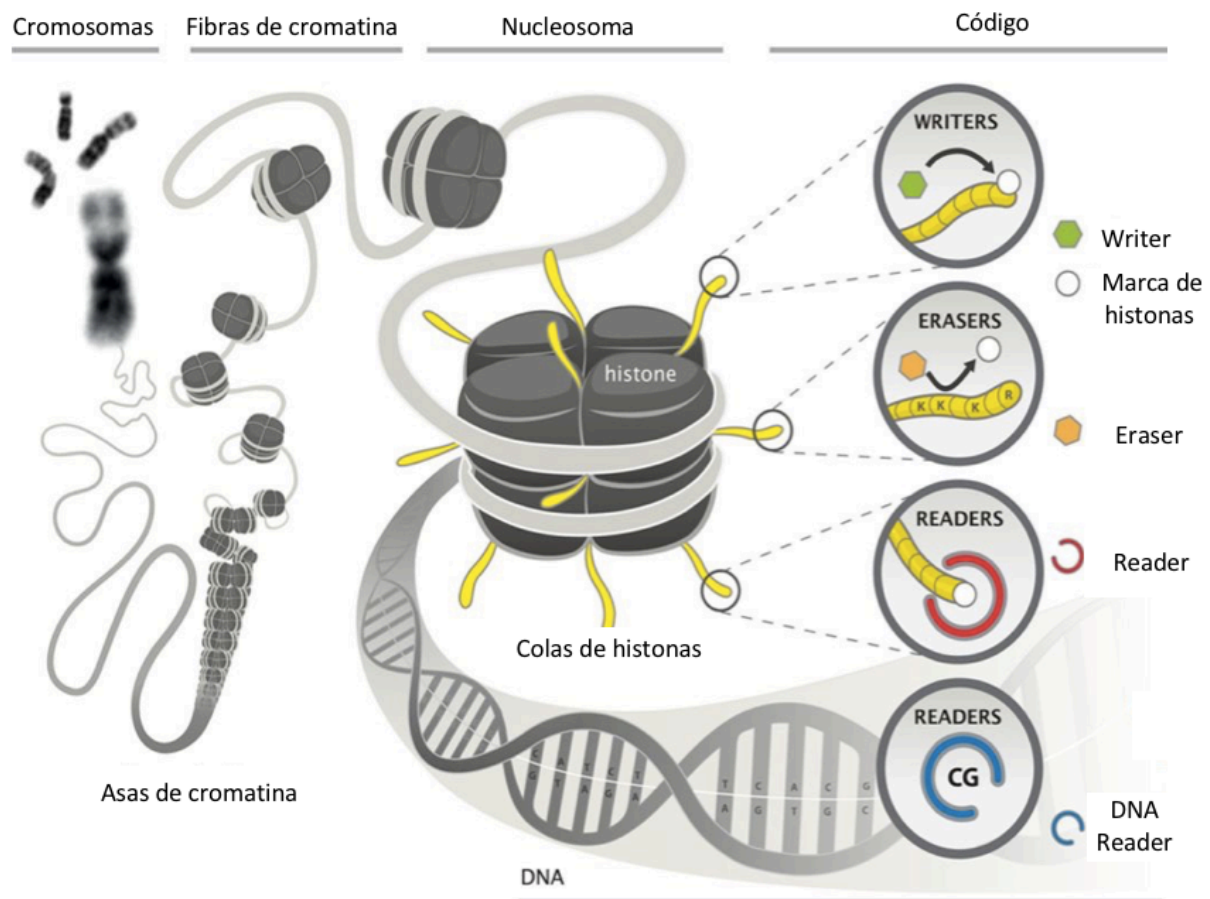
El reclutamiento de CTCF también puede verse influenciado por las modificaciones covalentes del DNA. Existe evidencia de que la proteína CTCF tiene poca afinidad de unión a secuencias de DNA metiladas (Hashimoto et al. 2017). Además, la direccionalidad en la secuencia del motivo de unión de CTCF determina la formación de asas de cromatina; para que estas se formen, la secuencia del motivo de unión debe estar orientada de manera convergente (de Wit et al. 2015) (Figura 4a).

Además de desempeñar un papel crucial en la organización del núcleo, CTCF funge como factor de transcripción y elemento regulatorio de muchas otras actividades importantes para el mantenimiento del genoma, como la reparación del DNA, la recombinación homóloga, el splicing alternativo, la impronta génica, la supresión de tumores y el efecto aislante que se denomina “insulator” (Arzate-Mejía, Recillas-Targa, y Corces 2018; Ohlsson, Bartkuhn, y Renkawitz 2010). Estudios recientes muestran que CTCF puede interactuar con cientos de moléculas endógenas de RNA y mantener de manera simultánea su unión con el DNA (Kung et al. 2015; Saldaña-Meyer et al. 2014). De igual manera, la proteína CTCF es capaz de contactar a una gran variedad de proteínas, entre las que destacan la RNA polimerasa II (RNAPII), la laminina, la poli(ADP-ribosa) polimerasa 1 (PARP1) y varias proteínas modificadoras de la cromatina (Zlatanova y Caiafa 2009).

## 2.2 Modificadores de la cromatina

Cada uno de los los mecanismos de regulación epigenética responde a las señales de un “código” dinámico que se establece en la cromatina dependiendo del contexto fisiológico y las señales extracelulares. En este código participan elementos que colocan, reconocen y retiran marcas o modificaciones químicas en los nucleótidos de DNA y las regiones amino terminales de las histonas (Figura 5).

Los “writers” o enzimas que colocan las modificaciones en el DNA son metiltransferasas (DNMTs) que catalizan la transferencia de un grupo metilo a partir del donador S-adenosil-L-metionina (SAM) a la quinta posición de los residuos de citosina (5mC). La metilación ocurre predominantemente en regiones enriquecidas con dinucleótidos de citosina y guanina (islas CpG), como los promotores génicos. La metilación del promotor de un gen causa su silenciamiento (Biswas y Rao 2018).



**Figura 5.** Niveles de compactación de la cromatina y elementos que establecen el código epigenético. Existen diferentes niveles de compactación de la cromatina, desde la formación de asas hasta la conformación de los nucleosomas. Las regiones amino terminales de las histonas y la secuencia desnuda de DNA son susceptibles a modificaciones químicas que son adicionadas, leídas y borradas por los elementos reguladores del código epigenético. Modificado de referencia (Keating, Plutzky, y El-Osta 2016).

Los writers de las marcas de histonas colocan modificaciones post-traduccionales (PTM) en las regiones amino terminal de las histonas. Algunas de las PTM conocidas son la metilación, la acetilación y la fosforilación en de residuos de lisina (K) o arginina (R). Existen enzimas que pueden agregar uno o hasta tres grupos metilo a un mismo de residuo de aminoácido. A las enzimas que agregan un solo grupo metilo se les conoce como distributivas, y a las que agregan de más de un grupo metilo como procesivas. La nomenclatura de las marcas de metilación de las histonas indica el tipo de histona que fue modificada, en cuál residuo y cuántos grupos fueron añadidos. Por ejemplo, la marca H3K4me3 indica que en el residuo de lisina 4 de la histona H3 fueron añadidos tres grupos metilo. Algunas marcas de histonas se han asociado con estados de activación o represión de la transcripción génica y/o con el nivel de compactación de la cromatina. Por ejemplo, la marca H3K4me3 se relaciona con regiones de eucromatina y activación de transcripcion génica, mientras que la marca H3K27me3 se relaciona con la formación de heterocromatina facultativa y la represión transcripcional. A esto se le denomina Código de Histonas (Hyun et al. 2017).

Los “readers” son un grupo de proteínas que reconocen e interpretan las modificaciones para permitir el ensamblaje de complejos proteicos. Estas enzimas cuentan con dominios especiales (dominios de unión a CpG metilado/MBD, bromodominios/BRD, cromodominios, etc) que se unen a diferentes modificaciones post-traduccionales en las histonas y covalentes en el DNA (Biswas y Rao 2018).

Los “erasers” son las proteínas que remueven las marcas y, por ende, cambian la señalización de la transcripción génica. Existe un eraser para cada tipo de modificación; es decir, desmetilasas, desacetilasas y fosforilasas, entre otras. Las proteínas que retiran la metilación del DNA son las metilcitosin dioxigenasas de la familia TET. Estas enzimas son dependientes de  $\alpha$ -cetoglutarato ( $\alpha$ -KG) y catalizan la oxidación sucesiva de la 5mC a sus derivados: la 5'-hidroxi-metilcitosina (5hmC), la 5'-formilcitosina (5fC) y la 5'-carboxilcitosina (5caC) (Breiling y Lyko 2015). Los nucleótidos oxidados son reconocidos como una señal de daño y restaurados por medio de las vías de reparación del DNA. Los derivados oxidados pueden ser escindidos por la timin DNA glicosilasa (TDG), que está acoplada a la vía de reparación de escisión de base (BER), desaminados por la citidina desaminasa (AID) o eliminados por dilución durante la replicación del DNA (X. Wu y Zhang 2017).

Los erasers de las marcas de histonas son desmetilasas (KTM) que pertenecen a dos familias enzimáticas: las Desmetilasas Específicas de Lisinas (LSD) y las desmetilasas Jumonji de dominio C (JmjC). Estas familias se distinguen entre otras cosas, por su mecanismo de acción y capacidad de desmetilación. Las enzimas de la familia LSD son amino-oxidasas dependientes de FAD que desmetilan residuos de lisina mono- y dimetilados mientras que las enzimas de la familia JmjC son oxigenasas dependientes de  $\alpha$ -KG capaces retirar la mono-, di- y trimetilación de los residuos de lisina modificados (Kampranis y Tschlis 2009).

### 3. Desmetilasa de histonas KDM4A

La desmetilasa de histonas KDM4A (también conocida como JMJD2A, JHDM3A y KIA0677) es un miembro de la familia de oxigenasas JmjC, que está evolutivamente conservada desde las levaduras hasta los mamíferos (Katoh y Katoh 2004). Esta desmetilasa se encuentra principalmente en el núcleo celular y tiene como blanco los residuos di- y trimetilados de las lisinas en la posición 9 y 36 de la histona H3 (H3K9me<sub>3/2</sub> y H3K36me<sub>3/2</sub>) (Guerra-Calderas et al. 2015). Estudios basados en la cinética y cristalografía de KDM4A sugieren que esta enzima tiene una mayor eficiencia desmetilando los residuos trimetilados y una preferencia por la marca de histonas H3K9 (Couture et al. 2007; Ng et al. 2007).

La metilación de la marca de histonas H3K9 es colocada por miembros de la familia de metiltransferasas SUV39. Las marcas H3K9me<sub>3/2</sub> se relacionan con el silenciamiento transcripcional y la formación de heterocromatina constitutiva (Shilatifard 2006; Hyun et al. 2017). En cambio, altos niveles de la marca H3K9me<sub>1</sub> se han detectado en el TSS de genes transcripcionalmente activos (Barski et al. 2007).

La metilación de H3K36 es colocada por miembros de la familia de metiltransferasas SET2 (Shilatifard 2006). La marca H3K36me<sub>3</sub> inhibe la transcripción de los genes en el sitio de inicio, pero facilita la elongación de la transcripción y previene su iniciación en el cuerpo de los genes. Además, se ha propuesto que esta marca tiene un papel importante en el splicing alternativo, la replicación, recombinación y reparación del DNA (Berry y Janknecht 2013).

Además de mediar la transcripción génica a través de la condensación de la cromatina, la desmetilasa KDM4A tiene otras funciones regulatorias en procesos como la replicación y senescencia celular, la respuesta al daño al DNA y la traducción (Guerra-Calderas et al. 2015; Black et al. 2010; Van Rechem et al. 2015). KDM4A regula también la progresión del ciclo celular incrementando la accesibilidad de la cromatina. Lo anterior debido a que antagoniza la unión de HP1, una proteína que media la formación de heterocromatina constitutiva (Black et al. 2010).

KDM4A interactúa físicamente con diferentes tipos de proteínas, como desacetilasas de histonas (HDACs), factores transcripcionales (ej. p53) y funge como cofactor del corepresor N-CoR durante la represión del gen *ASCL2* (Gray et al. 2005; Zhang, Yoon, y Wong 2005; Kim et al. 2012). Además, tiene sustratos no histónicos; por ejemplo, la proteína de dedos de zinc WIZ, la hormona somatomotropina coriónica humana (HCS), la proteína de cromodominio Y (CDYL) y la metiltransferasa G9a, entre otros (Ponnaluri et al. 2009; Johmura et al. 2016).

La abundancia de la proteína KDM4A es dependiente del ciclo celular. Su disponibilidad aumenta durante la fase G1, tiene un decremento durante la fase S y muestra una escasez

durante las fases G2/M; a pesar de que los niveles de mensajero se mantienen constantes durante el ciclo celular (Black et al. 2010). Estudios recientes han revelado que los niveles celulares de KDM4A se regulan post-transcripcionalmente mediante la ubiquitinación de sus residuos de aminoácidos. Los complejos de ubiquitin ligasa SCF, en particular SKIP1-CUI1-F-Box y FBXO22, regulan el recambio y la degradación de KDM4A durante el ciclo celular (Tan, Lim, y Harper 2011). Además, la disponibilidad de la proteína KDM4A es controlada por las ubiquitin ligasas RNF8 y RNF168 en respuesta al daño al DNA (Mallette et al. 2012)

La actividad de desmetilasa de KDM4A es regulada de manera endógena, a través de los niveles de oxígeno,  $\alpha$ -KG o de moléculas que compiten por su sitio catalítico como la 2-hidroxioglutarato (Chowdhury et al. 2011).

### 3.1 Dominios estructurales de KDM4A

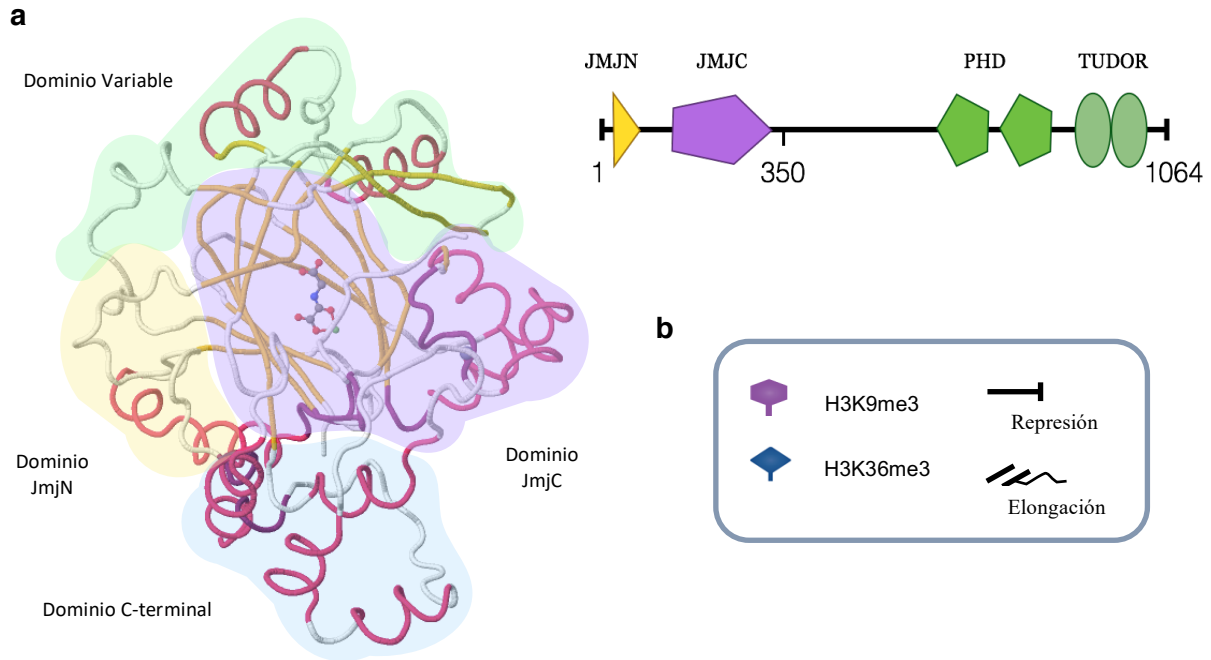
La desmetilasa KDM4A se compone en total de 1064 residuos de aminoácido y cuenta con cuatro dominios: un dominio JmjN, un dominio JmjC, un dominio carboxilo terminal (C-terminal) y un dominio estructural variable (Z. Chen et al. 2006) (Figura 6a). El centro catalítico de la enzima se forma a partir de los residuos de aminoácidos 1-350 y comprende a los dominios JmjN, JmjC y C-terminal (Hoffmann et al. 2012).

El dominio JmjN se forma a partir de los residuos de aminoácidos 14-56 y parece tener un papel importante en la conformación tridimensional del sitio catalítico y una función adaptadora para reclutar proteínas (Z. Chen et al. 2006). El dominio JmjC, que se forma con los residuos de aminoácidos 171-293, es responsable de la actividad enzimática de la proteína. El mecanismo de desmetilación es dependiente de  $\text{Fe}^{+2}$ , oxígeno y  $\alpha$ -KG y resulta en la liberación de formaldehído, succinato y  $\text{CO}_2$  (Hoffmann et al. 2012).

El dominio C-terminal, que se forma a partir de los residuos de aminoácidos 294-350, participa en la formación del sitio de unión al sustrato. Se ha visto que un motivo de dedos de zinc se forma entre un residuo de cisteína e histidina del dominio JmjC y dos residuos de cisteína del dominio C-terminal. La función de estos dedos de zinc es mantener una asociación estrecha entre ambos dominios (Z. Chen et al. 2006). Se cree que el dominio C-terminal tiene influencia en la localización subcelular de la desmetilasa KDM4A, aunque se desconoce cuál es el mecanismo involucrado y no se ha identificado hasta el momento una señal de localización nuclear (NLS) en su secuencia (Klose, Kallin, y Zhang 2006). A diferencia de KDM4A, en otras desmetilasas pertenecientes a la familia JmjC la NLS se encuentra en el dominio JmjN (Kamikawa y Donohoe 2014; Huang et al. 2013).

El dominio estructural variable comprende del residuo de aminoácidos 350-1064 cuenta con dos dominios Tudor y dos homeodominios de dedos de zinc (PHD) (Figura 6a). Los dominios PHD, que se forman con los residuos de aminoácidos 709-767 y 828-885 respectivamente, tienen una función incierta para la desmetilasa KDM4A, ya que reconocen a la marca

H3K4me2/3, un blanco de desmetilación distinto al de la enzima. La marca H3K4me3 se relaciona con regiones de eucromatina y activación de transcripción génica (Hyun et al. 2017); por lo que algunos investigadores creen que dicha marca podría reclutar a KDM4A, a través de sus dominios Tudor, a los promotores de genes activos para garantizar la desmetilación de la marca H3K9me3 (Berry y Janknecht 2013).



**Figura 6.** Dominios funcionales de la desmetilasa de histonas KDM4A. **a)** La familia a la que pertenece esta enzima comparte los dominios JmjN (amarillo), dominio JmjC (morado) y dominio C-terminal (azul). La desmetilasa KDM4A cuenta además con un dominio estructural variable con dominios dobles PHD y Tudor (verde). Dentro del área coloreada en morado se muestra una molécula de  $\alpha$ -KG. PDB 2PXJ. **b)** Blancos de desmetilación *in vivo* y su asociación biológica. Modificado de referencia (Z. Chen et al. 2006).

El doble dominio Tudor que se forma con los residuos de aminoácidos 897-954 y 955-1011 respectivamente, reconoce a las marcas de histonas H3K4me3 y H4K20me2/3 (Shiau et al. 2013). La marca H4K20me2 es un importante marcador de la replicación y reparación del DNA, mientras que la marca H4K20me3 parece estar involucrada con el silenciamiento de transposones y promotores génicos (Svobodová Kovaříková et al. 2018). Estudios recientes en proteínas similares sugieren que los dominios Tudor podrían actuar en conjunto para reconocer a la marca H3K9me3 (Xie, Jakoncic, y Qian 2012). Toda esta evidencia sugiere que los dominios PHD y Tudor podrían facilitar la unión de KDM4A a la cromatina; sin embargo, no hay evidencia de que estén mediando un mecanismo de reclutamiento (Labbé, Holowatyj, y Yang 2013). Algunos investigadores creen que el reconocimiento de los dominios PHD y Tudor influye, más bien, en la actividad de desmetilasa de KDM4A y la ayuda a mantener una asociación estrecha con otras proteínas de la cromatina (Guerra-Calderas et al. 2015).

## 3.2 La expresión de KDM4A en cáncer

El estudio de la desmetilasa KDM4A ha sido particularmente relevante en la investigación de más de 11 diferentes tipos de cáncer, entre los que destacan el cáncer de mama, de próstata, de pulmón y de colon (D. H. Lee et al. 2020). En todos los casos, el factor común es la sobreexpresión de KDM4A y la alteración de las marcas de histonas H3K9me2/3 y H3K36me3 (Guerra-Calderas et al. 2015).

La sobreexpresión de KDM4A se ha relacionado con la capacidad metastásica de varias líneas celulares (D. H. Lee et al. 2020). Ensayos de migración *in vitro* con células de cáncer de colon, han reportado que la inhibición de la actividad KDM4A disminuye la capacidad migratoria de las células tumorales (Kim et al. 2012). En cáncer de endometrio, se ha visto que el Knock Down (KD) de KDM4A reduce tanto la capacidad migrativa como invasiva de los tumores (H.-L. Wang et al. 2014). Resultados similares se han reportado en líneas celulares de cáncer de vejiga y endometrio (H.-L. Wang et al. 2014; F. Wang et al. 2019). Es por esto, que en los últimos años ha incrementado el interés por encontrar inhibidores específicos de la demetilasa KDM4A que puedan tener un potencial uso terapéutico en cáncer (H. J. Lee et al. 2017).

## 4. Antecedentes directos

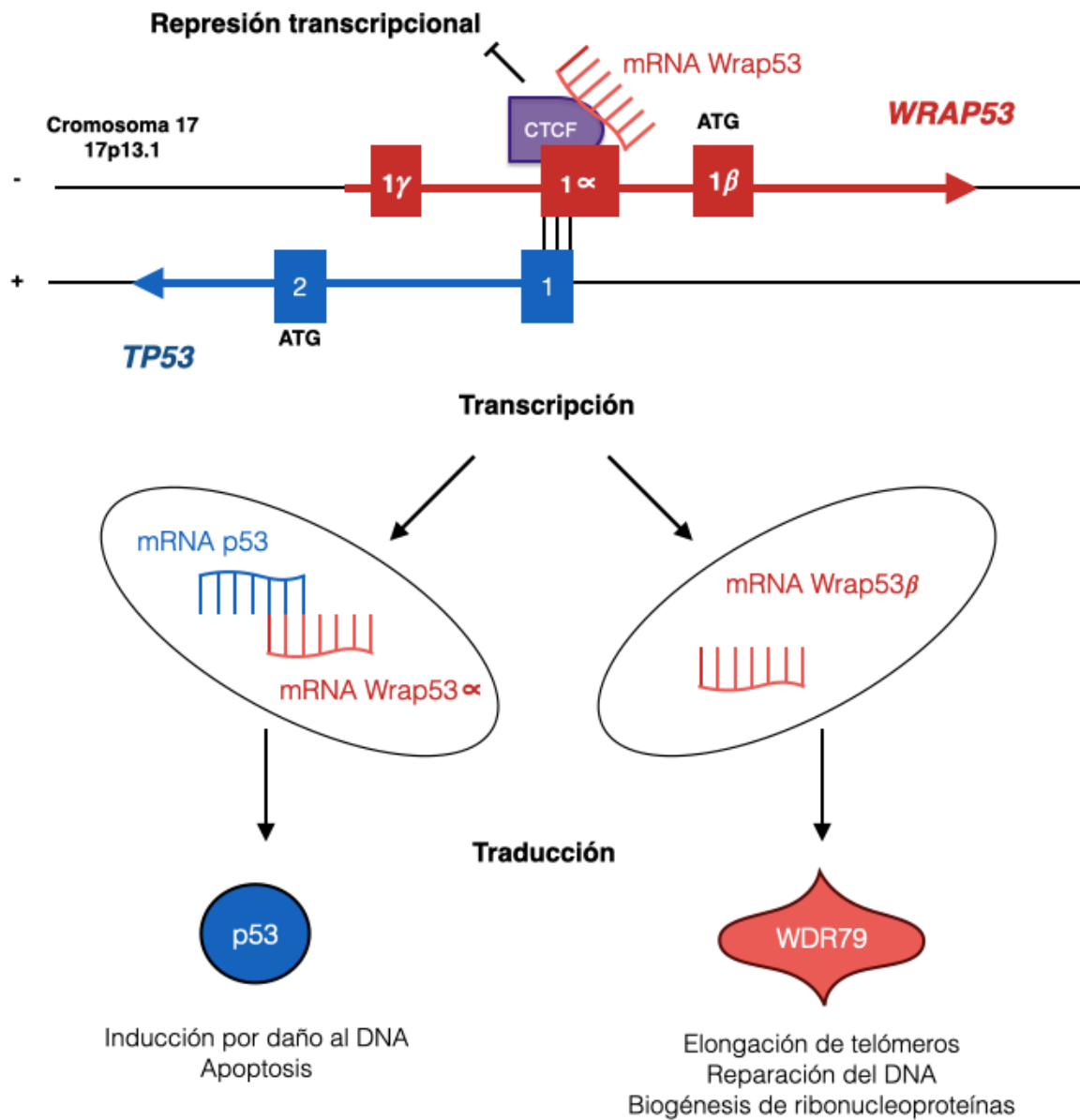
### 4.1. Genes involucrados en procesos tumorales

#### 4.1.1 El gen *WRAP53*

El gen *WRAP53* ha ganado relevancia en el estudio del cáncer a través de los años por su correlación con los altos niveles endógenos de mensajero y proteína del gen supresor de tumores *TP53*. La ausencia de la proteína p53 se relaciona con la proliferación descontrolada, el incremento en la sobrevivencia y la inestabilidad genómica de las células (Vousden y Lu 2002; Salah Mahmoudi et al. 2009).

En una primera caracterización del gen *WRAP53*, se encontró que su RNA mensajero es un transcrito antisentido de *TP53* (Figura 7). *WRAP53* cuenta con tres exones alternativos del inicio de la transcripción: 1 $\alpha$ , 1 $\beta$  y 1 $\gamma$ . El exón 1 $\gamma$  sobrelapa con el primer intrón de *TP53* y genera un transcrito de función desconocida hasta ahora. El exón 1 $\alpha$  sobrelapa en 227 pb con la secuencia del primer exón del gen *TP53*, siendo el único capaz de regular positivamente sus niveles de mensajero maduro. El mecanismo regulatorio es postranscripcional e involucra la formación de un dúplex de RNA a partir de los transcritos que incluyen el exón 1 $\alpha$  de *WRAP53* y primer intrón de *TP53*. La expresión de *WRAP53* $\alpha$  es necesaria para inducción de p53 después del daño al DNA y su posterior respuesta apoptótica (Salah Mahmoudi et al. 2009). Se piensa que la formación del dúplex protege al transcrito de *TP53* de la degradación. Este dúplex

se forma incluso con versiones mutantes de *TP53*, lo que destaca un potencial uso terapéutico del transcrito *WRAP53 $\alpha$*  (Farnebo, Bykov, y Wiman 2010).



**Figura 7.** Efecto regulatorio de los transcritos generados en el locus *TP53-WRAP53*. El promotor bidireccional de este locus regula la expresión de los transcritos antisentido de los genes *TP53* y *WRAP53*. El transcrito *Wrap53 $\alpha$*  regula positivamente la abundancia del transcrito de p53 y su inducción por daño al DNA. El transcrito *Wrap53 $\beta$*  codifica a la proteína WDR79 involucrada en diferentes procesos de replicación y reparación del DNA. El gen *TP53*, su transcrito y proteína se muestran en color azul. El gen *WRAP53*, sus isoformas y proteína se muestran en color rojo. Las líneas negras en locus *TP53-WRAP53* esquematizan la región de sobreposición de sus exones.

En estudios posteriores, se encontró que el factor de unión a CCCTC (CTCF) tiene un sitio de unión en el promotor de *TP53* que protege su expresión y mantiene la apertura de su cromatina al bloquear la incorporación de modificaciones represivas en las regiones amino terminal de las histonas (Soto-Reyes y Recillas-Targa 2010). De hecho, este sitio de unión de



CTCF tiene también un papel regulatorio en la expresión de *WRAP53*. La disminución en los niveles de CTCF reduce a su vez la expresión de todos los transcritos de *WRAP53*. Además, se demostró que CTCF contacta físicamente al mensajero de *WRAP53* $\alpha$  (naciente y/o maduro), a través de sus dedos de zinc (dedos 10-11) y su dominio carboxilo terminal. La unión entre CTCF y el mensajero de *WRAP53* es necesaria para el mantenimiento en la expresión del último (Saldaña-Meyer et al. 2014).

La proteína codificada a partir del exón 1 $\beta$  de *WRAP53*, conocida como WDR79 o TCAB1, es una subunidad de la enzima telomerasa, esencial para la elongación de los telómeros y el reclutamiento de la telomerasa a los Cuerpos de Cajal (Figura 7) (Venteicher et al. 2009). Los cuerpos de Cajal son estructuras nucleares que contienen factores involucrados en la maduración de ribonucleoproteínas, la formación del splicesoma, el procesamiento de los transcritos de histonas y el ensamblaje de la RNA polimerasa, entre otras funciones. Hoy en día, se sabe que la sobreexpresión de WDR79 promueve la transformación en una gran cantidad en líneas celulares, mientras que su depleción desencadena un proceso apoptótico en las células cancerosas que es independiente de p53 (S. Mahmoudi et al. 2011).

#### 4.1.2 El gen *JAG2*

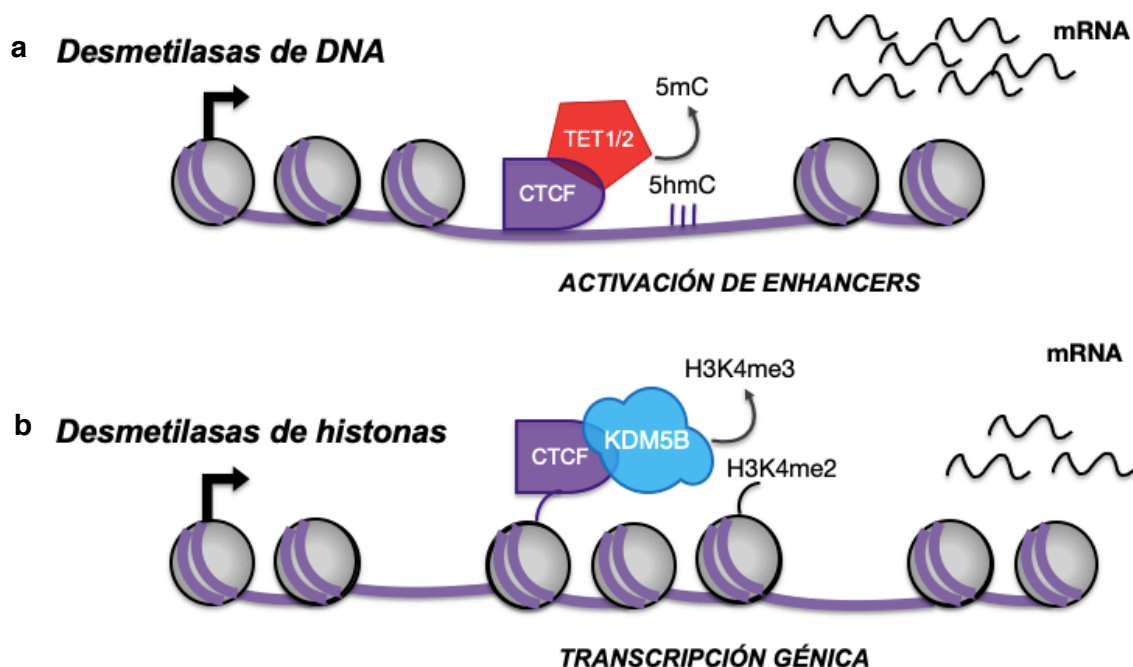
La proteína codificada por el gen *JAG2* es un ligando de la vía Notch, una vía evolutivamente conservada que desempeña un papel importante en la diferenciación de las células durante el desarrollo embrionario y la auto-renovación de órganos adultos (Lai 2004; Austin et al. 1995). La unión de *JAG2* con alguno de sus receptores (NOTCH1-NOTCH4), activa una vía de señalización que culmina con la escisión del propio receptor y la translocación de su dominio intracelular (NICD) al núcleo. Este dominio puede formar parte de complejos de activación de la transcripción de diferentes genes (Vaish, Kim, y Shim 2017; Cadigan y Waterman 2012).

La señalización de Notch es frecuentemente activada por hipoxia durante la progresión de los tumores, donde el aumento de *JAG2* en condiciones hipóxicas promueve la sobrevivencia de diferentes líneas celulares de cáncer de mama e induce un fenómeno conocido como la Transición Epitelio Mesénquima (EMT) (Xing et al. 2011). En tejidos de cáncer de mama, la sobreexpresión de *JAG2* se correlaciona con un incremento en el potencial metastásico de los tumores y menor sobrevivencia de los pacientes (Mittal et al. 2009; Xing et al. 2011). Los altos niveles de *JAG2* se relacionan también con el crecimiento y mantenimiento de las células madre cancerosas (Shan et al. 2017; Xing et al. 2011). Mientras que la subexpresión de *JAG2* disminuye la capacidad migratoria e invasiva de la variante celular Br4, conocida por hacer metástasis al cerebro (Nam et al. 2008). Aunque se desconoce el mecanismo exacto por el cual aumenta la expresión de *JAG2* en las células tumorales, algunos investigadores sugieren su potencial uso como marcador diagnóstico de cáncer de mama (Shaw et al. 2012).

## 4.2 Desmetilasas y su interacción con CTCF

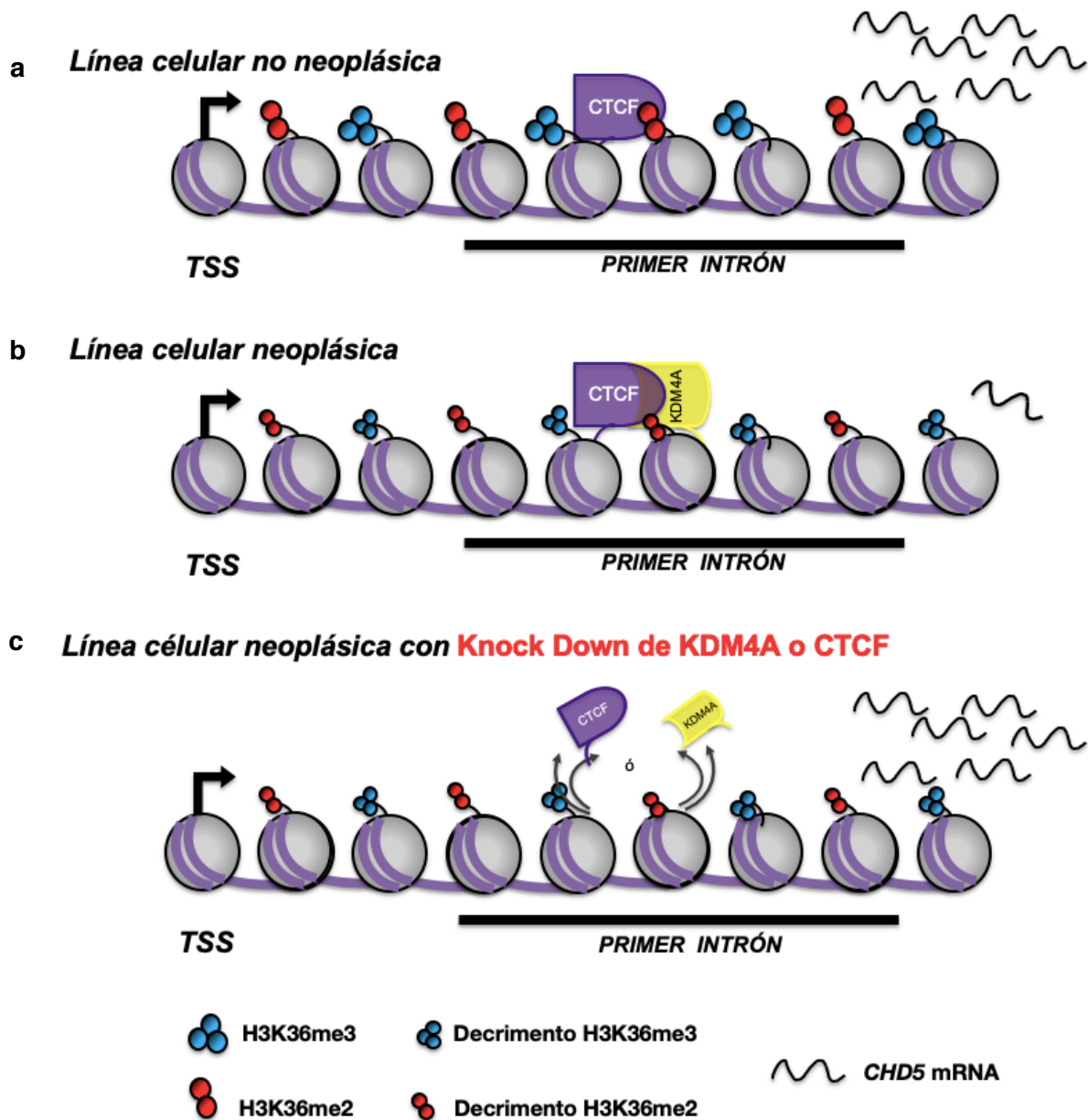
Los análisis de secuenciación masiva han mostrado que los patrones de distribución de las diferentes marcas de histonas son restrictivos a ciertas regiones del genoma y que no pueden ser explicados únicamente por la especificidad de los sitios activos de los writers y erasers del código epigenético (ENCODE Project Consortium 2012). Cada vez, se hace más evidente que los dominios no catalíticos y las interacciones con otras proteínas tienen un papel importante en el reclutamiento de las desmetilasas de histonas. Algunas hipótesis apuntan a que estas proteínas podrían ser reclutadas por su unión a moléculas con dominios de reconocimiento de secuencias específicas, como los RNAs no codificantes y los factores de transcripción (Dimitrova, Turberfield, y Klose 2015).

Hasta ahora, pocos estudios han demostrado la asociación entre CTCF y las desmetilasas de DNA e histonas. En algunos modelos murinos existe evidencia de que la proteína CTCF facilita el reclutamiento de las desmetilasas TET a sus sitios blanco, influenciando su actividad e intervención en procesos como la diferenciación celular. Se ha visto, por ejemplo, que la unión de CTCF a las regiones débilmente metiladas (LMR) favorece el reclutamiento de las TET y la conversión del DNA metilado (Feldmann et al. 2013; Z. Sun et al. 2013). De hecho, la co-localización de CTCF, TET1 y 5hmC ha sido corroborada a través de ensayos de co-inmunoprecipitación y se relaciona con la activación de enhancers y la diferenciación adipogénica a través de la hidroximetilación del DNA (Dubois-Chevalier et al. 2014) (Figura 8a).



**Figura 8.** La interacción entre CTCF y las desmetilasas. **a)** Mecanismo de activación de enhancers mediado por CTCF y las desmetilasas de DNA, TET 1 y TET2. **b)** Mecanismo de regulación de la transcripción génica mediada por CTCF y la desmetilasa de histonas KDM5B.

Solo se conocen dos casos en los que CTCF regula el reclutamiento de las desmetilasas de histonas. En células de cáncer de mama, se vio que la proteína CTCF es capaz de formar un complejo con la desmetilasa KDM5B, encargada de retirar las marcas H3K4me3/2. Los investigadores notaron que los sitios de union de KDM5B se sobrelapan con los de CTCF en la mayoría de las líneas de cáncer de mama y que este fenómeno se relaciona con una baja señal de la marca H3K4me3 y la expresión de genes asociados los fenotipos luminales (Yamamoto et al. 2014a) (Figura 8b).



**Figura 9.** Represión transcripcional del gen *CHD5* mediada por CTCF y KDM4A. El complejo protéico CTCF-KDM4A es reclutado al primer intrón del gen *CHD5* y promueve la desmetilación del residuo de histona H3K36me3. **a)** En células no neoplásicas, CTCF se localiza en el primer intrón de *CHD5* y las marcas H3K36me3/2 están enriquecidas. **b)** En células neoplásicas, KDM4A se encuentra sobrepresada y el complejo que forma con CTCF promueve la desmetilación de las marcas H3K36me3/2 y dirige la represión transcripcional. **c)** La subexpresión de cualquiera de las dos proteínas, CTCF o KDM4A, reactiva la expresión de *CHD5*. Modificado de referencia (Guerra-Calderas et al. 2018).

En nuestro grupo de trabajo se demostró la unión de CTCF con la desmetilasa KDM4A. Previamente existía evidencia de que la demetilasa KDM4A era reclutada al primer intrón del gen supresor de tumores *CHD5*; y de que su unión a dicha región debía ser el mecanismo principal por el cual se regulaba negativamente la transcripción de *CHD5*, ya que no existe metilación del DNA en su región promotora (Mulero-Navarro y Esteller 2008; Mallette y Richard 2012). Paralelamente, surgió evidencia de que el factor nuclear CTCF incrementa la frecuencia de desmetilación de KDM4A (Jeong et al. 2011); así que se realizaron estudios para determinar si el mecanismo por el cual KDM4A era reclutada al gen *CHD5* involucraba la participación de la proteína CTCF. Efectivamente, nuestro grupo demostró que KDM4A y CTCF se encuentran sobre-expresados y coexistiendo dentro del cuerpo del gen *CHD5* en líneas celulares derivadas de cáncer de mama y que su presencia favorece la subexpresión del gen supresor de tumores a través de la desmetilación de la marca de histonas H3K36me3/2 (Figura 9b). De hecho, el silenciamiento de KDM4A y/o CTCF guía a la reactivación del gen *CHD5* en líneas celulares neoplásicas (Figura 9c). En concordancia con el posible papel reclutador de CTCF, el Knock Out (KO) de KDM4A restaura la expresión de *CHD5* pero no afecta la localización de CTCF (Guerra-Calderas et al. 2018).

## **PLANTEAMIENTO DEL PROBLEMA**

Actualmente existe información detallada sobre los residuos metilados en las regiones amino terminal de las histonas y cuáles son las desmetilasas encargadas de remover estas modificaciones, pero poco se sabe sobre los mecanismos que regulan su reclutamiento e interacción con otras proteínas. Algunas hipótesis señalan la posibilidad de que estas enzimas sean reclutadas a través de su unión a proteínas con dominios de reconocimiento de secuencias consenso, como los factores de transcripción (Dimitrova, Turberfield, y Klose 2015). Tras la evidencia *in vitro* de que CTCF incrementa la frecuencia de desmetilación de KDM4A (Jeong et al. 2011) y de que esta desmetilasa es capaz de regular negativamente a genes supresores de tumores como *CHD5*, sin que exista metilación del DNA en su región promotora (Mulero-Navarro y Esteller 2008; Mallette y Richard 2012), nuestro grupo de trabajo realizó estudios para determinar el mecanismo por el cual KDM4A es reclutado y regula la expresión de *CHD5*. Se demostró que KDM4A y CTCF se encuentran sobre-expresados y coexistiendo dentro del cuerpo del gen *CHD5* en líneas celulares derivadas de cáncer de mama, y que su presencia está relacionada a la subexpresión del mismo (Guerra-Calderas et al. 2018). Este descubrimiento sugiere un nuevo mecanismo de represión transcripcional independiente del promotor del gen, en el que CTCF y KDM4A forman un complejo que promueve la desmetilación de H3K36me3/2. Dicho proceso podría estar regulando la expresión de otros genes supresores de tumores y oncogenes relacionados con la capacidad invasiva de la células neoplásicas, por lo que en el presente trabajo se desea conocer los sitios de unión e interacción de KDM4A y CTCF a nivel genómico, así como su participación en la regulación de la transcripción global en un modelo de cáncer de mama.

## HIPÓTESIS

El complejo formado por la desmetilasa de histonas KDM4A y el factor de transcripción CTCF regula la expresión de múltiples genes involucrados en procesos de migración e invasión celular en cáncer de mama.

## OBJETIVOS

### Objetivo general

Evaluar la participación genómica de la desmetilasa de histonas KDM4A y el factor de transcripción CTCF en la regulación de la transcripción de oncogenes y genes supresores involucrados en la capacidad migratoria e invasiva de la línea celular de cáncer de mama MCF7.

### Objetivos particulares

1. Caracterizar los sitios de unión de la desmetilasa de histonas KDM4A y el complejo proteico CTCF-KDM4A en el genoma de las células MCF7, mediante el análisis de datos previamente obtenidos de ensayos de inmunoprecipitación de la cromatina y secuenciación masiva (ChIP-seq) realizados por nuestro grupo de trabajo.
2. Estudiar el efecto del Knock Out de la desmetilasa de histonas KDM4A en la capacidad migratoria e invasiva de las células MCF7 mediante ensayos *in vitro*.
3. Evaluar la participación de la desmetilasa de histonas KDM4A en la regulación de la transcripción de dos de sus blancos génicos, a través de un modelo Knock Out para KDM4A en la línea celular MCF7 y un análisis de expresión con PCR cuantitativa (RT-qPCR).

## METODOLOGÍA

### I. Cultivo celular

Las líneas celulares provenientes de tejido mamario, MCF7 con fenotipo luminal y MDA-MB-231 con fenotipo basal, se emplearon como modelo epitelial neoplásico y metastásico, respectivamente. Todas las células fueron cultivadas como se plantea en la tabla A1 (véase apartado de metodología en extenso pág 63). Las células MCF7 Knock Out para KDM4A (MCF7-KDM4A<sup>KO</sup>) obtenidas previamente por el grupo de trabajo del Dr. Soto, se cultivaron en condiciones similares a las de la línea MCF7, salvo la administración de puromicina 3µg/mL

en el medio de cultivo como agente de selección positiva.

## II. Caracterización de las células MCF7 Knock Out de KDM4A

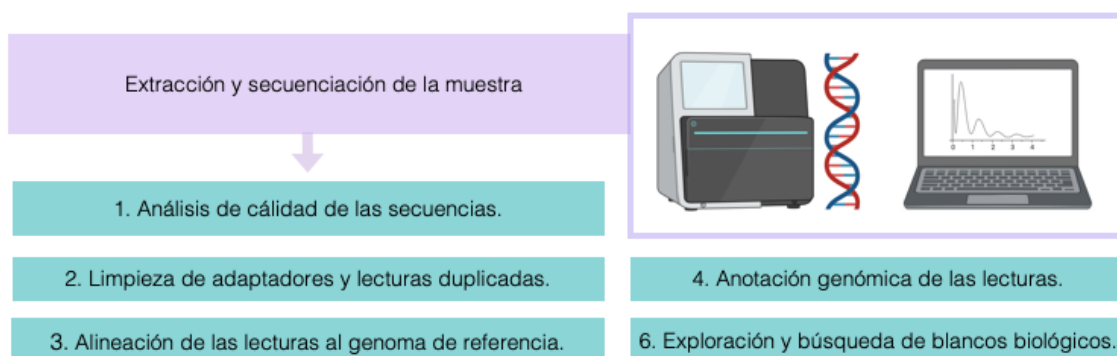
La caracterización de las células MCF7- KDM4A<sup>KO</sup> se realizó por medio de 1) análisis por citometría de flujo, que evaluó las subpoblaciones de células positivas a la señal de la proteína roja fluorescente (RFP), incorporada en los sitios de rompimiento de la doble cadena de DNA que genera el sistema CRISPR/Cas9 (por sus siglas en inglés, Clustered Regularly Interspaced Short Palindromic Repeats and associated protein Cas9); 2) un ensayo de fluorescencia que evidenció la expresión de RFP *in situ*; y 3) un Western Blot con tres réplicas biológicas, que constató la disminución en los niveles de expresión de la proteína KDM4A, tomando como control la expresión de actina.

## III. Ensayos de Inmunoprecipitación de la Cromatina (ChIP) y ChIP-ReChIP

Los ensayos de ChIP permiten estudiar el enriquecimiento selectivo de un antígeno/proteína en una región de cromatina particular. Los ensayos de ChIP y ChIP-ReChIP analizados por secuenciación masiva fueron previamente realizados por un miembro del laboratorio, M. en C. Lissania Guerra, mientras que los ensayos de ChIP puntuales se obtuvieron durante el desarrollo del presente trabajo. En ambos casos partimos de un cultivo celular de las líneas MCF7 o MCF7- KDM4A<sup>KO</sup> con confluencia del 80%. A estas células se les agregó formaldehído a una concentración final del 1%, para conseguir un entrecruzamiento entre el DNA y las proteínas. Pasado un tiempo, dicho entrecruzamiento se detuvo con una solución con glicina. Las células se lavaron con una solución amortiguadora y se lisaron con un buffer para poder obtener la cromatina. La cromatina se sonicó y se obtuvieron fragmentos de DNA de entre 200-500 pb. Seguido de esto, se inmunoprecipitó la modificación post-traduccional de las histonas o proteína de interés con ayuda de anticuerpos específicos (ej. KDM4A, CTCF, H3K9me3 y H3K36me3). El RNA y las proteínas se eliminaron con RNasa A y proteinasa K respectivamente. En el ensayo ChIP-ReChIP, se realizaron los mismos pasos que en el ensayo de ChIP convencional hasta el sonicado de las células. El siguiente paso fue la inmunoprecipitación de la primera proteína de interés (KDM4A o CTCF). Una vez obtenido este material se disoció el anticuerpo con DTT sin desacoplar las interacciones proteína-DNA. Después se realizó otra inmunoprecipitación para la segunda proteína de interés. Se realizaron dos ensayos de ChIP-ReChIP para alternar el orden de la primera proteína inmunoprecipitada (ChIP CTCF-KDM4A y ChIP KDM4A-CTCF). En cualquiera de los dos ensayos, se purificaron los fragmentos de DNA y se evaluaron por medio de una PCR cuantitativa para conocer la eficiencia de la amplificación y las veces de enriquecimiento de las proteínas o marcas de interés con respecto a la cantidad de DNA que había en el INPUT. Se consideró también el uso de oligos específicos para regiones de DNA que sirven como controles positivos y negativos de la inmunoprecipitación de las proteínas de interés. Finalmente, se realizó un análisis de varianza de una vía (ANOVA), seguido de un análisis de Dunnet para estimar diferencias significativa entre las condiciones experimentales y el control.

#### IV. Secuenciación de DNA de alto rendimiento de las muestras ChIP y ChIP-ReChIP (ChIP-Seq)

La técnica ChIP-seq permite identificar todos los sitios de unión a la cromatina de una proteína a nivel genómico. El material obtenido por triplicado de los ensayos ChIP o ChIP-ReChIP fue secuenciado por un método de síntesis a través de la tecnología de Illumina NextSeq 500 High Output. Con el apoyo del Dr. Nicolás Alcaraz, las lecturas fueron analizadas por diferentes criterios de calidad a través de la plataforma FastQC y los adaptadores y secuencias duplicadas son removidas. Para conocer las áreas del genoma donde las regiones inmunoprecipitadas mostraron un enriquecimiento, las secuencias obtenidas fueron mapeadas al genoma hg19 con el programa alineador BWA. Para crear tracks de señal y llamar los picos se utilizó la herramienta MACS2. Estas lecturas y picos de enriquecimiento se cargaron en el navegador UCSC donde fueron manualmente examinados. Posteriormente, se realizó un análisis de localización con el paquete ChIPSeekR, que provee datos estadísticos del enriquecimiento en regiones de importancia genómica como los promotores, intrones, exones y regiones intergénicas, etc. Esta información se correlacionó con los datos de expresión disponibles para dichos genes y se sometió a un análisis en el navegador STRING para indagar los procesos biológicos en los que están involucrados (Figura 10).



**Figura 10.** Representación esquemática del flujo de trabajo en el análisis bioinformático de los ensayos de secuenciación masiva. BioRender.

#### V. Reacción en cadena de la polimerasa (PCR)

La modalidad de PCR cuantitativa permite amplificar y estimar de manera simultánea la cantidad de una secuencia de DNA en una muestra. Las reacciones de PCR punto final se efectuaron a partir del DNA obtenido por ChIP o la lisis de las células de interés. El DNA se colocó en un “máster mix” que contiene los componentes necesarios para amplificar un templado particular de DNA con la ayuda de un termociclador. La amplificación de la región de interés se evaluó en un gel de agarosa al 2% (consultar la tabla de oligonucleótidos sintéticos utilizados en el apartado de metodología en extenso pág. 69). Las reacciones de PCR cuantitativa (qPCR) se evaluaron por medio de la señal del fluoróforo SYBR Green y la ayuda del sistema de Tiempo Real. Las reacciones se corrieron por triplicado y los valores promedio Ct se

utilizaron para la cuantificación con el método de doble delta CT ( $d\Delta CT$ ). Se consideró como parámetro normalizador la expresión del gen GAPDH.

## VI. Extracción y retrotranscripción del RNA

La retrotranscripción es la técnica con la que se obtienen moléculas de DNA codificantes (cDNA) para cuantificar de manera indirecta la expresión de un gen. El RNA se extrajo de las células MCF7 o MCF7-KDM4A<sup>KO</sup> con ayuda de una solución de fenol e isotiocianato de guanidina que mantiene la integridad del RNA y permite la separación del ADN y las proteínas. La integridad del RNA se evaluó con un gel de agarosa al 1%. El RNA obtenido se retrotranscribió por medio de una RT-PCR con oligonucleótidos oligo(dT) para obtener cDNA. Después, el nivel de expresión de los genes de interés (encontrados en los análisis de secuenciación masiva) se determinó por qPCR a través de oligonucleótidos sintéticos específicos.

## VII. Ensayo de invasión *in vitro*

El sistema Transwell permite evaluar la capacidad invasiva de las células *in vitro*. Consiste en una placa de 24 pozos y un inserto con la misma cantidad de pozos que poseen una membrana de PET con poros de 8 micrones por la que las células pueden migrar. En los ensayos de invasión este inserto está recubierto con una matriz o Matrigel, que simula la membrana basal de los tejidos *in vitro* y sirve como barrera para las células no invasivas. La cuantificación de la migración e invasión se obtuvo al estimar el número de células teñidas con calceína que fueron capaces de atravesar la matriz. El cálculo se obtuvo del conteo celular en varios campos del microscopio de fluorescencia con ayuda del programa Image J. Finalmente, se realizó un análisis ANOVA de una vía, seguido de un análisis de Dunnet para estimar diferencias significativas entre las condiciones experimentales y el control.

## VIII. Ensayo de migración *in vitro*

El ensayo de migración *in vitro* permite evaluar la capacidad migratoria de las células. El método consiste en hacer una herida con la ayuda de una punta de pipeta en una monocapa de células que se encuentra al 100% de confluencia. Estas células fueron incubadas 16 horas antes y durante todo el experimento con medio de cultivo a una concentración del 2% de suero fetal bovino, con la finalidad de sincronizarlas y evitar que proliferen. A partir del momento en el que se realizó la herida se empezó a contar el tiempo que ha transcurrido. Se realizaron varias mediciones de tiempo (ej. 0, 6, 24, 50 y 72 horas) en las que se tomaron fotografías para registrar el cambio en el área de la herida en una posición específica. El área de la herida se relacionó de manera inversa a la capacidad migratoria de las células. El cálculo del área de la herida se realizó a través de la herramienta TScratch para estandarizar las mediciones. Finalmente, se realizó un análisis ANOVA de dos vías para estimar diferencias significativas entre las condiciones experimentales y el control.



## RESULTADOS

### 1. Los sitios de unión de la desmetilasa KDM4A y el complejo CTCF-KDM4A coinciden con múltiples genes relacionados con la adhesión de las células MCF7

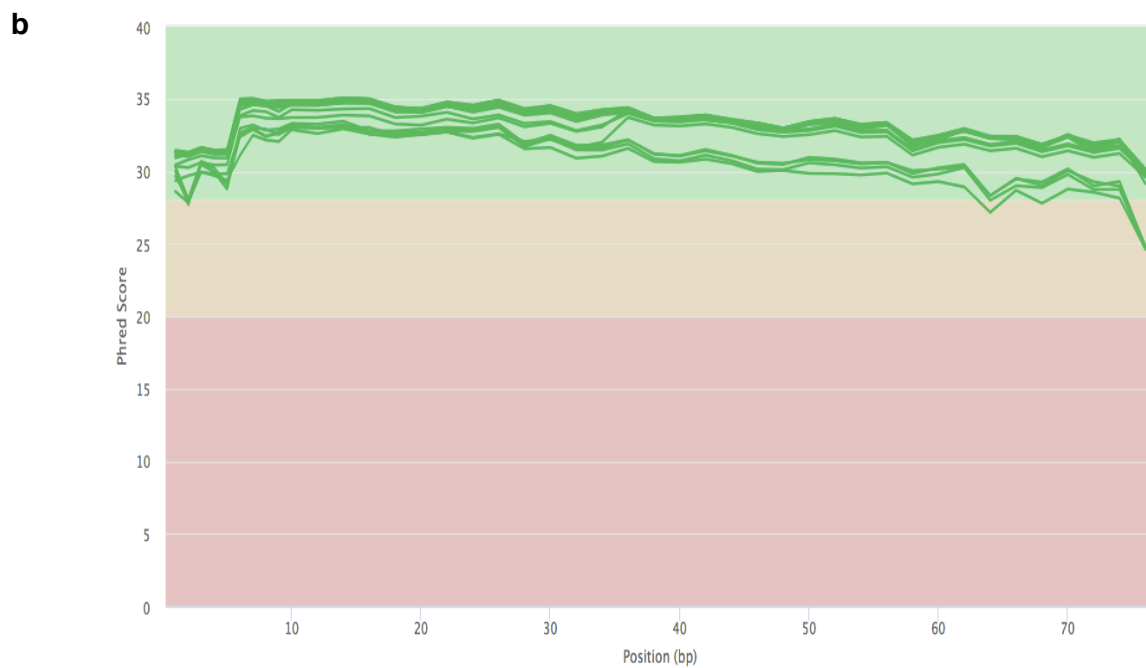
En años recientes, las tecnologías de secuenciación masiva de nueva generación (NGS) nos han permitido adaptar el fundamento de las técnicas de análisis puntual para estudiar perfiles globales de regulación, transcripción y expresión génica. La técnica ChIP-seq es un excelente ejemplo de esta premisa; ya que permite identificar todos los sitios de unión a la cromatina de una proteína y realizar inferencias acerca de su papel regulatorio en la transcripción génica y el recambio de distintas marcas epigenéticas (Weinmann y Farnham 2002). En el presente trabajo, se analizaron los datos de secuenciación masiva provenientes de los ensayos ChIP-seq y Re-ChIP-seq con el propósito de identificar los sitios de unión de la desmetilasa de histonas KDM4A y el complejo KDM4A-CTCF en el genoma de las células MCF7. La línea celular MCF7 es un modelo *in vitro* que se utiliza comunmente para el estudio del subtipo molecular luminal A de cáncer de mama (Holliday y Speirs 2011).

El análisis de los datos de secuenciación masiva involucra una serie de pasos bioinformáticos, que nos ayudan a filtrar e interpretar la información (descripción detallada en sección de Anexos pág. 76). El primer paso en el estudio de datos de secuenciación masiva es un análisis de calidad. Herramientas como FastQC nos permiten generar informes estadísticos del desempeño de los ensayos de secuenciación masiva a partir de diferentes parámetros bioinformáticos (Ewels et al. 2016). La tabla de estadísticas generales que ofrece FastQC muestra el número total de lecturas secuenciadas (M Reads), la cantidad de lecturas que se mapearon al genoma humano de referencia hg19 (M Reads Mapped) y el porcentaje del total que representa dicha cantidad de lecturas (%Mapped).

Dependiendo del comportamiento de la proteína bajo estudio, se recomienda cumplir con una cierta cantidad de lecturas. En el estudio de moléculas que se distribuyen ampliamente en el genoma, como las marcas de histonas, la Enciclopedia de los Elementos del DNA (ENCODE) recomienda tener más de 20 millones de lecturas (M Reads Mapped). En el caso de los factores de transcripción; es decir, proteínas que tienen sitios de unión discretos en el genoma, se recomienda contar con más de 10 millones de lecturas («Transcription Factor ChIP-seq Data Standards and Processing Pipeline – ENCODE» s. f.). El número de lecturas a través de las diferentes condiciones de nuestro estudio se encuentra en un rango de 4.1- 30 millones de lecturas, en el que la cantidad promedio de lecturas por cada condición supera los 5 millones de lecturas (Figura 11a). Un bajo número de lecturas puede disminuir la profundidad de lectura de un análisis; es decir, la relación que existe entre el número de bases secuenciadas y el tamaño del genoma, pero no descarta la ubicación de las lecturas con un mayor enriquecimiento. En este trabajo el porcentaje de lecturas mapeadas es alto. Esto quiere decir, que la mayoría de las lecturas fueron asociadas de manera exitosa con una región del genoma humano y es posible hacer un análisis con las lecturas más enriquecidas (Figura 11a).

**a**

	Sample Name	M Seqs	M Reads Mapped	% Mapped
Input ChIP	Input de Chip. R1	4.1	2.8	69.0%
	Input de Chip. R2	6.5	5.0	76.3%
CTCF	Chip de CTCF. R1	4.3	2.7	63.4%
	Chip de CTCF. R2	24.9	17.8	71.5%
KDM4A	Chip de KDM4A. R1	24.7	18.6	75.2%
	Chip de KDM4A. R2	18.6	15.1	81.2%
	Chip de KDM4A. R3	19.8	14.3	72.1%
Input reChIP	Input de Chip/RE Chip. R1	7.5	6.2	82.4%
	Input de Chip/RE Chip. R2	8.8	6.9	79.3%
ReChIP 1	Chip/RE Chip. CTCF - KDM4A. R1	20.0	12.0	60.1%
	Chip/RE Chip. CTCF - KDM4A. R2	19.7	14.7	74.6%
	Chip/RE Chip. CTCF - KDM4A. R3	23.3	16.9	72.6%
ReChIP 2	Chip/RE Chip. KDM4A - CTCF. R1	2.5	2.0	80.0%
	Chip/RE Chip. KDM4A - CTCF. R2	1.7	1.2	74.1%
	Chip/RE Chip. KDM4A - CTCF. R3	31.0	22.4	72.5%



**Figura 11.** Informe de FastQC sobre la calidad de las secuencias. **a)** Informe estadístico sobre la calidad de secuenciación de los ensayos de inmunoprecipitación de la cromatina (ChIP-seq). M seqs, Millones de lecturas de cada muestra; M reads Mapped, Millones de lecturas mapeadas; % Mapped, Porcentaje de lecturas mapeadas. **b)** Histograma de calidad de las secuencias. El eje “y” muestra los puntajes de calidad. Cuanto mayor sea el puntaje, mejor será llamada la base. El fondo del gráfico divide al eje Y en anotaciones de muy buena calidad (verde), anotaciones de calidad razonable (naranja) y anotaciones de mala calidad (rojo). FastQC.

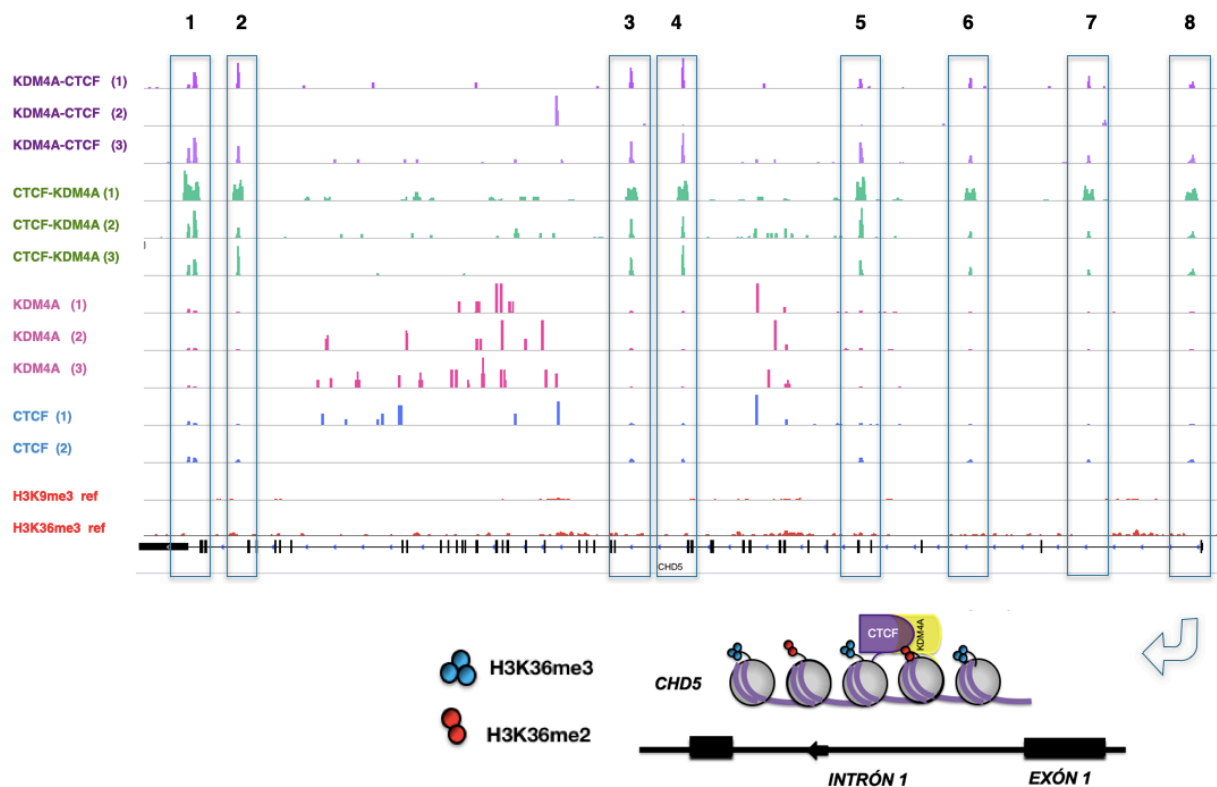
La herramienta FastQC provee también una visión del rango de valores de calidad de todas las bases para cada posición en una secuencia. El histograma de calidad que se muestra a continuación sugiere que todas las secuencias analizadas en este estudio se encuentran en un rango óptimo de calidad de acuerdo a las puntuaciones del formato Phred, un parámetro bioinformático que se asocia a la probabilidad de que una base en la secuencia haya sido leída de manera incorrecta durante la secuenciación (Figura 11b). Con base en lo anterior, podemos concluir que la calidad de las secuencias provenientes de los ensayos ChIP y ChIP-ReChIP en células MCF7 es adecuada para realizar un primer acercamiento al estudio genómico de la desmetilasa KDM4A y el complejo KDM4A-CTCF.

Una vez que se corrobora la calidad de las secuencias, se lleva a cabo una limpieza de los adaptadores que fueron agregados para crear las librerías y se eliminan todas las secuencias duplicadas. Posteriormente, las lecturas de cada condición son alineadas y anotadas a la región del genoma en la que se mapearon. Los programas de anotación genómica nos permiten conocer las características de los sitios de enriquecimiento de las lecturas, por ejemplo; en qué genes se encontraron alineadas dichas secuencias y a qué parte de las unidades transcripcionales se asocian (ej. Promotores, exones, etc.).

Con el objetivo de corroborar la evidencia conocida acerca de la unión de KDM4A y el complejo KDM4A-CTCF (Guerra-Calderas et al. 2018), se realizó una exploración manual del gen supresor de tumores *CHD5* con el visualizador IGV, donde los sitios de enriquecimiento de las lecturas asemejan montañas o picos. Observamos que efectivamente, los picos de unión de las proteínas KDM4A, CTCF y su complejo coinciden en el primer intrón del gen *CHD5* y más aún; que dentro del cuerpo del gen *CHD5* existen otros siete sitios intrónicos y exónicos de unión para estas proteínas que eran desconocidos (Figura 12).

Los programas de anotación genómica generan matrices de datos en las que la posición de cada lectura mapeada al genoma tiene un valor de certeza expresado en términos del q-value. El q-value es una clase de p-value ajustado a la Tasa de Descubrimiento Falso (FDR) que revela la probabilidad de obtener datos significativamente diferentes. De manera que, un valor pequeño de q-value indica que hay una alta probabilidad de que el enriquecimiento de una lectura es significativa. Los criterios de corte basados en el q value son “muy estrictos” para datos con q-value < 0.01, “estrictos” < 0.05 y “poco estrictos” < 0.1.

Para estudiar el enriquecimiento de CTCF, KDM4A y el complejo KDM4A-CTCF, en regiones de importancia genómica como los promotores, intrones, exones, regiones intergénicas y no traducidas (UTR), se realizó un análisis estadístico que consideró únicamente los datos que cumplieran con un criterio de corte estricto (q-value < 0.05). Los parámetros para discernir la ubicación de dichos elementos suelen ser señales consenso presentes en los promotores, marcos abiertos de lectura (ORF), etc. (Kaiser 2018).

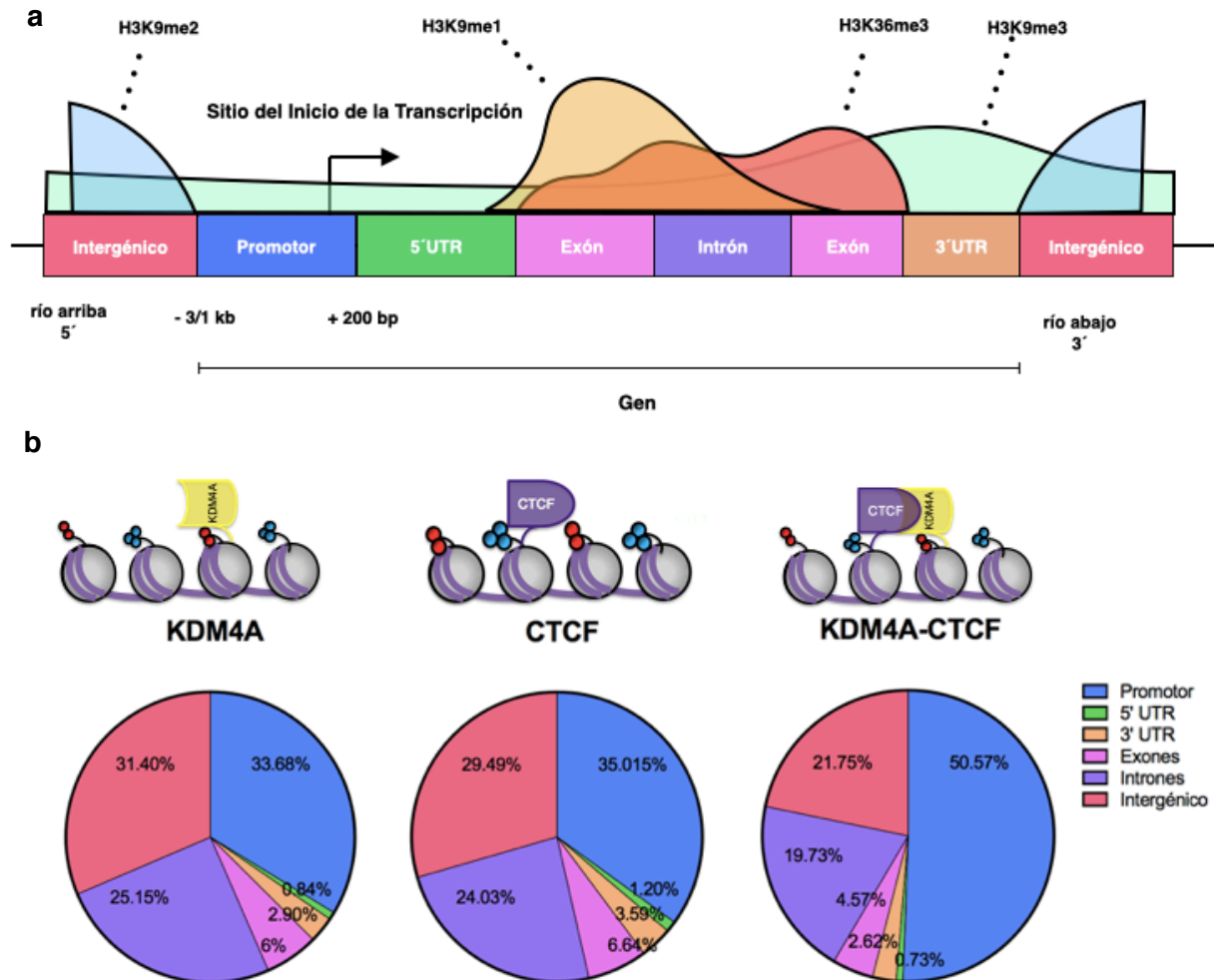


**Figura 12.** Distribución de KDM4A, CTCF y su complejo a lo largo del gen *CHD5*. Existen 8 sitios de unión del complejo KDM4A-CTCF en el cuerpo del gen *CHD5*. El octavo corresponde al sitio previamente caracterizado por Guerra-Caldera et al., 2018. Picos de enriquecimiento de KDM4A-CTCF (morado); CTCF-KDM4A (verde); KDM4A (rosa); CTCF (azul); H3K9me3 y H3K36me3 (rojo, secuencias de referencia obtenidas de ENCODE). IGV.

Los sitios de unión de KDM4A en la línea MCF7 eran desconocidos hasta ahora. Al empatar las posiciones de los picos con la estructura de los genes humanos, el análisis estadístico reveló que la proteína KDM4A se localiza primordialmente en las regiones promotoras (33.68%), intergénicas (31.40%), intrónicas (25.15%); y en menor frecuencia en regiones exónicas (6%) y regiones no codificantes 5' y 3' UTR (0.84% y 2.90%, respectivamente) (Figura 13b). De acuerdo con su papel de desmetilasa de las marcas de histonas H3K9me3 y H3K36me3, se esperaba un enriquecimiento de KDM4A en las regiones promotoras e intrónicas, como es el caso. Su alto enriquecimiento en las regiones intergénicas es un resultado interesante que podría estar asociado con el mecanismo de regulación que establece con la proteína HP1. En este sentido, la regulación de la marca de histonas H3K9 en las regiones intergénicas podría estar contribuyendo a la estructuración de la cromatina, más que a la regulación de la transcripción génica. En la figura 10a se esquematiza la distribución reportada de las marcas de histonas blanco de KDM4A (Figura 13a).

El análisis estadístico de los sitios de unión de CTCF se realizó únicamente a manera de control, ya que su distribución ha sido ampliamente reportada en diferentes líneas celulares incluyendo a MCF7. La proteína CTCF se localiza mayoritariamente en las regiones promotoras (35.01%), intergénicas (29.49%) e intrónicas (24.03%) y en menor frecuencia en

regiones exónicas (6.64%), 5' y 3' no codificantes (UTR; 1.20% y 3.59% respectivamente) (Figura 13b). El enriquecimiento de lecturas en estas regiones, refleja el papel arquitectónico que tiene la proteína CTCF delimitando compartimentos y regiones con señales regulatorias, como son los promotores génicos y las regiones intergénicas.



**Figura 13.** Distribución genómica de la desmetilasa de histonas KDM4A, CTCF y el complejo CTCF-KDM4A. **a)** Representación esquemática de la estructura de un gen y la distribución de las modificaciones de histonas a lo largo del mismo. Modificado de Barth et. al., 2010. Los genes humanos miden en promedio 10-15 kb, las regiones promotoras se encuentran entre 3-1 kb río arriba y 200 bp río abajo del del TSS y que todo lo que está después del sitio del final de la transcripción se consideran regiones intergénicas («Human Molecular Genetics» s. f.; Elkon y Agami 2017). **b)** Gráficos de pastel acerca de la distribución genómica de los sitios de unión de KDM4A, CTCF y el complejo CTCF-KDM4A.

Los sitios de unión del KDM4A-CTCF en la línea MCF7 también eran completamente desconocidos. El análisis estadístico mostró que el complejo KDM4A-CTCF, se localiza en las regiones promotoras (50.57%), intergénicas (21.75%) e intrónicas (19.73%); y con menor frecuencia en regiones exónicas (2.62%), 5' y 3' UTR (0.73% y 2.62% respectivamente) (Figura 13b) Aunque esta distribución muestra la misma tendencia que en los casos de KDM4A y CTCF, en la que estas proteínas se unen primordialmente a las regiones promotoras,

intergénicas e intrónicas; es claro que para el caso del complejo KDM4A-CTCF la unión a regiones promotoras es mucho mayor y representa prácticamente a la mitad de todos los sitios de unión. Esto podría sugerir, por un lado, que la interacción entre KDM4A y CTCF favorece el reclutamiento de la primera principalmente a las regiones promotoras; y por otro lado, que existe un mecanismo de reclutamiento alternativo de KDM4A a las regiones intrónicas e intergénicas donde disminuye su presencia en comparación con los sitios de unión estimados cuando se inmunoprecipita únicamente KDM4A.

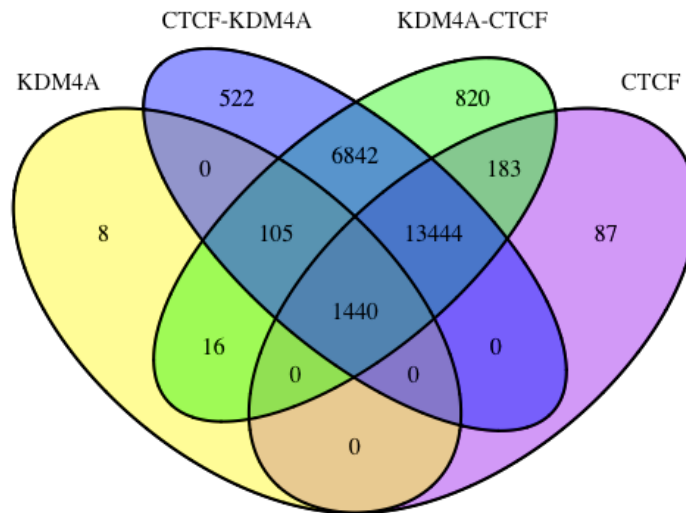
También, fue interesante descubrir que la unión al primer intrón representa alrededor del 30% del total de sitios de unión a regiones intrónicas, tanto para CTCF como KDM4A y su complejo. Esto es importante, ya que se ha visto que el primer intrón suele estar enriquecido de secuencias consenso y marcas de la cromatina que regulan la inclusión/exclusión de exones (Park, Hannenhalli, y Choi 2014).

Con el propósito de comparar la cantidad de genes a los que se unen KDM4A, CTCF y su complejo, se realizó un conteo génico que considero únicamente los datos que cumplieran con un criterio de corte muy estricto ( $q\text{-value} < 0.01$ ). Dado que el  $q\text{-value}$  se asigna a picos de enriquecimiento y puede existir más de un pico de enriquecimiento en un mismo gen, los genes duplicados de cada condición fueron eliminados.

El diagrama de Venn de la Figura 14a revela el número de genes que comparten el conjunto de KDM4A, CTCF, CTCF-KDM4A y KDM4A-CTCF. El universo del diagrama de Venn considera 23,467 genes en total, un número acorde con las estadísticas del Consorcio Internacional de Secuenciación del Genoma Humano, que reporta alrededor de 20,000-25,000 genes (Little 2005). El conjunto de KDM4A contiene en total 1,569 genes, de los cuáles 1,440 aparecen también en el resto de los conjuntos. Esto quiere decir que la desmetilasa KDM4A tiene múltiples sitios de unión a lo largo del genoma de las células MCF7 y coincide en la mayoría con la proteína CTCF. Lo que sugiere, una vez más, que CTCF podría influenciar de manera importante el reclutamiento de la desmetilasa KDM4A a sus sitios de unión. Como era previsto, el conteo génico de los ensayos de Re-ChIP es muy similar y coincide en más de 21,831 casos. Esta cantidad es muy alta dada la amplia distribución de la proteína CTCF en el genoma. El conjunto de CTCF tiene 15,154 y coincide en 14,884 casos con los genes encontrados en los ensayos de Re-ChIP. Esto indica que hay congruencia entre los sitios de unión encontrados en los ensayos de ChIP que involucran la inmunoprecipitación de CTCF.

El siguiente paso, fue la exploración a través del navegador IGV de los picos de enriquecimiento de KDM4A, CTCF y el complejo CTCF-KDM4A en un centenar de genes pertenecientes a la lista de intersección de nuestras cuatro condiciones (1,440 genes). En muchos casos, notamos que los picos de enriquecimiento de todos los ensayos ChIP-seq no coincidían de manera suficientemente exacta como para marcar una línea perpendicular en los tracks de señal. Esto quiere decir que, si bien, hay picos de enriquecimiento significativo en los mismos genes, hay ocasiones en las que la señal de los picos no proviene exactamente de la

misma región. A pesar de esto, se encontraron ejemplos muy claros de que el complejo KDM4A-CTCF se une una o más veces en la misma región, como fue el caso del gen supresor de tumores *CHD5* (Figura 12). En la tabla que se muestra a continuación, se presentan algunos ejemplos de oncogenes y genes supresores de tumores que han sido reportados en la literatura de cáncer de mama y que tienen una buena señal a lo largo de todas las condiciones (Tabla 1). Es importante destacar que para la mayoría de estos genes, no se ha descrito un mecanismo regulatorio que explique su sobreexpresión o subexpresión en cáncer de mama; y que este podría estar relacionado a la unión de KDM4A y el complejo CTCF-KDM4A.






**Figura 14.** Diagrama de Venn de los blancos génicos de KDM4A, CTCF, el complejo KDM4A-CTCF encontrados en los ensayos de secuenciación masiva. Conjunto de KDM4A (amarillo), CTCF-KDM4A (azul), KDM4A-CTCF (verde) y conjunto de CTCF (morado). RStudio.

Para indagar sobre los procesos biológicos en los que participan los genes a los que se une la desmetilasa KDM4A en la línea celular MCF7, se generó una matriz de datos que relaciona el nombre de cada gen con su valor q-value bajo un criterio de corte poco estricto (q-value < 0.1). Esta matriz fue sometida a un análisis de enriquecimiento funcional en el servidor de STRING (Szklarczyk et al. 2019). El servidor de STRING predice interacciones directas (físicas) e indirectas (funcionales) entre proteínas presentes en una matriz de acuerdo a su homología y lo reportado en la literatura (Szklarczyk et al. 2019). El resultado de este tipo de análisis es una red de interacción conectada por nodos de proteínas.

El análisis de enriquecimiento funcional predijo la existencia de una vía de 158 genes que podrían ser regulados por KDM4A y que están relacionados con la "adhesión de células homofílicas a través de moléculas de adhesión de membrana plasmática" (Término GO:0007156 de acuerdo a la clasificación funcional de Gene Ontology) (Figura 15). Es decir, una vía por la cual dos células son capaces de unirse a través de la misma molécula de adhesión presente en sus membranas plasmáticas. La red de interacción incluye proteínas clave en la

comunicación y migración celular, como moléculas de adhesión, receptores y antígenos de membrana (Figura 15). De acuerdo con Reactome, una de las bases de datos de las que se alimenta el servidor de STRING, los elementos de esta red pertenecen a 13 vías de señalización que están relacionadas con la adherencia y la comunicación celular («Home - Reactome Pathway Database» s. f.).

Blancos génicos	Nombre	Actividad biológica	Ref.
<b>Oncogenes</b> 	<i>GHR</i> <i>JAG2</i> <i>MAD1L1</i> <i>HDAC4</i>	Sobreexpresión asociada a formación de tumores. Promoción de la angiogénesis y el mantenimiento del cáncer. Mal pronóstico e insensibilidad a tratamientos con Taxol. Inhibición del gen resulta en la supresión de tumores.	(Gebre-Medhin et al. 2001) (Shan et al. 2017) (Q. Sun et al. 2013) (Hsieh et al. 2015)
<b>Genes supresores de tumores</b> 	<i>ASS1</i> <i>BIN1</i> <i>CSMD1</i> <i>DUOX1</i> <i>KCNQ1</i> <i>SKI</i>	Subexpresión en tumores sólidos Subexpresión asociada a metástasis y menor supervivencia Sobreexpresión inhibe la migración, invasión y adhesión celular. Subexpresión asociada a mayor proliferación y menor adhesión. Regulación del volumen celular y la secreción de potasio Inhibición de la progresión del cáncer a través de vía Hippo.	(Qiu et al. 2014) (Ghaneie et al. 2007) (Escudero-Esparza et al. 2016) (Fortunato et al. 2018) (vanTol et al. 2007) (Rashidian et al. 2015)
<b>Otros</b> 	<i>GADD45G</i> <i>RXRA</i> <i>SOX6</i>	Subexpresión asociada con la sensibilidad al daño del DNA Polimorfismos asociados con menor supervivencia. Expresado en fenotipo molecular luminal.	(Zhao et al. 2014) (Pande et al. 2013) (Kendrick et al. 2008)

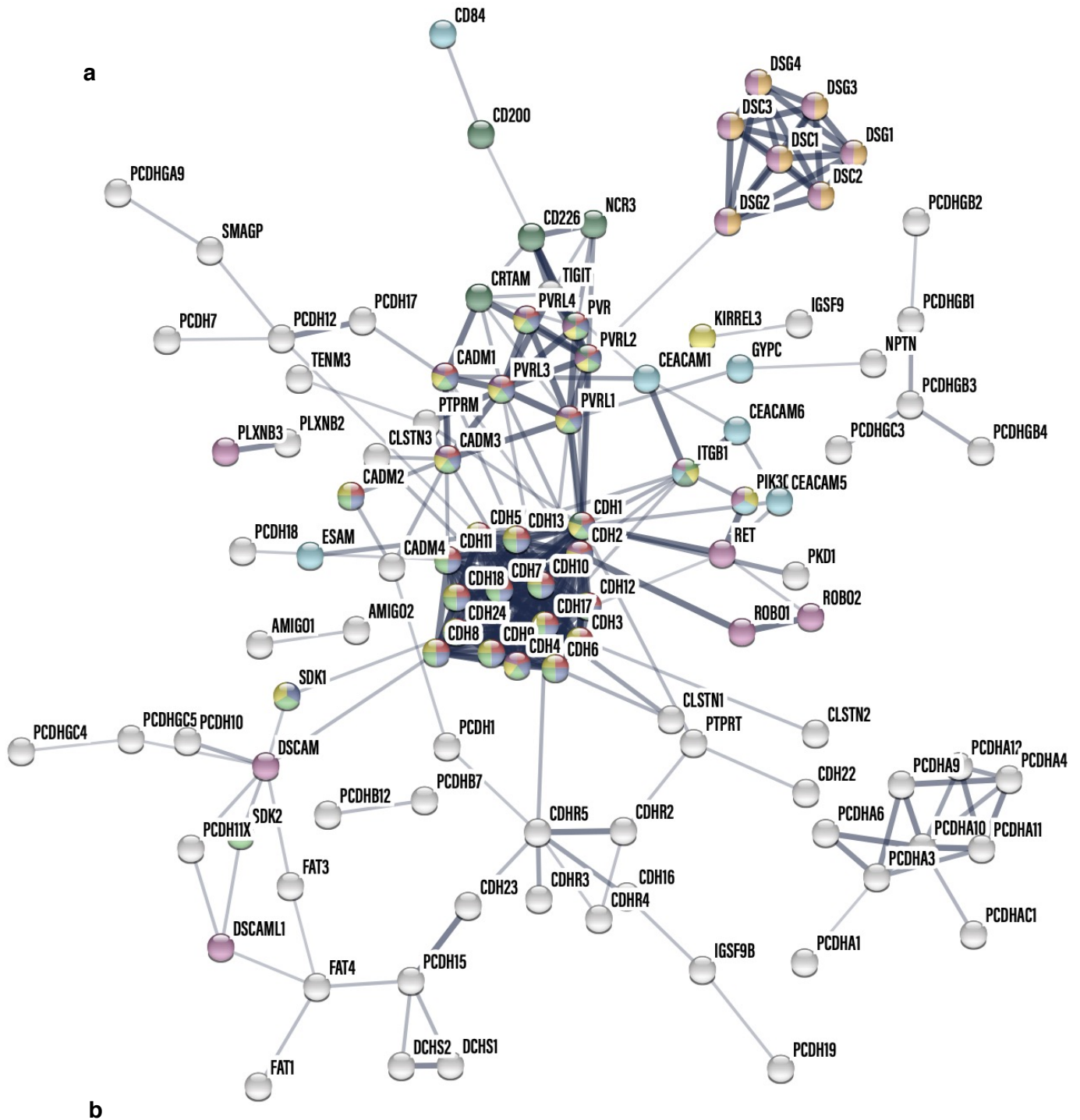
**Tabla 1.** Ejemplos de blancos génicos encontrados en los ensayos de secuenciación masiva CTCF-KDM4A.

En la Tabla 2, que se muestra a continuación se muestra el nombre de las proteínas que conforman cada vía. Algunas proteínas de la red pertenecen a más de una vía y otras no aparecen dentro de alguna vía conocida. La mayoría de las proteínas se comparten entre cuatro vías implicadas en la comunicación célula-célula: 1) Interacción de uniones adherentes, 2) Organización de unión célula-célula, 3) Organización de unión celular, 4) Comunicación célula-célula. Estas cuatro vías están implicadas en la comunicación celular mediante uniones adherentes, uniones estrechas, desmosomas y moléculas de reconocimiento de la matriz extracelular. El resto de las vías regulan procesos de comunicación más especializados, como la 5) Trans heterodimerización de la Nectina/Necl, 6) Esición apoptótica de proteínas de adhesión celular, 7) Interacciones inmunorreguladoras entre una célula linfocítica y no linfocítica, 8) Interacciones de la superficie celular en la pared vascular, 9) Formación de la matriz de fibronectina, 10) Formación de la envoltura cornificada, 11) Interacciones SDK, 12) Interacciones DSCAM y 13) Biología del desarrollo.










Estas 13 vías intervienen en procesos de comunicación celular que no son convencionalmente endoteliales y que en algunos casos son resultado de señales apoptóticas o de la respuesta inmune. Todo lo anterior sugiere que KDM4A regula la expresión de diversos genes relacionados con la comunicación y capacidad migratoria de las células MCF7.



## Análisis de Enriquecimiento Funcional en STRING



**Figura 15.** Vía de adhesión celular identificada por el análisis funcional de STRING. **a)** Se muestra el símbolo reportado para cada gen perteneciente a la vía. Los nodos representan proteínas y el grosor de las líneas indica el nivel de confianza de las asociaciones funcionales que predice el servidor. El color de los nodos se relaciona con las vías encontradas por Reactome (Tabla1). Los nodos desconectados fueron eliminados. **b)** Información sobre el Término GO asignado a la vía, su tamaño y su puntaje de enriquecimiento.

# ID del término	Description del término	Conteo de genes	Proteínas en la red
HSA-418990	 Interacción de uniones adherentes	24	CADM1,CADM2,CADM3,CDH1,CDH10,CDH11,CDH12,CDH13,CDH17,CDH18,CDH2,CDH24,CDH3,CDH4,CDH5,CDH6,CDH7,CDH8,CDH9,PVR,PVRL1,PVRL2,PVRL3,PVRL4
HSA-421270	 Organización de unión célula-célula	26	CADM1,CADM2,CADM3,CDH1,CDH10,CDH11,CDH12,CDH13,CDH17,CDH18,CDH2,CDH24,CDH3,CDH4,CDH5,CDH6,CDH7,CDH8,CDH9,PVR,PVRL1,PVRL2,PVRL3,PVRL4,SDK1,SDK2
HSA-446728	 Organización de unión celular	27	CADM1,CADM2,CADM3,CDH1,CDH10,CDH11,CDH12,CDH13,CDH17,CDH18,CDH2,CDH24,CDH3,CDH4,CDH5,CDH6,CDH7,CDH8,CDH9,ITGB1,PVR,PVRL1,PVRL2,PVRL3,PVRL4,SDK1,SDK2
HSA-1500931	 Comunicación célula-célula	29	CADM1,CADM2,CADM3,CDH1,CDH10,CDH11,CDH12,CDH13,CDH17,CDH18,CDH2,CDH24,CDH3,CDH4,CDH5,CDH6,CDH7,CDH8,CDH9,ITGB1, <b>KIRREL3</b> ,PIK3CB,PVR,PVRL1,PVRL2,PVRL3,PVRL4,SDK1,SDK2
HSA-420597	 Trans heterodimerización de la Nectina/Necl	7	CADM1,CADM3,PVR,PVRL1,PVRL2,PVRL3,PVRL4
HSA-351906	Escisión apoptótica de proteínas de adhesión celular	4	CDH1,DSG1,DSG2,DSG3
HSA-198933	 Interacciones inmunoregulatoras entre una célula linfoide y una célula no linfoide	8	<b>CD200,CD226</b> ,CDH1, <b>CRTAM</b> ,ITGB1, <b>NCR3</b> ,PVR,PVRL2
HSA-202733	 Interacciones de la superficie celular en la pared vascular	8	<b>CD84</b> ,CEACAM1, <b>CEACAM5</b> ,CEACAM6, <b>ESAM,GYPC</b> ,ITGB1,PIK3CB
HSA-1566977	Formación de la matriz de fibronectina	3	CEACAM1,CEACAM6,ITGB1
HSA-6809371	 Formación de la envoltura cornificada	7	<b>DSC1,DSC2,DSC3</b> ,DSG1,DSG2,DSG3,DSG4
HSA-373756	Interacciones SDK	2	SDK1,SDK2
HSA-376172	Interacciones DSCAM	2	DSCAM,DSCAML1
HSA-1266738	 Biología del desarrollo	17	CDH2,CDH4,DSC1,DSC2,DSC3,DSCAM,DSCAML1,DSG1,DSG2,DSG3,DSG4,ITGB1,PIK3CB,PLXNB3,RET,ROBO1,ROBO2

**Tabla 2.** Vías de comunicación y adhesión celular encontradas en la red de STRING. Los genes que son únicos para una vía aparecen en negritas. El color de los nodos en la figura 15 coincide con el color de la vía a la que pertenecen de acuerdo con Reactome.

## 2. Validación de un modelo Knock Out de KDM4A en células MCF7

Con el fin de estudiar el efecto del abatimiento en la expresión de la desmetilasa KDM4A en el comportamiento de las células MCF7, se implementó y validó en el laboratorio un sistema KO basado en la tecnología CRISPR/Cas9. El sistema consta de dos tipos de plásmidos. El primer plásmido contiene la secuencia codificante para la proteína CRISPR/Cas9 con una señal de señalización nuclear, la secuencia de la proteína verde fluorescente (GFP) para verificar la transfección del plásmido visualmente y una secuencia guía de RNA (sgRNA). La sgRNA dirige la especificidad y actividad de corte la enzima CRISPR/Cas9. El sistema KO para KDM4A está diseñado para causar rupturas de doble cadena (DSB) en tres regiones del gen: 1) entre el segundo intrón y el tercer exón, 2) en el tercer exón y 3) el octavo el exón (Figura 16a). Estos exones son de suma importancia para la expresión y formación del sitio catalítico de la enzima (Verrier et al. 2011; Sankar et al. 2017). El segundo plásmido contiene un molde

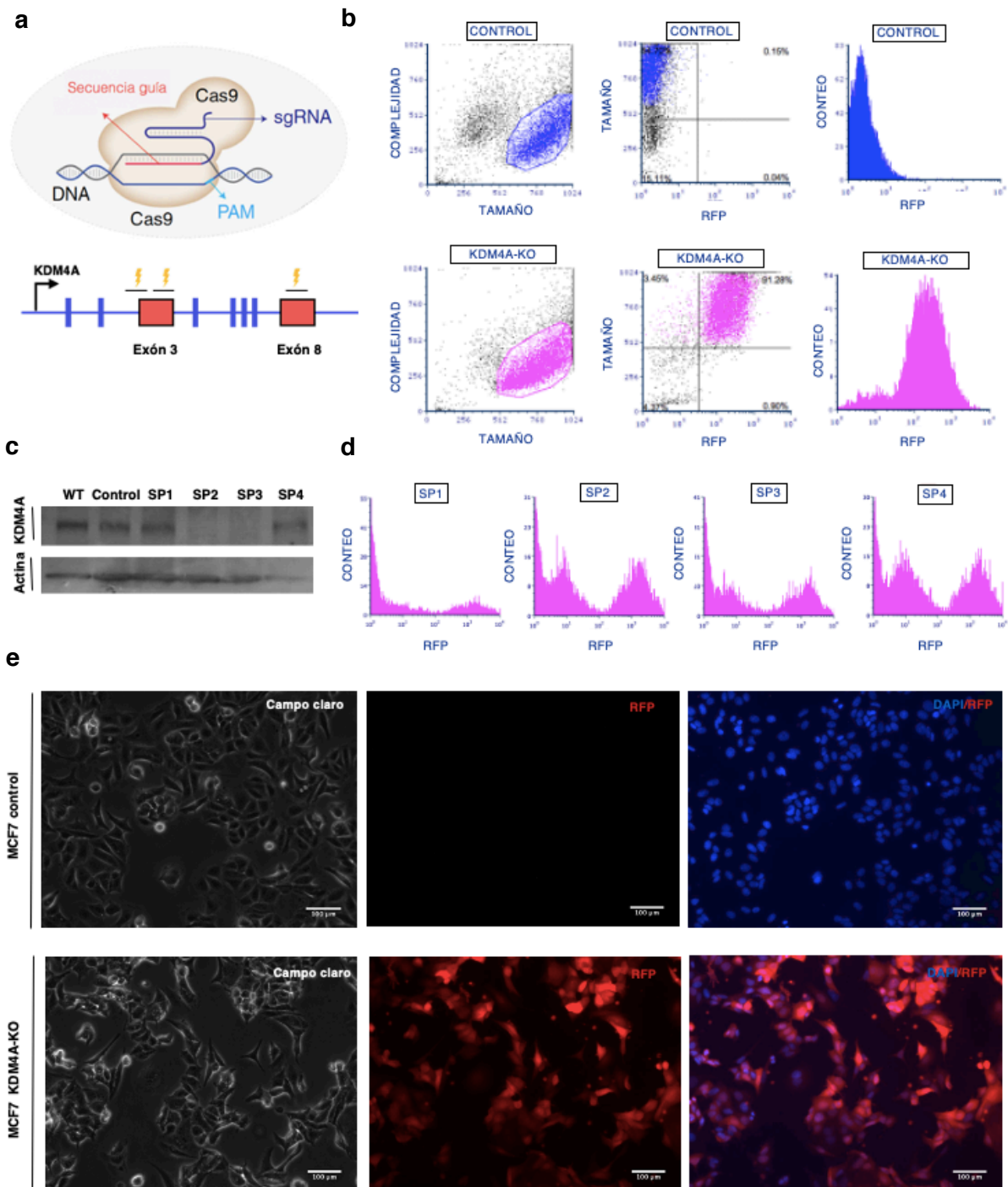
para la reparación directa por homología (HDR) a fin a los sitios de corte generados por la proteína CRISPR/Cas9 KO, un cassette de resistencia a puromicina para poder seleccionar las células y una secuencia codificante para RFP.

El éxito de la co-transfección de los plásmidos HDR y CRISPR/Cas9 KO se confirma con la detección RFP vía microscopía de fluorescencia. Dado que el proceso de transfección puede provocar cambios en la expresión génica y otros efectos adversos que pueden afectar la interpretación de resultados (Jacobsen, Calvin, y Lobenhofer 2009), la condición control también fue transfectada con un plásmido CRISPR/Cas9 KO, en el que no está presente el vector con RFP y el sgRNA es una secuencia aleatoria que no está dirigida a ningún gen en particular.

A través de un análisis de citometría con más de 10 mil eventos para cada condición, corroboramos que las características intrínsecas de la línea celular, como el tamaño y la complejidad no se ven afectados por la administración del sistema KO. Los eventos que se encuentran fuera de los compartimientos señalados con el polígono son partículas de menor tamaño que consideramos debris celular (Figura 16b). En un histograma que relaciona el número de células con la intensidad de fluorescencia de la proteína RFP (detectada a 588 nm), corroboramos que únicamente las células MCF7-KDM4A<sup>KO</sup> emiten una señal positiva. Este fenómeno hace que la curva se vea desplazada hacia la derecha. En un último gráfico se muestra que cerca del 91% de la población celular transfectada con el constructo CRISPR/Cas9 KO emite una señal positiva a RFP (Figura 16b).

Para evidenciar cualitativamente la diferencia en la expresión de RFP entre el control y las células KDM4A<sup>KO</sup>, se realizaron ensayos de fluorescencia. En estos ensayos se observa que ambas poblaciones celulares lucen morfológicamente iguales con una visión de campo claro; sin embargo, cuando se coloca el filtro para detectar la señal de RFP, únicamente las células KDM4A<sup>KO</sup> son positivas. La sobreposición de la señal nuclear del 4',6-diamidino-2-fenilindol (DAPI) y la señal de RFP, muestra que la señal de RFP se distribuye en el citoplasma de las KDM4A<sup>KO</sup>, lo cual es normal considerando que la proteína RFP no cuenta con una señal de localización nuclear (Figura 16e).

Con el propósito de obtener una sola población de células MCF7-KDM4A<sup>KO</sup>, se realizaron diluciones seriadas de la población parental y se seleccionaron los parches celulares que emitían una mayor señal de RFP. Cuatro subpoblaciones celulares fueron evaluadas por citometría y western blot: SP1, SP2, SP3 y SP4. El ensayo de Western Blot evidencia la presencia o ausencia de una proteína; y en este caso, reveló que dos subpoblaciones, SP2 y SP3, presentan una disminución en la expresión de KDM4A con respecto al control (Figura 16 c y d). El control de carga del ensayo western blot fue la proteína de expresión constitutiva actina (Kim et al., s. f.).



**Figura 16.** Sistema CRISPR/Cas9 y la comprobación del Knock Out de KDM4A en células MCF7. **a)** Representación esquemática del sistema CRISPR/Cas9. Secuencia guía de RNA (sgRNA), motivo adyacente de protoespaciador (PAM). Modificado de referencia (Adli 2018). **b)** Análisis citométrico de las poblaciones celulares KDM4A<sup>KO</sup>. Gráficas de puntos de dos parámetros lineales, complejidad vs tamaño; gráficas de puntos de dos parámetros, tamaño vs intensidad de RFP; histogramas logarítmicos de un parámetro, intensidad de RFP. Población control (azul); población MCF7 Knock Out (rosa). **c)** Detección por Western blot de la proteína KDM4A. Poblaciones Wild type (WT), control y subpoblaciones MCF7-KDM4A<sup>KO</sup> (SP1, SP2, SP3 y SP4). **d)** Análisis citométrico de las subpoblaciones celulares KDM4A<sup>KO</sup>. Histogramas logarítmicos de un parámetro, intensidad de RFP. **e)** Fluorescencia. Campos capturados en campo claro, filtro para RFP (rojo) y filtro para DAPI (azul), 20x.

### 3. La desmetilasa de histonas KDM4A regula la capacidad migratoria e invasiva de las células MCF7

Los análisis bioinformáticos mostraron que KDM4A se une a más de 150 genes presentes en el genoma de la línea celular MCF7 que están relacionados con la comunicación y migración celular; por lo que decidimos evaluar la capacidad migratoria e invasiva de las células MCF7-KDM4A<sup>KO</sup>.

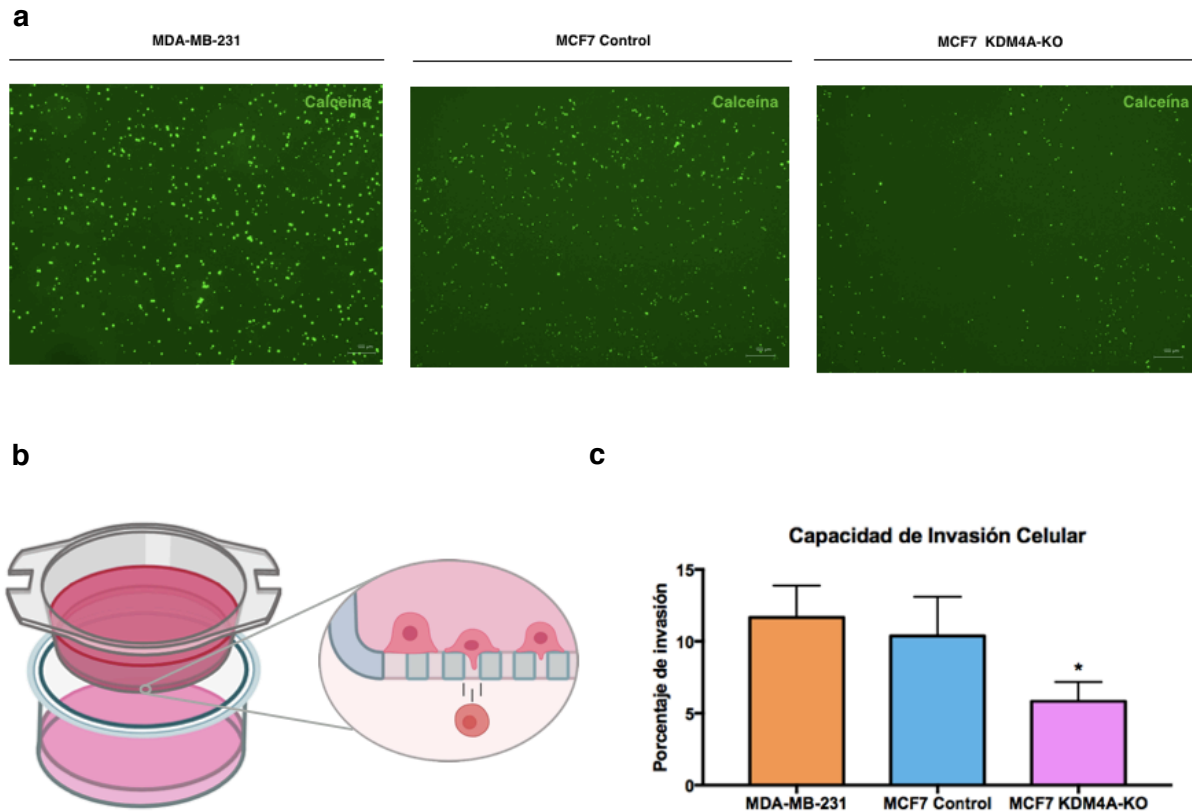
Evaluamos la capacidad invasiva de las células MCF7-KDM4A<sup>KO</sup> *in vitro* con ayuda del sistema Transwell (Figura 17b). Este sistema consiste en una placa de pozos, a la que se le pueden sobreponer insertos. Cada inserto posee una membrana con poros de 8 micrones por la que las células pueden migrar y un matrigel, que actúa de manera análoga a la membrana basal de los tejidos y representa una barrera para las células no invasivas. El método de valoración recaba los datos del conteo celular en diferentes campos del microscopio de fluorescencia y relaciona el número de células vivas se tiñeron con calceína y pudieron atravesar el matrigel, con el número de células que fueron sembradas en los insertos (Figura 17a). Esta relación proporciona un porcentaje de invasión. El ensayo reveló que la capacidad invasiva de las células MCF7 disminuye significativamente ( $p\text{-value} < 0.05$ ) con el KO de KDM4A (Figura 17c). En este modelo, las células de fenotipo basal MDA-MB-231 fueron empleadas como control positivo por su alta capacidad invasiva (Chavez, Garimella, y Lipkowitz 2010).

Con base en los resultados del ensayo anterior, en el que el KO de la desmetilasa KDM4A disminuye la capacidad invasiva de las células MCF7, decidimos evaluar *in vitro* si existe un cambio en la capacidad migratoria de las células MCF7-KDM4A<sup>KO</sup>. El primer paso en la realización de este ensayo es la generación de una herida en una monocapa de células 100% confluentes (Figura 18c). El área de la herida es monitoreada a través de distintos tiempos para relacionar el porcentaje del área desocupada con la capacidad migratoria de las células que se encuentran adyacentes a la lesión (Figura 18e).

En este ensayo, las células MDA-MB-231 fueron nuevamente empleadas como control positivo dada su alta capacidad migratoria; mientras que las dos subpoblaciones celulares, SP2 y SP3, representaron a la condición KO de KDM4A (Figura 18a). Tras calcular el área de libre a las 0, 6, 24, 50 y 72 horas después de haber hecho la herida, observamos que únicamente la línea celular MDA-MB-231 tiene un cambio significativo ( $p\text{-value} < 0.05$ ) en la capacidad migratoria con respecto al control de las células MCF7. La línea celular MDA-MB-231 cierra el área de la herida a las 24 horas, mientras que las células MCF7 control y MCF7-KDM4A<sup>KO</sup> mantienen la herida abierta aún después de 72 horas (Figura 18b).

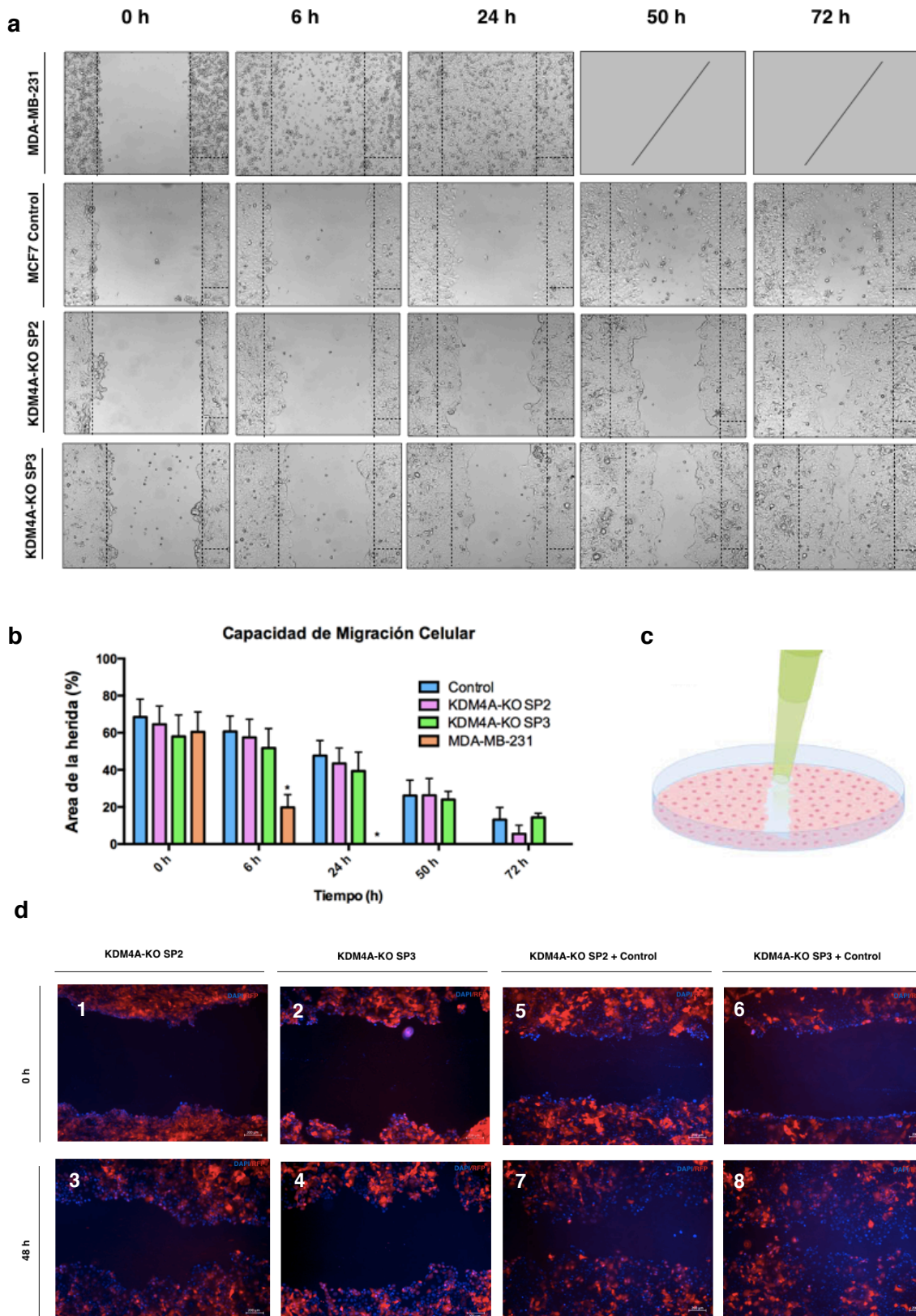
Para poder discernir si este fenómeno en el que las subpoblaciones KO de KDM4A, SP2 y SP3, cierran la herida de manera similar a las células control, se debía a una contaminación con células control o era una característica inherente de las células MCF7-KDM4A<sup>KO</sup>; se realizaron nuevos ensayos de la herida que fueron teñidos con DAPI y evaluados a través del filtro RFP

(Figura 18d). Los ensayos parecían indicar que las células más cercanas al cierre de la herida son aquellas que no incorporaron el sistema CRISPR/Cas9, pero la diferencia no era fácilmente perceptible (Figura 18d 1-4).



**Figura 17.** Ensayo de invasión celular. **a)** Fluorescencia. Campos capturados con el filtro de 488 nm, 5x. **b)** Representación esquemática de los insertos Transwell. **c)** Gráfica sobre la capacidad invasiva de las células MDA-MB-231 (naranja), MCF7 condición control (azul) y MCF7 KDM4A<sup>KO</sup> (rosa), estimada por el método de conteo celular. \*Análisis ANOVA de una vía seguido por una prueba Dunnet de comparación múltiple vs control. (p-value < 0.05). BioRender.

Para hacer más evidente la diferencia en la capacidad migratoria de las células, se mezclaron por partes iguales células provenientes de las poblaciones MCF7-KDM4A<sup>KO</sup> y de la población control. Los ensayos de fluorescencia enriquecidos con células de la población control mostraron que las células MCF7-KDM4A<sup>KO</sup> permanecen más alejadas al cierre de la herida, sugiriendo que existe una disminución muy sutil en la capacidad migratoria de las células MCF7-KDM4A<sup>KO</sup> con respecto al control (Figura 18d 5-8). Todas las evidencias anteriores sugieren que el abatimiento de la expresión de la desmetilasa KDM4A tiene un efecto en la capacidad migratoria e invasiva de las células MCF7.



**Figura 18.** Ensayo de migración *in vitro*. **a)** Visualización de las heridas en diferentes tiempos. Campos capturados en campo claro, 10x. **b)** Capacidad invasiva de las células MDA-MB-231 (naranja), MCF7 condición control (verde) y MCF7 KDM4A<sup>KO</sup> (azul), estimada con la medición del área de la herida.\* Prueba T de Student. p-value < 0.05. **c)** Representación esquemática del ensayo de migración *in vitro*. **d)** Fluorescencia. Campos de la herida capturados con sobreposición del filtro para RFP y DAPI, 10x. 1-4 Subpoblaciones SP2 y SP3. 5-8 Subpoblaciones KO mezcladas con células control. **e)** Representación esquemática de la migración celular. BioRender.

#### 4. La desmetilasa de histonas KDM4A regula la transcripción de los genes *JAG2* y *WRAP53*

Después de evaluar el efecto del KO de la desmetilasa KDM4A en la capacidad migratoria e invasiva de las células de cáncer mama, decidimos estudiar de manera puntual si la expresión de algunos de los genes blanco encontrados en los análisis masivos cambia cuando la desmetilasa KDM4A está ausente.

Se propuso el estudio de los oncogenes *JAG2* y *WRAP53* con base en la distribución de los picos de enriquecimiento del complejo KDM4A-CTCF en el análisis ChIP-seq. En dicho análisis, ambos genes fueron identificados bajo un criterio de corte estricto ( $q\text{-value} < 0.05$ ). En el visualizador IGV se observa que los picos de enriquecimiento de KDM4A, CTCF y su complejo coinciden en una o varias posiciones a lo largo de los genes *JAG2* y *WRAP53*. A manera de referencia, se agregaron al visualizador tres señales más obtenidas de las bases de datos de ENCODE: una señal de ChIP-seq para CTCF, y la señal de las marcas de histonas H3K9me3 y H3K36me3 (Figura 19).

En el gen *WRAP53* se encontró sólo un pico de enriquecimiento entre la región promotora y el primer exón (Figura 19a). En el caso del gen *JAG2* se observaron dos picos de enriquecimiento. El primer pico se encuentra entre el exón 20 y el intrón 19. El segundo pico se ubica entre el exón 19 y el intrón 19 (Figura 19b). Con base en la secuencia de este último se diseñaron los oligos de amplificación para qPCR. El enriquecimiento de la marca H3K9me3 en el pico analizado de ambos genes es casi nulo, mientras que solo en el gen *JAG2* la marca de histonas H3K36me3 parece estar enriquecida (Figura 19 a y b).

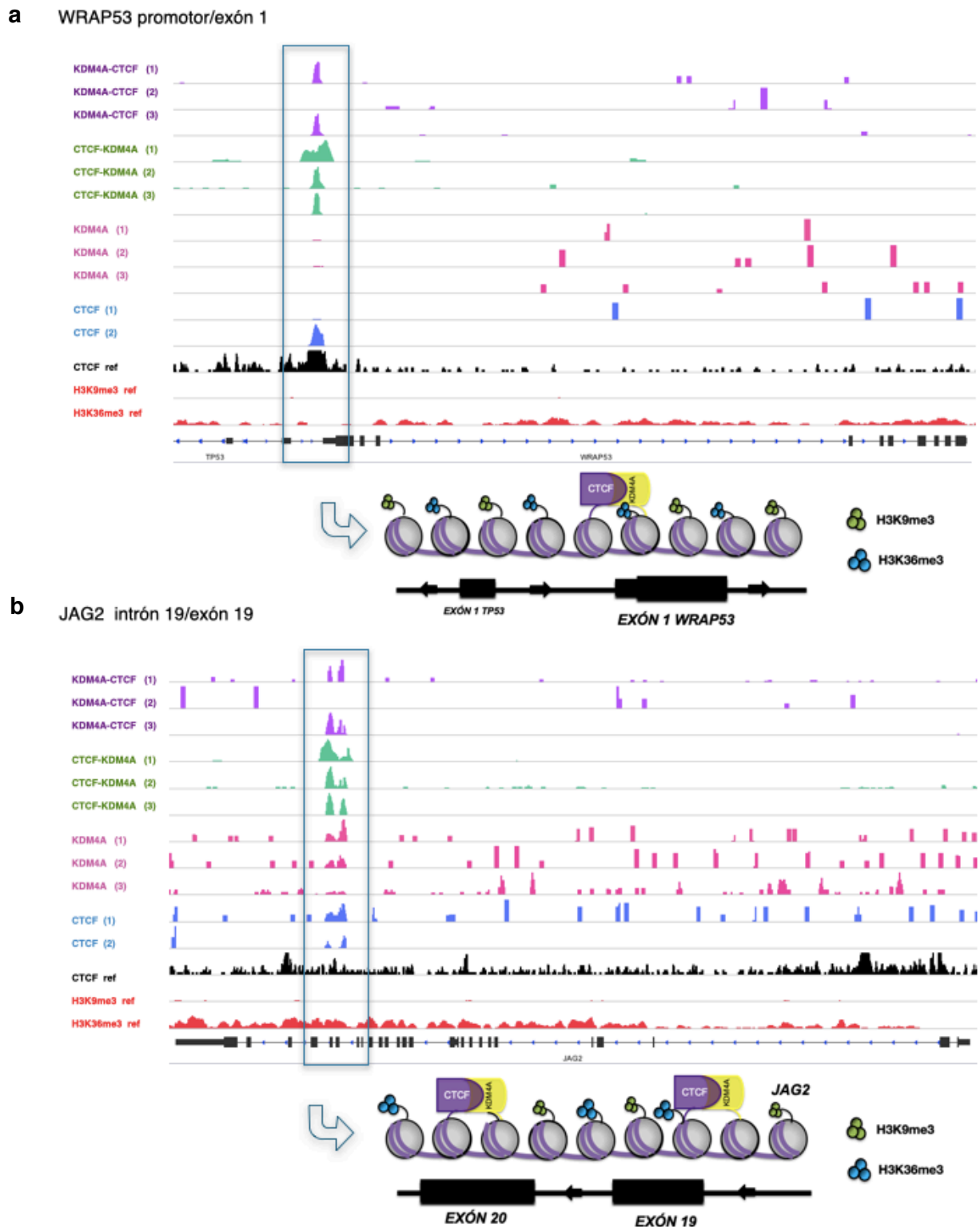
En los ensayos de ChIP se inmunoprecipitaron las proteínas KDM4A y las marcas de histonas H3K9me3 y H3K36me3. Observamos la presencia de KDM4A disminuye en las células KO, mientras que la presencia de las marcas H3K9me3 y H3K36me3 aumenta en ambos genes. De esta manera corroboramos que la ausencia de la desmetilasa de histonas KDM4A favorece el aumento en el enriquecimiento de sus marcas blanco (Figura 20 b y f).

Para saber si el aumento de las marcas H3K9me3 y H3K36me3 a causa del abatimiento de KDM4A tiene un efecto sobre la expresión de los genes *JAG2* y *WRAP53*, se realizaron análisis de expresión basados en la retrotranscripción de sus mensajeros maduros y su cuantificación por medio de qPCR. Los análisis de expresión mostraron que la expresión de ambos genes disminuye en células MCF7-KDM4A<sup>KO</sup>, siendo solo una tendencia para el gen *WRAP53* y un cambio significativo ( $p\text{-value} < 0.05$ ) en la expresión de *JAG2* con respecto al control (Figura 20 a y d). Para el cálculo de estos análisis se consideró como control la expresión relativa del gen constitutivo de la Gliceraldehído-3-fosfato deshidrogenasa (*GAPDH*).

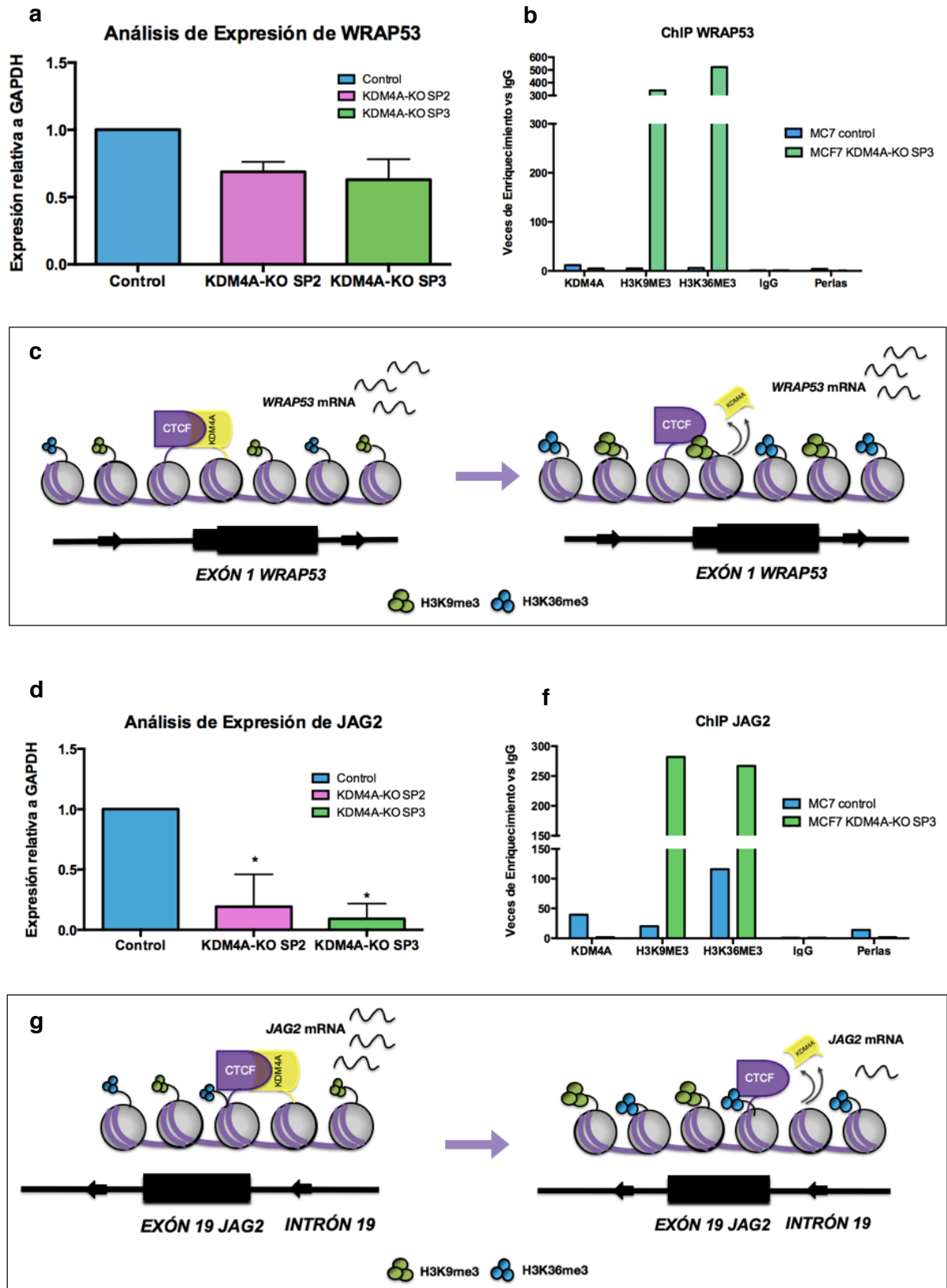
De acuerdo con los resultados presentados, podemos concluir que el abatimiento en la expresión de la desmetilasa KDM4A se relaciona con el aumento de las marcas de histonas



H3K9me3 y H3K36me3 en el promotor de *WRAP53* y el cuerpo del gen *JAG2*; así como con la disminución en la expresión de *JAG2* (Figura 20 c y g).



**Figura 19.** Distribución de KDM4A, CTCF y su complejo a lo largo de los genes *JAG2* y *WRAP53*. **a)** Gen *WRAP53*. **b)** Gen *JAG2*. Picos de enriquecimiento de KDM4A-CTCF (morado); CTCF-KDM4A (verde); KDM4A (rosa); CTCF (azul); H3K9me3 y H3K36me3 (rojo, secuencias de referencia obtenidas de ENCODE). El diámetro de las circunferencias se relaciona con la abundancia de la marca de histonas. IGV



**Figura 20.** Análisis sobre el enriquecimiento de las marcas de histonas H3K9me3 y H3K36me3 y la expresión de los genes *WRAP53* y *JAG2*. **a** y **d**) Análisis de expresión relativa a GAPDH. \*Análisis ANOVA de una vía seguido por una prueba Dunnet de comparación múltiple vs control.  $p$ -value < 0.05. **b** y **f**) Enriquecimiento de la proteína KDM4A y las marcas de histonas H3K9me3 y H3K36me3. **c** y **g**) Modelo sobre el efecto del Knock Out de la desmetilasa KDM4A en la expresión de los genes *WRAP53* y *JAG2*. El diámetro de las circunferencias se relaciona con la abundancia de la marca de histonas.

## DISCUSIÓN

La organización de la cromatina y la expresión génica dentro del núcleo son coordinados por diferentes mecanismos epigenéticos que actúan en respuesta a un código dinámico. El código que se establece a través de las modificaciones postraduccionales de las histonas depende de la actividad de las enzimas que se encargan de colocar y retirar dichas modificaciones. Las enzimas encargadas de remover uno o hasta tres grupos metilo de las regiones amino terminal de las histonas son las desmetilasas de la familia de oxigenasas Jumonji. En los últimos años, se ha visto que la expresión de estas enzimas juega un papel importante en el desarrollo de distintos tipos de cáncer. En particular, la sobreexpresión de la desmetilasa KDM4A ha sido reportada en tumores de cáncer colorrectal, de pulmón y de mama; donde se asocia con un mal pronóstico. Existe poca información acerca de cuáles son los blancos génicos de KDM4A y cómo es reclutada a ellos, ya que esta proteína carece de un dominio de reconocimiento de DNA. En nuestro grupo de trabajo, se descubrió que KDM4A y CTCF forman un complejo en el cuerpo del gen supresor de tumores *CHD5* que resulta en su subexpresión en líneas celulares de cáncer mama. Este hallazgo guió nuestros esfuerzos a evaluar la participación genómica de la desmetilasa de histonas KDM4A y el factor nuclear CTCF en la regulación de la transcripción de múltiples genes involucrados en procesos tumorales.

En el presente trabajo, se identificaron los sitios de unión de la desmetilasa de histonas KDM4A, el factor nuclear CTCF y el complejo proteico CTCF-KDM4A en el genoma de las células MCF7. La evaluación de los sitios de unión de CTCF se realizó únicamente a manera de control, ya que han sido ampliamente reportados en diferentes líneas celulares, incluyendo a MCF7. De acuerdo con un estudio realizado en 19 líneas celulares, la proteína nuclear CTCF se distribuye principalmente en regiones intrónicas (38%), intergénicas (45%) y promotoras (12%) (H. Wang et al. 2012). En otro estudio que recabó la información de 38 líneas celulares, se reportó que la proteína CTCF se distribuye principalmente en regiones intergénicas (53%), intrónicas (30%) y promotoras (12%) (H. Chen et al. 2012). Este último dato es más a fin a las proporciones encontradas en nuestro estudio, donde CTCF se distribuye más ampliamente en regiones intergénicas que en regiones intrónicas.

Los sitios de unión de KDM4A en la línea MCF7 eran desconocidos hasta ahora. Sólo dos investigaciones previas han reportado información acerca de los sitios de unión de KDM4A en líneas celulares de leucemia y en Células Troncales de Cáncer de Mama (BCSC). En líneas celulares de leucemia se observó que KDM4A se une primordialmente a las regiones adyacentes al TSS (Agger et al. 2016); mientras que en BCSC se observó que el mayor porcentaje de sitios de unión de KDM4A se ubica en regiones intrónicas (36%), intergénicas (32%) y promotoras (22.7%) (Metzger et al. 2017). Los hallazgos del presente trabajo revelaron que los sitios de unión KDM4A están mayoritariamente representados en las regiones promotoras, intergénicas e intrónicas del genoma de las células MCF7.

El enriquecimiento de KDM4A en las regiones promotoras e intrónicas de los genes es congruente con su papel como desmetilasa de la marca H3K9me3 en los promotores génicos, la marca H3K9me2 en las regiones intergénicas y la marca H3K36me3 que media la elongación de los transcritos en el cuerpo de los genes (Hyun et al. 2017). Mientras que la razón de su enriquecimiento en las regiones intergénicas del genoma es menos predecible. En otros organismos modelo como *Drosophila* se ha planteado que la interacción entre la proteína HP1 y KDM4A favorece su reclutamiento a regiones intergénicas, donde la marca H3K36me3 podría regular la estructura y replicación de los dominios de heterocromatina (Lin et al. 2012).

Los sitios de unión del complejo que forma KDM4A-CTCF mantienen la misma proporción que los de KDM4A, en la que las regiones promotoras están más representadas que las intergénicas y las intrónicas. Por el momento, no existe un antecedente sobre la preferencia de los sitios de unión de los complejos que forma CTCF con las desmetilasas de histonas. El único dato conocido proviene de la sobreposición de los sitios de unión de KDM5B y CTCF, donde se observó que el enriquecimiento en las regiones promotoras y no promotoras puede variar de acuerdo con la línea celular de cáncer de mama (Yamamoto et al. 2014b). Sin embargo; la sobreposición de la señal de enriquecimiento de dos proteínas no implica que ambas estén formando un complejo, tendrían que realizarse ensayos de ChIP-ReChIP y co-inmunoprecipitación de KDM5B y CTCF para estudiar mejor este fenómeno.

Es interesante notar que el porcentaje de unión a regiones intergénicas e intrónicas es menor cuando KDM4A está formando un complejo con CTCF; mientras que el porcentaje de unión en las regiones promotoras aumenta casi el doble. Esto podría indicar que la formación del complejo o su reclutamiento se ve favorecido en las regiones promotoras de los genes, y que KDM4A tiene otros mecanismos que permiten su reclutamiento al resto de las regiones génicas.

El diagrama de Venn que considera los sitios de KDM4A, CTCF y el complejo KDM4A-CTCF mostró que la mayoría de los blancos génicos de KDM4A coinciden con lo encontrado en las condiciones Re-ChIP. Entre estos blancos aparecen oncogenes (ej. *JAG2*, *MAD1L1*, *GHR*) y genes supresores de tumores (ej. *BINI*, *DUOX1*, *CSMD1*, *SKI*, etc). Esto aporta más evidencia de que el complejo CTCF-KDM4A podría estar regulando la expresión de múltiples genes involucrados en el desarrollo de cáncer de mama.

Al estudiar con más profundidad al conjunto de genes de KDM4A, notamos que algunos de los más enriquecidos de acuerdo con su q-value, han sido catalogados dentro de vías que regulan la comunicación y la migración celular. El análisis de Enriquecimiento Funcional en STRING reveló que cerca de 158 genes pertenecientes al conjunto de KDM4A están relacionados en una vía que regula la adhesión entre células a través de moléculas de membrana. Un fenómeno que esta estrechamente involucrado en la capacidad migratoria de las células de cáncer de mama (D.-M. Li y Feng 2011). Algunos de los genes implicados codifican cadherinas, canales de voltaje, receptores con función de cinasa y fosfatasa, receptores de inmunoglobulinas, proteínas adaptadoras, entre otros.

Aunque se conoce el efecto de la sobreexpresión de las desmetilasas de histonas de la familia JmjC en células de cáncer de mama, no se han descrito muchos blancos génicos que estén directamente relacionados con la migración celular. Un estudio con Células Troncales Embrionarias (ESC), reportó que las desmetilasas KDM4A y KDM4C median la expresión de la cinasa Flk y VE-cadherina, a través de la desmetilación de la marca H3K9 en el promotor de dichos genes. La expresión de Flk y VE-cadherina es clave en la vasculogénesis y la diferenciación endotelial de las ESC (L. Wu et al. 2015). Por su parte, un estudio con células embrionarias mostró que la sobreexpresión de la desmetilasa KDM6A promueve la transcripción de un centenar de genes involucrados en procesos de adhesión celular (Zhou et al. 2019).

En vista de estos hallazgos, realizamos ensayos para evaluar el efecto del abatimiento de KDM4A en capacidad migratoria e invasiva de las células MCF7. El análisis de invasión *in vitro* de reveló que las células MCF7-KDM4A<sup>KO</sup> tienen una capacidad de invasión significativamente menor a las células MCF7 control y a las células MDA-MB-231. Los ensayos de migración *in vitro* mostraron una disminución cualitativa en la capacidad de migración de las células MCF7-KDM4A<sup>KO</sup>.

Es importante destacar que la líneas celulares de tipo basal, como MDA-MB-231, expresan más altos niveles de KDM4A que las líneas celulares luminales que son menos migratorias e invasivas como MCF7 (Ye et al. 2015). En la literatura se ha reportado la capacidad de KDM4A para regular procesos de migración e invasión en la línea celular MDA-MB-231. Se sabe que, bajo condiciones de hipoxia, la depleción de KDM4A guía a la acumulación de la marca represiva H3K9me3 en el locus de HIF-1, un importante regulador de respuesta a hipoxia, cuya pérdida se correlaciona con la capacidad migratoria e invasiva de la línea celular MDA-MB-231 (Dobrynin et al. 2017). Respecto al papel regulatorio de KDM4A en la capacidad migratoria e invasiva de la línea celular MCF7, existe un estudio basado en un sistema KD que reporta hallazgos similares a los nuestros (B.-X. Li et al. 2011). Este grupo de trabajo probó que KDM4A promueve la progresión de cáncer de mama mediante la represión transcripcional del gen supresor de tumores *DIRAS3* (L.-L. Li et al. 2014). Toda la evidencia anterior sugiere que la desmetilasa de histonas KDM4A tiene un papel fundamental en la regulación de la migración y la invasión celular en cáncer de mama, que podría estar relacionado a la regulación transcripcional de múltiples moléculas de adhesión celular.

Finalmente, decidimos estudiar de manera puntual efecto de KDM4A sobre la transcripción de dos blancos génicos que fueron encontrados en nuestros análisis de secuenciación masiva, *WRAP53* y *JAG2*. El antecedente directo acerca del efecto del abatimiento de KDM4A en la expresión de *WRAP53* y *JAG2* proviene de una matriz de datos de RNA-seq previamente publicada de BCSC donde se administro un fuerte inhibidor de KDM4A llamado QC6352 (número de acceso en “Gene Expression Omnibus” GSE95293) (Metzger et al. 2017). En dicho estudio la expresión de *JAG* y *WRAP53* aumentaba tras la inhibición de la actividad catalítica de KDM4A. Este resultado contrasta con los hallazgos del presente trabajo, en los que la expresión de *JAG* y *WRAP53* sufre un decremento tras el KO de KDM4A en células MCF7.

Esta diferencia podría deberse por un lado, a que se evaluó el efecto del inhibidor de KDM4A en una línea celular diferente y a que el compuesto QC6352 inhibe por igual la actividad de varias desmetilasas de la familia JmjC (ej. KDM4A, B, C y D) cuyos blancos son también las marcas H3K9me3 y H3K36me3 (Metzger et al. 2017; Y. K. Chen et al. 2017).

En las células MCF7, el pico de enriquecimiento de KDM4A y el complejo KDM4A-CTCF en el gen *WRAP53* se encuentra en una región que comprende al promotor y el primer exón del gen. De acuerdo con el ensayo ChIP, la tendencia a la subexpresión del gen *WRAP53* podría deberse al aumento de la marca represiva H3K9me3 en su promotor. Es conocido que la trimetilación de la marca H3K9me en las regiones promotoras de genes media el silenciamiento transcripcional de los mismos (Barski et al. 2007).

Con respecto al gen *JAG2*, el pico de enriquecimiento de KDM4A y el complejo KDM4A-CTCF se encuentra en una región que comprende al intrón 19 y el exón 19. Los ensayos de ChIP mostraron que tanto la marca represiva H3K9me3 como la marca de elongación transcripcional H3K36me3 aumentaron en la región bajo estudio. Creemos, que la reducción significativa en la transcripción del gen *JAG2* podría atribuirse a un mecanismo mediado por la marca de histonas H3K9me3 más que por la presencia de la marca de histonas H3K36me3. La primera hipótesis es que la subexpresión de *JAG2* se debe a un aumento de la marca H3K9me3 en el cuerpo del gen. En otros estudios, se ha visto que el enriquecimiento de la marca H3K9me3 en la porción final y el cuerpo de los genes es un buen indicador de represión transcripcional, aunque los mecanismos aún no son muy claros (Rosenfeld, Xuan, y DeSalle 2009; Hahn et al. 2011). Algunos investigadores han señalado también que la marca H3K9me3 tiene otras funciones en el cuerpo de los genes, como el reconocimiento de exones constitutivos y alternativos durante el splicing alternativo (Bieberstein et al. 2016). La segunda hipótesis es que quizá el KO de KDM4A también aumentó los niveles H3K9me3 en el promotor de *JAG2* y esto podría estar regulando su sub-expresión; sin embargo, no se condujeron los experimentos necesarios para descartar ninguna de las hipótesis. Este hallazgo es interesante ya que demuestra que un aumento en una marca de elongación transcripcional como H3K36me3, no conlleva en todos los casos a la activación transcripcional, como fue el caso del gen *CHD5* (Guerra-Calderas et al. 2018). Sin duda, deben existir otras variables que influyen el efecto de una marca de histonas dependiendo del estado de la cromatina. En el caso de *CHD5*, no fue uno si no ocho picos de enriquecimiento de KDM4A y el complejo que podrían estar influenciando fuertemente sus bajos niveles de transcripción. Sería necesario entonces, realizar un análisis más robusto que relacione el número y la localización de los picos de enriquecimiento de KDM4A y el complejo KDM4A-CTCF, con el aumento de las marcas de histonas H3K36me3 y H3K9me3 para vislumbrar mejor el mecanismo de regulación transcripcional. Existe una última hipótesis acerca de la subexpresión de *JAG2* mediada por el KO de KDM4A. Podría ser posible que los bajos niveles de *JAG2* estén relacionados con la expresión de HIF-1, el activador transcripcional de *JAG2* que como se mencionó antes, se encuentra silenciado en células deficientes de KDM4A (Dobrynin et al. 2017; Pietras et al. 2011).

## CONCLUSIONES

Los hallazgos del presente trabajo señalan la participación de CTCF en el reclutamiento genómico de KDM4A y el papel dual del complejo KDM4A-CTCF en la regulación de la expresión de múltiples genes relacionados a la progresión del cáncer de mama, entre los que destacan los genes *JAG2* y *WRAP53*. La unión de KDM4A a la cromatina puede reprimir o promover la expresión génica a través de la remoción de las marcas H3K9me3 y H3K36me3 en el promotor y el cuerpo de los genes. El cambio en la expresión de KDM4A tiene un efecto sobre la capacidad migratoria e invasiva de las células MCF7, lo cual sugiere su participación dentro de procesos de metástasis celular en cáncer de mama.

## PERSPECTIVAS

1. Explorar la presencia de una secuencia consenso de unión de KDM4A y KDM4A-CTCF a través de un análisis de motivos de unión.
2. Realizar un ensayo RNA-seq con células MCF7 control y Knock Out de KDM4A para estudiar el efecto de su abatimiento en la expresión génica global y acotar una matriz de genes blanco.
3. Realizar ensayos de secuenciación masiva ChIP-seq dirigidos a las marcas de histonas H3K9me3 y H3K36me3 en células control y Knock Out de KDM4A para estudiar los mecanismos que regulan transcripcionalmente a sus genes blanco desde el promotor y el cuerpo del gen.
4. Implementar un modelo de sobreexpresión para KDM4A y estudiar su efecto en la expresión de los genes *JAG2* y *WRAP53*.
5. Realizar estudios de acomplamiento proteico para simular la interacción de KDM4A y CTCF, e indagar sobre los dominios que permiten su interacción.

## REFERENCIAS

- Adli, Mazhar. 2018. «The CRISPR Tool Kit for Genome Editing and Beyond». *Nature Communications* 9 (1): 1911. <https://doi.org/10.1038/s41467-018-04252-2>.
- Agger, Karl, Satoru Miyagi, Marianne Terndrup Pedersen, Susanne M. Kooistra, Jens Vilstrup Johansen, y Kristian Helin. 2016. «Jmjd2/Kdm4 Demethylases Are Required for Expression of Il3ra and Survival of Acute Myeloid Leukemia Cells». *Genes & Development* 30 (11): 1278-88. <https://doi.org/10.1101/gad.280495.116>.
- Almendo, V., y G. Fuster. 2011. «Heterogeneity of Breast Cancer: Etiology and Clinical Relevance». *Clinical & Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico* 13 (11):

- 767-73. <https://doi.org/10.1007/s12094-011-0731-9>.
- Arzate-Mejía, Rodrigo G., Félix Recillas-Targa, y Victor G. Corces. 2018. «Developing in 3D: The Role of CTCF in Cell Differentiation». *Development (Cambridge, England)* 145 (6). <https://doi.org/10.1242/dev.137729>.
- Austin, C. P., D. E. Feldman, J. A. Ida, y C. L. Cepko. 1995. «Vertebrate Retinal Ganglion Cells Are Selected from Competent Progenitors by the Action of Notch». *Development (Cambridge, England)* 121 (11): 3637-50.
- Barski, Artem, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, Zhibin Wang, Gang Wei, Iouri Chepelev, y Keji Zhao. 2007. «High-Resolution Profiling of Histone Methylations in the Human Genome». *Cell* 129 (4): 823-37. <https://doi.org/10.1016/j.cell.2007.05.009>.
- Baylin, Stephen B., y Peter A. Jones. 2011. «A Decade of Exploring the Cancer Epigenome - Biological and Translational Implications». *Nature Reviews. Cancer* 11 (10): 726-34. <https://doi.org/10.1038/nrc3130>.
- Beckmann, M. W., D. Niederacher, H. G. Schnürch, B. A. Gusterson, y H. G. Bender. 1997. «Multistep Carcinogenesis of Breast Cancer and Tumour Heterogeneity». *Journal of Molecular Medicine (Berlin, Germany)* 75 (6): 429-39. <https://doi.org/10.1007/s001090050128>.
- Berry, William L., y Ralf Janknecht. 2013. «KDM4/JMJD2 Histone Demethylases: Epigenetic Regulators in Cancer Cells». *Cancer Research* 73 (10): 2936-42. <https://doi.org/10.1158/0008-5472.CAN-12-4300>.
- Bieberstein, Nicole I., Eva Kozáková, Martina Huranová, Prasoon K. Thakur, Zuzana Krchňáková, Michaela Krausová, Fernando Carrillo Oesterreich, y David Staněk. 2016. «TALE-Directed Local Modulation of H3K9 Methylation Shapes Exon Recognition». *Scientific Reports* 6: 29961. <https://doi.org/10.1038/srep29961>.
- Biswas, Subhankar, y C. Mallikarjuna Rao. 2018. «Epigenetic Tools (The Writers, The Readers and The Erasers) and Their Implications in Cancer Therapy». *European Journal of Pharmacology* 837 (octubre): 8-24. <https://doi.org/10.1016/j.ejphar.2018.08.021>.
- Black, Joshua C., Andrew Allen, Capucine Van Rechem, Emily Forbes, Michelle Longworth, Katrin Tschöp, Claire Rinehart, et al. 2010. «Conserved Antagonism between JMJD2A/KDM4A and HP1 $\gamma$  during Cell Cycle Progression». *Molecular Cell* 40 (5): 736-48. <https://doi.org/10.1016/j.molcel.2010.11.008>.
- Bray, Freddie, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, y Ahmedin Jemal. 2018. «Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries». *CA: A Cancer Journal for Clinicians* 68 (6): 394-424. <https://doi.org/10.3322/caac.21492>.
- Breiling, Achim, y Frank Lyko. 2015. «Epigenetic Regulatory Functions of DNA Modifications: 5-Methylcytosine and Beyond». *Epigenetics & Chromatin* 8: 24. <https://doi.org/10.1186/s13072-015-0016-6>.
- Bulyk, M. L., X. Huang, Y. Choo, y G. M. Church. 2001. «Exploring the DNA-Binding Specificities of Zinc Fingers with DNA Microarrays». *Proceedings of the National Academy of Sciences of the United States of America* 98 (13): 7158-63. <https://doi.org/10.1073/pnas.111163698>.
- Burcin, M., R. Arnold, M. Lutz, B. Kaiser, D. Runge, F. Lottspeich, G. N. Filippova, V. V. Lobanenkova, y R. Renkawitz. 1997. «Negative Protein 1, Which Is Required for Function of the Chicken Lysozyme Gene Silencer in Conjunction with Hormone Receptors, Is Identical to the Multivalent Zinc Finger Repressor CTCF». *Molecular and Cellular Biology* 17 (3): 1281-88. <https://doi.org/10.1128/mcb.17.3.1281>.
- Byler, Shannon, Sarah Goldgar, Sarah Heerboth, Meghan Leary, Genevieve Housman, Kimberly Moulton, y Sibaji Sarkar. 2014. «Genetic and Epigenetic Aspects of Breast



- Cancer Progression and Therapy». *Anticancer Research* 34 (3): 1071-77.
- Cadigan, Ken M., y Marian L. Waterman. 2012. «TCF/LEFs and Wnt Signaling in the Nucleus». *Cold Spring Harbor Perspectives in Biology* 4 (11). <https://doi.org/10.1101/cshperspect.a007906>.
- «Cancer - Latest research and news | Nature». s.f. Accedido 13 de enero de 2020. <https://www.nature.com/subjects/cancer/>.
- Chatterjee, Aniruddha, Euan J. Rodger, y Michael R. Eccles. 2018. «Epigenetic Drivers of Tumorigenesis and Cancer Metastasis». *Seminars in Cancer Biology* 51: 149-59. <https://doi.org/10.1016/j.semcancer.2017.08.004>.
- Chavez, Kathryn J., Sireesha V. Garimella, y Stanley Lipkowitz. 2010. «Triple Negative Breast Cancer Cell Lines: One Tool in the Search for Better Treatment of Triple Negative Breast Cancer». *Breast Disease* 32 (1-2): 35-48. <https://doi.org/10.3233/BD-2010-0307>.
- Chen, Hebing, Yao Tian, Wenjie Shu, Xiaochen Bo, y Shengqi Wang. 2012. «Comprehensive Identification and Annotation of Cell Type-Specific and Ubiquitous CTCF-Binding Sites in the Human Genome». *PloS One* 7 (7): e41374. <https://doi.org/10.1371/journal.pone.0041374>.
- Chen, YoungK., Tiziana Bonaldi, Alessandro Cuomo, Joselyn R. Del Rosario, David J. Hosfield, Toufike Kanouni, Shih-chu Kao, et al. 2017. «Design of KDM4 Inhibitors with Antiproliferative Effects in Cancer Models». *ACS Medicinal Chemistry Letters* 8 (8): 869-74. <https://doi.org/10.1021/acsmedchemlett.7b00220>.
- Chen, Zhongzhou, Jianye Zang, Johnathan Whetstone, Xia Hong, Foteini Davrazou, Tatiana G. Kutateladze, Michael Simpson, et al. 2006. «Structural Insights into Histone Demethylation by JMJD2 Family Members». *Cell* 125 (4): 691-702. <https://doi.org/10.1016/j.cell.2006.04.024>.
- Chowdhury, Rasheduzzaman, Kar Kheng Yeoh, Ya-Min Tian, Lars Hillringhaus, Eleanor A. Bagg, Nathan R. Rose, Ivanhoe K. H. Leung, et al. 2011. «The Oncometabolite 2-Hydroxyglutarate Inhibits Histone Lysine Demethylases». *EMBO Reports* 12 (5): 463-69. <https://doi.org/10.1038/embo.2011.43>.
- Couture, Jean-François, Evys Collazo, Patricia A. Ortiz-Tello, Joseph S. Brunzelle, y Raymond C. Trievel. 2007. «Specificity and Mechanism of JMJD2A, a Trimethyllysine-Specific Histone Demethylase». *Nature Structural & Molecular Biology* 14 (8): 689-95. <https://doi.org/10.1038/nsmb1273>.
- Cremer, Thomas, y Marion Cremer. 2010. «Chromosome Territories». *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889. <https://doi.org/10.1101/cshperspect.a003889>.
- Cutter, Amber R., y Jeffrey J. Hayes. 2015. «A Brief Review of Nucleosome Structure». *FEBS Letters* 589 (20PartA): 2914-22. <https://doi.org/10.1016/j.febslet.2015.05.016>.
- De La Rosa-Velázquez, Inti A., Héctor Rincón-Arango, Luis Benítez-Bribiesca, y Félix Recillas-Targa. 2007. «Epigenetic Regulation of the Human Retinoblastoma Tumor Suppressor Gene Promoter by CTCF». *Cancer Research* 67 (6): 2577-85. <https://doi.org/10.1158/0008-5472.CAN-06-2024>.
- DeSantis, Carol E., Jiemin Ma, Mia M. Gaudet, Lisa A. Newman, Kimberly D. Miller, Ann Goding Sauer, Ahmedin Jemal, y Rebecca L. Siegel. 2019. «Breast Cancer Statistics, 2019». *CA: A Cancer Journal for Clinicians* 69 (6): 438-51. <https://doi.org/10.3322/caac.21583>.
- Dimitrova, Emilia, Anne H. Turberfield, y Robert J. Klose. 2015. «Histone Demethylases in Chromatin Biology and Beyond». *EMBO Reports* 16 (12): 1620-39. <https://doi.org/10.15252/embr.201541113>.
- Dobrynin, Grzegorz, Tom E. McAllister, Katarzyna B. Leszczynska, Shaliny Ramachandran, Adam J. Krieg, Akane Kawamura, y Ester M. Hammond. 2017. «KDM4A Regulates

- HIF-1 Levels through H3K9me3». *Scientific Reports* 7 (1): 11094. <https://doi.org/10.1038/s41598-017-11658-3>.
- Dubois-Chevalier, Julie, Frédéric Oger, H  l  ne Dehondt, Fran  ois F. Firmin, C  line Gheeraert, Bart Staels, Philippe Lefebvre, y J  r  me Eeckhoutte. 2014. «A Dynamic CTCF Chromatin Binding Landscape Promotes DNA Hydroxymethylation and Transcriptional Induction of Adipocyte Differentiation». *Nucleic Acids Research* 42 (17): 10943-59. <https://doi.org/10.1093/nar/gku780>.
- Elkon, Ran, y Reuven Agami. 2017. «Characterization of Noncoding Regulatory DNA in the Human Genome». *Nature Biotechnology* 35 (8): 732-46. <https://doi.org/10.1038/nbt.3863>.
- ENCODE Project Consortium. 2012. «An Integrated Encyclopedia of DNA Elements in the Human Genome». *Nature* 489 (7414): 57-74. <https://doi.org/10.1038/nature11247>.
- Escudero-Esparza, Astrid, Michael Bartoschek, Chrysostomi Gialeli, Marcin Okroj, Sioned Owen, Karin Jirstr  m, Akira Orimo, Wen G. Jiang, Kristian Pietras, y Anna M. Blom. 2016. «Complement Inhibitor CSMD1 Acts as Tumor Suppressor in Human Breast Cancer». *Oncotarget* 7 (47): 76920-33. <https://doi.org/10.18632/oncotarget.12729>.
- Esparza-L  pez, Jos  , Elizabet Escobar-Arriaga, Santos Soto-Germes, y Mar  a de Jes  s Ibarra-S  nchez. 2017. «Breast Cancer Intra-Tumor Heterogeneity: One Tumor, Different Entities». *Revista De Investigacion Clinica; Organo Del Hospital De Enfermedades De La Nutricion* 69 (2): 66-76. <https://doi.org/10.24875/ric.17002177>.
- Ewels, Philip, M  ns Magnusson, Sverker Lundin, y Max K  ller. 2016. «MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report». *Bioinformatics (Oxford, England)* 32 (19): 3047-48. <https://doi.org/10.1093/bioinformatics/btw354>.
- Farnebo, Marianne, Vladimir J. N. Bykov, y Klas G. Wiman. 2010. «The P53 Tumor Suppressor: A Master Regulator of Diverse Cellular Processes and Therapeutic Target in Cancer». *Biochemical and Biophysical Research Communications* 396 (1): 85-89. <https://doi.org/10.1016/j.bbrc.2010.02.152>.
- Feldmann, Angelika, Robert Ivanek, Rabih Murr, Dimos Gaidatzis, Lukas Burger, y Dirk Sch  ubeler. 2013. «Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions». *PLoS Genetics* 9 (12): e1003994. <https://doi.org/10.1371/journal.pgen.1003994>.
- Felsenfeld, Gary, y Mark Groudine. 2003. «Controlling the Double Helix». *Nature* 421 (6921): 448-53. <https://doi.org/10.1038/nature01411>.
- Flavahan, William A., Elizabeth Gaskell, y Bradley E. Bernstein. 2017. «Epigenetic Plasticity and the Hallmarks of Cancer». *Science (New York, N.Y.)* 357 (6348). <https://doi.org/10.1126/science.aal2380>.
- Fortunato, Rodrigo S., Luciana R. Gomes, Veridiana Munford, Carolina Fittipaldi Pessoa, Annabel Quinet, Fabio Hecht, Gustavo S. Kajitani, Cristiane Bedran Milito, Denise P. Carvalho, y Carlos Frederico Martins Menck. 2018. «DUOX1 Silencing in Mammary Cell Alters the Response to Genotoxic Stress». *Oxidative Medicine and Cellular Longevity* 2018: 3570526. <https://doi.org/10.1155/2018/3570526>.
- Foulds, L. 1954. «The Experimental Study of Tumor Progression: A Review». *Cancer Research* 14 (5): 327-39.
- Fraser, Peter, y Wendy Bickmore. 2007. «Nuclear Organization of the Genome and the Potential for Gene Regulation». *Nature* 447 (7143): 413-17. <https://doi.org/10.1038/nature05916>.
- Gandhi, Nishant, y Gokul M. Das. 2019. «Metabolic Reprogramming in Breast Cancer and Its Therapeutic Implications». *Cells* 8 (2). <https://doi.org/10.3390/cells8020089>.
- Gebre-Medhin, M., L. G. Kindblom, H. Wennbo, J. T  rnell, y J. M. Meis-Kindblom. 2001. «Growth Hormone Receptor Is Expressed in Human Breast Cancer». *The American*

- Journal of Pathology* 158 (4): 1217-22. [https://doi.org/10.1016/S0002-9440\(10\)64071-0](https://doi.org/10.1016/S0002-9440(10)64071-0).
- Ghaneie, Arezoo, Vlasta Zemba-Palko, Hiromichi Itoh, Kaori Itoh, Daitoku Sakamuro, Seigo Nakamura, Alejandro Peralta Soler, y George C. Prendergast. 2007. «Bin1 Attenuation in Breast Cancer Is Correlated to Nodal Metastasis and Reduced Survival». *Cancer Biology & Therapy* 6 (2): 192-94. <https://doi.org/10.4161/cbt.6.2.3587>.
- Goldberg, Aaron D., C. David Allis, y Emily Bernstein. 2007. «Epigenetics: A Landscape Takes Shape». *Cell* 128 (4): 635-38. <https://doi.org/10.1016/j.cell.2007.02.006>.
- Gray, Steven G., Antonio H. Iglesias, Fernando Lizcano, Raul Villanueva, Sandra Camelo, Hisaka Jingu, Bin T. Teh, et al. 2005. «Functional Characterization of JMJD2A, a Histone Deacetylase- and Retinoblastoma-Binding Protein». *The Journal of Biological Chemistry* 280 (31): 28507-18. <https://doi.org/10.1074/jbc.M413687200>.
- Guerra-Calderas, Lissania, Rodrigo González-Barrios, Luis A. Herrera, David Cantú de León, y Ernesto Soto-Reyes. 2015. «The Role of the Histone Demethylase KDM4A in Cancer». *Cancer Genetics* 208 (5): 215-24. <https://doi.org/10.1016/j.cancergen.2014.11.001>.
- Guerra-Calderas, Lissania, Rodrigo González-Barrios, Carlos César Patiño, Nicolás Alcaraz, Marisol Salgado-Albarrán, David Cantú de León, Clementina Castro Hernández, et al. 2018. «CTCF-KDM4A Complex Correlates with Histone Modifications That Negatively Regulate CHD5 Gene Expression in Cancer Cell Lines». *Oncotarget* 9 (24): 17028-42. <https://doi.org/10.18632/oncotarget.24798>.
- Hahn, Maria A., Xiwei Wu, Arthur X. Li, Torsten Hahn, y Gerd P. Pfeifer. 2011. «Relationship between Gene Body DNA Methylation and Intragenic H3K9me3 and H3K36me3 Chromatin Marks». *PloS One* 6 (4): e18844. <https://doi.org/10.1371/journal.pone.0018844>.
- Hanahan, D., y R. A. Weinberg. 2000. «The Hallmarks of Cancer». *Cell* 100 (1): 57-70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9).
- Hanahan, Douglas, y Robert A. Weinberg. 2011. «Hallmarks of Cancer: The next Generation». *Cell* 144 (5): 646-74. <https://doi.org/10.1016/j.cell.2011.02.013>.
- Happel, Nicole, y Detlef Doenecke. 2009. «Histone H1 and Its Isoforms: Contribution to Chromatin Structure and Function». *Gene* 431 (1-2): 1-12. <https://doi.org/10.1016/j.gene.2008.11.003>.
- Hashimoto, Hideharu, Dongxue Wang, John R. Horton, Xing Zhang, Victor G. Corces, y Xiaodong Cheng. 2017. «Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA». *Molecular Cell* 66 (5): 711-720.e3. <https://doi.org/10.1016/j.molcel.2017.05.004>.
- Hegde, Madhuri R., y Michael R. Crowley. 2019. «4 - Genome and Gene Structure\*». En *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics (Seventh Edition)*, editado por Reed E. Pyeritz, Bruce R. Korf, y Wayne W. Grody, 53-77. Academic Press. <https://doi.org/10.1016/B978-0-12-812537-3.00004-4>.
- Heger, Peter, Birger Marin, Marek Bartkuhn, Einhard Schierenberg, y Thomas Wiehe. 2012. «The Chromatin Insulator CTCF and the Emergence of Metazoan Diversity». *Proceedings of the National Academy of Sciences of the United States of America* 109 (43): 17507-12. <https://doi.org/10.1073/pnas.1111941109>.
- Hoffmann, Inga, Martin Roatsch, Martin L. Schmitt, Luca Carlino, Martin Pippel, Wolfgang Sippl, y Manfred Jung. 2012. «The Role of Histone Demethylases in Cancer Therapy». *Molecular Oncology* 6 (6): 683-703. <https://doi.org/10.1016/j.molonc.2012.07.004>.
- Holliday, Deborah L., y Valerie Speirs. 2011. «Choosing the Right Cell Line for Breast Cancer Research». *Breast Cancer Research: BCR* 13 (4): 215. <https://doi.org/10.1186/bcr2889>.
- «Home - Reactome Pathway Database». s.f. Accedido 16 de febrero de 2020.

- <https://reactome.org/>.
- Hsieh, Tsung-Hua, Chia-Yi Hsu, Cheng-Fang Tsai, Cheng-Yu Long, Chee-Yin Chai, Ming-Feng Hou, Jau-Nan Lee, Deng-Chyang Wu, Shao-Chun Wang, y Eing-Mei Tsai. 2015. «MiR-125a-5p Is a Prognostic Biomarker That Targets HDAC4 to Suppress Breast Tumorigenesis». *Oncotarget* 6 (1): 494-509. <https://doi.org/10.18632/oncotarget.2674>.
- Huang, Xiaobin, Lingna Zhang, Hongyan Qi, Jimin Shao, y Jing Shen. 2013. «Identification and Functional Implication of Nuclear Localization Signals in the N-Terminal Domain of JMJD5». *Biochimie* 95 (11): 2114-22. <https://doi.org/10.1016/j.biochi.2013.08.002>.
- Huisinga, Kathryn L., Brent Brower-Toland, y Sarah C. R. Elgin. 2006. «The Contradictory Definitions of Heterochromatin: Transcription and Silencing». *Chromosoma* 115 (2): 110-22. <https://doi.org/10.1007/s00412-006-0052-x>.
- «Human Molecular Genetics». s. f. CRC Press. Accedido 12 de enero de 2020. <https://www.crcpress.com/Human-Molecular-Genetics/Strachan-Read/p/book/9780815345893>.
- Hyun, Kwangbeom, Jongcheol Jeon, Kihyun Park, y Jaehoon Kim. 2017. «Writing, Erasing and Reading Histone Lysine Methylations». *Experimental & Molecular Medicine* 49 (4): e324. <https://doi.org/10.1038/emm.2017.11>.
- International Human Genome Sequencing Consortium. 2004. «Finishing the Euchromatic Sequence of the Human Genome». *Nature* 431 (7011): 931-45. <https://doi.org/10.1038/nature03001>.
- Jabbari, Kamel, Maharshi Chakraborty, y Thomas Wiehe. 2019. «DNA Sequence-Dependent Chromatin Architecture and Nuclear Hubs Formation». *Scientific Reports* 9 (1): 14646. <https://doi.org/10.1038/s41598-019-51036-9>.
- Jacobsen, Linda, Susan Calvin, y Edward Lobenhofer. 2009. «Transcriptional Effects of Transfection: The Potential for Misinterpretation of Gene Expression Data Generated from Transiently Transfected Cells». *BioTechniques* 47 (1): 617-24. <https://doi.org/10.2144/000113132>.
- Jeong, Young Sun, Jung Sun Park, Yong Ko, y Yong-Kook Kang. 2011. «JHDM3A Module as an Effector Molecule in Guide-Directed Modification of Target Chromatin». *The Journal of Biological Chemistry* 286 (6): 4461-70. <https://doi.org/10.1074/jbc.M110.176040>.
- Johmura, Yoshikazu, Jia Sun, Kyoko Kitagawa, Keiko Nakanishi, Toshiya Kuno, Aya Naiki-Ito, Yumi Sawada, et al. 2016. «SCF(Fbxo22)-KDM4A Targets Methylated P53 for Degradation and Regulates Senescence». *Nature Communications* 7 (febrero): 10574. <https://doi.org/10.1038/ncomms10574>.
- Kaiser, Marie I. 2018. «ENCODE and the Parts of the Human Genome». *Studies in History and Philosophy of Biological and Biomedical Sciences* 72 (diciembre): 28-37. <https://doi.org/10.1016/j.shpsc.2018.10.008>.
- Kamikawa, Yasunao F., y Mary E. Donohoe. 2014. «The Localization of Histone H3K27me3 Demethylase Jmjd3 Is Dynamically Regulated». *Epigenetics* 9 (6): 834-41. <https://doi.org/10.4161/epi.28524>.
- Kampranis, Sotirios C., y Philip N. Tsichlis. 2009. «Histone Demethylases and Cancer». *Advances in Cancer Research* 102: 103-69. [https://doi.org/10.1016/S0065-230X\(09\)02004-1](https://doi.org/10.1016/S0065-230X(09)02004-1).
- Kanwal, Rajnee, Karishma Gupta, y Sanjay Gupta. 2015. «Cancer Epigenetics: An Introduction». *Methods in Molecular Biology (Clifton, N.J.)* 1238: 3-25. [https://doi.org/10.1007/978-1-4939-1804-1\\_1](https://doi.org/10.1007/978-1-4939-1804-1_1).
- Katoh, Masuko, y Masaru Katoh. 2004. «Identification and Characterization of JMJD2 Family Genes in Silico». *International Journal of Oncology* 24 (6): 1623-28.
- Keating, Samuel T., Jorge Plutzky, y Assam El-Osta. 2016. «Epigenetic Changes in Diabetes

- and Cardiovascular Risk». *Circulation Research* 118 (11): 1706-22. <https://doi.org/10.1161/CIRCRESAHA.116.306819>.
- Kelly, Andrew D., y Jean-Pierre J. Issa. 2017. «The Promise of Epigenetic Therapy: Reprogramming the Cancer Epigenome». *Current Opinion in Genetics & Development* 42 (febrero): 68-77. <https://doi.org/10.1016/j.gde.2017.03.015>.
- Kendrick, Howard, Joseph L. Regan, Fiona-Ann Magnay, Anita Grigoriadis, Costas Mitsopoulos, Marketa Zvelebil, y Matthew J. Smalley. 2008. «Transcriptome Analysis of Mammary Epithelial Subpopulations Identifies Novel Determinants of Lineage Commitment and Cell Fate». *BMC Genomics* 9 (diciembre): 591. <https://doi.org/10.1186/1471-2164-9-591>.
- Kim, Tae-Dong, Fang Jin, Sook Shin, Sangphil Oh, Stan A. Lightfoot, Joseph P. Grande, Aaron J. Johnson, Jan M. van Deursen, Jonathan D. Wren, y Ralf Janknecht. s. f. «Histone demethylase JMJD2A drives prostate tumorigenesis through transcription factor ETV1». *The Journal of Clinical Investigation* 126 (2): 706-20. <https://doi.org/10.1172/JCI78132>.
- Kim, Tae-Dong, Sook Shin, William L. Berry, Sangphil Oh, y Ralf Janknecht. 2012. «The JMJD2A Demethylase Regulates Apoptosis and Proliferation in Colon Cancer Cells». *Journal of Cellular Biochemistry* 113 (4): 1368-76. <https://doi.org/10.1002/jcb.24009>.
- Klose, Robert J., Eric M. Kallin, y Yi Zhang. 2006. «JmJc-Domain-Containing Proteins and Histone Demethylation». *Nature Reviews. Genetics* 7 (9): 715-27. <https://doi.org/10.1038/nrg1945>.
- Kung, Johnny T., Barry Kesner, Jee Young An, Janice Y. Ahn, Catherine Cifuentes-Rojas, David Colognori, Yesu Jeon, et al. 2015. «Locus-Specific Targeting to the X Chromosome Revealed by the RNA Interactome of CTCF». *Molecular Cell* 57 (2): 361-75. <https://doi.org/10.1016/j.molcel.2014.12.006>.
- Labbé, Roselyne M., Andreana Holowatyj, y Zeng-Quan Yang. 2013. «Histone Lysine Demethylase (KDM) Subfamily 4: Structures, Functions and Therapeutic Potential». *American Journal of Translational Research* 6 (1): 1-15.
- Labi, V., y M. Erlacher. 2015. «How Cell Death Shapes Cancer». *Cell Death & Disease* 6 (marzo): e1675. <https://doi.org/10.1038/cddis.2015.20>.
- Lai, Eric C. 2004. «Notch Signaling: Control of Cell Communication and Cell Fate». *Development (Cambridge, England)* 131 (5): 965-73. <https://doi.org/10.1242/dev.01074>.
- Lee, Dong Hoon, Go Woon Kim, Yu Hyun Jeon, Jung Yoo, Sang Wu Lee, y So Hee Kwon. 2020. «Advances in Histone Demethylase KDM4 as Cancer Therapeutic Targets». *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, enero. <https://doi.org/10.1096/fj.201902584R>.
- Lee, Hyo Jeong, Bo-Kyoung Kim, Kyoung Bin Yoon, Yong-Chul Kim, y Sun-Young Han. 2017. «Novel Inhibitors of Lysine (K)-Specific Demethylase 4A with Anticancer Activity». *Investigational New Drugs* 35 (6): 733-41. <https://doi.org/10.1007/s10637-017-0496-2>.
- Li, Bei-Xu, Ming-Chang Zhang, Cheng-Liang Luo, Peng Yang, Hui Li, Hong-Mei Xu, Hong-Fei Xu, Yi-Wen Shen, Ai-Min Xue, y Zi-Qin Zhao. 2011. «Effects of RNA Interference-Mediated Gene Silencing of JMJD2A on Human Breast Cancer Cell Line MDA-MB-231 in Vitro». *Journal of Experimental & Clinical Cancer Research: CR* 30 (octubre): 90. <https://doi.org/10.1186/1756-9966-30-90>.
- Li, Dong-Mei, y Yu-Mei Feng. 2011. «Signaling Mechanism of Cell Adhesion Molecules in Breast Cancer Metastasis: Potential Therapeutic Targets». *Breast Cancer Research and Treatment* 128 (1): 7-21. <https://doi.org/10.1007/s10549-011-1499-x>.
- Li, Li-Liang, Ai-Min Xue, Bei-Xu Li, Yi-Wen Shen, Yu-Hua Li, Cheng-Liang Luo, Ming-

- Chang Zhang, et al. 2014. «JMJD2A Contributes to Breast Cancer Progression through Transcriptional Repression of the Tumor Suppressor ARHI». *Breast Cancer Research: BCR* 16 (3): R56. <https://doi.org/10.1186/bcr3667>.
- Lin, Chia-Hui, Ariel Paulson, Susan M. Abmayr, y Jerry L. Workman. 2012. «HP1a Targets the Drosophila KDM4A Demethylase to a Subset of Heterochromatic Genes to Regulate H3K36me3 Levels». *PLoS ONE* 7 (6). <https://doi.org/10.1371/journal.pone.0039758>.
- Little, Peter F. R. 2005. «Structure and Function of the Human Genome». *Genome Research* 15 (12): 1759-66. <https://doi.org/10.1101/gr.4560905>.
- Mahmoudi, S., S. Henriksson, L. Farnebo, K. Roberg, y M. Farnebo. 2011. «WRAP53 Promotes Cancer Cell Survival and Is a Potential Target for Cancer Therapy». *Cell Death & Disease* 2 (enero): e114. <https://doi.org/10.1038/cddis.2010.90>.
- Mahmoudi, Salah, Sofia Henriksson, Martin Corcoran, Cristina Méndez-Vidal, Klas G. Wiman, y Marianne Farnebo. 2009. «Wrap53, a Natural P53 Antisense Transcript Required for P53 Induction upon DNA Damage». *Molecular Cell* 33 (4): 462-71. <https://doi.org/10.1016/j.molcel.2009.01.028>.
- Mallette, Frédérick A., Francesca Mattioli, Gaofeng Cui, Leah C. Young, Michael J. Hendzel, Georges Mer, Titia K. Sixma, y Stéphane Richard. 2012. «RNF8- and RNF168-Dependent Degradation of KDM4A/JMJD2A Triggers 53BP1 Recruitment to DNA Damage Sites». *The EMBO Journal* 31 (8): 1865-78. <https://doi.org/10.1038/emboj.2012.47>.
- Mallette, Frédérick A., y Stéphane Richard. 2012. «JMJD2A Promotes Cellular Transformation by Blocking Cellular Senescence through Transcriptional Repression of the Tumor Suppressor CHD5». *Cell Reports* 2 (5): 1233-43. <https://doi.org/10.1016/j.celrep.2012.09.033>.
- Marshall, Amy D., Charles G. Bailey, y John E. J. Rasko. 2014. «CTCF and BORIS in Genome Regulation and Cancer». *Current Opinion in Genetics & Development* 24 (febrero): 8-15. <https://doi.org/10.1016/j.gde.2013.10.011>.
- Merkenschlager, Matthias, y Elphège P. Nora. 2016. «CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation». *Annual Review of Genomics and Human Genetics* 17: 17-43. <https://doi.org/10.1146/annurev-genom-083115-022339>.
- Metzger, Eric, Stella S. Stepputtis, Juliane Strietz, Bogdan-Tiberius Preca, Sylvia Urban, Dominica Willmann, Anita Allen, et al. 2017. «KDM4 Inhibition Targets Breast Cancer Stem-like Cells». *Cancer Research* 77 (21): 5900-5912. <https://doi.org/10.1158/0008-5472.CAN-17-1754>.
- Mittal, Suruchi, Deepa Subramanyam, Devaveena Dey, Rekha V. Kumar, y Annapoorni Rangarajan. 2009. «Cooperation of Notch and Ras/MAPK Signaling Pathways in Human Breast Carcinogenesis». *Molecular Cancer* 8 (diciembre): 128. <https://doi.org/10.1186/1476-4598-8-128>.
- Mothoneos, Jenny, y Eleonora Pelosi. 2017. *Living with Advanced Cancer: A Guide for People with Cancer, Their Families and Friends*. Woolloomooloo: Cancer Council of New South Wales.
- Mulero-Navarro, Sonia, y Manel Esteller. 2008. «Chromatin Remodeling Factor CHD5 Is Silenced by Promoter CpG Island Hypermethylation in Human Cancer». *Epigenetics* 3 (4): 210-15. <https://doi.org/10.4161/epi.3.4.6610>.
- Nakahashi, Hirotaka, Kyong-Rim Kieffer Kwon, Wolfgang Resch, Laura Vian, Marei Dose, Diana Stavreva, Ofir Hakim, et al. 2013. «A Genome-Wide Map of CTCF Multivalency Redefines the CTCF Code». *Cell Reports* 3 (5): 1678-89. <https://doi.org/10.1016/j.celrep.2013.04.024>.
- Nam, Do-Hyun, Hye-Min Jeon, Shiyeon Kim, Mi Hyun Kim, Young-Ju Lee, Min Su Lee,

- Hyunggee Kim, et al. 2008. «Activation of Notch Signaling in a Xenograft Model of Brain Metastasis». *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 14 (13): 4059-66. <https://doi.org/10.1158/1078-0432.CCR-07-4039>.
- Ng, Stanley S., Kathryn L. Kavanagh, Michael A. McDonough, Danica Butler, Ewa S. Pilka, Benoit M. R. Lienard, James E. Bray, et al. 2007. «Crystal Structures of Histone Demethylase JMJD2A Reveal Basis for Substrate Specificity». *Nature* 448 (7149): 87-91. <https://doi.org/10.1038/nature05971>.
- Oh, Tae Gyu, Shu-Ching M. Wang, y George E. O. Muscat. 2017. «Therapeutic Implications of Epigenetic Signaling in Breast Cancer». *Endocrinology* 158 (3): 431-47. <https://doi.org/10.1210/en.2016-1716>.
- Ohlsson, Rolf, Marek Bartkuhn, y Rainer Renkawitz. 2010. «CTCF Shapes Chromatin by Multiple Mechanisms: The Impact of 20 Years of CTCF Research on Understanding the Workings of Chromatin». *Chromosoma* 119 (4): 351-60. <https://doi.org/10.1007/s00412-010-0262-0>.
- Pande, Mala, Patricia A. Thompson, Kim-Anh Do, Aysegul A. Sahin, Christopher I. Amos, Marsha L. Frazier, Melissa L. Bondy, y Abenaa M. Brewster. 2013. «Genetic Variants in the Vitamin D Pathway and Breast Cancer Disease-Free Survival». *Carcinogenesis* 34 (3): 587-94. <https://doi.org/10.1093/carcin/bgs369>.
- Park, Seung Gu, Sridhar Hannenhalli, y Sun Shim Choi. 2014. «Conservation in First Introns Is Positively Associated with the Number of Exons within Genes and the Presence of Regulatory Epigenetic Signals». *BMC Genomics* 15 (junio): 526. <https://doi.org/10.1186/1471-2164-15-526>.
- Perou, C. M., T. Sørlie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, et al. 2000. «Molecular Portraits of Human Breast Tumours». *Nature* 406 (6797): 747-52. <https://doi.org/10.1038/35021093>.
- Pietras, Alexander, Kristoffer von Stedingk, David Lindgren, Sven Pålman, y Håkan Axelson. 2011. «JAG2 Induction in Hypoxic Tumor Cells Alters Notch Signaling and Enhances Endothelial Cell Tube Formation». *Molecular Cancer Research: MCR* 9 (5): 626-36. <https://doi.org/10.1158/1541-7786.MCR-10-0508>.
- Ponnaluri, V. K. Chaithanya, Divya Teja Vavilala, Sandeep Putty, William G. Gutheil, y Mridul Mukherji. 2009. «Identification of Non-Histone Substrates for JMJD2A-C Histone Demethylases». *Biochemical and Biophysical Research Communications* 390 (2): 280-84. <https://doi.org/10.1016/j.bbrc.2009.09.107>.
- Qiu, Fuming, Yun-Ru Chen, Xiyong Liu, Cheng-Ying Chu, Li-Jiuan Shen, Jinghong Xu, Shikha Gaur, et al. 2014. «Arginine Starvation Impairs Mitochondrial Respiratory Function in ASS1-Deficient Breast Cancer Cells». *Science Signaling* 7 (319): ra31. <https://doi.org/10.1126/scisignal.2004761>.
- Rashidian, Juliet, Erwan Le Scolan, Xiaodan Ji, Qingwei Zhu, Melinda M. Mulvihill, Daniel Nomura, y Kunxin Luo. 2015. «Ski Regulates Hippo and TAZ Signaling to Suppress Breast Cancer Progression». *Science Signaling* 8 (363): ra14. <https://doi.org/10.1126/scisignal.2005735>.
- Rosenfeld, Jeffrey A., Zhenyu Xuan, y Rob DeSalle. 2009. «Investigating Repetitively Matching Short Sequencing Reads: The Enigmatic Nature of H3K9me3». *Epigenetics* 4 (7): 476-86. <https://doi.org/10.4161/epi.4.7.9809>.
- Saldaña-Meyer, Ricardo, Edgar González-Buendía, Georgina Guerrero, Varun Narendra, Roberto Bonasio, Félix Recillas-Targa, y Danny Reinberg. 2014. «CTCF Regulates the Human P53 Gene through Direct Interaction with Its Natural Antisense Transcript, Wrap53». *Genes & Development* 28 (7): 723-34. <https://doi.org/10.1101/gad.236869.113>.

- Sankar, Aditya, Susanne Marije Kooistra, Javier Martin Gonzalez, Claes Ohlsson, Matti Poutanen, y Kristian Helin. 2017. «Maternal Expression of the Histone Demethylase Kdm4a Is Crucial for Pre-Implantation Development». *Development (Cambridge, England)* 144 (18): 3264-77. <https://doi.org/10.1242/dev.155473>.
- Scully, Olivia Jane, Boon-Huat Bay, George Yip, y Yingnan Yu. 2012. «Breast Cancer Metastasis». *Cancer Genomics & Proteomics* 9 (5): 311-20.
- Shan, Naing Lin, Joseph Wahler, Hong Jin Lee, Min Ji Bak, Soumyasri Das Gupta, Hubert Maehr, y Nanjoo Suh. 2017. «Vitamin D Compounds Inhibit Cancer Stem-like Cells and Induce Differentiation in Triple Negative Breast Cancer». *The Journal of Steroid Biochemistry and Molecular Biology* 173: 122-29. <https://doi.org/10.1016/j.jsbmb.2016.12.001>.
- Sharma, Shikhar, Theresa K. Kelly, y Peter A. Jones. 2010. «Epigenetics in Cancer». *Carcinogenesis* 31 (1): 27-36. <https://doi.org/10.1093/carcin/bgp220>.
- Shaw, Jacqueline A., Karen Page, Kevin Blighe, Natasha Hava, David Guttery, Becky Ward, James Brown, et al. 2012. «Genomic Analysis of Circulating Cell-Free DNA Infers Breast Cancer Dormancy». *Genome Research* 22 (2): 220-31. <https://doi.org/10.1101/gr.123497.111>.
- Shiau, Carrie, Michael J. Trnka, Alen Bozicevic, Idelisse Ortiz Torres, Bassem Al-Sady, Alma L. Burlingame, Geeta J. Narlikar, y Danica Galonić Fujimori. 2013. «Reconstitution of Nucleosome Demethylation and Catalytic Properties of a Jumonji Histone Demethylase». *Chemistry & Biology* 20 (4): 494-99. <https://doi.org/10.1016/j.chembiol.2013.03.008>.
- Shilatifard, Ali. 2006. «Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression». *Annual Review of Biochemistry* 75: 243-69. <https://doi.org/10.1146/annurev.biochem.75.103004.142422>.
- Siegel, Rebecca L., Kimberly D. Miller, y Ahmedin Jemal. 2018. «Cancer Statistics, 2018». *CA: A Cancer Journal for Clinicians* 68 (1): 7-30. <https://doi.org/10.3322/caac.21442>.
- Song, Qingxuan, Sofia D. Merajver, y Jun Z. Li. 2015. «Cancer Classification in the Genomic Era: Five Contemporary Problems». *Human Genomics* 9 (octubre): 27. <https://doi.org/10.1186/s40246-015-0049-8>.
- Soto-Reyes, E., y F. Recillas-Targa. 2010. «Epigenetic Regulation of the Human P53 Gene Promoter by the CTCF Transcription Factor in Transformed Cell Lines». *Oncogene* 29 (15): 2217-27. <https://doi.org/10.1038/onc.2009.509>.
- Sun, Qian, Xianyu Zhang, Tong Liu, Xiaolong Liu, Jingshu Geng, Xiaohui He, Yang Liu, y Da Pang. 2013. «Increased Expression of Mitotic Arrest Deficient-like 1 (MAD1L1) Is Associated with Poor Prognosis and Insensitive to Taxol Treatment in Breast Cancer». *Breast Cancer Research and Treatment* 140 (2): 323-30. <https://doi.org/10.1007/s10549-013-2633-8>.
- Sun, Zhiyi, Jolyon Terragni, Terragni Jolyon, Janine G. Borgaro, Yiwei Liu, Ling Yu, Shengxi Guan, et al. 2013. «High-Resolution Enzymatic Mapping of Genomic 5-Hydroxymethylcytosine in Mouse Embryonic Stem Cells». *Cell Reports* 3 (2): 567-76. <https://doi.org/10.1016/j.celrep.2013.01.001>.
- Svobodová Kovaříková, Alena, Soňa Legartová, Jana Krejčí, y Eva Bártoová. 2018. «H3K9me3 and H4K20me3 Represent the Epigenetic Landscape for 53BP1 Binding to DNA Lesions». *Aging* 10 (10): 2585-2605. <https://doi.org/10.18632/aging.101572>.
- Szklarczyk, Damian, Annika L Gable, David Lyon, Alexander Junge, Stefan Wyder, Jaime Huerta-Cepas, Milan Simonovic, et al. 2019. «STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets». *Nucleic Acids Research* 47 (Database issue): D607-13. <https://doi.org/10.1093/nar/gky1131>.



- Tan, Meng-Kwang Marcus, Hui-Jun Lim, y J. Wade Harper. 2011. «SCF(FBXO22) Regulates Histone H3 Lysine 9 and 36 Methylation Levels by Targeting Histone Demethylase KDM4A for Ubiquitin-Mediated Proteasomal Degradation». *Molecular and Cellular Biology* 31 (18): 3687-99. <https://doi.org/10.1128/MCB.05746-11>.
- Tanaka, Takuji, Masahito Shimizu, Takahiro Kochi, y Hisataka Moriwaki. 2013. «Chemical-Induced Carcinogenesis». *Journal of Experimental & Clinical Medicine* 5 (6): 203-9. <https://doi.org/10.1016/j.jecm.2013.10.009>.
- Tang, Yuan, Yue Wang, Mohammad F. Kiani, y Bin Wang. 2016. «Classification, Treatment Strategy, and Associated Drug Resistance in Breast Cancer». *Clinical Breast Cancer* 16 (5): 335-43. <https://doi.org/10.1016/j.clbc.2016.05.012>.
- «Transcription Factor ChIP-seq Data Standards and Processing Pipeline – ENCODE». s. f. Accedido 29 de enero de 2020. [https://www.encodeproject.org/chip-seq/transcription\\_factor/](https://www.encodeproject.org/chip-seq/transcription_factor/).
- Vaish, Vivek, Joohwee Kim, y Minsub Shim. 2017. «Jagged-2 (JAG2) Enhances Tumorigenicity and Chemoresistance of Colorectal Cancer Cells». *Oncotarget* 8 (32): 53262-75. <https://doi.org/10.18632/oncotarget.18391>.
- Van Rechem, Capucine, Joshua C. Black, Myriam Boukhali, Martin J. Aryee, Susanne Gräslund, Wilhelm Haas, Cyril H. Benes, y Johnathan R. Whetstine. 2015. «Lysine Demethylase KDM4A Associates with Translation Machinery and Regulates Protein Synthesis». *Cancer Discovery* 5 (3): 255-63. <https://doi.org/10.1158/2159-8290.CD-14-1326>.
- vanTol, Brenna L., Sergey Missan, Julie Crack, Shasta Moser, William H. Baldrige, Paul Linsdell, y Elizabeth A. Cowley. 2007. «Contribution of KCNQ1 to the Regulatory Volume Decrease in the Human Mammary Epithelial Cell Line MCF-7». *American Journal of Physiology. Cell Physiology* 293 (3): C1010-1019. <https://doi.org/10.1152/ajpcell.00071.2007>.
- Venteicher, Andrew S., Eladio B. Abreu, Zhaojing Meng, Kelly E. McCann, Rebecca M. Terns, Timothy D. Veenstra, Michael P. Terns, y Steven E. Artandi. 2009. «A Human Telomerase Holoenzyme Protein Required for Cajal Body Localization and Telomere Synthesis». *Science (New York, N.Y.)* 323 (5914): 644-48. <https://doi.org/10.1126/science.1165357>.
- Verrier, Laure, Fabrice Escaffit, Catherine Chailleux, Didier Trouche, y Marie Vandromme. 2011. «A New Isoform of the Histone Demethylase JMJD2A/KDM4A Is Required for Skeletal Muscle Differentiation». *PLoS Genetics* 7 (6): e1001390. <https://doi.org/10.1371/journal.pgen.1001390>.
- Virani, Shama, Shami Virani, Justin A. Colacino, Jung H. Kim, y Laura S. Rozek. 2012. «Cancer Epigenetics: A Brief Review». *ILAR Journal* 53 (3-4): 359-69. <https://doi.org/10.1093/ilar.53.3-4.359>.
- Vousden, Karen H., y Xin Lu. 2002. «Live or Let Die: The Cell's Response to P53». *Nature Reviews. Cancer* 2 (8): 594-604. <https://doi.org/10.1038/nrc864>.
- Wang, Fengyi, Yuancui Li, Fengzhi Shan, Qinglei Zhang, Lei Wang, Bin Sheng, y Guangzhou Cheng. 2019. «Upregulation of JMJD2A Promotes Migration and Invasion in Bladder Cancer through Regulation of SLUG». *Oncology Reports* 42 (4): 1431-40. <https://doi.org/10.3892/or.2019.7246>.
- Wang, Hao, Matthew T. Maurano, Hongzhu Qu, Katherine E. Varley, Jason Gertz, Florencia Pauli, Kristen Lee, et al. 2012. «Widespread Plasticity in CTCF Occupancy Linked to DNA Methylation». *Genome Research* 22 (9): 1680-88. <https://doi.org/10.1101/gr.136101.111>.
- Wang, Hong-Li, Mei-Mei Liu, Xin Ma, Lei Fang, Zong-Feng Zhang, Tie-Fang Song, Jia-Yin Gao, et al. 2014. «Expression and Effects of JMJD2A Histone Demethylase in

- Endometrial Carcinoma». *Asian Pacific Journal of Cancer Prevention: APJCP* 15 (7): 3051-56. <https://doi.org/10.7314/apjcp.2014.15.7.3051>.
- Weigelt, Britta, y Jorge S. Reis-Filho. 2009. «Histological and Molecular Types of Breast Cancer: Is There a Unifying Taxonomy?» *Nature Reviews. Clinical Oncology* 6 (12): 718-30. <https://doi.org/10.1038/nrclinonc.2009.166>.
- Weinmann, Amy S., y Peggy J. Farnham. 2002. «Identification of Unknown Target Genes of Human Transcription Factors Using Chromatin Immunoprecipitation». *Methods (San Diego, Calif.)* 26 (1): 37-47. [https://doi.org/10.1016/S1046-2023\(02\)00006-3](https://doi.org/10.1016/S1046-2023(02)00006-3).
- Wit, Elzo de, Erica S. M. Vos, Sjoerd J. B. Holwerda, Christian Valdes-Quezada, Marjon J. A. M. Versteegen, Hans Teunissen, Erik Splinter, Patrick J. Wijchers, Peter H. L. Krijger, y Wouter de Laat. 2015. «CTCF Binding Polarity Determines Chromatin Looping». *Molecular Cell* 60 (4): 676-84. <https://doi.org/10.1016/j.molcel.2015.09.023>.
- Wooden, Jessica, y Pawel Ciborowski. 2014. «Chromatin Immunoprecipitation for Human Monocyte Derived Macrophages». *Methods (San Diego, Calif.)* 70 (2-3): 89-96. <https://doi.org/10.1016/j.ymeth.2014.08.018>.
- Wu, Liangtang, Kishore K. Wary, Sergei Revskoy, Xiaopei Gao, Kitman Tsang, Yulia A. Komarova, Jalees Rehman, y Asrar B. Malik. 2015. «Histone Demethylases KDM4A and KDM4C Regulate Differentiation of Embryonic Stem Cells to Endothelial Cells». *Stem Cell Reports* 5 (1): 10-21. <https://doi.org/10.1016/j.stemcr.2015.05.016>.
- Wu, Xiaoji, y Yi Zhang. 2017. «TET-Mediated Active DNA Demethylation: Mechanism, Function and Beyond». *Nature Reviews. Genetics* 18 (9): 517-34. <https://doi.org/10.1038/nrg.2017.33>.
- Xie, Si, Jean Jakoncic, y Chengmin Qian. 2012. «UHRF1 Double Tudor Domain and the Adjacent PHD Finger Act Together to Recognize K9me3-Containing Histone H3 Tail». *Journal of Molecular Biology* 415 (2): 318-28. <https://doi.org/10.1016/j.jmb.2011.11.012>.
- Xing, F., H. Okuda, M. Watabe, A. Kobayashi, S. K. Pai, W. Liu, P. R. Pandey, et al. 2011. «Hypoxia-Induced Jagged2 Promotes Breast Cancer Metastasis and Self-Renewal of Cancer Stem-like Cells». *Oncogene* 30 (39): 4075-86. <https://doi.org/10.1038/onc.2011.122>.
- Yamamoto, Shoji, Zhenhua Wu, Hege G. Russnes, Shinji Takagi, Guillermo Peluffo, Charles Vaske, Xi Zhao, et al. 2014a. «JARID1B Is a Luminal Lineage-Driving Oncogene in Breast Cancer». *Cancer Cell* 25 (6): 762-77. <https://doi.org/10.1016/j.ccr.2014.04.024>.
- . 2014b. «JARID1B Is a Luminal Lineage-Driving Oncogene in Breast Cancer». *Cancer Cell* 25 (6): 762-77. <https://doi.org/10.1016/j.ccr.2014.04.024>.
- Ye, Qin, Andreana Holowatyj, Jack Wu, Hui Liu, Lihong Zhang, Takayoshi Suzuki, y Zeng-Quan Yang. 2015. «Genetic Alterations of KDM4 Subfamily and Therapeutic Effect of Novel Demethylase Inhibitor in Breast Cancer». *American Journal of Cancer Research* 5 (4): 1519-30.
- Zhang, Dianzheng, Ho-Guen Yoon, y Jiemin Wong. 2005. «JMJD2A Is a Novel N-CoR-Interacting Protein and Is Involved in Repression of the Human Transcription Factor Achaete Scute-like Homologue 2 (ASCL2/Hash2)». *Molecular and Cellular Biology* 25 (15): 6404-14. <https://doi.org/10.1128/MCB.25.15.6404-6414.2005>.
- Zhao, Lei, Haihui Gu, Jianfeng Chang, Junyu Wu, Daliang Wang, Su Chen, Xiaomei Yang, y Baohua Qian. 2014. «MicroRNA-383 Regulates the Apoptosis of Tumor Cells through Targeting Gadd45g». *PloS One* 9 (11): e110472. <https://doi.org/10.1371/journal.pone.0110472>.
- Zhou, Chuan, Yizhi Wang, Jingcheng Zhang, Jianmin Su, Quanli An, Xin Liu, Min Zhang, Yongsheng Wang, Jun Liu, y Yong Zhang. 2019. «H3K27me3 Is an Epigenetic Barrier While KDM6A Overexpression Improves Nuclear Reprogramming Efficiency».

*FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 33 (3): 4638-52. <https://doi.org/10.1096/fj.201801887R>.  
Zlatanova, Jordanka, y Paola Caiafa. 2009. «CTCF and Its Protein Partners: Divide and Rule?»  
*Journal of Cell Science* 122 (Pt 9): 1275-84. <https://doi.org/10.1242/jcs.039990>.

## ANEXOS

### 1. Metodología en extenso

1.1 Cultivo celular.....	63
1.2 Análisis de citometría.....	64
1.3 Ensayo de fluorescencia.....	64
1.4 Extracción y cuantificación de proteínas.....	65
1.5 Western blot.....	66
1.6 Reacción en cadena de la polimerasa (PCR): cuantitativa.....	69
1.7 Extracción y retrotranscripción del RNA.....	70
1.8 Ensayos de Inmunoprecipitación de la Cromatina (ChIP).....	72
1.9 Análisis de muestras de Secuenciación de DNA de alto rendimiento de las muestras ChIP y ChIP-ReChIP (ChIP-Seq).....	76
1.10 Ensayo de invasión <i>in vitro</i> .....	76
1.11 Ensayo de migración <i>in vitro</i> .....	78

#### 1.1 Cultivo celular

Propósito: Mantener una línea celular en un ambiente artificial (*in vitro*) que provee los nutrientes y condiciones necesarias para su crecimiento y proliferación.

Las líneas celulares MCF7, MDA-MB-231 [ATCC (American Type Culture Collection)] fueron cultivadas con medio suplementado con Suero Fetal Bovino (SFB) al 10%. Las células MCF7-KDM4A<sup>KO</sup> obtenidas previamente por el grupo de trabajo del Dr. Soto, se cultivan en condiciones similares a las de la línea MCF7; salvo la administración de puromicina 3µg/mL al medio de cultivo<sup>1</sup>. Las condiciones de cultivo celular se muestran en la tabla A1.

Línea celular	Medio suplementado	Condiciones
MCF7 (HTB-22)	DMEM/F12 (cat.11320-033, Gibco). SFB 10% (cat. 10500064, Gibco).	Aire 95%; CO2 5%; 37°C
MCF7-KDM4A <sup>KO</sup>	DMEM/F12 (cat.11320-033, Gibco). SFB 10% (cat. 10500064, Gibco). Puromicina 3µg/mL (cat. P8833, Sigma).	Aire 95%; CO2 5%; 37°C
MDA-MB-231 (HTB-26)	DMEM/F12 (cat.11320-033, Gibco). SFB 10% (cat. 10500064, Gibco).	Aire 95%; CO2 5%; 37°C

**Tabla A1.** Condiciones de cultivo celular. Se mencionan las condiciones programas en la incubadora celular y el medio suplementado utilizado en el cultivo de cada línea celular.

## 1.2 Análisis de citometría

Propósito: Identificar subpoblaciones celulares positivas a la señal de interés.

### a) Ensayo de citometría

1. Descartar medio de cultivo celular y lavar 2 veces PBS\*.
2. Adicionar de 1mL de Verseno<sup>2\*\*</sup> e incubar a 37°C durante 5 min.
3. Recuperación del Verseno y células en suspensión.
4. Centrifugación durante 5 min a 1200 rpm.
5. Descartar el sobrenadante y resuspender en el botón celular en 1mL de PBS.
6. Trasferir 500 µL de la suspensión celular a un tubo de citometría y realizar el análisis correspondiente. La señal de RFP se excita a 488 nm o 532 nm en línea de láser y se detecta a 588 nm.

\* PBS. NaCl 137 mM, KCl 2.7 mM, Na<sub>2</sub>HPO<sub>4</sub> 10 mM, KH<sub>2</sub>PO<sub>4</sub> 1.8mM. pH 7.2

\*\*Verseno. EDTA 0.53mM en PBS

### b) Obtención de subpoblaciones celulares

7. Realizar diluciones seriadas con los 500 µL restantes de la suspensión celular en una nueva caja de cultivo celular con 24 pozos.
8. Incubar las células a 37°C y CO<sub>2</sub> 5% durante 48 horas
9. Repetir los pasos 1-6 con la suspensión celular de cada pozo. Cada pozo representa una subpoblación celular diferente.
10. Conservar las subpoblaciones celulares con la mayor cantidad de células positivas a la señal de RFP, de acuerdo con el análisis citométrico

### c) Evaluación y procesamiento de los datos

El ensayo se realizó en un BD FACSCalibur. Todos los datos del ensayo fueron analizados con el programa DeNovoSoftware FCS Express versión 6.

## 1.3 Ensayo de fluorescencia

Propósito: Localizar una proteína/antígeno específico *in situ*.

---

<sup>1</sup> La puomicina es un aminoglucósido que se utiliza comúnmente como marcador de selección, ya que inhibe la síntesis de proteínas de las células eucariotas que no poseen la secuencia de la N-acetiltransferasa (pac) de *Streptomyces*.

<sup>2</sup> El Verseno favorece la disociación no enzimática de las células adherentes en cultivo, debido a que el EDTA es un agente quelante de los elementos que requieren las células para desempeñar dicha función.

a) Inmunofluorescencia

1. Esterilizar seis cubreobjetos y colocarlos en cajas de 60 mm.
2. Sembrar 500,000 células sobre cada portaobjetos e incubar a 37°C y CO<sub>2</sub> 5% durante 24 horas.
3. Descartar el medio de cultivo y lavar 2 veces con PBS.
4. Fijar las células de cada pozo con 1mL de formaldehído al 4% en PBS e incubar a temperatura ambiente y agitación suave durante 10 minutos.
5. Descartar la solución de formaldehído y realizar 3 lavados con PBS de 5 minutos cada uno.
6. Mantener las cajas de 60mm en una cámara húmeda para evitar su secado. Remover el cubreobjetos de las cajas de cultivo y colocarlo verticalmente sobre un papel absorbente para remover el exceso de PBS.
7. Montar preparación con 20 µL de medio de montaje para fluorescencia con DAPI<sup>3</sup> (cat. H-1200, Vectashield Vector) sobre un portaobjetos limpio.
8. Almacenar los portaobjetos a 4°C cubiertos de la luz.

b) Captura y procesamiento de imágenes

Las fluorescencias fueron observadas en un microscopio invertido Axiovert 40 C/40 CFL ZEISS con el objetivo 20x. Las imágenes fueron procesadas con el programa ZEISS ZEN edición azul.

## 1.4 Extracción y cuantificación de proteínas

Propósito: Extraer y cuantificar proteínas totales a partir de cultivos celulares.

a) Extracción de proteína total

1. Descartar medio de cultivo celular y lavar 2 veces con PBS. Mantener las cajas en hielo.
2. Agregar 150 µL de buffer de lisis\* por caja celular y mantenerla en hielo o a -20°C durante 15 min.
3. Recuperación del buffer de lisis y células lisadas en un tubo limpio con ayuda de una espátula.
4. Centrifugar 14,000 rpm durante 30 min a 4 °C.
5. Transferir sobrenadante a un tubo limpio y descartar el pellet.
6. Almacenar las proteínas a -80 °C.

\* Buffer de lisis: NP40 1%, Tris-HCL pH8 50Mm, NaCL 150Mm, Inhibidores de proteasas 1X (cat.5871, Cell signaling)

---

<sup>3</sup> El DAPI es un maracador fluorescente que tiñe las cadenas de DNA en regiones enriquecidas por adenina y timina.

b) Cuantificación de proteínas con Ácido Bicinconínico (BCA)<sup>1</sup>

7. Realizar una curva de dilución de Albúmina de Suero Bovino (BSA) con las siguientes concentraciones: 10µg/ µL, 5 µg/ µL, 2.5 µg/ µL, 1.25 µg/ µL.
8. Diluir las muestras de proteína total en buffer de lisis en una proporción 1:10.
9. Depositar en una placa de 96 pozos, 10 µL de las siguientes series de soluciones por triplicado: a) buffer de lisis. b) curva de dilución BSA. c) las muestras de proteína 1:10.
10. Preparar la mezcla de reactivos del kit de cuantificación de proteínas Sigma cat. B-9643. Se debe considerar que por cada 100 partes del reactivo A (BCA) se utilizan 2 partes del Reactivo B (CuSO<sub>4</sub>). A cada pozo se le agregan 200µL de la mezcla y la placa se cubre de la luz.
11. Incubar la placa a temperatura ambiente durante 30 min. y leer la placa con ayuda de un espectrofotómetro.

c) Evaluación y procesamiento de los datos

Las lecturas se registraron en el fotómetro para microplacas Multiskan FC (cat. 51119000, Thermo Scientific) a una absorbancia de 562 nm. Con base en las lecturas de la curva de BSA se realiza una regresión lineal para obtener la ecuación de la recta. A partir de la ecuación de la recta se calcula la concentración de proteína total de cada una de las muestras

## 1.5 Western blot

Propósito: Evidenciar la presencia/ausencia o el cambio en la expresión de una proteína.

a) Electroforesis de proteínas

1. Montar el módulo de fundición (cat. 165-8022, BioRad) y preparar los componentes del gel de gradiente: a) solución concentradora al 5% y b) separadora al 8%. En la tabla A2 se muestran las recomendaciones del proveedor.
2. Verter entre los cristales de 1.5 mm que se encuentran en el módulo de fundición, la solución separadora hasta cubrir 2/3 de la altura del cristal. Dejar polimerizar a temperatura ambiente durante 15 min.
3. Verter entre los cristales de 1.5 mm la solución concentradora hasta cubrir el filo de los cristales. Colocar el peine entre los dos cristales y dejar polimerizar a temperatura ambiente durante 15 min.
4. Montar la cámara de electroforesis y verter buffer de electroforesis\*.
5. Alicuotar en un tubo limpio un volumen equivalente a 60 µg de proteína a partir de una muestra de proteína total previamente cuantificada.
6. Agregar a la alícuota un volumen de solución Laemmli\*\* que respete la proporción 1:4.

---

<sup>4</sup> El método de cuantificación con BCA combina la reducción del Cu<sup>+2</sup> que tiene lugar cuando las proteínas se encuentran en un medio alcalino (reacción de Biuret) con la detección colorimétrica altamente sensible y selectiva del catión Cu<sup>+</sup> por el BCA.

7. Las muestras de proteínas con Laemmli<sup>5</sup> se incuban a 90°C durante 10 min.
8. Retirar el peine del gel de acrilamida y cargar en los pozos: a) marcador de peso molecular para proteínas (cat. 26616, Thermo Scientific). b) muestras de proteína con Laemmli.
9. Electroforesis a 80V durante 1.5 horas.

	Gel Concentrador		Gel Separador	
	4%	7.5%	12%	X%
30% Acrylamide/bis	1.98 ml	3.75 ml	6.0 ml	0.5 × X ml
0.5M Tris-HCl, pH 6.8	3.78 ml	—	—	—
1.5M Tris-HCl, pH 8.8	—	3.75 ml	3.75 ml	3.75 ml
10% SDS	150 µl	150 µl	150 µl	150 µl
diH <sub>2</sub> O	9 ml	7.28 ml	5.03 ml	11.03-(0.5 × X) ml
TEMED	15 µl	7.5 µl	7.5 µl	7.5 µl
10% APS	75 µl	75 µl	75 µl	75 µl
Total Volume	15 ml	15 ml	15 ml	15 ml

**Tabla A2.** Componentes del gel de gradiente y sus concentraciones de acuerdo con las recomendaciones de BioRad.

\* Buffer de electroforesis. Tris 25 mM, glicina 192 mM, SDS 0.1%. pH 8.3

\*\* Laemmli. SDS 4%, glicerol 20%, Tris 0.125M, bromofenol azul 0.004%, β mercaptoetanol 5% pH 6.8

b) Transferencia húmeda de proteínas

10. Retirar el gel de los cristales y mantener a temperatura ambiente en solución TBS\*.
11. Montar el cassette de transferencia sobre una superficie humedecida con buffer de transferencia\*\* en el siguiente orden: a) esponja, b) papel filtro (cat. 1703967, BioRad), c) gel de acrilamida, d) la membrana de nitrocelulosa (cat. 10600000, Amersham), e) papel filtro, f) esponja.
12. Colocar el cassette dentro de la placa con electrodos.
13. Colocar dentro de la cámara de electroforesis la placa con electrodos y un refrigerante. Verter buffer de transferencia hasta cubrir la cámara.
14. Transferencia a 45 V en cuarto frío durante 16 horas.
15. Teñir la membrana de nitrocelulosa con rojo Ponceau<sup>6\*\*\*</sup> para verificar la transferencia de proteínas.

<sup>5</sup> La solución Laemmli desnaturaliza las proteínas al romper sus enlaces covalentes y reducir sus enlaces disulfuro, además de servir como buffer de carga.

<sup>6</sup> El Rojo Ponceau es un colorante que se une a los grupos amino positivamente cargados y las regiones no polares de las proteínas.



\* TBS. Tris-Cl 50 mM, pH 7.5. NaCl 150 mM

\*\*Buffer de transferencia. Tris 25 mM, glicina 192 mM, SDS 0.1%, metanol 20%.

\*\*\* Rojo Ponceau 0.1% Pounceau S, 5% ácido acético.

c) Bloqueo e inmunodetección

16. Lavar de la membrana con TBS en agitación hasta desteñir la membrana.
17. Incubar la membrana con solución de bloqueo a temperatura ambiente y agitación durante 1 h.
18. Incubar la membrana con el anticuerpo de interés, en agitación a 4°C durante 16 horas. La solución con anticuerpo se diluye en buffer de bloqueo. Las soluciones con anticuerpos se preparan conforme a lo indicado en la tabla A3.
19. Lavar 3 veces la membrana con TBS-Tween\* en agitación durante 10 min.
20. Incubar la membrana con el anticuerpo secundario de interés, a temperatura ambiente y en agitación durante 1 hora.
21. Lavar 3 veces la membrana con TBS-Tween en agitación durante 10 min.
22. Marcar con un plumón (LI-COR cat. 926-91000) la ubicación del marcador de peso y humedecer las membranas con 500 µL del sustrato quimioluminiscente del kit Super Signal Thermo Scientific cat. 34094.<sup>7</sup>
23. Revelado en cuarto oscuro con películas para autoradiografía (cat. Z370398-50EA, KODAK).

Antígeno	No. Catálogo	Concentración	Solución de bloqueo
KDM4A	Rabbit C37E5, Cell signaling	1:500	Albumina 5%
Actina	Goat sc-1615, Santa Cruz	1:5,000	Leche 5%
Anti-rabbit	sc-2004, Santa Cruz	1:5,000	Leche 0.5%
Anti-goat	sc-2020, Santa Cruz	1:10,000	Leche 5%

**Tabla A3.** Anticuerpos utilizados en los ensayos de Western Blot. Todas las soluciones de bloqueo se diluyen en TBS-Tween.

\* TBS-Tween. Tris-Cl 50 mM, pH 7.5. NaCl 150 mM, 0.1% Tween.

d) Evaluación y procesamiento de los datos

El tiempo de exposición de las placas es de 3 y 10 min. para Actina y KDMA respectivamente. La placas reveladas se escanean con una impresora multifuncional (cat. L375, Epson) y los archivos se guardan en formato de imagen TIFF.

<sup>7</sup> La oxidación del sustrato quimioluminiscente de la Peroxidasa de Rábano Picante (HRP) que esta acoplada los anticuerpos secundarios permite la detección de un antígeno.

## 1.6 Reacción en cadena de la polimerasa (PCR): cuantitativa

Propósito: Amplificar y cuantificar de manera simultánea la cantidad de una secuencia de DNA en una muestra.

### a) PCR cuantitativa

1. Realizar la reacción del PCR de acuerdo con las instrucciones de uso del reactivo SYBR Green Thermo Scientific cat. K0221. Preparar el mix de reacción como se indica en la tabla A4.
2. Servir el mix de reacción en los tubos de reacción.
3. Servir en los tubos de reacción el DNA y agua en el caso de los controles.

Componente	Concentración final
Agua	Aforar
Primer Forward	0.3 uM
Primer Reverse	0.3 uM
SYBR (2X)	12.5 µL
Templado de DNA	<500 ng
Final	25 µL

**Tabla A4.** Componentes de la reacción de PCR cuantitativa.

4. El programa de amplificación comprende 3 etapas:
  - Desnaturalización: 95°C durante 5 min.
  - Amplificación (40 ciclos): 95°C durante 0.15 min. y Temperatura de alineamiento (Tm) durante 1 min.
  - Curva de Melting: 95°C durante 0.15 min, 60°C durante 1 min y 95°C durante 0.15 min.
5. Registrar los datos de la amplificación con un sistema de tiempo real.

En la tabla A5 se muestra las características de los oligonucleótidos empleados en las reacciones de PCR cuantitativa. Los oligonucleótidos utilizados en los ensayos de expresión fueron diseñados para amplificar únicamente los transcritos maduros; es por esto que comprenden más de dos regiones exónicas.

b) Evaluación y procesamiento de los datos

El ensayo se realizó en un sistema de Tiempo real Applied Biosystems, StepOnePlus. Las reacciones se corren por triplicado y los valores promedio  $Ct^8$  se utilizan para la cuantificación con el método de doble delta CT ( $d\Delta CT$ ).

Blanco Génico	Ensayo	Región	Secuencia 5'-3'
<i>WRAP53</i>	ChIP	Promotor	F- GAC AGG TCT GAA GCC TG. R- CGG GAC GTG AAA GGT TAG
<i>JAG2</i>	ChIP	Intrón 19	F- GGC TCC CCA CTT AAT CAC TT R- GTG GCC ACA ACT ACT TCC TG
<i>RB</i>	ChIP Ctl. Negativo de KDMA	Exón 27	F- CTAACACTGGCATGTTCAAAGC R- GGTGTAGGGGAGGGG
<i>ASCL2</i>	ChIP Ctl. Positivo de KDMA	-1922 río arriba del TSS	F- CGAAGACCGGGTAGAGA R- TCTGGGTGTGGCTGCATAC
<i>WRAP53</i>	Expresión	Exones 9-10-11. Todos los transcritos	F- GAT CTG GAC CCG ACC GGG CA R-CAC GCT GAC CGG AGG CAG TG
<i>JAG2</i>	Expresión	Exones 25-26	F- CTT CCA CAG GTC TGC TGG TG R- CTC CTC TCC CGC TCT TTC CT
<i>GAPDH</i>	Expresión Ctl.	Exones 8-9	F- TCT TAC TCC TTG GAG GCC AT R-TGT CAA GCT CAT TTC CTG GT

**Tabla A5.** Oligonucleótidos empleados en las reacciones de PCR cuantitativa.

## 1.7 Extracción y retrotranscripción del RNA

Propósito: Extraer RNA mensajero maduro a partir de una población celular y obtener una molécula de DNA codificante (cDNA) para cuantificar de manera indirecta la expresión de un gen.

a) Extracción de RNA

1. Descartar medio de cultivo celular, lavar 2 veces con PBS y mantener las cajas en hielo.
2. Agregar 1 mL de trizol<sup>9</sup> (cat. 15596018, Thermo Scientific) y distribuirlo con movimientos circulares en la caja.
3. Incubar durante 5 min. a 4°C, recuperar el trizol con las células lisadas y trasladarlo a un tubo limpio.

<sup>8</sup> El valor Ct es la intersección entre una curva de amplificación y la línea de umbral. Es una medida relativa de la concentración de una secuencia blanco en la reacción de PCR.

<sup>9</sup> El Trizol es una solución con un pH ácido que lisa las células y mantiene la integridad del RNA.

4. Agregar 200  $\mu\text{L}$  cloroformo por cada mililitro de trizol utilizado en la lisis. Mezclar vigorosamente e incubar de 3 min.
5. Centrifugar a 12000 rpm a 4°C durante 15 min.
6. Recuperar la fase acuosa en un tubo limpio y agregar 500  $\mu\text{L}$  de isopropanol por cada mililitro de trizol utilizado en la lisis. Incubar a 4°C durante 10 min.
7. Centrifugar a 12000 rpm a 4°C durante 10 min. y descartar el sobrenadante. El RNA forma un pellet blanco.
8. Resuspender el pellet en 1mL de alcohol al 75% por cada mililitro de trizol utilizado en la lisis.
9. Mezclar vigorosamente y centrifugar a 8000 rpm a 4°C durante 5 min.
10. Descartar el sobrenadante y resuspender el pellet en 20  $\mu\text{L}$  de agua.
11. Cuantificar el RNA y hacer alícuotas que tengan una concentración de 0.33ng/ $\mu\text{L}$ .
12. Almacenar a -80 °C.

b) Retrotranscripción

13. Realizar la retrotranscripción del RNA de acuerdo con las instrucciones del kit GeneAmp Thermo Scientific cat. N8080143. Preparar el mix de reacción como se indica en la tabla A6.
14. Adicionar todos los componentes excepto el Oligo dT y la muestra de RNA.

<b>Componente</b>	<b>Concentración final</b>
Agua DEPC	Aforar
PCR Buffer 10x	1x
MgCl <sub>2</sub> 25 mM	5mM
dNTPs 10 mM ATP, GTP, CTP, TTP	1mM
DTT	5mM
Inhibidor de RNAsas (20U/ $\mu\text{L}$ )	1 U/ $\mu\text{L}$
MulV Reverso transcriptasa (50U/ $\mu\text{L}$ )	2.5U/ $\mu\text{L}$
Oligo d(T)16	2.5 mM
Templado de RNA	1 $\mu\text{g}$
Final	20 $\mu\text{L}$

**Tabla A6.** Componentes de la reacción de retrotranscripción.

15. El programa de retrotranscripción comprende 2 etapas:  
 Alinaemiento: 65°C durante 5 min. y 4°C durante 2 min.  
 Extensión: 37°C durante 40 min. 95°C durante 5 min y 20°C durante 5 minutos.

16. Agregar el oligo dT y la muestra de RNA en los tubos de reacción e incubar solos durante la etapa de alineamiento.
17. Agregar al tubo de reacción el mix de reacción previamente preparado e incubar durante la etapa de extensión.
18. Agregar 40  $\mu\text{L}$  de agua estéril al cDNA, de manera que el volumen final sea de 60  $\mu\text{L}$ .

c) Evaluación y procesamiento de los datos

El ensayo se realizó en un termociclador Veriti Applied Biosystems. La expresión génica se evalúa de manera indirecta por medio de la PCR cuantitativa de las moléculas de cDNA. El análisis se realiza en función a la expresión relativa al gen GAPDH. Finalmente, se lleva a cabo un análisis ANOVA de una vía seguido por una prueba Dunnet de comparación múltiple vs control con el programa PRISM.

## 1.8 Ensayos de Inmunoprecipitación de la Cromatina (ChIP)

Propósito: Estudiar el enriquecimiento selectivo de un antígeno/proteína en una región de cromatina particular.

a) Extracción de la cromatina

1. Cultivar las células de interés en 6 cajas de 150 mm a una confluencia del 80 %.
2. Revisar que el volumen del medio en cada caja sea de 18 mL, si no es así, agregar el volumen faltante con PBS 1X.
3. Añadir 2 mL de buffer de entrecruzamiento\* para obtener una concentración final de formaldehído al 1 %. Mantener las células en agitación por 10 minutos.
4. Añadir 1 mL de glicina 2.5 M para obtener una concentración final de 125 mM. Mantener las células en agitación por 5 minutos.
5. Retirar el medio de las cajas y hacer 3 lavados de 3 min con PBS y en agitación.
6. Remover el PBS y agregar 500  $\mu\text{L}$  buffer de lisis con inhibidores\*\* a una de las cajas. Incubar esta caja durante 5 minutos a 4°C y guardar las demás cajas a 4°C también.
7. Cosechar todo el contenido de la primera caja con una espátula y transferirlo a la siguiente caja. Los pasos 6 y 7 se repiten hasta coleccionar el contenido de la última caja.
8. Almacenar el contenido de la última caja en un tubo limpio a 4°C.

\* Buffer de entrecruzamiento. Formaldehído 11%, NaCl 100 mM, EGTA 0.5 mM, HEPES 50mM pH 8.

\*\*Buffer de lisis. NP40 1%, NaCl 150mM, Tris-HCL 50 mM pH 8, Inhibidores de proteasas 1x.

b) Evaluación de la cromatina

9. Tomar una alícuota de 100  $\mu\text{L}$  de cromatina en un tubo limpio. Esta alícuota es el control de la extracción de la cromatina.

10. Sonicar<sup>10</sup> la cromatina con pulsos de 30 segundos, pulso 0.5 - 0.5, y amplitud de 37% (cat. EF2810E, Daigger).
11. Tomar una alícuota de 60 µL de la cromatina sonicada y agregar 5 µg de RNasa A (Thermo Scientific cat. EN0531) e incubar a 37°C durante una hora.
12. Agregar 5 µg de proteinasa K (Thermo Scientific cat. AM2546) a la alícuota e incubar a a 55°C durante 2 horas.
13. Adicionar 100 µL de agua estéril a la alícuota e incubar a 65°C durante 16 horas para eliminar el entrecruzamiento<sup>11</sup>.
14. Realizar una extracción fenol/cloroformo y precipitación del DNA con alcohol (pasos 14-21). Llevar la muestra a un volumen final de 300 µl con agua y agregar un volumen igual de fenol cloroformo alcohol isoamílico (P3803, Sigma).
15. Mezclar ambos componentes hasta obtener una fase blanca y centrifugar 15 min. a 14 000 rpm a 4°C. Recuperar la fase acuosa y agregar un volumen de cloroformo (300 µL) para remover el fenol.
16. Mezclar hasta obtener una sola fase y centrifugar 15 min a 14 000 rpm 4°C. Recuperar la fase acuosa que contiene el DNA.
17. Añadir 2 volúmenes de ET-OH al 100 % (600 µL) y 0.66 volúmenes de acetato de amonio o de sodio al 5 M (200 µL) a la fase acuosa para precipitar el DNA
18. Mezclar hasta homogenizar e incubar a -80°C durante 2 horas. Centrifugar la muestra durante 30 min. a 14 000 rpm y descartar el sobrenadante.
19. Agregar 400 µL de etanol al 75%, mezclar y centrifugar 7 min a 4°C.
20. Descartar el sobrenadante y conservar el pellet.
21. Secar el pellet a 37°C durante 7 min y resuspenderlo en 20 µL de agua esteril.
22. Evaluar la calidad y fragmentación del DNA mediante la técnica de electroforesis con un gel de agarosa al 1%. Posterior a la sonicación se esperan fragmentos de DNA de entre 200 y 500 pb, lo equivalente a la distancia de uno o dos nucleosomas. Esta longitud es lo suficientemente corta para tener una buena especificidad y resolución en el enriquecimiento y suficientemente larga para tener un alineamiento eficiente si el material se desea secuenciar (Wooden y Ciborowski 2014).

c) Inmunoprecipitación de las proteínas

23. Tomar una alícuota de la cromatina almacenada a 4°C y cuantificar las proteínas mediante el método de Lowry<sup>12</sup> (cat. 23240, Thermo Scientific).

---

<sup>10</sup> Sonicar. Se recomienda que la longitud de los fragmentos de 200-500 pb, lo equivalente a la distancia de uno o dos nucleosomas. Esta longitud es lo suficientemente corta para tener una buena especificidad y resolución en el enriquecimiento y suficientemente larga para tener un alineamiento eficiente de la secuencia.

<sup>11</sup> El formaldehído genera enlaces covalentes entre los grupos amino o imina de las proteínas y los ácidos nucleicos. El entrecruzamiento de la cromatina es reversible al aumentar la temperatura.

<sup>12</sup> El método de Lowry combina la reacción de Biuret con la oxidación de los residuos aromáticos de las proteínas.

24. Realizar la inmuprecipitación (IP) de las proteínas de interés de acuerdo con las instrucciones del kit One Day ChIP Diagenode cat. C01010080, el volumen de la cromatina no debe ser mayor al 28% en cada IP. El volumen de cada IP se establece con base en la concentración de proteínas en la muestra de cromatina. La cantidad de proteína por IP debe ser equivalente a la proteína total extraída de un cultivo celular de 2 millones de células. En el caso de la línea celular MCF7 esta cantidad de células corresponde aproximadamente a 724 µg de proteína.
25. Repartir cada IP en tubos limpios y añadir la cantidad adecuada de inhibidores para ese volumen de muestra (1x). Se debe considerar que los controles del INPUT y las perlas requieren el mismo volumen de muestra que una IP.
26. Adicionar a cada IP la cantidad estandarizada para el anticuerpo primario de interés (Tabla A7), invertir los tubos varias veces y dejarlos en agitación en rotor a 4°C durante 16 horas.

Antígeno	No. Catálogo	Concentración
KDM4A	Mouse ab105953, Abcam	5 µg/IP
H3K9me3	Rabbit pAB-056-050, Diagenode	2.7 µg/IP
H3K36me3	Rabbit pAb-058-050, Diagenode	1.9 µg/IP
IgG (Control negativo)	Rabbit IgG kch-504-250, Diagenode	2 µg/IP

**Tabla A7.** Anticuerpos primarios utilizados en los ensayos ChIP.

d) Purificación de DNA

27. Lavar 42 µL de perlas con 525µL de ChIP buffer 1x por cada IP. Centrifugar a 2000 rpm durante 3 minutos a 4°C. Repetir el lavado una vez más sin perturbar el pellet de perlas.
28. Agregar nuevamente 525µL de ChIP buffer 1x por cada IP.
29. Alicuotar 500 µL de la solución perlas en tubos limpios y centrifugar a 2000 rpm durante 2 minutos a 4°C.
30. Descartar el sobrenadante sin perturbar el pellet de perlas y mantenerlo en hielo.
31. Centrifugar las IP a 14,000 rpm, durante 10 minutos a 4°C. Los complejos anticuerpo-cromatina se encuentran en el sobrenadante.
32. Transferir el sobrenadante a uno de los tubos con el pellet de perlas.
33. Incubar de una hora a 3 horas a 4°C en movimiento.
34. Agregar 1 mL de etanol absoluto al INPUT. Mezclar de 2 a 3 veces por inversión e incubar durante 10 minutos en hielo.
35. Centrifugar el INPUT a 10,000 rpm durante 10 min a 4°C.
36. Descartar el sobrenadante del INPUT sin perturbar el pellet, lavarlo con 1 mL de OH-75%.
37. Centrifugar el INPUT a 9,000 rpm durante 10 minutos a 4°C. Mantener el pellet a temperatura ambiente.

38. Centrifugar las IP a 2500-3000 rpm durante 2 minutos a 4°C, para depositar en el fondo las perlas.
39. Descartar el sobrenadante sin perturbar el pellet y lavarlo con 1mL de ChIP buffer 1X en rotación a 4°C durante 5 minutos. Centrifugar a 2750 rpm por 5 minutos en frío. Repetir esta operación 5 veces. Se obtiene el pellet con los complejos perlas-anticuerpo-cromatina.
40. Agregar al INPUT 100 µL de agua estéril, resuspenderlo e incubarlo a temperatura ambiente.
41. Añadir 100 µL de suspensión purificadora (slurry) a cada IP. Incubar a temperatura ambiente durante 1 minuto.
42. Añadir 100 µL de suspensión purificadora al INPUT e incubarlo a temperatura ambiente durante 1 minuto.
43. Mezclar por inversión todos los tubos. Colocar los tubos en agua hirviendo durante 10 minutos.
44. Agregar 3 µg de Proteinasa K e incubarlo a 55°C durante por lo menos 30 min.
45. Incubar todos los tubos durante 10 minutos en agua hirviendo y centrifugar a 14,000 rpm durante 1 minuto a 4°C.
46. Transferir 70 µL del sobrenadante a un tubo limpio.
47. Añadir 130 µL de agua estéril (Todos los tubos menos al INPUT), mezclar vigorosamente y centrifugar a 14,000 rpm durante 1 minuto a 4°C.
48. Tomar 130 µL del sobrenadante y trasladarlo al tubo donde se encuentran los 70 µL previos. El volumen final de cada muestra es de 200 µL.
49. Realizar diluciones seriales del INPUT (ej. 0.5%, 2.5%, 5%, 10%).
50. Conservar las muestras de DNA a -20°C.
51. Evaluar el enriquecimiento de las proteínas o marcas de histonas mediante PCR cuantitativa.

e) Evaluación y procesamiento de los datos

Los fragmentos de DNA obtenidos del ensayo ChIP se evalúan por medio de una PCR cuantitativa para conocer el enriquecimiento de las proteínas o marcas de histonas en una región de interés. La calidad de los ensayos ChIP a través de controles negativos y positivos de para la inmunoprecipitación de las proteínas de interés. En el caso de KDM4A, el control positivo fue una región localizada a -1922 bp río arriba del TSS de *ASCL2* (Tan, Lim, y Harper 2011). El primer intrón del gen *CHD5* fue el control positivo de la inmunoprecipitación de CTCF (Guerra-Calderas et al. 2018). El control negativo de ambas proteínas fue el exón 27 del gen *RB* (De La Rosa-Velázquez et al. 2007). Finalmente, se estiman las veces de enriquecimiento de una secuencia al comparar la cantidad de DNA recuperado de la inmunoprecipitación y de los controles negativos; es decir, las perlas y los anticuerpos IgG inespecíficos. El análisis involucra los siguientes cálculos: 1) Eficiencia de amplificación (EA) y porcentaje de recuperación. Con base en los Ct de la curva de INPUT se realiza una regresión lineal para obtener la ecuación de la recta. A partir de la pendiente en la ecuación de la recta se calcula la eficiencia de amplificación de los oligonucleótidos.  $AE = 10^{(-1 / pendiente)}$ . El porcentaje de recuperación del INPUT se calcula con la ecuación:  $\%INPUT = AE^{(Ct INPUT - Ct IP)} \times Fd \times 100\%$ ; donde Fd es el factor de dilución del INPUT. 2) Ocupancia (Oc). La ocupación relativa se puede calcular como una relación entre la señal del IP y el control negativo:  $Oc = 2^{(Ct Control negativo - Ct IP)}$ . Los



valores de recuperación u ocupancia se grafican con el programa PRISM.

### **1.9 Análisis de muestras de Secuenciación de DNA de alto rendimiento de las muestras ChIP y ChIP-ReChIP (ChIP-Seq)**

Propósito: Estudiar los sitios del genoma donde las proteínas inmunoprecipitadas, KDM4A y CTCF, muestran un enriquecimiento.

#### a) Flujo de trabajo

1. Secuenciación por síntesis de los fragmentos de DNA obtenidos de los ensayos ChIP y ChIP-ReChIP a través de la tecnología de Illumina NextSeq 500 High Output. Modalidad de “lectura única 1x75” con más de 25 millones de fragmentos secuenciados por muestra.
2. Analizar la calidad de las secuencias a través de la plataforma FastQC.
3. Remover de adaptadores, secuencias duplicadas y mapear las secuencias al genoma hg19 con el programa alineador BWA.
4. Crear tracks de señal y llamar los picos con la herramienta MACS2.
5. Identificar los picos de enriquecimiento de las proteínas de interés en el navegador UCSC.
6. Realizar anotación genómica y análisis de localización con el paquete ChIPSeekeR.
7. Filtrar y enlistar los picos más significativos de acuerdo a su q-value con el programa MACS2.
8. Crear de una matriz de datos que relaciona lista de los picos significativos de los análisis ChIP-seq y sus coordenadas, para obtener una lista de blancos génicos con RStudio. Criterio poco estricto q-value < 0.05
9. Creación de un diagrama de Venn con base en la matriz de datos del paso 8 y la implementación del programa RStudio. Criterio de corte muy estricto q-value < 0.01.
10. Búsqueda bibliográfica e inspección manual del enriquecimiento simultáneo de CTCF, KDM4A y su complejo en el visualizador IGV. Criterio de corte estricto q-value < 0.05.
11. Explorar la red de interacción de los blancos génicos/proteínas con picos de enriquecimiento significativos con el navegador STRING. Criterio poco estricto q-value < 0.1.

### **1.10 Ensayo de invasión *in vitro***

Propósito: Este sistema permite evaluar la propiedad migratoria e invasiva de las células *in vitro*. Consiste en una placa de 24 pozos y un inserto con la misma cantidad de pozos que poseen una membrana de PET con poros de 8 micrones por la que las células pueden migrar. El inserto esta recubierto una matriz o Matrigel, que simula la membrana basal de las células *in vitro* y sirve como barrera para las células no invasivas.

a) Ensayo de invasión

1. Realizar el ensayo de invasión de acuerdo con las instrucciones del sistema Transwell Corning cat. 354165. Cultivar las células de interés en una placa convencional de 24 pozos a una confluencia del 80 %.
2. Atemperar las placas de invasión.
3. Rehidratar las placas de invasión agregando 500 µL de PBS por la pared del inserto.
4. Incubar las placas de invasión 2 horas a 37°C en un ambiente sin CO<sub>2</sub>.
5. Retirar el PBS de las placas invirtiéndolas cuidadosamente sobre una charola con sanitas.
6. Agregar 750 µl de quimioatrayente\* a cada pozo de la parte basal por la ranura especial que se encuentra en el inserto.
7. Tripsinizar y contar las células presentes en cada pozo de la placa convencional. Se debe considerar que en cada pozo del inserto se deben depositar alrededor de 25 mil células resuspendidas en 500 µL de medio DMEM sin SFB.
8. Agregar por la parte apical del inserto 500 µL de la solución con células resuspendidas.
9. Incubar la placa de invasión a 37°C, 5% CO<sub>2</sub> durante 20-22 horas.
10. Remover el medio de la parte apical del sistema de invasión, invirtiendo el inserto sobre una charola con sanitas.
11. Agregar en una nueva placa de 24 pozos 500 µL de solución balanceada HBSS con calceína AM<sup>13</sup>\* en cada pozo.
12. Transferir el inserto a la placa con calceína e incubar a 37°C, 5% CO<sub>2</sub> durante 1 hora.
13. Evaluar la fluorescencia de los pozos a 494/517 nm y contar el número de células que invadieron la matriz.

\*Quimioatrayente. Medio DMEM con SFB al 5%.

\*\* Solución balanceada HBSS con calceína. AM Calceína 4 µg/mL (Calceína 1mM en DMSO).

b) Evaluación y procesamiento de los datos

Los pozos fueron observados en un microscopio invertido Axiovert 40 C/40 CFL ZEISS con el objetivo 5x y procesadas con el programa ZEISS ZEN edición azul. El cálculo se obtiene del conteo celular en varios campos del microscopio de fluorescencia con ayuda del programa Image J. El porcentaje de invasión se obtiene a través de las siguientes fórmulas: 1) *No. células por inserto* =  $(\text{No. Promedio de células de campo/área del campo en cm}^2) \times \text{área del inserto}$ . 2)  $\% \text{Invasión} = (\text{No. células por inserto} / \text{No. Células sembradas}) \times 100$ . Finalmente, se realiza un análisis ANOVA de una vía, seguido de un análisis de Dunnet con el programa PRISM para conocer si existe una diferencia significativa entre las condiciones experimentales y el control.

---

<sup>13</sup> La calceína AM es una molécula permeable a la membrana celular que fluoresce cuando es convertida por esterasas intracelulares a calceína. La Calceína AM es un indicador de viabilidad celular.

### 1.11 Ensayo de migración *in vitro*

Propósito: Evaluar la capacidad migratoria de las células *in vitro*. El área de la herida a través del tiempo se relaciona de manera inversa a la capacidad migratoria de las células.

#### a) Ensayo de migración

1. Cultivar las células de interés en una placa de 6 pozos.
2. Cambiar el medio de cultivo habitual por uno con SFB al 2% cuando las células alcancen una confluencia del 100%.
3. Incubar la placa a 37°C, 5% CO<sub>2</sub> durante 16 horas
4. Realizar una herida en la monocapa celular con una punta de micropipeta de 200uL.
5. Descartar medio de cultivo celular y lavar 2 veces con PBS
6. Agregar 2 mL de medio de cultivo con SFB al 2%.
7. Incubar la placa a 37°C, 5% CO<sub>2</sub> durante el tiempo que dure el monitoreo del ensayo.
8. Tomar fotografías de la misma región de la herida a las 0, 6, 8, 16, 24, etc. horas de haber realizado la herida. Evaluar por lo menos 4 regiones de cada pozo.
9. Medir el cambio en el área de la herida de cada región a través del tiempo.

#### b) Evaluación y procesamiento de los datos

Las heridas fueron observadas en un un microscopio invertido Axiovert 40 C/40 CFL ZEISS con el objetivo 10x y las imágenes procesadas con el programa ZEISS ZEN edición azul. Deben tomarse varias fotografías de una misma región de la herida a través del tiempo para registrar el cambio en su área. La medición del área de la herida se estima a través del programa TScratch. Finalmente, se realiza una prueba T de Student con el programa PRISM para conocer si existe una diferencia significativa entre las condiciones experimentales y el control.

## 2. Información Académica

Durante mis estudios de maestría tuve la oportunidad de participar y colaborar en las siguientes actividades:

- a) Colaborar en la escritura de dos artículos de revisión:
  1. Hernández-Romero IA, Guerra-Calderas L, Salgado-Albarrán M, Maldonado-Huerta T, Soto-Reyes E. The Regulatory Roles of Non-coding RNAs in Angiogenesis and Neovascularization From an Epigenetic Perspective. *Front Oncol.* 2019 Oct 24;9:1091. doi: 10.3389/fonc.2019.01091. eCollection 2019.
  2. Guerra-Calderas L, Hernández-Romero IA, Salgado-Albarrán M, Wegman-Ostrosky T and Soto-Reyes E. A new kid is on the block: when CTCF meets demethylases. *Cancer Genetics.* 2020 Sometido.
- b) Congreso XLIV Congreso Nacional de Genética Humana, en la que expuse los resultados obtenidos durante mi proyecto de maestría.
- c) V Simposio del Posgrado en Ciencias Naturales e Ingeniería de la Universidad Autónoma Metropolitana, Unidad Cuajimalpa, donde obtuve un reconocimiento especial y un premio por la presentación de los resultados obtenidos durante mi proyecto de maestría.
- d) Epigenomics meeting 2019: epigenetic regulation in human diseases, en la que expuse los resultados obtenidos durante mi proyecto de maestría.



# The Regulatory Roles of Non-coding RNAs in Angiogenesis and Neovascularization From an Epigenetic Perspective

Itzel Alejandra Hernández-Romero<sup>†</sup>, Lissania Guerra-Calderas<sup>†</sup>,  
Marisol Salgado-Albarrán, Tatiana Maldonado-Huerta and Ernesto Soto-Reyes\*

Natural Sciences Department, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico

## OPEN ACCESS

### Edited by:

Erika Ruiz-García,  
National Institute of Cancerology  
(INCan), Mexico

### Reviewed by:

Fahd Al-Mulla,  
Genatak, Kuwait  
Shao-Chun Wang,  
China Medical University, Taiwan

### \*Correspondence:

Ernesto Soto-Reyes  
esotoreyes@correo.cua.uam.mx

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Molecular and Cellular Oncology,  
a section of the journal  
Frontiers in Oncology

Received: 20 June 2019

Accepted: 03 October 2019

Published: 24 October 2019

### Citation:

Hernández-Romero IA,  
Guerra-Calderas L,  
Salgado-Albarrán M,  
Maldonado-Huerta T and  
Soto-Reyes E (2019) The Regulatory  
Roles of Non-coding RNAs in  
Angiogenesis and Neovascularization  
From an Epigenetic Perspective.  
Front. Oncol. 9:1091.  
doi: 10.3389/fonc.2019.01091

Angiogenesis is a crucial process for organ morphogenesis and growth during development, and it is especially relevant during the repair of wounded tissue in adults. It is coordinated by an equilibrium of pro- and anti-angiogenic factors; nevertheless, when affected, it promotes several diseases. Lately, a growing body of evidence is indicating that non-coding RNAs (ncRNAs), such as miRNAs, circRNAs, and lncRNAs, play critical roles in angiogenesis. These ncRNAs can act *in cis* or *trans* and alter gene transcription by several mechanisms including epigenetic processes. In the following pages, we will discuss the functions of ncRNAs in the regulation of angiogenesis and neovascularization, both in normal and disease contexts, from an epigenetic perspective. Additionally, we will describe the contribution of Next-Generation Sequencing (NGS) techniques to the discovery and understanding of the role of ncRNAs in angiogenesis.

**Keywords:** angiogenesis, non-coding RNA, epigenetics, neovascularization, next generation sequencing, miRNAs, lncRNAs, circRNA

## INTRODUCTION

In the vascular network, blood vessels act as channels for nutrients, oxygen delivery, and metabolic waste evacuation. The growth of new capillary vessels, known as angiogenesis, plays key roles in embryonic development and in tissue homeostasis and remodeling in adults, as well as in cancer initiation and progression (1, 2). The balance between pro- and anti-angiogenic factors (such as VEGF, PDGF, and TSP-1/2) coordinates angiogenesis and other neovascularization mechanisms such as intussusceptive angiogenesis, vasculogenesis, lymphangiogenesis, vessel co-option, and vasculogenic mimicry (3–5).

Over the last few decades, the study of angiogenesis has helped researchers to understand vascular physiology and its implications for several diseases. For instance, in atherosclerosis, ischemia, and retinopathy, excessive or insufficient vascular growth can affect the behavior of endothelial and smooth muscle cells (6, 7). Studies of the neovascularization processes have also provided molecular targets for the development of therapies to delay cancer progression, since it is well-known that angiogenesis is an essential process that is altered in tumors (8).

Nowadays, the study of the molecular mechanisms involved in angiogenesis is being built on different experimental approaches, such as cell migration, proliferation, and metabolic assays or histological and tri-dimensional models, that approach specific stages of angiogenesis; however, only pieces of the puzzle have been elucidated (9). With advances in high-throughput genomic

technologies such as microarrays, next-generation sequencing (NGS), and bioinformatic analyses, a genome-wide perspective of the elements involved in the angiogenic process is now being taken. Some of the newest players revealed by these approaches are non-coding RNAs (ncRNAs), which have gained relevance in the field of epigenetics (10–12). Therefore, in this review, we will describe the epigenetic regulatory functions of ncRNAs in physiological angiogenesis and vascular diseases, as well as the contribution of NGS technologies to the discovery of new roles for ncRNAs that are associated with angiogenesis.

## AN OVERVIEW OF EPIGENETICS

In 1939, the term “epigenetics” was coined by Conrad Hal Waddington (13). Today, one of the most accepted definitions of the term explains that “epigenetics is the study of the heritable changes in gene expression that cannot be explained by alterations in the DNA sequence” (14). Among the epigenetic components that coordinate nucleus organization and gene transcription are DNA methylation, histone post-translational modifications (PTMs), and histone positioning, but recently, ncRNAs have been incorporated as epigenetic modifiers, because many of these can function as scaffolding elements to transport proteins with epigenetic functions (15). Each of these processes is stimulated by the signals derived from a dynamic epigenetic code that is established on the chromatin depending on the physiological and extracellular context. The writers, readers, and erasers of this code are proteins that place, recognize, or remove chemical modifications of DNA nucleotides and within the amino-terminal regions of histones. Most chromatin “writers” are methyltransferases that catalyze the transfer of methyl groups. DNA methylation occurs predominantly in regions enriched in CpG sites. The occurrence of methylation at the promoter regions of genes is associated with gene silencing. PTMs alter the regulation of gene transcription by changing the structure of chromatin depending on the particular residue that is modified (16, 17). The “readers” are proteins that recognize and associate with the epigenetic modifications, interpret them, and, in many cases, promote the assembly of protein complexes. The erasers remove the modifications and, therefore, alter signaling

**Abbreviations:** BDNF, Brain-Derived Neurotrophic Factor; BRG1, Brahma related gene-1; CAD, Coronary Artery Disease; circRNA, circular RNAs; DEGs, Differentially Expressed Genes; DNMT, DNA methyltransferase; ElciRNAs, Exon-intron circular RNAs; EPCs, Endothelial Progenitor Cells; EZH2, Enhancer of Zeste Homolog 2; GRO-Seq, Global run-on sequencing; HDAC, Histone deacetylase; HDL, High-density lipoprotein; HF, Heart Failure; HIF1, Hypoxia Inducible Factors 1; HUVEC, Human Umbilical Vein Endothelial Cells; IH, Infantile hemangioma; lincRNAs, intergenic lincRNAs; lncRNAs, long non-coding RNAs; LOXL2, Lysyl oxidase-like 2; MeCP2, Methyl-CpG-binding protein 2; miRNAs, microRNAs; MMP, Matrix metalloproteinase; mRNA, messenger RNA; ncRNAs, non-coding RNAs; NGS, Next-Generation Sequencing; PB-EPCs, Peripheral Blood EPCs; PRC, Polycomb Repressive Complex; pre-miRNA, precursor hairpin miRNA; REST, Repressor Element-1 Silencing Transcription; RNA-seq, RNA sequencing; SIRT1, NAD-dependent deacetylase sirtuin1; smRNA-seq, small RNA-seq; SNPs, Single Nucleotide Polymorphisms; SUZ12, Suppressor of Zeste 12 Protein Homolog 2; TE, Transcription Factor; TGF- $\beta$ , Transforming Growth Factor; TSS, Transcription Start Sites; UC-EPCs, Umbilical Cord EPCs; UHRF1, E3 ubiquitin ligase with PHD and RING finger domain 1; VASH1, Angiogenesis inhibitor vasohibin 1; VEGF, Vascular Endothelial Growth Factor.

components that contribute to the regulation of gene expression. Recently, it has been reported that ncRNAs can mediate the binding of epigenetic proteins to their target sequences. Though they do not function alone as “classic” epigenetic modifiers, they play a vital role in both the recruitment and transcriptional regulation of epigenetic modifiers (18). In fact, multiple chromatin-remodeling enzymes have been shown to directly contact ncRNAs, including Enhancer of Zeste Homolog 2 (EZH2) and Suppressor of Zeste 12 Protein Homolog (SUZ12) (writer and eraser within the Polycomb repressive complex 2/PRC2, respectively), and nuclear architectural proteins like Yin Yang 1 and CTCF, among others (19–22). The incorporation of ncRNAs as epigenetic elements has opened up new fields of study in which they have been shown to regulate gene expression. In the following pages, we will provide an overview of the ncRNAs involved in angiogenesis, focusing on those involved in epigenetic processes.

## MiRNAs AND THEIR EPIGENETIC TARGETS IN NEOVASCULARIZATION AND ANGIOGENIC PROCESSES

MicroRNAs (miRNAs) are short ncRNAs with a length of 19–23 nucleotides that are conserved in animals, plants, and some viruses (23–25). MiRNAs are transcribed as long pri-microRNAs (pri-miRNA) and are subsequently processed to ~70-nucleotide precursor hairpins (pre-miRNA) by the RNase Drosha (26). Pre-miRNAs are then exported to the cytoplasm and recognized by the RNase DICER, which removes the loop linking the 3' and 5' ends of the hairpin, producing a ~20-nucleotide miRNA duplex (27). Later, one of these strands is fused into the RNA Induced Silencing Complex (RISC), where both the miRNA and its messenger RNA (mRNA) target interact (28).

MiRNAs have two main functions: post-transcriptional gene regulation and RNA silencing. They act by pairing bases with a complementary sequence located in the 3'UTR region of target mRNA (29, 30). Consequently, these mRNAs are regulated by one or more mechanisms that include the inhibition of mRNA translation to proteins by ribosomes and by mRNA strand cleavage into two fragments and poly(A) tail shortening that results in mRNA disruption (29, 31). In the last 10 years, the field of miRNA biology has ignited, revealing amazing functions in angiogenesis. These miRNAs have been termed angiomiRs, and they target key angiogenesis molecular drivers, such as metalloproteinases, hypoxia inducible factor 1 (HIF1), cytokines, and growth factors, such as EGFL7, FGF11, PDGFRB, and the vascular endothelial growth factor (VEGF) family (32–34).

MiRNAs are not considered epigenetic components, but some of them are modulated by epigenetic mechanisms. This mainly affects their regulatory region through the incorporation of DNA methylation, repressive histone marks, or the loss of transcriptional factors, as has been reported for *miR-125b1* and *miR-124* (35, 36). Others, known as Epi-miRNAs, can also regulate the gene expression of epigenetic elements, DNA methyltransferases (DNMTs) (such as *miR-152*, *miR-30*, and *miR-148a/b*), histone deacetylases (HDACs) (such as *miR-140*,

*miR-1*, and *miR-449a*), and the Polycomb Group of genes (such as *miR-101* and *miR-26a*) (37–44), and some of them have been considered angiomiRs (39, 40). MiRNAs and their identified epigenetic targets in angiogenesis are listed in **Table 1**.

### **MiR-30a-3p**

Transforming Growth Factor (TGF- $\beta$ ) is a relevant cytokine that functions in the process of vascular homeostasis and is involved in the vascular development of endothelial cells. It has been reported that the administration of TGF- $\beta$  to endothelial cells leads to decreased *miR-30a-3p* expression. The absence of this microRNA results in increased levels of methyl-CpG-binding protein 2 (MeCP2), a protein associated with silencing of *SIRT1* (45). *SIRT1* is necessary for the migration of endothelial cells to occur throughout sprouting angiogenesis, and the loss of this enzyme induces abnormal angiogenesis *in vivo* (52). Conversely, increased levels of *miR-30a-3p* expression lead to the activation of *SIRT1* expression (**Figure 1A**). Further experiments revealed that MeCP2 enhanced the methylation status of the *SIRT1* promoter, probably by DNMT1 recruitment, leading to a reduction in *SIRT1* expression and endothelial angiogenic defects (53).

### **MiR-101**

The microRNA *miR-101* acts as a tumor suppressor, promoting apoptosis and inhibiting cell proliferation, angiogenesis, invasion, and metastasis. *MiR-101* performs its regulatory functions by targeting an abundant range of epigenetic molecular effectors, such as *DNMT3A*, *EZH2*, and *HDAC9* (54, 55). In endothelial cells, high levels of VEGF are associated with the downregulation of *miR-101*, allowing an increase in *EZH2* (46). *EZH2* is associated with the formation of heterochromatin and can affect multiple target genes such as *Vasohibin 1* (*VASH1*), which functions as a negative feedback modulator of angiogenesis in vascular endothelial cells (56, 57) (**Figure 1B**). The overexpression *miR-101* leads to *EZH2* repression and the activation of *VASH1* transcription. This evidence, taken together, suggests that *miR-101* is involved in multiple processes such as cellular growth attenuation, migration, and invasion mechanisms and the ability of endothelial cells to form capillary-like structures in glioblastomas (47).

### **MiR-20a**

*MiR-20a* belongs to the *miR-17-92* cluster and has been linked to breast cancer cells with a high angiogenic profile. High levels of *miR-20a* are correlated with complex vascular structures and larger vessels, suggesting that *miR-20a* could be used as a potential new angiogenic target (58). Additionally, overexpression of *miR-20a* affects the mRNA stability of the lysine acetyltransferase, p300. In mouse myocardium cells, p300 is a key factor that regulates angiogenic and hypertrophic programs, influencing the expression of many related genes, such as *Hif1*, *Vegfc*, *Vegfa*, *Angpt1*, and *Egln3*. Interestingly, high p300 levels induce an increase in the expression of *miR-20a*, providing a feedback inhibition loop for p300 that prevents its pro-angiogenic effects (48).

### **MiR-137**

*MiR-137* has a tumor suppressor gene function that has been reported for several neoplasms (49, 59, 60). It was also reported that this miRNA can inhibit angiogenesis and cell proliferation by *EZH2* downregulation in glioblastomas. Overexpression of *miR-137* reduces the mRNA and protein levels of *EZH2*, while downregulation of *miR-137* is associated with poor prognosis in affected patients (49).

### **MiR-124**

The miRNA *miR-124* is highly conserved, from nematodes to humans. Three human genes encoding *miR-124* have previously been characterized (*miR124a-1*, *miR-124a-2*, and *miR-124a-3*) and the majority have been shown to be deregulated in neoplasms (61). Also, it has been shown that expression of *miR-124* is elevated after treatment with certain drugs such as niclosamide. In this case, it is associated with the inhibition of vasculogenic mimicry formation, particularly by reducing levels of phosphorylated STAT3 (62).

Some reports propose that *miR-124* suppresses the E3 ubiquitin ligase with PHD and RING finger domain 1 (*UHRF1*) expression, a factor involved in the recruitment of epigenetic components in bladder cancer tissues. Also, *UHRF1* is known to enhance malignancy, inducing cellular proliferation, migration, and angiogenesis (63). *MiR-124* overexpression resulted in *UHRF1* suppression through the competitive binding of its 3'-UTR region. In addition, *miR-124* overexpression attenuated tumor growth and cell proliferation *in vivo* and invasion, migration, and vasculogenic mimicry *in vitro*. Further, it reduced VEGF protein levels and levels of the matrix metalloproteinases MMP-2 and MMP-9. A matrigel assay in a three-dimensional culture revealed reductions in tubular channel formation when *miR-124* was over-expressed in bladder cancer cell lines compared to the control group, suggesting that *miR-124* indirectly regulates vasculogenic mimicry in bladder cancer (44).

### **MiR-214**

Originating from intron 14 of the *Dynamina-3* gene (*DNM3*), the primary transcript of *miR-214* produces four different miRNAs (*miR-199-3p*, *miR-199-5p*, *miR214-3p*, and *miR-214-5p*) (64). During the endothelial differentiation of embryonic stem cells, the Brain-Derived Neurotrophic Factor (BDNF) promotes angiogenesis, *in vitro* and *in vivo*, by increasing levels of *miR-214*. The *miR-214* inhibits *EZH2* at the post-transcriptional level, leading to reductions in *EZH2* occupancy at the *NOS3* promoter (50). Also, *miR-214* controls the BDNF-mediated upregulation of neuropilin 1, VEGF-R, and Crk-associated substrate kinase (50, 65). Thus, *miR-214* is a downstream player within the BDNF signaling pathway that regulates important angiogenic targets.

### **MiR-200b**

*miR-200b* is part of the miR-200 family, which is organized into two main groups according to seed sequence. The miRNAs in group A are *miR-141* and *miR-200a*, while the miRNAs in group B are *miR-200b*, *miR-200c*, and *miR-429* (66). Particularly, *miR-200b* has been indicated to have a role in the process of angiogenesis. Studies of malignant neoplasms demonstrated that

**TABLE 1** | Summary of MicroRNAs and their epigenetic targets in angiogenesis and vascular disease.

Common name	Function	Mechanism	Model or disease	References
MiR-30a-3p	Required for endothelial cell migration during sprouting angiogenesis	Base-pairing with matching sequences within <i>MeCP2</i> mRNA	Human umbilical vein endothelial cells	(45)
MiR-101	Inhibits cellular proliferation, migration, invasion and attenuates formation of capillary-like structures	Base-pairing with matching sequences within <i>EZH2</i> mRNA	Human brain microvascular endothelial cells and glioblastoma	(46, 47)
MiR20-a	Inhibits angiogenic and hypertrophic programs	Base-pairing with matching sequences within <i>p300</i> mRNA	Mouse myocardium	(48)
MiR-137	Inhibits cellular proliferation and angiogenesis	Base-pairing with matching sequences within <i>EZH2</i> mRNA	Glioblastoma and xenografts of severe combined immunodeficiency mice	(49)
MiR-124	Inhibits cellular proliferation, migration, invasion and formation of capillary-like structures	Base-pairing with matching sequences within <i>UHRF1</i> mRNA	Bladder cancer	(44)
MiR-214	Promote angiogenesis and endothelial differentiation	Base-pairing with matching sequences within <i>EZH2</i> mRNA	Embryonic stem cells	(50)
MiR-200b	Inhibits the formation of capillary-like structures	Possibly base-pairing with matching sequences within <i>p300</i> mRNA	Diabetic retinopathy	(51)

*miR-200b* controls the epithelial to mesenchymal transition by downregulating p300 (67–70). In addition, p300 activates *HIF1*, which is a transcriptional regulator of *VEGF-A*, and triggers the development of abundant blood vessels (71–73). Since *miR-200b* negatively regulates p300, this miRNA has antiangiogenic properties (51).

In sum, these studies suggest that miRNAs have the capacity to indirectly affect epigenetic pathways in endothelial cells and influence the angiogenic response. This opens up the possibility of considering miRNAs as therapeutic targets or biomarkers, an exciting prospect since therapies for both vascular diseases and cancer are needed. In several diseases, miRNAs have proven to be excellent biomarkers as a result of their high circulating levels. Indeed, analysis of oncogenic and suppressor miRNAs that are found in primary tumors against non-neoplastic cells revealed exosome-mediated sorting mechanisms related to cancer progression (74, 75). It is unknown whether similar mechanisms could be utilized by Epi-miRNAs during the evolution of vascular diseases. Recently, the attention of the scientific community has been focused on other, widely-studied ncRNAs known as long non-coding RNAs (lncRNAs), which have master regulatory functions in angiogenesis.

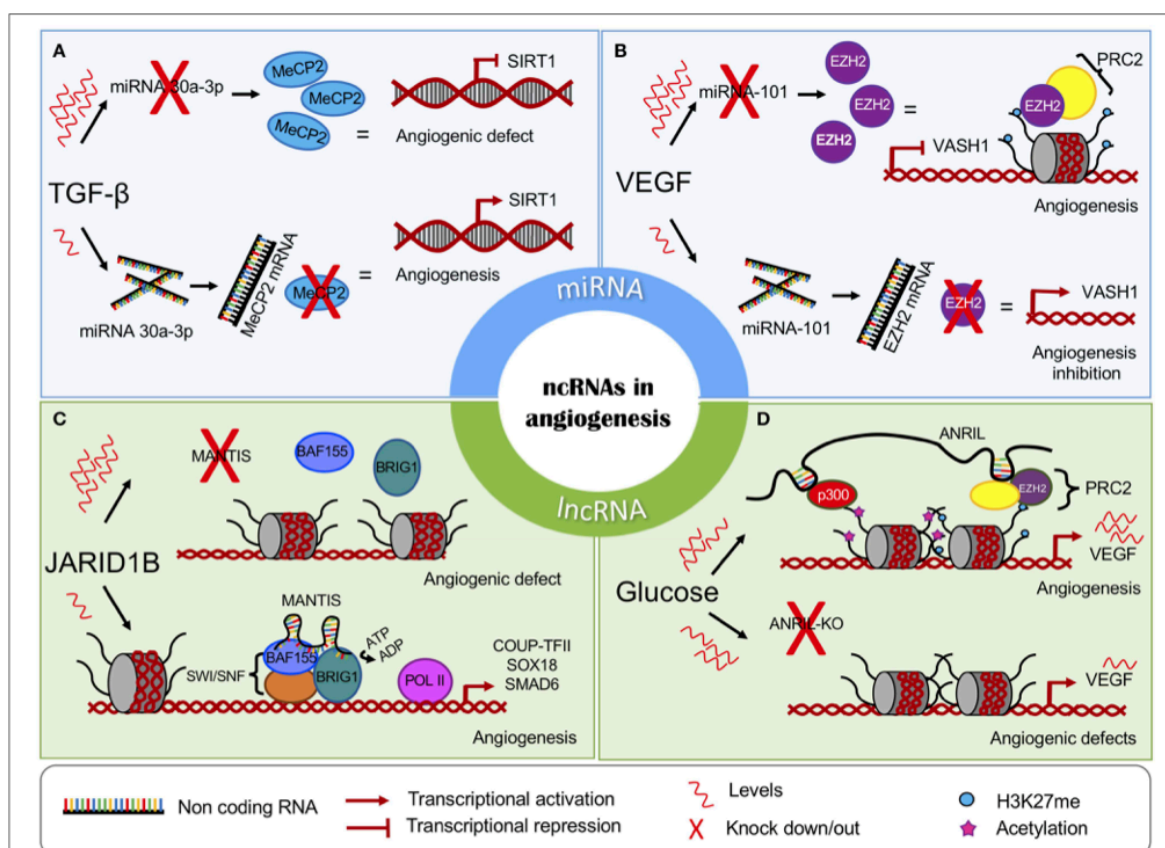
## LONG NON-CODING RNAs AS SCAFFOLDS FOR EPIGENETIC PARTNERS IN NEOVASCULARIZATION

lncRNAs are all ncRNAs larger than 200 nucleotides and are classified according to their proximity to protein-coding genes as intergenic, intronic, bidirectional, sense, and antisense lncRNAs. Massive analyses have revealed that lncRNAs are originated using the same mechanisms as protein-coding genes; however, contrary to protein-coding genes, lncRNAs show a preference for having

two-exon transcripts, and most of them lack any protein coding-potential. Also, lncRNAs show tissue-specific expression patterns and are predominantly located in the nucleus rather than the cytoplasm. In fact, there are several lines of evidence that suggest that lncRNAs are significantly more enriched in chromatin than miRNAs (76).

lncRNAs can indirectly modulate DNA methylation at CpG sites, which in turn regulates gene transcription. For example, Tsix recruits DNMT3a to methylate and silence the *XIST* promoter. *XIST* is an important effector involved in the inactivation of the X chromosome (77). Likewise, the lncRNA *Kcnq1ot1* recruits the *de novo* DNA demethylase DNMT1 to control the methylation status of ubiquitously imprinted genes during mouse development (78). lncRNAs can act as guides or scaffolds, facilitating interaction between several proteins, such as those that are part of chromatin-modifying complexes, causing gene activation or repression, depending on the interaction partners involved (79, 80). The polycomb repressive complexes PRC1 and PRC2, the transcriptional repressor element-1 silencing transcription factor REST, its cofactor (REST/CoREST), other epigenetic components like the mixed lineage leukemia protein and the H3K9 methyltransferase G9a, physically interact with lncRNAs (78, 80, 81). In addition, many lncRNAs such as *HOTAIR*, *Xist*, *Kcnq1ot1*, and *Breaveheart* interact with PRC2, implying that these ncRNAs play a role in recruiting this complex through its subunits (*EZH2*, *SUZ12*, *EED*, *RBBP4*, and *AEBP2*) or through a bridging protein (such as *JARID2*) to their target genes (82, 83). Likewise, the expression of many angiogenesis-related genes involved in the VEGF signaling pathway is regulated through lncRNAs (such as *H19*, *MEG3*, and *HOTAIR*), and recently, researchers discovered that some of them perform their regulatory function by influencing the expression and activity of several epigenetic modulators (20, 22). lncRNAs and their identified epigenetic targets in angiogenesis are listed in **Table 2**.





**FIGURE 1 |** The chromatin regulatory role of non-coding RNAs in angiogenesis. The miRNAs can interfere with the expression of key epigenetic players, leading to the (A) induction or (B) inhibition of angiogenesis. The lncRNAs regulate (C) the activity and (D) recruitment of chromatin-modifying complexes. MiRNAs are represented in blue panels, and lncRNAs are represented in green.

## MANTIS

MANTIS is a recently discovered lncRNA required for endothelial cell function and proper angiogenesis. MANTIS is induced in the endothelium of glioblastoma tumors and is overexpressed during vascular regeneration in atherosclerosis regression. It alters angiogenic sprouting, tube formation, and epithelial cell migration. Loss of MANTIS expression is reported during pulmonary arterial hypertension, and its downregulation also led to the reduced expression of many angiogenesis-related mRNAs (80).

In endothelial cells, MANTIS is upregulated following the knockdown of the histone demethylase JARID1B. JARID1B loss triggers increased H3K4me3 levels at transcription start sites (TSS) of the MANTIS gene, facilitating gene expression. Interestingly, in patients with idiopathic pulmonary arterial hypertension, a disease characterized by endothelial dysfunction, MANTIS expression is downregulated, while JARID1B is upregulated (80).

Novel studies have revealed that MANTIS functions as a scaffold and regulates the activity of Brahma related gene-1 (BRG1), the catalytic subunit of the SWI/SNF chromatin remodeling complex. The MANTIS-BRG1 interaction allows for increased binding of BAF155, which is a core component of the SWI/SNF complex, enhancing BRG1 ATPase activity and chromatin relaxation at the TSS of the transcription factor *COUP-TFII*, which, in turn, recruits RNA Pol II binding and transcription of the pro-angiogenic genes *SOX18* and *SMAD6*. The knockdown of MANTIS reduces BRG1 ATPase activity (80) (Figure 1C).

## ANRIL

ANRIL is an antisense lncRNA from the INK4 locus. It encodes two cyclin-dependent kinase inhibitors, p15 (INK4b) and p16 (INK4a), and a protein known as ARF. All of the genes cooperate in tumor suppressor networks. When these genes are silenced, proatherosclerotic cellular mechanisms are enhanced,

**TABLE 2** | Summary of lncRNAs and their epigenetic targets in angiogenesis and vascular disease.

Common name	Function	Mechanism	Model or disease	References
MANTIS	Promotes angiogenic sprouting and proper endothelial function	Interplay with BRG1 and favors ATPase activity in chromatin remodeling	Human umbilical vein endothelial cells, glioblastoma, atherosclerosis, and idiopathic pulmonary, arterial hypertension	(80)
ANRIL	Induces the formation of capillary-like structures	Recruitment of EZH2 and p300 to the VEGF promoter	Diabetic retinopathy	(84)
GATA6-AS	Promotes angiogenic sprouting	Binds to LOXL2 and regulates H3K4 trimethylation of angiogenesis- and hypoxia-related genes	Human endothelial cell-based xenograft model	(85)

such as increased adhesion and diminished apoptosis (86). In fact, ANRIL expression is correlated with the risk of some vascular diseases such as coronary atherosclerosis and carotid arteriosclerosis (87).

It has been shown that ANRIL recruits PRC2 or PRC1 to different target genes by directly interacting with their subunits EZH2, SUZ12, and CBX7 (86, 88, 89). In a diabetic retinopathy cellular model, high glucose levels upregulated ANRIL and VEGF expression. In turn, ANRIL positively regulated EZH2, EED, and p300 levels. Furthermore, ANRIL recruits EZH2 and histone acetyl-transferase p300 to the VEGF promoter, enhancing its expression and angiogenic effects. It was shown that ANRIL silencing prevented the formation of capillary-like structures in spite of the angiogenic influence of high glucose levels (84) (Figure 1D). Moreover, ANRIL silencing also promoted *miR-200b* expression, a previously described miRNA that has been shown to be involved in regulating VEGF (90).

### GATA6-AS

GATA6-AS is the hypoxia-regulated long non-coding antisense transcript of *GATA6* and promotes angiogenesis by negatively regulating lysyl oxidase-like 2 (*LOXL2*). *LOXL2* catalyzes the oxidative deamination of lysines and hydroxylysines, which results in the generation of non-methylated H3K4 and gene silencing. Thus, GATA6-AS silencing leads to increased *LOXL2* activity and transcriptional repression. In the nucleus, the physical interaction between GATA6-AS and *LOXL2* positively regulates the expression of several angiogenesis- and hypoxia-related genes, such as periostin and cyclooxygenase-2. It has been shown that GATA6-AS silencing in epithelial cells significantly prevented TGF- $\beta$ 2-induced endothelial to mesenchymal transition and augmented angiogenic sprouting in xenograft models *in vivo* (85).

Like epi-miRNAs, the epi-lncRNAs are excellent candidates biomarkers due to their easy collection and tissue specificity. Although there are few examples of epi-lncRNAs in angiogenesis, the implications behind these interactions provide an interesting view of the mechanisms in which lncRNAs regulate not only the recruitment but also the activity of chromatin modifiers. Another layer of complexity is added if we consider that lncRNAs have many alternative splice forms, including the non-linear, circular RNAs (circRNAs).

## CIRCULAR RNAs IN NEOVASCULARIZATION

Circular RNAs (circRNA) are single-stranded RNAs that are widely conserved in all life domains and form a covalent closed loop (91). The discovery of this type of RNA has occurred fairly recently, and before their discovery, the RNAs were considered the result of errors within the process of gene transcription. These circRNAs are produced by a back-splicing process of pre-mRNA, in which a downstream splice donor is linked to an upstream acceptor (92, 93). The splice forms can circularize from exonic, intronic, or a combination of both regions (EiCiRNAs) (94).

In cancer-derived cell lines, it has been reported that changes in DNMTs and the hypermethylation of the CpG islands of some genes that host circRNA can induce gene silencing of both linear RNA and circRNA, suggesting an epigenetic mechanism that produces two molecular "hits" (95). Because circRNA lack 5' and 3' ends, these cannot be degraded by exonucleases. Instead, circRNA levels may be regulated by endonucleases and exosomal deportation (96). These molecules are stable, abundant and specific to certain cell types, having distinct transcriptional patterns for specific tissues and multiple isoforms in eukaryotic cells (97). CircRNAs have been linked to different biological processes, including cell proliferation, senescence, and apoptosis, among others. The study of circRNA has increased in recent years, since they have been shown to be related to both physiological and pathological processes (98). In fact, circRNAs have been proposed as potential biomarkers for neurological disorders, infectious diseases, cancer, and preeclampsia as a result of their availability in circulating body fluids (99–102).

The circRNAs have transcriptional and post-transcriptional regulatory functions. EiCiRNAs such as EIF3J associate with ribonucleoproteins like U1 and the Pol II at the promoters of their parental genes to enhance their own expression (94). Similar to EiCiRNAs, some circRNAs (such as ciANKRD52) can positively regulate their own expression through interaction with the Pol II complex (103). Other circRNAs regulate alternative splicing or serve as sponges to bind, store, or sequester miRNAs and other protein complexes containing transcription factors and RNA binding proteins (94, 104, 105). Due to the ability of circRNA to bind to miRNAs, they have been referred to as miRNA sponges (106). Despite their recent discovery, some evidence suggests that circRNAs are implicated in angiogenesis (e.g., circRNA-MYLK)

and many cardiovascular diseases, such as atherosclerosis (e.g., circR-284), myocardial infarction (e.g., ciRS-7), and coronary artery disease (CAD) (e.g., circ\_0124644), among others (107, 108). However, to our knowledge, no study has shown that circRNAs have an epigenetic regulatory role in angiogenesis. Similar to the lncRNA ANRIL, a circularized and anti-sense splice variant of the INK4/ARF locus (cANRIL) has been associated with atherosclerotic vascular disease (109). Moreover, in the cytoplasm, the binding of circANRIL to the rRNA-processing machinery impairs its function and causes nucleolar fragmentation and stress signaling (110). These findings suggest that, just like their longer-sized isoform, the variant cANRIL may have a role in the epigenetic regulation of vascular disease.

The study of ncRNA has opened up a new research field, and this has been extended to the genome scale. This type of experimental approach has become common practice in both the research laboratory and at the clinical level. Therefore, along with a growing array of genomic analysis machinery, bioinformatics platforms have also been developed, thus generating a new set of tools for the study and analysis of ncRNA.

## CONTRIBUTION OF NGS TECHNOLOGIES TO THE DISCOVERY OF NEW ncRNAs

In recent years, increasing quantities of data have been obtained from NGS technologies such as mass RNA sequencing (RNA-seq), small RNA-seq (smRNA-seq), and single-cell RNA-seq, among others. These technologies have revealed that the human genome encodes for more than 90,000 non-coding RNAs and that these play an important role in several diseases (111). Using publicly available genomic information, it is now possible to discover and characterize novel disease-associated ncRNAs. In the next section, we will describe some of the key discoveries that have been made thanks to NGS data, in which ncRNAs are shown to have roles in angiogenesis and neovascularization processes.

The study of the ncRNAs involved in molecular processes associated with neovascularization and angiogenesis in several diseases can be carried out by using RNA-seq approaches, especially where angiogenesis or neovascularization is one of the causes, risk factors, or consequences of the disorders. Some of the diseases studied in this manner have been ischemia stroke, CAD, hemangioma, and heart failure (HF). Furthermore, angiogenesis and neovascularization are strongly related to endothelial functioning and the transcriptional programming of endothelial progenitor cells (EPCs). Thus, the study of the molecular mechanisms involved in the regulation of EPCs is of great interest. Nevertheless, only a few studies have been conducted on human umbilical vein endothelial cells (HUVEC) or other endothelial models to understand the role of ncRNAs using NGS technologies. In this section, we will provide a compilation of some studies aiming to identify or characterize ncRNAs involved in vascular processes.

First, in 2012 Cheng et al. performed smRNA-seq on umbilical cord blood EPCs (UC-EPCs), which was known for its enrichment in EPCs, and compared the expression profiles

against EPCs derived from peripheral blood in adults (PB-EPCs) to understand the underlying mechanisms involved the functional differences between these two models. They identified specific patterns of miRNAs (miRNome) in UC-EPCs and PB-EPCs in which 54 miRNAs were overexpressed in UC-EPC and 50 miRNAs were overexpressed in PB-EPCs. For instance, UC-EPCs expressed miRNAs involved in angiogenesis such as *miR-31* and *mir-18a*, while PB-EPCs are enriched in tumor-suppressive miRNA expression such as that of *miR-10a* and *mir-26a* (112).

A study performed by Wang and colleagues in 2014 revealed that there was cooperation between VEGF and miRNAs in CAD progression. They performed smRNA-seq and identified EPC-specific miRNome that was related to angiogenic processes, which suggests that miRNAs in EPCs with a poor capacity to enhance angiogenesis might have higher levels of miRNAs targeting VEGF. Indeed, they identified anti-VEGF miRNAs such as *miR-361-5p* that were enriched in EPCs and in the plasma of patients with CAD (113).

Also, atherosclerosis appears to be one of the factors leading to CAD. In 2018, Mao and colleagues conducted a study to identify miRNAs linked with carotid atherosclerosis. They performed a differential expression analysis to identify genes that were specifically associated with either primary or advanced atherosclerotic plaque tissues. Using public databases, they predicted 23 miRNAs that targeted the differentially expressed genes, such as *miR-126*, *miR-155*, *miR-19A*, and *miR-19B*, which can play a regulating role in neovascularization and angiogenesis (114).

Furthermore, a study from Liu et al. (115) identified differentially expressed ncRNAs that were predicted to be involved in the regulation of high-density lipoprotein (HDL) metabolism, the deregulation of which is believed to be one of the main causes of CAD. To this end, they treated HUVEC cells with HDL from healthy subjects and patients with CAD and hypercholesterolemia. After RNA-seq analysis, 41 ncRNAs were identified, and researchers were able to show that the ncRNAs, along with protein-coding genes such as *DGKA* and *UBE2V1*, have critical functions in vascular cells (115).

Additionally, it is well-known that endothelial cell metabolism is sensitive to hypoxia, which is an adverse effect of atherosclerotic lesions in humans. In 2018, Moreau et al. investigated the lncRNA profiles of HUVEC cells using global run-on sequencing (GRO-Seq). GRO-Seq is a sequencing method that measures active transcription, identifying newly synthesized RNA, and providing sufficient resolution to map the position and orientation of transcripts detected. This group aimed to discover changes in the expression patterns of lncRNAs in HUVEC cells exposed to hypoxia and demonstrated that hypoxia affects the transcription of ~1,800 lncRNAs. Among the most relevant lncRNAs identified were *MALAT1*, *HYMAI*, *LOC730101*, *KIAA1656*, and *LOC339803*, which were differentially expressed in human atherosclerotic lesions compared to normal vascular tissue (116).

In contrast, heart and circulatory system diseases often involve changes in vascular smooth muscle or cardiac cells. In 2018, Cheng et al. used RNA-seq to identify circRNAs in human aortic valves. They recognized 1,412 specific circRNAs, most of which

originated from exons of their host genes. Furthermore, after performing a gene ontology enrichment analysis, they found that the host genes were associated with pathways regulating aortic valve function (ECM-receptor interaction pathway, ErbB signaling pathway, and vascular smooth muscle contraction pathway) (117). In addition, Bell et al. identified novel lncRNAs in human vascular smooth muscle cells in 2014. This work expanded our knowledge of the relevance of lncRNAs in the control of smooth muscle cells. The researchers performed an RNA-seq experiment examining expression patterns in human coronary artery smooth muscle cells. Their analysis revealed 31 novel lncRNAs. They discovered and characterized a novel vascular cell-enriched lncRNA that they named *SENCR*. They performed RNA-seq after knockdown of *SENCR* and observed that expression of Myocardin and genes involved in the contraction of smooth muscle were reduced, while expression of other promigratory genes was enhanced (118). These results have enhanced our understanding of vascular cells and should be further studied in order to discern lncRNAs in vascular diseases. Finally, in 2015, Di Salvo et al. analyzed the expression profiles of cells derived from 22 human hearts from patients with Heart Failure (HF) vs. non-HF donor hearts. Initially, they discovered 2,085 lncRNAs, and subsequent analyses revealed 48 differentially expressed lncRNAs in HF patients. Among these, AP000783.2, RP11-403B2.6, and RP11-60A24.3 were identified (119).

Angiogenesis and neovascularization processes affect the prognosis of patients who have suffered from brain stroke ischemia. Thus, the identification of ncRNAs involved in these processes might be useful for their further use as drug targets or biomarkers for the disease. Therefore, Zhang et al. (120) aimed to uncover which ncRNAs have altered expression profiles after cerebrovascular dysfunction in ischemic stroke. Using bulk RNA-seq, they profiled lncRNA signatures in primary brain microvascular endothelial cells after oxygen-glucose deficiency. This approach allowed for the identification of 362 differentially expressed lncRNAs. The top three lncRNAs that were upregulated were *Snhg12*, *Malat1*, and *lnc-OGD 1006*, while the top three down-regulated lncRNAs were *281008D09Rik*, *Peg13*, and *lnc-OGD 3916* (120).

Another disease model that has been studied in order to identify ncRNAs involved in angiogenesis and neovascularization is infantile hemangioma (IH), which is a type of vascular tumor in infants. Li et al. investigated whether ncRNAs have a role in IH pathogenesis in 2018. The researchers used a bulk RNA-seq approach to examine global ncRNAs expression profiles in IH patients compared to their matched, normal-skin controls. In this study, researchers identified 256 lncRNAs and 142 miRNAs that were differentially expressed. They also found more than a thousand sponge modulators involved in miRNA-, lncRNA-, and mRNA-mediated interactions. These findings suggest the presence of an endogenous ncRNA regulatory network associated with the development of IH and other vascular diseases (121).

Overall, the studies described above have shown that NGS technologies can be very effective in identifying and

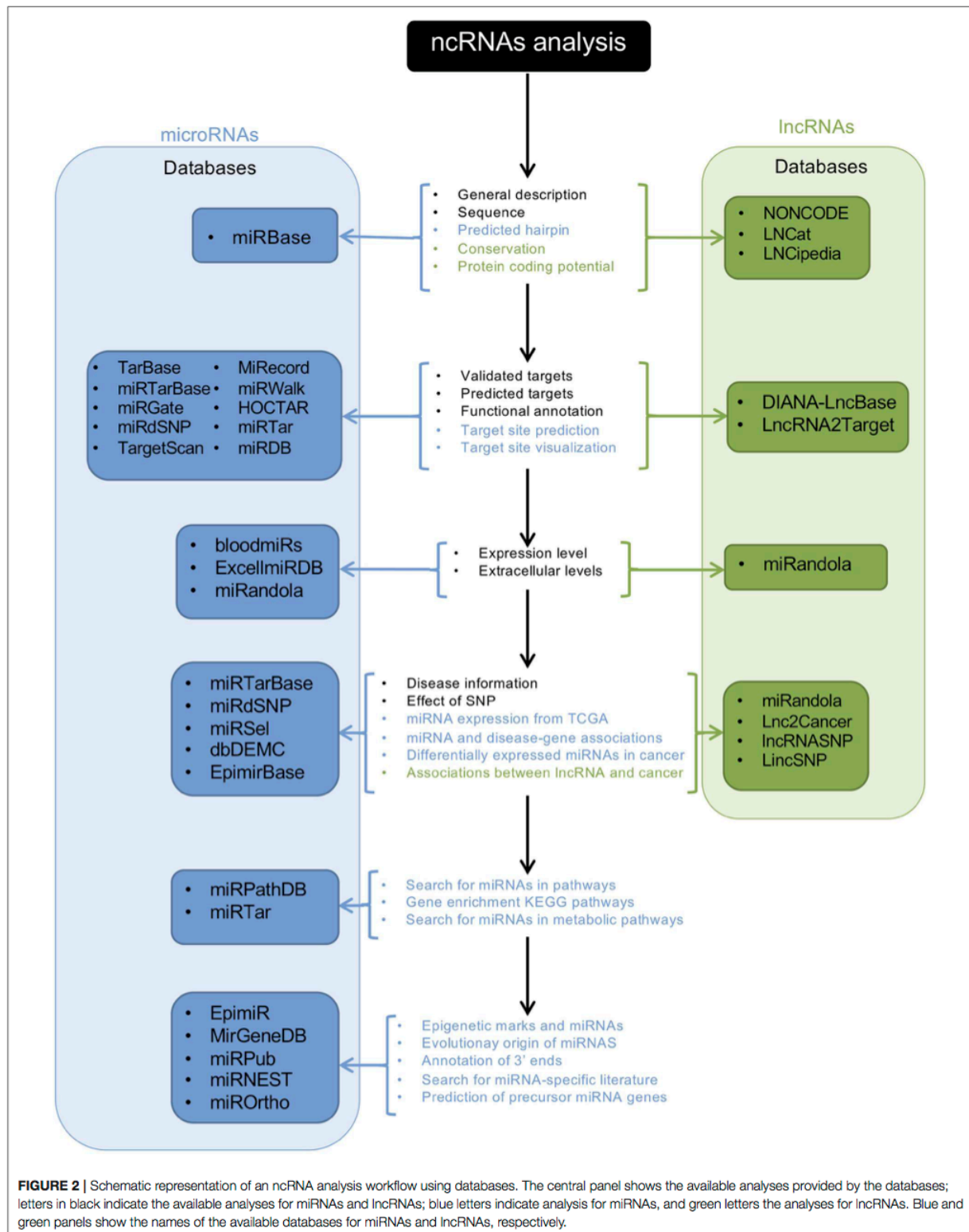
characterizing ncRNAs. This type of technology has helped researchers to understand the regulatory role of ncRNAs in angiogenic and neovascularization processes. However, studies in this field are just emerging, and additional research will be required to expand our knowledge and translated into clinical use.

## CURRENT APPROACHES USED TO DISCOVER NEW ncRNAs

After the development of NGS technologies, ncRNAs have been discovered, and multiple efforts have been made to organize, collect, provide, and unify all available information regarding ncRNAs so that it can be accessed by the research community. Furthermore, new methods have developed to predict and identify novel ncRNAs. Here we present some of the cutting-edge bioinformatics approaches currently being used to study ncRNAs and give some examples of how they are used in the study of neovascularization processes (Figure 2). For a detailed explanation, see the following reference (122).

Transcriptome-wide association studies can be performed to identify expression-trait associations where ncRNAs might be involved. This method can identify single-nucleotide polymorphisms (SNPs) located in transcribed regions of ncRNA genes that can be related to a specific phenotype. A second bioinformatic approach is the use of tools for the prediction of primary, secondary, and tertiary ncRNA structures to obtain information about their potential function. This method has been used for circRNAs, smRNAs, and lncRNAs. The third approach to studying ncRNAs is the use of biological networks. These types of analyses enhance our understanding of the function of ncRNAs by integrating expression, regulatory, and protein-protein interaction networks. ncRNAs are highly connected in these networks and can influence more than one target gene in order to produce a specific phenotype. These approaches can identify disease-specific regulatory modules where ncRNAs play an important role (122).

Though the effective methods described above can be used to discover and understand the biological functions of ncRNAs, they have not been adequately exploited to reveal the roles of ncRNAs in angiogenesis or neovascularization. So far, only a few studies have used advanced bioinformatics tools for this purpose. For example, in 2018, Li et al. detected novel circRNAs related with IH using RNA-seq data. The best experimental approach for the detection of circRNAs is the use of deep sequencing of RNA treated with RNase R (which leaves a circRNA-enriched sample). The availability of tools to predict novel circRNAs from RNA-seq data is of great value, given that RNA-seq data are much more highly available (122). Thus, Li et al. used circRNAFinder, a tool able to predict circRNAs from bulk RNA-seq experiments, and identified 249 circRNA candidates differentially expressed between IH and matched normal skin samples. The circRNAs *hsa\_circRNA001885* and *hsa\_circRNA006612* were further investigated by this group, providing novel insights about the disease (123).



As shown previously, the development of tools used to predict and identify novel ncRNAs is invaluable. The increasing number of RNA-seq experiments and access to databases will increasingly facilitate the discovery of novel ncRNAs, and the characterization of ncRNAs will become increasingly straightforward. For instance, ANGIOGENES is a database that has been created to store information related to angiogenic processes. It depicts experimental data obtained from RNA-seq

experiments in endothelial cells. This allows for the *in-silico* detection of genes expressed in several endothelial cell types from different tissues. ANGIOGENES uses publicly-available RNA-seq experiments and identifies endothelial cell-specific ncRNAs in human, mouse, and zebrafish. The database facilitates further analyses using GO enrichment terms and is available online (124). In addition to ANGIOGENES, EndoDB is another database that retrieves information about endothelial cells from

**TABLE 3 |** Databases and tools for the ncRNAs study.

Database	ncRNAs	Website	Species	References
<b>ANNOTATION RESOURCES</b>				
miRbase	microRNAs	<a href="http://www.mirbase.org/">http://www.mirbase.org/</a>	All	(126, 127)
NONCODE	lncRNAs	<a href="http://www.noncode.org/">http://www.noncode.org/</a>	All	(111)
LNCipedia	lncRNAs	<a href="https://lncipedia.org/">https://lncipedia.org/</a>	Human	(128)
LNCact	lncRNAs	<a href="http://biocc.hrbmu.edu.cn/LNCat/">http://biocc.hrbmu.edu.cn/LNCat/</a>	Human	(129)
<b>TARGET RESOURCES</b>				
TarBase	microRNAs	<a href="http://diana.imis.athena-innovation.gr/DianaTools/">http://diana.imis.athena-innovation.gr/DianaTools/</a>	Human/mouse, fruit fly, worm, and zebrafish	(130)
miRTarBase	microRNAs	<a href="http://mirtarbase.mbc.nctu.edu.tw">http://mirtarbase.mbc.nctu.edu.tw</a>	Human, mouse, virus	(131)
miRGate	microRNAs	<a href="http://mirgate.bioinfo.cnio.es">http://mirgate.bioinfo.cnio.es</a>	Human, rat, mouse	(132)
miRdSNP	microRNAs	<a href="http://mirdsnp.ccr.buffalo.edu">http://mirdsnp.ccr.buffalo.edu</a>	Human	(133)
TargetScan	microRNAs	<a href="http://www.targetscan.org/vert_72/">http://www.targetscan.org/vert_72/</a>	Human	(134)
CSmiRTar	microRNAs	<a href="http://cosbi.ee.ncku.edu.tw/CSmiRTar/">http://cosbi.ee.ncku.edu.tw/CSmiRTar/</a>	Human, mouse	(135)
MiRecords	microRNAs	<a href="http://c1.accurascience.com/miRecords/">http://c1.accurascience.com/miRecords/</a>	Human, rat, mouse, fly, worm, chicken	(136)
miRSel	microRNAs	<a href="https://services.bio.ifi.lmu.de/mirsel/">https://services.bio.ifi.lmu.de/mirsel/</a>	Human, mouse	(137)
miRWalk	microRNAs	<a href="http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/">http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/</a>	Human, mouse	(138)
miRPathDB	microRNAs	<a href="https://mpd.bioinf.uni-sb.de/">https://mpd.bioinf.uni-sb.de/</a>	Human, mouse	(139)
HOCTARdb	microRNAs	<a href="http://hoctar.tigem.it/">http://hoctar.tigem.it/</a>	Human	(140)
miRTar	microRNAs	<a href="http://mirtar.mbc.nctu.edu.tw/human/">http://mirtar.mbc.nctu.edu.tw/human/</a>	Human	(141)
miRDB	microRNAs	<a href="http://www.mirdb.org/">http://www.mirdb.org/</a>	Human, rat, mouse, dog, chicken	(142)
DIANA-LncBase	microRNAs lncRNAs	<a href="http://www.microrna.gr/LncBase">www.microrna.gr/LncBase</a>	Human, mouse	(143)
LncRNA2Target	lncRNAs	<a href="http://123.59.132.21/lncrna2target">http://123.59.132.21/lncrna2target</a>	Human, mouse	(144)
<b>CELL TYPE SPECIFIC RESOURCES</b>				
bloodmiRs	microRNAs	<a href="http://134.245.63.235/ikmb-tools/bloodmiRs/">http://134.245.63.235/ikmb-tools/bloodmiRs/</a>	Human	(145)
ExcellmiRDB	microRNAs	<a href="http://www.excellmirdb.brfjaisalmer.com/%27%">http://www.excellmirdb.brfjaisalmer.com/%27%</a>	Human	(146)
miRandola	microRNAs	<a href="http://mirandola.iit.cnr.it/">http://mirandola.iit.cnr.it/</a>	Human	(147)
miREnvironment	microRNAs	<a href="http://www.cuilab.cn/miren">http://www.cuilab.cn/miren</a>	Human	(148)
HMED	microRNAs	<a href="http://bioinfo.life.hust.edu.cn/smallRNA/">http://bioinfo.life.hust.edu.cn/smallRNA/</a>	Human	(149)
<b>DISEASE-RELATED RESOURCES</b>				
dbDEMOC	microRNAs	<a href="http://www.picb.ac.cn/dbDEMOC/">http://www.picb.ac.cn/dbDEMOC/</a>	Human	(150)
EpimiRBase	microRNAs	<a href="https://www.epimirbase.eu/">https://www.epimirbase.eu/</a>	Human	(151)
HMDD	microRNAs	<a href="http://www.cuilab.cn/hmdd">http://www.cuilab.cn/hmdd</a>	Human	(152)
OncomiRDB	microRNAs	<a href="http://lifeome.net/database/oncomirdb/">http://lifeome.net/database/oncomirdb/</a>	Human	(153)
LncRNADisease	lncRNAs	<a href="http://www.cuilab.cn/lncmadisease">http://www.cuilab.cn/lncmadisease</a>	Human	(154)
Lnc2Cancer	lncRNAs	<a href="http://www.bio-bigdata.com/lnc2cancer/">http://www.bio-bigdata.com/lnc2cancer/</a>	Human	(155)
lncRNASNP	lncRNAs	<a href="http://bioinfo.life.hust.edu.cn/lncRNASNP/">http://bioinfo.life.hust.edu.cn/lncRNASNP/</a>	Human	(156)
LincSNP	lncRNAs	<a href="http://bioinfo.hrbmu.edu.cn/LincSNP/">http://bioinfo.hrbmu.edu.cn/LincSNP/</a>	Human	(157)
<b>OTHER RESOURCES</b>				
EpimiR	microRNAs	<a href="http://210.46.85.180:8080/EpimiR/">http://210.46.85.180:8080/EpimiR/</a>	Human	(158)
MirGeneDB	microRNAs	<a href="http://mirgenedb.org/">http://mirgenedb.org/</a>	All	(25)
miRBaseTracker	microRNAs	<a href="http://mirbasetracker.org/">http://mirbasetracker.org/</a>	All	(127)
mirPub	microRNAs	<a href="http://www.microrna.gr/mirpub/">http://www.microrna.gr/mirpub/</a>	All	(159)
miRNEST	microRNAs	<a href="http://rhesus.amu.edu.pl/mirnest/copy/">http://rhesus.amu.edu.pl/mirnest/copy/</a>	All	(160)
miROrtho	microRNAs	<a href="http://cegg.unige.ch/mirortho">http://cegg.unige.ch/mirortho</a>	All	(161)

different platforms for several species (125). Other databases are available for the study of ncRNAs; nevertheless, these are not specialized in angiogenesis or neovascular processes. Databases and tools used for the study of ncRNAs are listed in Table 3.

We know that endothelial cells are heterogeneous; for instance, they function differently depending on vessel type (162). To uncover the molecular mechanisms controlling this heterogeneity, single-cell RNA sequencing analyses (scRNA-seq) have the potential to enhance our understanding of vascular biology. scRNA-seq is currently being used to study and assess cellular heterogeneity. Particularly with respect to cancer research, this approach has proved to be valuable (163–165); nevertheless, its use in vascular research is just beginning. Recently published studies have mostly focused on protein-coding genes (166, 167). The participation of ncRNAs, along with epigenetic factors, in regulating the metabolic activities of endothelial cells from a single-cell perspective in vascular development and diseases is not yet clear.

## CONCLUDING REMARKS

ncRNAs comprise a new frontier in genetic regulation that has impacts on several research areas. Undoubtedly, the study of angiogenesis and neovascularization has been enhanced through the integration of the study of ncRNAs and epigenetics. Further, ncRNAs are involved in the regulation of several angiogenic targets through epigenetic mechanisms. On the basis of this relationship, a new field of opportunity has emerged in which biomarkers and specific therapies may be identified that can improve the treatment of different vascular diseases and cancers. NGS platforms allow for the global analysis of ncRNA expression and can be used to compare different

physiological and pathological processes. Most of the pathways and mechanisms controlling the ncRNA-mediated regulation of angiogenesis remain unexplored. It is likely that new research strategies implementing an epigenetic perspective will facilitate future discoveries.

## AUTHOR CONTRIBUTIONS

IH-R, LG-C, MS-A, TM-H, and ES-R wrote the manuscript. MS-A and IH-R did the artwork. TM-H and LG-C combined their information to make the tables. All authors contributed to manuscript revision and read and approved the submitted version.

## FUNDING

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT) through the Fondo Sectorial de Investigación en Salud y Seguridad Social (FOSISS, Grant No. 0261181), Fondo CB-SEP-CONACyT (284748), and UAM-PTC-704. ES-R was supported by the Natural Science Department at UAM Cuajimalpa Unit.

## ACKNOWLEDGMENTS

IH-R and TM-H are masters students and LG-C and MS-A are doctoral students from Programa de Maestría y Doctorado en Ciencias Bioquímicas, UNAM and received a fellowship from CONACyT (IH-R: CVU 886138, TM-H: CVU 924685, LG-C: CVU 588391, and MS-A: CVU 659273). MS-A was also a beneficiary of the German Academic Exchange Service (DAAD Grant No. 91693321).

## REFERENCES

- Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. *Nature*. (2011) 473:298–307. doi: 10.1038/nature10144
- Song X, Shan D, Chen J, Jing Q. miRNAs and lncRNAs in vascular injury and remodeling. *Sci China Life Sci.* (2014) 57:826–35. doi: 10.1007/s11427-014-4698-y
- Helkin A, Maier KG, Gahtan V. Thrombospondin-1, -2 and -5 have differential effects on vascular smooth muscle cell physiology. *Biochem Biophys Res Commun.* (2015) 464:1022–7. doi: 10.1016/j.bbrc.2015.07.044
- Xie H, Cui Z, Wang L, Xia Z, Hu Y, Xian L, et al. PDGF-BB secreted by preosteoclasts induces angiogenesis during coupling with osteogenesis. *Nat Med.* (2014) 20:1270–8. doi: 10.1038/nm.3668
- Xu X, Zong Y, Gao Y, Sun X, Zhao H, Luo W, Jia S. VEGF induce vasculogenic mimicry of choroidal melanoma through the PI3k Signal Pathway. *Biomed Res Int.* (2019) 2019:3909102. doi: 10.1155/2019/3909102
- Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature.* (2000) 407:249–57. doi: 10.1038/35025220
- Tímár J, Döme B, Fazekas K, Janovics A, Paku S. Angiogenesis-dependent diseases and angiogenesis therapy. *Pathol Oncol Res.* (2001) 7:85–94. doi: 10.1007/BF03032573
- Bieniasz-Krzywiec P, Martín-Pérez R, Ehling M, García-Caballero M, Pinioti S, Pretto S, et al. Podoplanin-expressing macrophages promote lymphangiogenesis and lymphoinvasion in breast cancer. *Cell Metab.* (2019) 30:1–20. doi: 10.1016/j.cmet.2019.07.015
- Nowak-Sliwinska P, Alitalo K, Allen E, Anisimov A, Aplin AC, Auerbach R, et al. Consensus guidelines for the use and interpretation of angiogenesis assays. *Angiogenesis.* (2018) 21:425–532. doi: 10.1007/s10456-018-9613-x
- Celletti FL, Waugh JM, Amabile PG, Brendolan A, Hilfiker PR, Dake MD. Vascular endothelial growth factor enhances atherosclerotic plaque progression. *Nat Med.* (2001) 7:425–9. doi: 10.1038/86490
- Li M, Liu C, Bin J, Wang Y, Chen J, Xiu J, et al. Mutant hypoxia inducible factor-1 $\alpha$  improves angiogenesis and tissue perfusion in ischemic rabbit skeletal muscle. *Microvasc Res.* (2011) 81:26–33. doi: 10.1016/j.mvr.2010.09.008
- Xue K, Zhao X, Zhang Z, Qiu B, Tan QSW, Ong KH, et al. Sustained delivery of anti-VEGFs from thermogel depots inhibits angiogenesis without the need for multiple injections. *Biomater Sci.* (2019). doi: 10.1039/C9BM01049A. [Epub ahead of print].
- Nicoglou A. Waddington's epigenetics or the pictorial meetings of development and genetics. *Hist Philos Life Sci.* (2018) 40:61. doi: 10.1007/s40656-018-0228-8
- Felsenfeld G. A brief history of epigenetics. *Cold Spring Harb Perspect Biol.* (2014) 6:a018200. doi: 10.1101/cshperspect.a018200
- Richard JLC, Eichhorn PJA. Deciphering the roles of lncRNAs in breast development and disease. *Oncotarget.* (2018) 9:20179–212. doi: 10.18632/oncotarget.24591
- Felsenfeld G, Groudine M. Controlling the double helix. *Nature.* (2003) 421:448–53. doi: 10.1038/nature01411
- Jenuwein T, Allis CD. Translating the histone code. *Science.* (2001) 293:1074–80. doi: 10.1126/science.1063127

18. Biswas S, Rao CM. Epigenetic tools. (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur J Pharmacol.* (2018) 837:8–24. doi: 10.1016/j.ejphar.2018.08.021
19. Iliopoulos D, Lindahl-Allen M, Polytharchou C, Hirsch HA, Tschichl PN, Struhl K. Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Mol Cell.* (2010) 39:761–72. doi: 10.1016/j.molcel.2010.08.013
20. Song Y, Wang R, Li L-W, Liu X, Wang Y-F, Wang Q-X, et al. Long non-coding RNA HOTAIR mediates the switching of histone H3 lysine 27 acetylation to methylation to promote epithelial-to-mesenchymal transition in gastric cancer. *Int J Oncol.* (2019) 54:77–86. doi: 10.3892/ijo.2018.4625
21. Tang W, Zhou W, Xiang L, Wu X, Zhang P, Wang J, et al. The p300/YY1/miR-500a-5p/HDAC2 signalling axis regulates cell proliferation in human colorectal cancer. *Nat Commun.* (2019) 10:663. doi: 10.1038/s41467-018-08225-3
22. Wang F, Tang Z, Shao H, Guo J, Tan T, Dong Y, et al. Long noncoding RNA HOTTIP cooperates with CCCTC-binding factor to coordinate HOXA gene expression. *Biochem Biophys Res Commun.* (2018) 500:852–9. doi: 10.1016/j.bbrc.2018.04.173
23. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* (2004) 116:281–97. doi: 10.1016/S0092-8674(04)00045-5
24. Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet.* (2011) 12:861–74. doi: 10.1038/nrg3074
25. Fromm B, Billipp T, Peck LE, Johansen M, Tarver JE, King BL, et al. A uniform system for the annotation of vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu Rev Genet.* (2015) 49:213–42. doi: 10.1146/annurev-genet-120213-092023
26. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet.* (2010) 11:597–610. doi: 10.1038/nrg2843
27. Lund E, Dahlberg JE. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb Symp Quant Biol.* (2006) 71:59–66. doi: 10.1101/sqb.2006.71.050
28. Catalanotto C, Cogoni C, Zardo G. MicroRNA in control of gene expression: an overview of nuclear functions. *Int J Mol Sci.* (2016) 17:1712. doi: 10.3390/ijms17101712
29. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell.* (2009) 136:215–33. doi: 10.1016/j.cell.2009.01.002
30. Wang X-J, Reyes JL, Chua N-H, Gaasterland T. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.* (2004) 5:R65. doi: 10.1186/gb-2004-5-9-r65
31. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem.* (2010) 79:351–79. doi: 10.1146/annurev-biochem-060308-103103
32. Chen H-X, Xu X-X, Tan B-Z, Zhang Z, Zhou X-D. MicroRNA-29b inhibits angiogenesis by targeting VEGFA through the MAPK/ERK and PI3K/Akt signaling pathways in endometrial carcinoma. *Cell Physiol Biochem.* (2017) 41:933–46. doi: 10.1159/000460510
33. Tzeng H-E, Chang A-C, Tsai C-H, Wang S-W, Tang C-H. Basic fibroblast growth factor promotes VEGF-C-dependent lymphangiogenesis via inhibition of miR-381 in human chondrosarcoma cells. *Oncotarget.* (2016) 7:38566–78. doi: 10.18632/oncotarget.9570
34. Hua Z, Lv Q, Ye W, Wong C-KA, Cai G, Gu D, et al. MiRNA-directed regulation of VEGF and other angiogenic factors under hypoxia. *PLoS ONE.* (2006) 1:e116. doi: 10.1371/journal.pone.0000116
35. Soto-Reyes E, González-Barrios R, Cisneros-Soberanis F, Herrera-Goepfert R, Pérez V, Cantú D, et al. Disruption of CTCF at the miR-125b1 locus in gynecological cancers. *BMC Cancer.* (2012) 12:40. doi: 10.1186/1471-2407-12-40
36. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setién F, et al. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res.* (2007) 67:1424–29. doi: 10.1158/0008-5472.CAN-06-4218
37. Amodio N, Rossi M, Raimondi L, Pitari MR, Botta C, Tagliaferri P, et al. miR-29s: a family of epi-miRNAs with therapeutic implications in hematologic malignancies. *Oncotarget.* (2015) 6:12837–61. doi: 10.18632/oncotarget.3805
38. Chen Z, Wang X, Liu R, Chen L, Yi J, Qi B, et al. KDM4B-mediated epigenetic silencing of miRNA-615-5p augments RAB24 to facilitate malignancy of hepatoma cells. *Oncotarget.* (2016) 8:17712–25. doi: 10.18632/oncotarget.10832
39. Garzon R, Liu S, Fabbri M, Liu Z, Heaphy CEA, Callegari E, et al. MicroRNA-29b induces global DNA hypomethylation and tumor suppressor gene reexpression in acute myeloid leukemia by targeting directly DNMT3A and 3B and indirectly DNMT1. *Blood.* (2009) 113:6411–8. doi: 10.1182/blood-2008-07-170589
40. Malumbres M. miRNAs and cancer: an epigenetics view. *Mol Aspects Med.* (2013) 34:863–74. doi: 10.1016/j.mam.2012.06.005
41. Tuddenham L, Wheeler G, Ntounia-Fousara S, Waters J, Hajhosseini MK, Clark I, et al. The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Lett.* (2006) 580:4214–7. doi: 10.1016/j.febslet.2006.06.080
42. Varambally S, Cao Q, Mani R-S, Shankar S, Wang X, Ateeq B, et al. Genomic Loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science.* (2008) 322:1695–9. doi: 10.1126/science.1165395
43. Verrando P, Capovilla M, Rahmani R. Trans-nonachlor decreases miR-141-3p levels in human melanocytes *in vitro* promoting melanoma cell characteristics and shows a multigenerational impact on miR-8 levels in *Drosophila*. *Toxicology.* (2016) 368–369:129–41. doi: 10.1016/j.tox.2016.09.003
44. Wang X, Wu Q, Xu B, Wang P, Fan W, Cai Y, et al. MiR-124 exerts tumor suppressive functions on the cell proliferation, motility and angiogenesis of bladder cancer by fine-tuning UHRF1. *FEBS J.* (2015) 282:4376–88. doi: 10.1111/febs.13502
45. Volkmann I, Regalla K, Nils P, Jan F, Seema D, Angelika H, et al. MicroRNA-mediated epigenetic silencing of sirtuin1 contributes to impaired angiogenic responses. *Circulation Res.* (2013) 113:997–1003. doi: 10.1161/CIRCRESAHA.113.301702
46. Smits M, Mir SE, Nilsson RJA, van der Stoop PM, Niers JM, Marquez VE, et al. Down-regulation of miR-101 in endothelial cells promotes blood vessel formation through reduced repression of EZH2. *PLoS ONE.* (2011) 6:e16282. doi: 10.1371/journal.pone.0016282
47. Smits M, Nilsson J, Mir SE, van der Stoop PM, Hulleman E, Niers JM, et al. miR-101 is down-regulated in glioblastoma resulting in EZH2-induced proliferation, migration, and angiogenesis. *Oncotarget.* (2010) 1:710–20. doi: 10.18632/oncotarget.205
48. Shehadeh LA, Sharma S, Pessanha M, Wei JQ, Liu J, Yuan H, et al. MicroRNA-20a constrains p300-driven myocardial angiogenic transcription by direct targeting of p300. *PLoS ONE.* (2013) 8:e79133. doi: 10.1371/journal.pone.0079133
49. Sun J, Zheng G, Gu Z, Guo Z. MiR-137 inhibits proliferation and angiogenesis of human glioblastoma cells by targeting EZH2. *J Neurooncol.* (2015) 122:481–9. doi: 10.1007/s11060-015-1753-x
50. Descamps B, Saif J, Benest AV, Biglino G, Bates DO, Chamorro-Jorganes A, et al. BDNF (Brain-Derived Neurotrophic Factor) promotes embryonic stem cells differentiation to endothelial cells via a molecular pathway, including microRNA-214, EZH2. (Enhancer of Zeste Homolog 2), and eNOS. (Endothelial Nitric Oxide Synthase). *Arterioscler Thromb Vasc Biol.* (2018) 38:2117–25. doi: 10.1161/ATVBAHA.118.311400
51. McArthur K, Feng B, Wu Y, Chen S, Chakrabarti S. MicroRNA-200b regulates vascular endothelial growth factor-mediated alterations in diabetic retinopathy. *Diabetes.* (2011) 60:1314–23. doi: 10.2337/db10-1557
52. Potente M, Ghaeni L, Baldessari D, Mostoslavsky R, Rossig L, Dequiedt F, Haendeler J, et al. SIRT1 controls endothelial angiogenic functions during vascular growth. *Genes Dev.* (2007) 21:2644–58. doi: 10.1101/gad.435107
53. Wang C, Wang F, Li Z, Cao Q, Huang L, Chen S. MeCP2-mediated epigenetic regulation in senescent endothelial progenitor cells. *Stem Cell Res Ther.* (2018) 9:87. doi: 10.1186/s13287-018-0828-y
54. Jin Q, He W, Chen L, Yang Y, Shi K, You Z. MicroRNA-101-3p inhibits proliferation in retinoblastoma cells by targeting EZH2 and HDAC9. *Exp Ther Med.* (2018) 16:1663–70. doi: 10.3892/etm.2018.6405
55. Wei X, Xiang T, Ren G, Tan C, Liu R, Xu X, et al. miR-101 is down-regulated by the hepatitis B virus x protein and induces aberrant DNA methylation by targeting DNA methyltransferase 3A. *Cell Signal.* (2013) 25:439–46. doi: 10.1016/j.cellsig.2012.10.013



56. Huang W, Ren Y, Liu H. Vasohibin 1 inhibits Adriamycin resistance in osteosarcoma cells via the protein kinase B signaling pathway. *Oncol Lett.* (2018) 15:5983–8. doi: 10.3892/ol.2018.8074
57. Lu C, Han HD, Mangala LS, Ali-Fehmi R, Newton CS, Ozburn L, et al. Regulation of tumor angiogenesis by EZH2. *Cancer Cell.* (2010) 18:185–97. doi: 10.1016/j.ccr.2010.06.016
58. Luengo-Gil G, Gonzalez-Billalabeitia E, Perez-Henarejos SA, Navarro Manzano E, Chaves-Benito A, Garcia-Martinez E, et al. Angiogenic role of miR-20a in breast cancer. *PLoS ONE.* (2018) 13:e0194638. doi: 10.1371/journal.pone.0194638
59. Zhang W, Chen J-H, Shan T, Aguilera-Barrantes I, Wang L-S, Huang TH-M, et al. miR-137 is a tumor suppressor in endometrial cancer and is repressed by DNA hypermethylation. *Lab Invest.* (2018) 98:1397–407. doi: 10.1038/s41374-018-0092-x
60. Ding F, Zhang S, Gao S, Shang J, Li Y, Cui N, et al. MiR-137 functions as a tumor suppressor in pancreatic cancer by targeting MRGFP. *J Cell Biochem.* (2018) 119:4799–807. doi: 10.1002/jcb.26676
61. Ben Gacem R, Ben Abdelkrim O, Ziadi S, Ben Dhiab M, Trimeche M. Methylation of miR-124a-1, miR-124a-2, and miR-124a-3 genes correlates with aggressive and advanced breast cancer disease. *Tumour Biol.* (2014) 35:4047–56. doi: 10.1007/s13277-013-1530-4
62. Li X, Yang Z, Han Z, Wen Y, Ma Z, Wang Y. Niclosamide acts as a new inhibitor of vasculogenic mimicry in oral cancer through upregulation of miR-124 and downregulation of STAT3. *Oncol Rep.* (2018) 39:827–33. doi: 10.3892/or.2017.6146
63. Gao S-P, Sun H-F, Li L-D, Fu W-Y, Jin W. UHRF1 promotes breast cancer progression by suppressing KLF17 expression by hypermethylating its promoter. *Am J Cancer Res.* (2017) 7:1554–65. doi: 10.1158/1538-7445.SABCS16-P1-08-11
64. Stevens HC, Deng L, Grant JS, Pinel K, Thomas M, Morrell NW, et al. Regulation and function of miR-214 in pulmonary arterial hypertension. *Pulm Circ.* (2016) 6:109–17. doi: 10.1086/685079
65. Namba T, Koike H, Murakami K, Aoki M, Makino H, Hashiya N, et al. Angiogenesis induced by endothelial nitric oxide synthase gene through vascular endothelial growth factor expression in a rat hindlimb ischemia model. *Circulation.* (2003) 108:2250–7. doi: 10.1161/01.CIR.0000093190.53478.78
66. Sinha M, Ghatak S, Roy S, Sen CK. microRNA-200b as a switch for inducible adult angiogenesis. *Antioxid Redox Signal.* (2015) 22:1257–72. doi: 10.1089/ars.2014.6065
67. Adam L, Zhong M, Choi W, Qi W, Nicoloso M, Arora A, et al. miR-200 expression regulates epithelial-to-mesenchymal transition in bladder cancer cells and reverses resistance to epidermal growth factor receptor therapy. *Clin Cancer Res.* (2009) 15:5060–72. doi: 10.1158/1078-0432.CCR-08-2245
68. Li Y, VandenBoom TG, Kong D, Wang Z, Ali S, Philip PA, et al. Up-regulation of miR-200 and let-7 by natural agents leads to the reversal of epithelial-to-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells. *Cancer Res.* (2009) 69:6704–12. doi: 10.1158/0008-5472.CAN-09-1298
69. Mees ST, Mardin WA, Wendel C, Baeumer N, Willscher E, Senninger N, et al. EP300—a miRNA-regulated metastasis suppressor gene in ductal adenocarcinomas of the pancreas. *Int J Cancer.* (2010) 126:114–24. doi: 10.1002/ijc.24695
70. Feng B, Cao Y, Chen S, Chu X, Chu Y, Chakrabarti S. miR-200b mediates endothelial-to-mesenchymal transition in diabetic cardiomyopathy. *Diabetes.* (2016) 65:768–79. doi: 10.2337/db15-1033
71. Arany Z, Huang LE, Eckner R, Bhattacharya S, Jiang C, Goldberg MA, et al. An essential role for p300/CBP in the cellular response to hypoxia. *Proc Natl Acad Sci USA.* (1996) 93:12969–73. doi: 10.1073/pnas.93.23.12969
72. Detmar M, Brown LF, Berse B, Jackman RW, Elicker BM, Dvorak HF, et al. Hypoxia regulates the expression of vascular permeability factor/vascular endothelial growth factor. (VFP/VEGF) and its receptors in human skin. *J Invest Dermatol.* (1997) 108:263–8. doi: 10.1111/1523-1747.ep12286453
73. Liu Y, Cox SR, Morita T, Kourebanas S. Hypoxia regulates vascular endothelial growth factor gene expression in endothelial cells. Identification of a 5' enhancer. *Circ Res.* (1995) 77:638–43. doi: 10.1161/01.RES.77.3.638
74. Al-Mulla F. Vesicular control of metastasis: entrap the “virtuous” and free the “wicked” microRNA. *Transl Cancer Res.* (2017) 6:S999-S1002-S1002. doi: 10.21037/tcr.2017.06.49
75. Teng Y, Ren Y, Hu X, Mu J, Samykutty A, Zhuang X, et al. MVP-mediated exosomal sorting of miR-193a promotes colon cancer progression. *Nat Commun.* (2017) 8:14448. doi: 10.1038/ncomms14448
76. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* (2012) 22:1775–89. doi: 10.1101/gr.132159.111
77. Sun BK, Deaton AM, Lee JT. A Transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol Cell.* (2006) 21:617–28. doi: 10.1016/j.molcel.2006.01.028
78. Mohammad F, Mondal T, Guseva N, Pandey GK, Kanduri C. Kcnq101 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development.* (2010) 137:2493–9. doi: 10.1242/dev.048181
79. Ding X, Jia X, Wang C, Xu J, Gao S-J, Lu C. A DHX9-lncRNA-MDM2 interaction regulates cell invasion and angiogenesis of cervical cancer. *Cell Death Differ.* (2019) 26:1750–65. doi: 10.1038/s41418-018-0242-0
80. Leisegang MS, Fork C, Josipovic I, Richter FM, Preussner J, Hu J, et al. Long Noncoding RNA MANTIS facilitates endothelial angiogenic function. *Circulation.* (2017) 136:65–79. doi: 10.1161/CIRCULATIONAHA.116.026991
81. Miao H, Wang L, Zhan H, Dai J, Chang Y, Wu F, et al. A long noncoding RNA distributed in both nucleus and cytoplasm operates in the PYCARD-regulated apoptosis by coordinating the epigenetic and translational regulation. *PLoS Genet.* (2019) 15:e1008144. doi: 10.1371/journal.pgen.1008144
82. Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. *RNA.* (2015) 21:2007–22. doi: 10.1261/rna.053918.115
83. Brockdorff N. Noncoding RNA and polycomb recruitment. *RNA.* (2013) 19:429–42. doi: 10.1261/rna.037598.112
84. Thomas AA, Feng B, Chakrabarti S. ANRIL: a regulator of VEGF in diabetic retinopathy. *Invest Ophthalmol Vis Sci.* (2017) 58:470–80. doi: 10.1167/iovs.16-20569
85. Neumann P, Jaé N, Knau A, Glaser SF, Fouani Y, Rossbach O, et al. The lncRNA GATA6-AS epigenetically regulates endothelial gene expression via interaction with LOXL2. *Nat Commun.* (2018) 9:237. doi: 10.1038/s41467-017-02431-1
86. Nie F, Sun M, Yang J, Xie M, Xu T, Xia R, et al. Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. *Mol Cancer Ther.* (2015) 14:268–77. doi: 10.1158/1535-7163.MCT-14-0492
87. Congrains A, Kamide K, Oguro R, Yasuda O, Miyata K, Yamamoto E, et al. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. *Atherosclerosis.* (2012) 220:449–55. doi: 10.1016/j.atherosclerosis.2011.11.017
88. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor suppressor gene. *Oncogene.* (2011) 30:1956–62. doi: 10.1038/onc.2010.568
89. Yap KL, Li S, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell.* (2010) 38:662–74. doi: 10.1016/j.molcel.2010.03.021
90. Ruiz MA, Feng B, Chakrabarti S. Polycomb repressive complex 2 regulates MiR-200b in retinal endothelial cells: potential relevance in diabetic retinopathy. *PLoS ONE.* (2015) 10:e0123987. doi: 10.1371/journal.pone.0123987
91. Danan M, Schwartz S, Edelman S, Sorek R. Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* (2012) 40:3131–42. doi: 10.1093/nar/gkr1009
92. Barrett SP, Salzman J. Circular RNAs: analysis, expression and potential functions. *Development.* (2016) 143:1838–47. doi: 10.1242/dev.128074

93. Lei K, Bai H, Wei Z, Xie C, Wang J, Li J, Chen Q. The mechanism and function of circular RNAs in human diseases. *Exp Cell Res.* (2018) 368:147–58. doi: 10.1016/j.yexcr.2018.05.002
94. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, et al. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol.* (2015) 22:256–64. doi: 10.1038/nsmb.2959
95. Ferreira HJ, Davalos V, de Moura MC, Soler M, Perez-Salvia M, Bueno-Costa A, et al. Circular RNA CpG island hypermethylation-associated silencing in human cancer. *Oncotarget.* (2018) 9:29208–19. doi: 10.18632/oncotarget.25673
96. Dou Y, Cha DJ, Franklin JL, Higginbotham JN, Jeppesen DK, Weaver AM, et al. Circular RNAs are down-regulated in KRAS mutant colon cancer cells and can be transferred to exosomes. *Sci Rep.* (2016) 6:37982. doi: 10.1038/srep37982
97. Li X, Yang L, Chen L-L. The biogenesis, functions, and challenges of circular RNAs. *Mol Cell.* (2018) 71:428–42. doi: 10.1016/j.molcel.2018.06.034
98. Dang R-Y, Liu F-L, Li Y. Circular RNA hsa\_circ\_0010729 regulates vascular endothelial cell proliferation and apoptosis by targeting the miR-186/HIF-1 $\alpha$  axis. *Biochem Biophys Res Commun.* (2017) 490:104–110. doi: 10.1016/j.bbrc.2017.05.164
99. Bai Y, Rao H, Chen W, Luo X, Tong C, Qi H. Profiles of circular RNAs in human placenta and their potential roles related to preeclampsia. *Biol Reprod.* (2018) 98:705–12. doi: 10.1093/biolre/iy034
100. Lü L, Sun J, Shi P, Kong W, Xu K, He B, et al. Identification of circular RNAs as a promising new class of diagnostic biomarkers for human breast cancer. *Oncotarget.* (2017) 8:44096–107. doi: 10.18632/oncotarget.17307
101. Qian Z, Liu H, Li M, Shi J, Li N, Zhang Y, et al. Potential diagnostic power of blood circular RNA expression in active pulmonary tuberculosis. *EBioMed.* (2018) 27:18–26. doi: 10.1016/j.ebiom.2017.12.007
102. Yang P, Qiu Z, Jiang Y, Dong L, Yang W, Gu C, et al. Silencing of cZNF292 circular RNA suppresses human glioma tube formation via the Wnt/ $\beta$ -catenin signaling pathway. *Oncotarget.* (2016) 7:63449–55. doi: 10.18632/oncotarget.11523
103. Zhang Y, Zhang X-O, Chen T, Xiang J-F, Yin Q-F, Xing Y-H, et al. Circular intronic long noncoding RNAs. *Mol Cell.* (2013) 51:792–806. doi: 10.1016/j.molcel.2013.08.017
104. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, et al. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell.* (2014) 56:55–66. doi: 10.1016/j.molcel.2014.08.019
105. Schneider T, Hung L-H, Schreiner S, Starke S, Eckhof H, Rossbach O, et al. CircRNA-protein complexes: IMP3 protein component defines subfamily of circRNPs. *Sci Rep.* (2016) 6:31313. doi: 10.1038/srep31313
106. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature.* (2013) 495:333–8. doi: 10.1038/nature11928
107. He R, Liu P, Xie X, Zhou Y, Liao Q, Xiong W, et al. circGFRA1 and GFRA1 act as ceRNAs in triple negative breast cancer by regulating miR-34a. *J Exp Clin Cancer Res.* (2017) 36:145. doi: 10.1186/s13046-017-0614-1
108. Zhong Z, Huang M, Lv M, He Y, Duan C, Zhang L, et al. Circular RNA MYLK as a competing endogenous RNA promotes bladder cancer progression through modulating VEGFA/VEGFR2 signaling pathway. *Cancer Lett.* (2017) 403:305–17. doi: 10.1016/j.canlet.2017.06.027
109. Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* (2010) 6:e1001233. doi: 10.1371/journal.pgen.1001233
110. Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, et al. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun.* (2016) 7:12429. doi: 10.1038/ncomms12429
111. Fang S, Zhang L, Guo J, Niu Y, Wu Y, Li H, et al. NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res.* (2018) 46:D308–14. doi: 10.1093/nar/gkx1107
112. Cheng C-C, Lo H-H, Huang T-S, Cheng Y-C, Chang S-T, Chang S-J, et al. Genetic module and miRNome trait analyses reflect the distinct biological features of endothelial progenitor cells from different anatomic locations. *BMC Genom.* (2012) 13:447. doi: 10.1186/1471-2164-13-447
113. Wang H-W, Lo H-H, Chiu Y-L, Chang S-J, Huang P-H, Liao K-H, et al. Dysregulated miR-361-5p/VEGF axis in the plasma and endothelial progenitor cells of patients with coronary artery disease. *PLOS ONE.* (2014) 9:e98070. doi: 10.1371/journal.pone.0098070
114. Mao Z, Wu F, Shan Y. Identification of key genes and miRNAs associated with carotid atherosclerosis based on mRNA-seq data. *Medicine.* (2018) 97:e9832. doi: 10.1097/MD.0000000000009832
115. Liu X, Wang TT, Li Y, Shi MM, Li HM, Yuan HX, et al. High density lipoprotein from coronary artery disease patients caused abnormal expression of long non-coding RNAs in vascular endothelial cells. *Biochem Biophys Res Commun.* (2017) 487:552–9. doi: 10.1016/j.bbrc.2017.04.082
116. Moreau PR, Örd T, Downes NL, Niskanen H, Bouvy-Liivrand M, Aavik E, et al. Transcriptional profiling of hypoxia-regulated non-coding RNAs in human primary endothelial cells. *Front Cardiovasc Med.* (2018) 5:159. doi: 10.3389/fcvm.2018.00159
117. Chen J, Wang J, Jiang Y, Gu W, Ni B, Sun H, et al. Identification of circular RNAs in human aortic valves. *Gene.* (2018) 642:135–44. doi: 10.1016/j.gene.2017.10.016
118. Bell RD, Long X, Lin M, Bergmann JH, Nanda V, Cowan SL, et al. Identification and initial functional characterization of a human vascular cell-enriched long noncoding RNA. *Arterioscler Thromb Vasc Biol.* (2014) 34:1249–59. doi: 10.1161/ATVBAHA.114.303240
119. Di Salvo TG, Guo Y, Su YR, Clark T, Brittain E, Absi T, et al. Right ventricular long noncoding RNA expression in human heart failure. *Pulm Circ.* (2015) 5:135–61. doi: 10.1086/679721
120. Zhang J, Yuan L, Zhang X, Hamblin MH, Zhu T, Meng F, et al. Altered long non-coding RNA transcriptomic profiles in brain microvascular endothelium after cerebral ischemia. *Exp Neurol.* (2016) 277:162–70. doi: 10.1016/j.expneurol.2015.12.014
121. Li J, Li Q, Chen L, Gao Y, Zhou B, Li J. Competitive endogenous RNA networks: integrated analysis of non-coding RNA and mRNA expression profiles in infantile hemangioma. *Oncotarget.* (2018) 9:11948–63. doi: 10.18632/oncotarget.23946
122. Wolfien M, Brauer DL, Bagnacani A, Wolkenhauer O. Workflow development for the functional characterization of ncRNAs. *Methods Mol Biol.* (2019) 1912:111–32. doi: 10.1007/978-1-4939-8982-9\_5
123. Wang J, Liu K, Liu Y, Lv Q, Zhang F, Wang H. Evaluating the bias of circRNA predictions from total RNA-Seq data. *Oncotarget.* (2017) 8:110914–21. doi: 10.18632/oncotarget.22972
124. Müller R, Weirick T, John D, Militello G, Chen W, Dimmeler S, et al. ANGIOGENES: knowledge database for protein-coding and noncoding RNA genes in endothelial cells. *Sci Rep.* (2016) 6:32475. doi: 10.1038/srep32475
125. Khan S, Taverna F, Rohlenova K, Treps L, Geldhof V, de Rooij L, et al. EndoDB: a database of endothelial cell transcriptomics data. *Nucleic Acids Res.* (2019) 47:D736–44. doi: 10.1093/nar/gky997
126. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* (2019) 47:D155–62. doi: 10.1093/nar/gky1141
127. Van Peer G, Lefever S, Anckaert J, Beckers A, Rihani A, Van Goethem A, et al. miRBase tracker: keeping track of microRNA annotation changes. *Database.* (2014) 2014:bau080. doi: 10.1093/database/bau080
128. Volders P-J, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, et al. LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* (2019) 47:D135–9. doi: 10.1093/nar/gky1031
129. Xu J, Bai J, Zhang X, Lv Y, Gong Y, Liu L, et al. A comprehensive overview of lncRNA annotation resources. *Brief Bioinform.* (2017) 18:236–49. doi: 10.1093/bib/bbw015
130. Sethupathy P, Corda B, Hatzigeorgiou AG. TarBase: a comprehensive database of experimentally supported animal microRNA targets. *RNA.* (2006) 12:192–7. doi: 10.1261/rna.2239606
131. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* (2018) 46:D296–302. doi: 10.1093/nar/gkx1067

132. Andrés-León E, González Peña D, Gómez-López G, Pisano DG. miRGate: a curated database of human, mouse and rat miRNA-mRNA targets. *Database*. (2015) 2015:bav035. doi: 10.1093/database/bav035
133. Bruno AE, Li L, Kalabus JL, Pan Y, Yu A, Hu Z. miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes. *BMC Genom.* (2012) 13:44. doi: 10.1186/1471-2164-13-44
134. Bonome T, Levine DA, Shih J, Randonovich M, Pise-Masison CA, Bogomolny F, et al. A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer. *Cancer Res.* (2008) 68:5478–86. doi: 10.1158/0008-5472.CAN-07-6595
135. Wu W-S, Tu B-W, Chen T-T, Hou S-W, Tseng JT. CSmiRTar: Condition-specific microRNA targets database. *PLoS ONE.* (2017) 12:e0181231. doi: 10.1371/journal.pone.0181231
136. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.* (2009) 37:D105–110. doi: 10.1093/nar/gkn851
137. Naeem H, Küffner R, Csaba G, Zimmer R. miRsel: automated extraction of associations between microRNAs and genes from the biomedical literature. *BMC Bioinform.* (2010) 11:135. doi: 10.1186/1471-2105-11-135
138. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat Methods.* (2015) 12:697. doi: 10.1038/nmeth.3485
139. Backes C, Kehl T, Stöckel D, Fehlmann T, Schneider L, Meese E, et al. miRPathDB: a new dictionary on microRNAs and target pathways. *Nucleic Acids Res.* (2017) 45:D90–6. doi: 10.1093/nar/gkw926
140. Gennarino VA, Sardiello M, Mutarelli M, Dharmalingam G, Maselli V, Lago G, et al. HOCTAR database: a unique resource for microRNA target prediction. *Gene.* (2011) 480:51–8. doi: 10.1016/j.gene.2011.03.005
141. Hsu JB-K, Chiu C-M, Hsu S-D, Huang W-Y, Chien C-H, Lee T-Y, et al. miRTar: an integrated system for identifying miRNA-target interactions in human. *BMC Bioinform.* (2011) 12:300. doi: 10.1186/1471-2105-12-300
142. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 43:D146–52. doi: 10.1093/nar/gku1104
143. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. *Nucleic Acids Res.* (2016) 44:D231–8. doi: 10.1093/nar/gkv1270
144. Cheng L, Wang P, Tian R, Wang S, Guo Q, Luo M, et al. LncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res.* (2019) 47:D140–4. doi: 10.1093/nar/gky1051
145. Juzenas S, Venkatesh G, Hübenal M, Hoepfner MP, Du ZG, Paulsen M, et al. A comprehensive, cell specific microRNA catalogue of human peripheral blood. *Nucleic Acids Res.* (2017) 45:9290–301. doi: 10.1093/nar/gkx706
146. Barupal JK, Saini AK, Chand T, Meena A, Beniwal S, Suthar JR, et al. ExcellmiRDB for translational genomics: a curated online resource for extracellular microRNAs. *OMICS.* (2015) 19:24–30. doi: 10.1089/omi.2014.0106
147. Russo F, Bella SD, Nigita G, Macca V, Laganà A, Giugno R, et al. miRandola: extracellular circulating MicroRNAs database. *PLOS ONE.* (2012) 7:e47786. doi: 10.1371/journal.pone.0047786
148. Yang Q, Qiu C, Yang J, Wu Q, Cui Q. miREnvironment database: providing a bridge for microRNAs, environmental factors and phenotypes. *Bioinformatics.* (2011) 27:3329–30. doi: 10.1093/bioinformatics/btr556
149. Gong J, Wu Y, Zhang X, Liao Y, Sibanda VL, Liu W, et al. Comprehensive analysis of human small RNA sequencing data provides insights into expression profiles and miRNA editing. *RNA Biol.* (2014) 11:1375–85. doi: 10.1080/15476286.2014.996465
150. Yang Z, Wu L, Wang A, Tang W, Zhao Y, Zhao H, Teschendorff AE. dbDEMOC 2.0: updated database of differentially expressed miRNAs in human cancers. *Nucleic Acids Res.* (2017) 45:D812–8. doi: 10.1093/nar/gkw1079
151. Mooney C, Becker BA, Raoof R, Henshall DC. EpimiRBase: a comprehensive database of microRNA-epilepsy associations. *Bioinformatics.* (2016) 32:1436–8. doi: 10.1093/bioinformatics/btw008
152. Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, et al. HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.* (2019) 47:D1013–7. doi: 10.1093/nar/gky1010
153. Wang D, Gu J, Wang T, Ding Z. OncomiRDB: a database for the experimentally verified oncogenic and tumor-suppressive microRNAs. *Bioinformatics.* (2014) 30:2237–8. doi: 10.1093/bioinformatics/btu155
154. Bao Z, Yang Z, Huang Z, Zhou Y, Cui Q, Dong D. LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.* (2019) 47:D1034–7. doi: 10.1093/nar/gky905
155. Gao Y, Wang P, Wang Y, Ma X, Zhi H, Zhou D, et al. Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. *Nucleic Acids Res.* (2019) 47:D1028–33. doi: 10.1093/nar/gky1096
156. Miao Y-R, Liu W, Zhang Q, Guo A-Y. lncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. *Nucleic Acids Res.* (2018) 46:D276–80. doi: 10.1093/nar/gkx1004
157. Ning S, Yue M, Wang P, Liu Y, Zhi H, Zhang Y, et al. LincSNP 2.0: an updated database for linking disease-associated SNPs to human long non-coding RNAs and their TFBSs. *Nucleic Acids Res.* (2017) 45:D74–8. doi: 10.1093/nar/gkw945
158. Dai E, Yu X, Zhang Y, Meng F, Wang S, Liu X, et al. EpimiR: a database of curated mutual regulation between miRNAs and epigenetic modifications. *Database.* (2014) 2014:bau023. doi: 10.1093/database/bau023
159. Vergoulis T, Kanellos I, Kostoulas N, Georgakilas G, Sellis T, Hatzigeorgiou A, et al. mirPub: a database for searching microRNA publications. *Bioinformatics.* (2015) 31:1502–4. doi: 10.1093/bioinformatics/btu819
160. Szcześniak MW, Makiłowska I. miRNEST 2.0: a database of plant and animal microRNAs. *Nucleic Acids Res.* (2014) 42:D74–7. doi: 10.1093/nar/gkt1156
161. Gerlach D, Kriventseva EV, Rahman N, Vejnar CE, Zdobnov EM. miOrtho: computational survey of microRNA genes. *Nucleic Acids Res.* (2009) 37:D111–7. doi: 10.1093/nar/gkn707
162. Potente M, Mäkinen T. Vascular heterogeneity and specialization in development and disease. *Nat Rev Mol Cell Biol.* (2017) 18:477–94. doi: 10.1038/nrm.2017.36
163. Liu SJ, Nowakowski TJ, Pollen AA, Lui JH, Horlbeck MA, Attenello FJ, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* (2016) 17:67. doi: 10.1186/s13059-016-0932-1
164. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell.* (2019) 177:1330–45.e18. doi: 10.1016/j.cell.2019.03.005
165. Xiao Z, Dai Z, Locasale JW. Metabolic landscape of the tumor microenvironment at single cell resolution. *Nat Commun.* (2019) 10:3763. doi: 10.1038/s41467-019-11738-0
166. Lukowski SW, Patel J, Andersen SB, Sim S-L, Wong HY, Tay J, et al. Single-cell transcriptional profiling of aortic endothelium identifies a hierarchy from endovascular progenitors to differentiated cells. *Cell Rep.* (2019) 27:2748–58.e3. doi: 10.1016/j.celrep.2019.04.102
167. Zhao Q, Eichten A, Parveen AA, Adler C, Huang Y, Wang W, et al. Single-cell transcriptome analyses reveal endothelial cell heterogeneity in tumors and changes following anti-angiogenic treatment. *Cancer Res.* (2018) 78:2370–82. doi: 10.1158/0008-5472.CAN-17-2728

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

Copyright © 2019 Hernández-Romero, Guerra-Calderas, Salgado-Albarrán, Maldonado-Huerta and Soto-Reyes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

**Cancer Genetics**  
**A new kid is on the block: when CTCF meets demethylases.**  
 --Manuscript Draft--

<b>Manuscript Number:</b>	
<b>Article Type:</b>	Review Article
<b>Section/Category:</b>	Other areas not listed elsewhere
<b>Keywords:</b>	CTCF Histone demethylases DNA demethylases TET KDM.
<b>Corresponding Author:</b>	Ernesto Soto-Reyes, Ph.D. Universidad Autonoma Metropolitana Cuajimalpa Mexico, MEXICO
<b>First Author:</b>	Lissania Guerra-Calderas, M. Sc.
<b>Order of Authors:</b>	Lissania Guerra-Calderas, M. Sc. Itzel Hernández-Romero Marisol Salgado-Albarrán, M. Sc. Talia Wegman-Ostrosky, Ph.D. Ernesto Soto-Reyes, Ph.D.
<b>Abstract:</b>	<p>&lt;p class="Newparagraph" style="margin: 0cm 0cm 0.0001pt; text-indent: 36pt; line-height: 32px; font-size: 12pt; font-family: "Times New Roman", serif; text-align: justify;"&gt;&lt;span lang="EN-US" style="font-family: Arial, sans-serif;"&gt;Architectural proteins are essential epigenetic regulators of the chromatin loop assemble and chromosomal arrangement in the nucleus. The displacement or deficit in the levels of proteins like CTCF, the master weaver of the genome, leads to serious embryonic defects and the development of diseases such as cancer. Growing evidence point out the fact that the nuclear factor CTCF, interplays with multiple protein partners, as well as its isoforms, its paralogous protein BORIS and other chromatin-modifying elements related to the histone and DNA demethylation. However, the mechanisms and outcomes underlying these associations are still not clear. In this review, we will focus on the interactions that have been reported between CTCF and DNA and histone demethylases as new protein partners, and we will also discuss how CTCF can influence their activity and recruitment to DNA.&lt;/span&gt;&lt;/p&gt;&lt;/p&gt;</p>
<b>Suggested Reviewers:</b>	Lucia Altucci lucia.altucci@unicampania.it  Amaia Lujambio amaia.lujambio@mssm.edu  Sónia Melo smelo@ipatimup.pt
<b>Opposed Reviewers:</b>	

**A new kid is on the block: when CTCF meets demethylases.**

Lissania Guerra-Calderas<sup>1</sup>, Itzel Alejandra Hernández-Romero<sup>1</sup>, Marisol Salgado-Albarrán<sup>1,2</sup>, Talia Wegman-Ostrosky<sup>3</sup> and Ernesto Soto-Reyes<sup>1\*</sup>.

*<sup>1</sup>Natural Sciences Department, Universidad Autónoma Metropolitana-Cuajimalpa (UAM-C), México City 05300, México.*

*<sup>2</sup>Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University of Munich, Munich, Germany*

*<sup>3</sup>Dirección de Investigación, Instituto Nacional de Cancerología, Mexico City, Mexico,*

\*Corresponding author: Ernesto Soto-Reyes. UAM-C, Av. Vasco de Quiroga 4871, Col. Santa Fe, Del. Cuajimalpa de Morelos. 05348, CDMX, México.

Email: esotoreyes@cua.uam.mx.

**Abstract.**

Architectural proteins are essential epigenetic regulators of the chromatin loop assemble and chromosomal arrangement in the nucleus. The displacement or deficit in the levels of proteins like CTCF, the master weaver of the genome, leads to serious embryonic defects and the development of diseases such as cancer. Growing evidence point out the fact that the nuclear factor CTCF, interplays with multiple protein partners, as well as its isoforms, its paralogous protein BORIS and other chromatin-modifying elements related to the histone and DNA demethylation. However, the mechanisms and outcomes underlying these associations are still not clear. In this review, we will focus on the interactions that have been reported between CTCF and DNA and histone demethylases as new protein partners, and we will also discuss how CTCF can influence their activity and recruitment to DNA.

Keywords: CTCF, Histone demethylases, DNA demethylases, TET, KDM.

**Introduction.**

Chromatin organization and gene expression within the nucleus are coordinated by different epigenetic mechanisms that act in response to a dynamic remodeling code. Although, part of the effort of many research groups has been focused on the characterization and description of the epigenetic processes that take place in the cell. However, the information about how these epigenetic components are recruited to specific regions of the genome is limited. Some of the proteins with epigenetic function lack DNA binding domains, thus the mechanisms by which they are recruited to specific sites are poorly understood.

Architectural proteins, such as CTCF, are the main actors in the recruitment of different protein complexes to chromatin [1], acting as polyvalent transcription factors [2]. The nuclear factor CTCF was originally characterized as a transcriptional repressor [3]; however, later on, it was also found to act as a transcriptional activator [4,5]. CTCF can bind to tens of thousands of genomic sites, some of them are tissue-specific, and about 5,000 sites are ultra-conserved among the species and tissues. The most conserved sites are usually located at chromatin domain boundaries regulatory regions as enhancers and gene promoters others lays within gene bodies [6–8]. When bound to these sites along the genome, CTCF can depict several functions, such as insulator [9], transcriptional pausing [10] and formation of chromatin loops in conjunction with cohesin, among others [11]. One of the interrogators that remains open is how can CTCF achieve different functions? The response to this question may lie in the context-dependent interactions of CTCF with several protein partners. As an example, CTCF can function as scaffold protein for tissue-specific transcriptional activators, repressors, essential proteins such as cohesin, RNA polymerase II and chromatin remodeling enzymes, like histone deacetylases [10,12–16]. Recently, it has been reported that CTCF can interact with proteins that participate in the demethylation of DNA and histones, opening a wide range of CTCF functions similar to the swiss army knife [17–20].

Histone and DNA demethylase enzymes remove the methyl group attached covalently to DNA cytosines or lysins and arginines found in histone proteins. It was long thought that DNA and histone methylation was a stable and irreversible modification, but in 2004 the first histone lysine demethylase was characterized and in 2009 the first DNA demethylase was discovered. To date,

about ~20 histone demethylases and 3 DNA demethylases have been described [21,22]. *In vitro* studies have shown that the simple binding of these enzymes to their substrate is enough for a demethylation reaction, suggesting that these enzymes must be finely controlled by regulatory mechanisms in order to prevent aberrant demethylation [23–25]. This regulation can be carried out at different stages like controlling their expression levels, their catalytic activity and it has also been suggested that the recruitment mechanisms of demethylases have a critical role in regulating these enzymes, however, it is not yet clearly understood how the demethylases are directed towards specific sites in chromatin avoiding aberrant demethylation [25]. One of the current proposals is that certain transcriptional factors might be responsible for recruiting these epigenetic components such as histone and DNA demethylases, where one of these factors is CTCF. In this review, we will address the participation of histones and DNA demethylases, CTCF isoforms and paralog proteins as new protein partners.

#### **The protagonist: the architectonic factor CTCF.**

The architectural protein CTCF is constituted by 727 amino acids and contains a DNA-binding domain that is composed by 11 Zinc Fingers (ZF), flanked by 10 amino acids at the C-terminal region and 267 amino acids at the N-terminal region [26].

It was found that CTCF binds to three spaced repetitions of the sequence the CCCTC, located upstream of the chicken *c-myc* promoter and therefore was named CCCTC binding factor (CTCF) [3]. The CTCF binding motif contains a 12-to-15 bp consensus sequence 5'-NCA-NNA-G(G/A)N-GGC-(G/A)(C/G)(T/C)-3' [27]. This binding sequence is recognized by CTCF using a combination of its 11 ZF motifs. CTCF binds to about 55,000 DNA sites in different cell types and in



total 77,811 distinct binding sites across all the cell lineages. Furthermore, CTCF can interact with itself and with cohesin to form chromatin loop structure, where the amino-terminal region, the ZF1 and ZF2 of CTCF are indispensable for the retention of cohesin. (Figure 1A) [11,28]. These types of interactions are highly dynamic, where DNA methylation promotes the dissociation of CTCF [29,30]. CTCF's ability to bind to several sequences is due to the use of different combinations of its ZFs, because this peculiar characteristic, CTCF deserve the status of a "multivalent protein" [4].

#### **CTCF isoforms and CTCFL (BORIS) the paralog of CTCF.**

The majority of the previous reports have been focused on the study of the CTCF canonical isoform, but a recent study in 2019, revealed that a short isoform of CTCF (CTCF-s) that lacks exons three and four (corresponding to ZF1 to ZF3), can compete with the canonical CTCF binding site. Also, the binding of CTCF-s, disrupts CTCF-cohesin binding leading to disturbance in chromatin loop formation and disrupting the long-range interaction between enhancers and promoters (Figure 1B) [31]. A question that remains open is whether the lack of ZF1 to ZF3 in the CTCF-s, affects its interaction with multiple proteins (Figure 1B).

Similarly, *CTCF* has a paralog gene called *CTCF-Like (CTCFL)*, that encodes a protein named BORIS (Brother of the Regulator of Imprinted Sites). *CTCF* gene was proposed to have duplicated in early mammals, originating *CTCFL*. This gene is normally expressed in testis and is overexpressed in some cancer types, such as ovarian cancer [32,33]. CTCF and BORIS share a similar 11 ZF domain, which has as DNA binding capabilities, but different N- and C-terminal domains. Due to their protein similarity, CTCF and BORIS can potentially

bind to DNA with a similar sequence [34]; but unlike CTCF, which binds to Low Methylated Regions (LMRs), BORIS binds to methylated DNA, suggesting that BORIS may occupy DNA sites that CTCF is unable to bind <sup>35</sup>. This proposes that CTCF and BORIS compete for a subset of the genomic sites in cancer cells [35]. In this regard, it has been reported that CTCF and BORIS homodimer and heterodimer occupancy at clustered CTCF binding motifs affects transcriptional regulation and, potentially, the establishment of chromatin loops, since BORIS alone is insufficient to recruit the cohesin complex, which is indispensable for the chromatin loop formation [35,36]. In contrast, a recent report shows that BORIS may mediate chromatin loops, which in turn leads to the recognition of super-enhancers that drive the ectopic expression of several genes [37].

Together, the data above mentioned suggests that despite that CTCF and BORIS have a very similar DNA binding sequence, DNA methylation status of the consensus binding region could significantly affect which protein may be recruited (Figure 1C).

Due to its multivalence properties and its great plasticity to bind countless sequences, identifying and understanding the elements that lead to the different functions of CTCF has been incredibly challenging. CTCF can be involved in multiple gene regulation processes, such as repression, activation, insulating function, alternative splicing and gene printing among others, due to its association with other proteins, new CTCF functions will be discovered. These interactions, in particular with DNA and histone demethylases, will be discussed in the following sections.

### **CTCF functions.**

CTCF depicts different functions depending on the location of the binding site and the epigenetic context. Among its known functions are the 3D chromatin organization, imprinting regulation, insulator protein and mRNA splicing. Regarding the 3D chromatin structure, CTCF helps the establishment of chromatin loops and also anchors DNA to cellular structures like the nuclear lamina [38]. Also, CTCF defines the boundaries between euchromatic and heterochromatic DNA. In genomic imprinting, CTCF has an essential role in controlling the *IGF2* (Insulin like Growth Factor 2) and the *H19* genes [39,40]. Locally, CTCF can act as a protein insulator by blocking the contact between enhancers and promoters, thus controlling the activity of enhancers of certain domains. At genomic level, CTCF can participate in the formation of Topologically Associating Domains (TADs) and also may act as a chromatin barrier by stopping the spread of heterochromatic structures [41,42].

In addition to controlling the 3D chromatin structure and functioning as an insulator, CTCF is also involved in alternative splicing, where CTCF bound downstream of alternative exons promotes RNA Polymerase II (RNAPII) pausing, allowing the required temporal context for co-transcriptional spliceosome assembly at weak upstream splice sites [10].

Part of the remaining questions is: How can CTCF achieve countless functions, which are frequently seemingly unconnected. The response to this query may lie in the context-dependent interactions of CTCF with several protein partners.

#### **CTCF and its partners: protein-protein interactions.**

It is still unclear what determines which protein partner of CTCF is chosen for each function, but numerous possibilities can be considered. One suggested

mechanism is the combinatorial use of the 11 ZF to recognize and bind to a variety of DNA sequences [43]; for instance, a study in 2013 revealed that ZF 4–7 functions as an anchor to bind around eighty percent of the target sequences that have the central motif, this study also found that unconserved flanking sequences can be recognized by groups ZF 1–2 and ZF 8–11, which helps to stabilize the CTCF-DNA binding [44], as a consequence, the remaining free (unbound) ZFs and the amino and carboxyl-terminal regions might serve as a platform for interaction with other proteins (Figure 1B). A second probable mechanism are the post-translational modifications of CTCF and/or the partner's, which could be used as a protein code to a wide range of proteins that depict different functions in different cellular conditions, where phosphorylation of CTCF alters its interaction with the partner protein, such as RNAPII [13].

CTCF physically binds to itself forming homodimers which cause the bound DNA to form loops. This homodimerization is ZF-dependent, whereas the association with other proteins depends on the remaining ZFs as well as the amino and carboxyl termini regions (Figure 1A) [45–47].

Another well-described interaction is the association of cohesin to several of the CTCF binding sites. In this case, CTCF serves as a guide or barricade protein to help position a slipping cohesin molecule on the chromatin template [16,52].

CTCF can accomplish several functions by sharing chromatin binding sites and interacting directly or indirectly with many other proteins [7,47,53], among these proteins are chromatin-associated proteins, like H2A.Z and Cohesins, DNA binding proteins, like YY1 and Kaiso, miscellaneous proteins, like Lamin A and RNA Polymerase II (RNAPII), and multifunctional proteins, like Poly (ADP-

Ribose) Polymerase 1 (PARP1) enzyme, IPO13 and Nucleophosmine, but also CTCF can interact with its paralog protein, BORIS among several other proteins, described in Table 1 (Figure 2A) [13,16,36,46,54–59].

Taken together, the previous data show that CTCF can interact with a wide range of proteins; furthermore, the composition of these CTCF-protein complexes leads to a different effect on gene transcription on a specific site of the genome. Nevertheless, the interaction of CTCF with DNA and histones demethylases has been poorly addressed, but some pieces of evidence show that CTCF can interact with DNA and histone demethylases increasing its biological functions.

#### **The histone lysine demethylases.**

The methylation of histones can happen on arginine and lysine residues found in the N-terminal region of histones. Arginine residues can be mono- or dimethylated, this chemical modification is generally associated with transcriptional activation [76,77]. In contrast, lysine residues can be mono-, di- or trimethylated and are associated with either transcriptional activation or repression, depending on the level of methylation and lysine location [78][78]. As an example, di- and trimethylation at H3K4 is related to enhanced gene expression; whereas trimethylation at H3K9 and H3K27 is associated to transcriptional repression [79].

Previously, it was long thought that histone methylation was stable and irreversible, since the half-lives of the canonical and methylated histones are very similar; however, the search for proteins with the ability to demethylate methylated lysine residues began in the 1960s and until now there have been characterized more than 20 enzymes capable of removing this covalent modification [80].

In 1964, the Paik group reported an enzyme that could demethylate mono and dimethylated lysines [81]. Years later the same research group partially purified a protein that could demethylate histone residues [82,83]. However, the molecular identity of this enzyme remained unclear for several decades. It was not until 2004 that the Yang Shi group identified and characterized the first histone lysine demethylase, LSD1 (Lysine Specific Demethylase 1A) [21], later renamed KDM1A [84]. This enzyme removes the methyl group from lysins 4 and 9 of histone 3 (H3K4me<sub>2/1</sub>, H3K9me<sub>2/1</sub>) [21,85] by an oxidation reaction, proposing that KDM1A is relevant in the dynamism of posttranslational modifications of histones and, therefore, in gene transcription [21,86].

Currently, the methylation of the lysins found in the histones can be actively removed by true histone demethylases that are classified depending on the chemical mechanism of demethylation, into two families: the amine oxidase and the oxygen, iron- and  $\alpha$ -ketoglutarate dependent proteins [80].

The amino oxidase like family is composed by two members: KDM1A, the first histone lysine demethylase described by Shi and cols. in 2004 [21] and the KDM1B enzyme.

KDM1A demethylates H3K4me<sub>2</sub>, H3K4me<sub>1</sub>, H3K9me<sub>2</sub>, and H3K9me<sub>1</sub>; therefore, it has a dual role as co-activator or as co-repressor. Additionally to demethylating histone proteins, KDM1A also demethylates several non-histone proteins like AGO2, DNMT1, E2F1, ER $\alpha$ , HIF-1 $\alpha$ , HSP90, MEFD2, MYPT1, p53, RB1 and STAT3 [87–90]. The non-histone demethylation activity of KDM1A has different roles for example, when KDM1A demethylates STAT3 and p53 it enhances their functional activity but it does not change their expression levels, KDM1A stabilizes HIF-1, MEFD2 and E2F1, and on the contrary, this

demethylase destabilizes MYPT1 [91]. The fact that KDM1A can demethylate several non-histone substrates strongly suggests that this enzyme has participation in different cellular processes like cell cycle progression and apoptosis [92].

Regarding KDM1B, it demethylates H3K4me2 and H3K4me1, although non-histone substrates for KDM1B have not yet been described. A study in 2015 found that KDM1B possesses a surprising ubiquitin ligase E3 activity. This enzyme reduces lung cancer cell growth through the promotion of ubiquitylation and subsequent degradation of OGT. This study redefined the functions of KDM1B and uncovered a relationship among histone demethylation and ubiquitin-dependent pathway [93]. Nevertheless, further research is required in order to define the non-histone substrates and new roles for KDM1B [92].

The oxidase family, have a common amine oxidase-like domain and are FAD-dependent enzymes. This family can remove mono- and dimethyl groups. The catalytic site of these enzymes is located in the interphase of the FAD-binding domain and the amine oxidase-like domain [94]. The amine oxidases cannot demethylate trimethylated lysine, due to the FAD-dependent catalytic mechanism [95,96][95,96].

The oxygenase family is the largest one, more than 20 JmjC (Jumonji) domain-containing enzymes have been described, and these are subclassified depending on the substrate that they demethylate. Like the oxidase family, non-histone substrates have been identified for some oxygenase enzymes, for example KDM2A has been reported to demethylate the p65 subunit of the NF- $\kappa$ B factor, this active demethylation inhibits NF- $\kappa$ B activity [97]. KDM3A has shown to be able to remove the methyl group of p53-K372me1 and reduce its

transcriptional activity [78]. Similarly, a study demonstrated that KDM4A, KDM4B and KDM4C are able to demethylate *in vitro* non-histone peptides from CDYL1, CSB, G9a and WIZ proteins, proposing that these enzymes could possess an extra demethylation role of non-histone substrates *in vivo* [98]. Several JmjC demethylases have many potentially non-histone substrates and it is important to further investigate other non-histone-mediated pathways.

Unlike the oxidase family, the oxygen, iron- and  $\alpha$ -ketoglutarate dependent proteins act by a different reaction mechanism, this mechanism allows the removal of a trimethylated modification, which cannot be performed by KDM1A and KDM1B [94][99,100]. The  $\alpha$ -ketoglutarate and iron- dependant family also called as Jumonji-type demethylases, belong to the cupin family of oxygenases, these proteins enclose an  $\text{Fe}^{2+}$  ion in their catalytic domain, and a they also use  $\alpha$ -ketoglutarate as co-substrate [94].

*In vitro* studies have shown that the simple interaction of histone demethylases to their substrate is enough for the demethylation reaction to be carried out [23,24], suggesting that these enzymes must be finely controlled by regulatory mechanisms *in vivo* in order to prevent aberrant demethylation [25].

#### **Histone Lysine Demethylase Regulation.**

The group of Shi and colleagues proposes that the activity of demethylases is regulated in a modular and step-wise way, involving protein-protein interactions with DNA-binding factors, as well as other chromatin-remodeling enzymes, the recognition chromatin environment by extra non-enzymatic 'reader' domains that are present in the demethylases and/or their related proteins [25,101].



One mechanism of the regulation of KDMs is at gene expression level. In fact, many demethylases present restricted patterns of embryonic and adult expression [25,102,103].

Another mechanism for the regulation of demethylases is the control of their catalytic activity [104–106]; the binding of several DNA-binding transcription factors to these enzymes [85,107–109] and its post-translational modifications can diminish the demethylases specificity for their substrates, as well as their catalytic activity [41,110].

Also, endogenous inhibitors like fumarate, oxaloacetate and succinate, as well as exogenous inhibitors like 3,5-dimethoxyphenylacrylic acid, N-oxalylglycine, QC6352 and PKF118-310, can affect the catalytic activity of the histone demethylases enzymes [111–115].

Moreover, it has been suggested that the chromatin environment provides certain selectivity to demethylases since it controls the accessibility of these proteins to their target sites (Lan et al., 2008). And thus, the mechanisms of demethylases recruitment to their target sites have a critical role in their regulation; however, it is still not clearly understood how the demethylases are directed towards specific chromatin sites and avoid aberrant demethylation [25].

Further, transcriptional factors have been proposed to play an important role in mediating the recruitment of different epigenetic components such as histone demethylases. One of the factors that have been associated is the multivalent protein CTCF.

#### **Lysine demethylases as new partners of CTCF.**

Once in the cell, the KDMs activity needs to be efficiently directed to the target chromatin. This targeting implies both the recognition of chromatin status

by additional domains, "non-enzymatic readers" [101] present in demethylases and/or their associated proteins, as well as possible associations with non-coding RNAs, and also the specific recognition of a given locus (for example, of the promoter elements). In this sense, the majority of the histone lysine demethylases lack a DNA binding domain thus the recognition of a specific locus is mediated by protein-protein interactions with DNA binding factors and other chromatin-modifying enzymes. These inputs let a network-based evaluation of the chromatin status and permit demethylases to function as bright tuners for methylation state [25]. Until now, few studies have demonstrated the association between CTCF and histone demethylases; only two KDMs partners have been found. The first one was reported in 2014 by the Polyak group, they found by co-immunoprecipitation, that CTCF formed a complex with the H3K4me3 and H3K4me2 specific demethylase KDM5B [116]. Also, when conducting ChIP-seq assays, they discovered that KDM5B sites overlap with those of CTCF in most mammary cancer cell lines, and this overlapping phenomenon correlates with a lower H3K4me3 signal compared to those non-overlapping sites [18]. This suggests that CTCF takes part in this fine-tuned regulation of KDM5B and the regulation of basal/stem cell genes, like *ACTG2*, *APOE*, *CTGF*, *FN1* and *TGFβ2*, among others. (Figure 2B).

Another CTCF histone demethylase partner is KDM4A. The first clue that CTCF could be a KDM4A partner was reported in 2011 by the Kang group, they performed transfection and immunofluorescence assays and observed that the demethylation frequency of KDM4A was enhanced by the presence of CTCF [117]. This study opened the window to another report in 2018, where our group observed by co-immunoprecipitation, that CTCF and KDM4A form a protein

complex; furthermore, they showed by ChIP-qPCR and ChIP-Re/ChIP-qPCR, that CTCF and KDM4A coexist in the first intron of *CHD5*, the promoter of *WRAP53*, and the region located at -1922 bp of the *ASCL2* TSS (Figure 2A). The coexistence of CTCF and KDM4A correlates with the reduction of H3K36me3/2 histone modifications at the first intron of *CHD5* and is associated with its transcriptional downregulation. Furthermore, the CTCF or KDM4A depletion mediated by siRNA, leads the *CHD5* reactivation expression, proposing that both proteins are involved in the negative regulation of this gene. Finally, they showed that the knockout of KDM4A by CRISPR/Cas9 restored the expression of *CHD5* and H3K36me3 and H3K36me2 histone marks, without disturbing CTCF localization [20] (Figure 2B). Nevertheless, currently it is still unknown if this complex is related to repression or activation of genome wide and if CTCF might be also one of the key proteins driving the specificity of KDM4A.

To our knowledge, there are no studies evaluating the association between CTCF and other histone demethylases. Nevertheless, ChIP-seq studies show overlapping between KDM5A, KDM5C, KDM1A and CTCF, suggesting that CTCF could be involved in their regulation; however further studies are required to determine which is the participation of CTCF in the modulation of these enzymes.

### **The TET enzymes.**

DNA methylation is an epigenetic process involving a methyl group transfer to the C5 position of the cytosine to form 5-methylcytosine (5mC) [118]. 5mC can be found nearly completely within CpG dinucleotides in mammalian somatic cells [119]. About 70–80% of cytosines in CpG islands is methylated on both strands in human somatic cells. In general, CpG methylation is very

predominant in repetitive sequences and in gene bodies, but infrequent at CpG islands within housekeeping promoters [119].

DNA methylation can have different functions; for example, the methylation of promoters and enhancers is negatively correlated with transcription [119,120]. DNA methylation is also involved in numerous other essential processes, such as genomic imprinting, X chromosome inactivation, retrotransposon elements suppression [119], and is essential for normal development [121,122]. According to their functional importance, DNA methylation patterns are not random, well-regulated and tissue-specific [123,124].

Similarly to histones, DNA can be demethylated, this process can be accomplished either passively, by simply not methylating the new DNA strand after replication; or actively, by a replication-independent process that involves the TET proteins coupled with the base excision repair (BER) pathway [125].

Two replication-independent (active) demethylation mechanisms have been described that link the TET proteins activity with BER. The first mechanism, involves TETs that oxidize 5hmC to produce 5fC and 5caC, which in turn can be excised by TDG; their following replacement with a new non-methylated cytosine results in demethylation [126–130].

The second mechanism, which includes AID (Activation-Induced cytidine Deaminase) and APOBEC (Apolipoprotein B mRNA editing enzyme), is still debatable and has been reported by one laboratory in transfected human cells and in mouse brain [131,132]. They proposed that 5hmC is deaminated to 5-hydroxyuracil (5hmU) by AID and APOBEC family enzymes, removed by SMUG1 (Single-strand selective Monofunctional Uracil DNA Glycosylase 1) or TDG

glycosylases and ultimately replaced by cytosine. This sequential deamination and removal of 5hmC is similar to the deamination of 5mC and the removal of the resulting T:G mismatches by TDG, a mechanism previously proposed to occur in zebrafish [133,134], human cancers [134], primordial germ cells [135] and fused cells undergoing reprogramming [136]. In support of these mechanisms, TDG can remove T:G mismatches and 5hmU:G mismatches *in vitro* [132,137].

Synchronized fluctuations of DNA methylation, demethylation, nucleosome positioning, and CTCF chromatin binding have an important role for establishing cell-type-specific chromatin states during differentiation. A recent study, attempted to elucidate the molecular mechanisms that link these processes, they performed a Double Knockout (DKO) of the *Tet1* and *Tet2* dioxygenases in mouse Embryonic Stem Cells (mESC). They determined the changes in DNA methylation, CTCF binding, nucleosome positioning, and gene expression in DKO mESCs, and developed biophysical models to foresee differential CTCF binding. They found that a quantity of CTCF sites shown direct correlations with the CpG modification state: CTCF was preferentially lost from sites that were marked with 5hmC in Wild Type (WT) cells but not from 5fC-enriched sites. They also found that some CTCF sites function as divergence or bifurcation points, and these points define the differential methylation landscape. CTCF loss from; for example boundaries of chromatin loops, promoters and TADs; was linked with DNA methylation and DNA demethylation dissemination and can be related to down-regulation of adjacent genes [138]. The results of the Teif group revealed a hierarchical interaction among cytosine modifications, nucleosome positions, and DNA sequence that controls differential CTCF binding and regulates gene expression. The Teif group

results support another study that linked to the down-regulation of nearby super-enhancers after the removal of cohesin [139].

In summary, DNA methylation is involved in vital processes of the cell and it is also a dynamic mark that is highly regulated during different cell stages. One of these mechanisms of regulation is active demethylation, which is performed by TET enzymes.

The first evidences of the TETs discovery were in 2007, where, the identification of the *Trypanosoma cruzi* enzymes JBP1 and JBP2, responsible of gene silencing through hydroxylation and glycosylation of a thymine methyl group (named as J Base) gave insights about the kind of “eraser” proteins in charge of removing DNA methylation [140]. Therefore, in 2009, when looking for mammalian homologs of the trypanosome thymidine hydroxylases, Tahiliani and colleagues [22] identified three human TET proteins, TET1, TET2 and TET3, and additionally revealed that TET1 can catalyze conversion of 5mC to 5hmC *in vitro* and in cultured cells, and it was shown that the other two members TET2 and TET3, were also able to demethylate lysine residues [128]. In 2009, additional base modifications were found, while comparing the abundance of 5mC in granulose and Purkinje cells a new modified nucleotide was discovered to be enriched at the brain, the 5hmC [141]. It was though that 5hmC could alter the local chromatin environment through the recruitment or displacement of proteins that bind to it selectively [142]. Many CpG promoters contain 5hmC around 500-2,000 bases upstream and downstream of the TSS, but not at the TSS itself [143].

The TET protein family was first described when the *TET1* gene was identified as a fusion partner of the Mixed-Lineage Leukemia (MLL) gene in acute

myeloid leukemia [144]. Nevertheless, TET proteins were not at a central stage until they were found to oxidize 5mC to 5-hydroxymethyl-cytosine (5hmC) as a step of the DNA demethylation mechanism [22,145]. Subsequent reports revealed that TET proteins further oxidize 5hmC to 5-formyl-cytosine (5fC) and 5-carboxyl-cytosine (5caC), which are removed through the Base Excision Repair (BER) pathway, thereby finishing the demethylation process [127,128].

TET proteins have fundamental roles in epigenetic reprogramming, embryogenesis, development, and tumorigenesis; it is well known that their inactivation contributes to the DNA hypermethylation observed in cancer [146,147]. All TET proteins contain a C-terminal catalytic domain made of a Cys-rich region and a Double-Stranded  $\beta$ -Helix (DSBH) fold, characteristic of iron (II)/ $\alpha$ -ketoglutarate Fe(II) $\alpha$ -kg-dependent dioxygenases. Additionally, TET1 and TET3 contain an N-terminal CXXC ZF domain with high affinity for unmethylated CpG dinucleotides [22,125]. In vertebrates, TET2 do not have the CXXC finger domain because it experienced a chromosomal inversion event that detached this domain as a separate gene, called IDAX (Inhibition of the Dishevelled and Axin Complex) [148].

The TET proteins utilize a base-flipping mechanism to target the methylated cytosine and acquire the necessary architecture for accommodating 5mC derivatives for chemical modification [149]. While, 5mC oxidation reaction mechanisms by the TET enzymes is very similar to the demethylation processes carried out by the Jumonji C-containing proteins, like histone demethylases TET proteins must be finely controlled by regulatory mechanisms.

#### **TET enzymes regulation.**

The activity of TET demethylases is regulated by the availability of substrates, cofactors (like Vitamin C) and the factors targeting the demethylation machinery. Like other proteins, demethylases are subject to post-translational modifications, such as ubiquitination, acetylation, phosphorylation, GlcNAcylation and PARylation or degradation [125].

The expression of TET proteins expression is strongly regulated at the transcriptional level. For instance, in mESCs both *Tet1* and *Tet2* genes are positively regulated by Oct4, and their mRNA levels decline radically following mESC differentiation. Differentially, *Tet3* is significantly upregulated during differentiation [150]. It has become very clear that, in addition to dynamic transcriptional modulation, TET protein recruitment and activity are regulated at several levels that together control the outcome on DNA methylation patterns. The TET enzymes activity can be stimulated or inhibited in the presence of different cofactors (like Vitamin C), metabolites and the factors targeting the demethylation machinery. For example, in cancer, the aberrant accumulation of 2-hydroxyglutarate (2-HG), succinate, and fumarate inhibits TET protein enzymatic [151,152]. In comparison, increasing levels of vitamin C has been shown to stimulate TET protein enzymatic activity in cultured cells and mice tissues [153–156]. Although the exact mechanism is unknown, it is probable that vitamin C cooperates directly with TET the catalytic domain and allows a local reducing environment which, in turn, increases recycling efficacy of the Fe(II) cofactor [156].

Furthermore, interaction with several binding partners both stably and transiently is likely to alter the genomic location and stability of TET proteins [157,158]. Apart from their catalytic activity regulation, TET1 and TET3 are more



likely recruited to genomic target sites through direct binding of their respective CXXC domains to DNA. *In vitro* binding assays and *in vivo* chromatin immunoprecipitation assays confirm that these domains can bind CpG-rich oligonucleotides with a slight preference for unmethylated versus methylated substrates [147,159–162]. Furthermore, the detected enrichment of TET1 at CpG islands may consequently reflect the higher number of potential binding sites in these elements. TET1 has been shown to be related with distal regulatory elements and genome-wide diffuse 5hmC deposition [163,164]; also, TET3-mediated cytosine oxidation in the developing embryo happens at a wide range of genomic components [158]. Therefore, the CXXC domain may be involved in the recruitment of TET1 and TET3 to CpG sites regardless of local CpG density, and this is consistent with a model in which TET proteins, such as DNMT, bind along of the genome.

The recognition of unmethylated CpGs by the CXXC demethylases domain is not the only mechanism to explain TET recruitment [160]; since the active participation of an isoform lacking the CXXC and N-terminus domains (Tet1s) and even the interaction between TET2 and IDAX point to the existence of another TET recruitment interactions [165]. Some researchers inferred that these domains facilitate but do not determine TET genomic localization as protein partners could do [166].

In contrast to TET1 and TET3, the structure of TET2 does not expose any obvious domains reported to bind directly to DNA, and it is therefore potentially recruited to genomic DNA by a CXXC domain-independent mechanism [165,167]. Considering the potential regulatory role of TET2 at distal regulatory components, it is possible that TET2 recruitment can be facilitated through direct

binding of DNA targeting partners. In fact it has been revealed, that the TET2 protein, both endogenous or overexpressed can bind tissue-specific TFs like Early B cell factor 1 (EBF1) [168] and WT1 [169,170]. Dynamic expression of DNA-binding factors and interaction with TET2 can likely concede tissue-specific and temporal modulation of TET activity on a limited set of genomic *loci* [157].

Furthermore, the recruitment by DNA-binding factors can also result in TET2 protein degradation. The TET2-interacting protein IDAX/CXXC4, which is thought to enclose the ancestral CXXC domain of TET2, binds unmethylated DNA and is positioned mainly at CpG islands. Analysis of cell lines with combined ectopic expression of these proteins proposed that IDAX/CXXC4 cooperates with and activates caspase-dependent degradation of TET2 by a mechanism dependent on DNA binding [165]. Therefore, it is possible that IDAX/CXXC4 may have evolved to negatively regulate TET2 activity on CpG islands with little or no effect outside of them. Although this represents a probable model, IDAX/CXXC4 and TET2 were not co-expressed in the cells in which the studies were performed, and the interaction remains to be validated on endogenous model [157,165].

In summary, the TET enzymes can interact with DNA using their DNA-binding domain (CXXC) and form protein complexes with other epigenetic components to modify gene transcription. Nevertheless, there is new evidence that suggests that TET enzymes could be recruited to DNA by a CXXC-independent mechanism. In this regard, CTCF can play an important role, which will be further described.

**DNA demethylases as new partners of CTCF.**

Since 2013, it was proposed that CTCF binds to LMRs could be mediating the turnover of DNA methylation through TET recruitment [171]. The first evidence was a regularly oscillating 5hmC pattern observed around the binding sites of CTCF in mESC; which suggests that accessibility and 5hmC deposition could be sensitive to CTCF binding [172]. The specific genomic co-localization of CTCF, TET1, or TET2 and 5hmC was probed by co-immunoprecipitation assays on 3T3-L1 and HEK293T and correlated with enhancer activation on differentiated cells through the facilitation of hydroxymethylation of DNA 5hmC. Concluding that CTCF promotes TET-mediated DNA hydroxymethylation of enhancers driving adipocyte differentiation (Figure 2C) [17]. The same year, the Rippe group studied the relationship between DNA demethylation, CTCF binding and nucleosome repositioning in Embryonic Stem Cells (ESC). They found a 5mC/5hmC-dependent competition for DNA-binding between CTCF and nucleosomes during stem cell development; where TET1 binding facilitates the eviction and lability of nucleosomes that can be displaced by CTCF, driving a differential binding site selection during stem cell development [173]. Nevertheless, the relationship between CTCF and TET demethylases is not only relevant on cell differentiation processes, since a study in 2016 described an epigenetic splicing switch based on the relative ratio of 5mC to its oxidation derivatives. The Oberdoerffer group, revealed that dynamic TET1 and TET2-catalyzed DNA demethylation regulates CTCF-dependent alternative pre-mRNA splicing in human lymphocytes. Their study found that CTCF directly interacts with 5caC-containing DNA *in vitro*, and that CTCF-binding sites in CD45+ cells are enriched with 5hmC and 5caC, both marks strongly associated with alternative exon inclusion [19](Figure 2C).

A recent study describes that the 5caC directs the new CTCF binding within genomic DNA, the Oberdoerffer group revealed that 5caC stabilizes the binding of CTCF with DNA. Completely, the CTCF binding profile in a cellular model that amasses genomic 5caC recognized ~ 13,000 new CTCF sites. Since CTCF has several functions in gene expression, these findings have wide-ranging allegations and suggest that induced 5caC is a potential mechanism to achieve differential CTCF binding in cells [174].

More studies are needed to better describe the exact functions that DNA hydroxymethylation plays in transcriptional regulatory events, and also additional explorations will be required to define the basis of CTCF binding associated with 5caC *in vivo*.

Taken together, the above information, suggests that CTCF could interact, directly or indirectly with histone and DNA demethylases, it is still unknown if these complexes are related with the repression, activation or other transcriptional processes.

#### **Conclusions and final remarks.**

CTCF has become one of the most important nuclear factors reported in the literature, where more than 2,000 publications are related to this protein. Part of the relevance of this protein lies in its ability to interact with multiple protein partners. The incorporation of new partners such as CTCF isoforms (CTCF-s), paralog proteins (BORIS) and also proteins with demethylation activity (KDMs and TET), take CTCF study to another level of biological complexity. Currently, it remains unclear if the CTCF-s isoform or BORIS interact or form a protein complex with DNA or histone demethylases, and the implications of these

protein complexes in biological processes. Finally, the only thing we can be sure of, is that CTCF is the new kid on the block that still has many secrets to discover.

### **Acknowledges**

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT) through the Fondo Sectorial de Investigación en Salud y Seguridad Social (FOSISS, Grant No. 0261181), Fondo CB-SEP-CONACyT (284748), Apoyo para proyectos de investigación científica, desarrollo tecnológico e innovación en salud ante la contingencia por COVID-19 (Grant No. 312021), and UAM-PTC-704. ES-R was supported by the Natural Science Department at UAM Cuajimalpa Unit. LG-C and MS-A are doctoral students, and IH-R is a master student, from Programa de Maestría y Doctorado en Ciencias Bioquímicas, UNAM, and received a fellowship from CONACyT (LG-C: CVU 588391, IH-R: CVU 886138 and MS-A: CVU 659273). MS-A was also a beneficiary of the German Academic Exchange Service (DAAD ref. no. 91693321).

### **Declaration of interest statement.**

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

### **References.**

[1] Heath H, Ribeiro de Almeida C, Sleutels F, Dingjan G, van de Nobelen S, Jonkers I, et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J* 2008;27:2839–50. <https://doi.org/10.1038/emboj.2008.214>.

[2] Holwerda SJB, de Laat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond, B, Biol Sci*

2013;368:20120369. <https://doi.org/10.1098/rstb.2012.0369>.

- [3] Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* 1990;5:1743–53.
- [4] Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol* 1996;16:2802–13.
- [5] Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH, et al. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol* 1993;13:7612–24.
- [6] Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 2007;128:1231–45. <https://doi.org/10.1016/j.cell.2006.12.048>.
- [7] Lee B-K, Iyer VR. Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation. *J Biol Chem* 2012;287:30906–13. <https://doi.org/10.1074/jbc.R111.324962>.
- [8] Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* 2012;488:116–20. <https://doi.org/10.1038/nature11243>.
- [9] Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999;98:387–96.
- [10] Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, et al.

CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing.

Nature 2011;479:74–9. <https://doi.org/10.1038/nature10442>.

[11] Pugacheva EM, Kubo N, Loukinov D, Tajmul M, Kang S, Kovalchuk AL, et al.

CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention.

Proc Natl Acad Sci USA 2020;117:2020–31. <https://doi.org/10.1073/pnas.1911708117>.

[12] Mukhopadhyay R, Yu W, Whitehead J, Xu J, Lezcano M, Pack S, et al. The

binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. Genome Res 2004;14:1594–602.

<https://doi.org/10.1101/gr.2408304>.

[13] Chernukhin I, Shamsuddin S, Kang SY, Bergström R, Kwon Y-W, Yu W, et al.

CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. Mol Cell Biol 2007;27:1631–48.

<https://doi.org/10.1128/MCB.01993-06>.

[14] Kim YW, Kim A. Histone acetylation contributes to chromatin looping between the locus control region and globin gene by influencing hypersensitive site formation.

Biochim Biophys Acta 2013;1829:963–9.

<https://doi.org/10.1016/j.bbagr.2013.04.006>.

[15] Ferrari R, de Llobet Cucalon LI, Di Vona C, Le Dilly F, Vidal E, Lioutas A, et

al. TFIIC Binding to Alu Elements Controls Gene Expression via Chromatin Looping and Histone Acetylation. Mol Cell 2019. <https://doi.org/10.1016/j.molcel.2019.10.020>.

[16] Li Y, Haarhuis JHI, Cacciatore ÁS, Oldenkamp R, van Ruiten MS, Willems L,

et al. The structural basis for cohesin-CTCF-anchored loops. Nature 2020.

<https://doi.org/10.1038/s41586-019-1910-z>.

[17] Dubois-Chevalier J, Oger F, Dehondt H, Firmin FF, Gheeraert C, Staels B, et al.

A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and

- transcriptional induction of adipocyte differentiation. *Nucleic Acids Res* 2014;42:10943–59. <https://doi.org/10.1093/nar/gku780>.
- [18] Yamamoto S, Wu Z, Russnes HG, Takagi S, Peluffo G, Vaske C, et al. JARID1B is a luminal lineage-driving oncogene in breast cancer. *Cancer Cell* 2014;25:762–77. <https://doi.org/10.1016/j.ccr.2014.04.024>.
- [19] Marina RJ, Sturgill D, Bailly MA, Thenoz M, Varma G, Prigge MF, et al. TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. *EMBO J* 2016;35:335–55. <https://doi.org/10.15252/embj.201593235>.
- [20] Guerra-Calderas L, González-Barríos R, Patiño CC, Alcaraz N, Salgado-Albarrán M, León DC de, et al. CTCF-KDM4A complex correlates with histone modifications that negatively regulate CHD5 gene expression in cancer cell lines. *Oncotarget* 2018;9:17028–42. <https://doi.org/10.18632/oncotarget.24798>.
- [21] Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004;119:941–53. <https://doi.org/10.1016/j.cell.2004.12.012>.
- [22] Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009;324:930–5. <https://doi.org/10.1126/science.1170116>.
- [23] Ng SS, Kavanagh KL, McDonough MA, Butler D, Pilka ES, Lienard BMR, et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* 2007;448:87–91. <https://doi.org/10.1038/nature05971>.
- [24] Couture J-F, Collazo E, Ortiz-Tello PA, Brunzelle JS, Trievel RC. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nat Struct Mol Biol* 2007;14:689–95. <https://doi.org/10.1038/nsmb1273>.



- [25] Lan F, Nottke AC, Shi Y. Mechanisms involved in the regulation of histone lysine demethylases. *Curr Opin Cell Biol* 2008;20:316–25.  
<https://doi.org/10.1016/j.ceb.2008.03.004>.
- [26] Vostrov AA, Taheny MJ, Quitschke WW. A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter. *J Biol Chem* 2002;277:1619–27.  
<https://doi.org/10.1074/jbc.M109748200>.
- [27] Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X. Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Mol Cell* 2017;66:711-720.e3. <https://doi.org/10.1016/j.molcel.2017.05.004>.
- [28] Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* 2012;22:1680–8. <https://doi.org/10.1101/gr.136101.111>.
- [29] Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. *Cell* 2016;167:233-247.e17.  
<https://doi.org/10.1016/j.cell.2016.08.056>.
- [30] Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* 2017;6.  
<https://doi.org/10.7554/eLife.25776>.
- [31] Li J, Huang K, Hu G, Babarinde IA, Li Y, Dong X, et al. An alternative CTCF isoform antagonizes canonical CTCF occupancy and changes chromatin architecture to promote apoptosis. *Nat Commun* 2019;10:1–13. <https://doi.org/10.1038/s41467-019-08949-w>.
- [32] Hore TA, Deakin JE, Graves JAM. The Evolution of Epigenetic Regulators CTCF and BORIS/CTCF in Amniotes. *PLOS Genetics* 2008;4:e1000169.

<https://doi.org/10.1371/journal.pgen.1000169>.

[33] Salgado-Albarrán M, González-Barrios R, Guerra-Calderas L, Alcaraz N, Estefanía Sánchez-Correa T, Castro-Hernández C, et al. The epigenetic factor BORIS (CTCF) controls the androgen receptor regulatory network in ovarian cancer. *Oncogenesis* 2019;8:41. <https://doi.org/10.1038/s41389-019-0150-2>.

[34] Bergmaier P, Weth O, Dienstbach S, Boettger T, Galjart N, Mernberger M, et al. Choice of binding sites for CTCFL compared to CTCF is driven by chromatin and by sequence preference. *Nucleic Acids Res* 2018;46:7097–107.

<https://doi.org/10.1093/nar/gky483>.

[35] Lobanenkov VV, Zentner GE. Discovering a binary CTCF code with a little help from BORIS. *Nucleus* 2018;9:33–41.

<https://doi.org/10.1080/19491034.2017.1394536>.

[36] Pugacheva EM, Rivero-Hinojosa S, Espinoza CA, Méndez-Catalá CF, Kang S, Suzuki T, et al. Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome Biol* 2015;16.

<https://doi.org/10.1186/s13059-015-0736-8>.

[37] Debruyne DN, Dries R, Sengupta S, Seruggia D, Gao Y, Sharma B, et al. BORIS promotes chromatin regulatory interactions in treatment-resistant cancer cells. *Nature* 2019;572:676–80. <https://doi.org/10.1038/s41586-019-1472-0>.

[38] Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 2008;453:948–51. <https://doi.org/10.1038/nature06947>.

[39] Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 2000;405:482–5.

<https://doi.org/10.1038/35013100>.

- [40] Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000;405:486–9. <https://doi.org/10.1038/35013106>.
- [41] Cuddapah S, Jothi R, Schones DE, Roh T-Y, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 2009;19:24–32. <https://doi.org/10.1101/gr.082800.108>.
- [42] Barrington C, Georgopoulou D, Pezic D, Varsally W, Herrero J, Hadjur S. Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nat Commun* 2019;10:2908. <https://doi.org/10.1038/s41467-019-10725-9>.
- [43] Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* 2001;17:520–7.
- [44] Nakahashi H, Kwon K-RK, Resch W, Vian L, Dose M, Stavreva D, et al. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep* 2013;3:1678–89. <https://doi.org/10.1016/j.celrep.2013.04.024>.
- [45] Hou C, Zhao H, Tanimoto K, Dean A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A* 2008;105:20398–403. <https://doi.org/10.1073/pnas.0808506106>.
- [46] Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 2004;13:291–8.
- [47] Zlatanova J, Caiafa P. CTCF and its protein partners: divide and rule? *J Cell Sci* 2009;122:1275–84. <https://doi.org/10.1242/jcs.039990>.
- [48] Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin

- mediates transcriptional insulation by CCCTC-binding factor. *Nature* 2008;451:796–801. <https://doi.org/10.1038/nature06634>.
- [49] Rubio ED, Reiss DJ, Welsh PL, Distèche CM, Filippova GN, Baliga NS, et al. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci USA* 2008;105:8309–14. <https://doi.org/10.1073/pnas.0801273105>.
- [50] Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular *c-myc* and *H19/Igf2* insulators. *EMBO J* 2008;27:654–66. <https://doi.org/10.1038/emboj.2008.1>.
- [51] Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 2008;132:422–33. <https://doi.org/10.1016/j.cell.2008.01.011>.
- [52] Xiao T, Wallace J, Felsenfeld G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol* 2011;31:2174–83. <https://doi.org/10.1128/MCB.05093-11>.
- [53] Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Curr Opin Genet Dev* 2007;17:400–7. <https://doi.org/10.1016/j.gde.2007.08.005>.
- [54] Chen S, Luperchio TR, Wong X, Doan EB, Byrd AT, Roy Choudhury K, et al. A Lamina-Associated Domain Border Governs Nuclear Lamina Interactions, Transcription, and Recombination of the *Tcrb* Locus. *Cell Rep* 2018;25:1729-1740.e6. <https://doi.org/10.1016/j.celrep.2018.10.052>.
- [55] Defossez P-A, Kelly KF, Filion GJP, Pérez-Torrado R, Magdinier F, Menoni H, et al. The human enhancer blocker CTC-binding factor interacts with the transcription

factor Kaiso. *J Biol Chem* 2005;280:43017–23.

<https://doi.org/10.1074/jbc.M510802200>.

[56] Schwalie PC, Ward MC, Cain CE, Faure AJ, Gilad Y, Odom DT, et al. Co-binding by YY1 identifies the transcriptionally active, highly conserved set of CTCF-bound regions in primate genomes. *Genome Biol* 2013;14:R148.

<https://doi.org/10.1186/gb-2013-14-12-r148>.

[57] Wang R, Shen J, Huang P, Zhu X. CCCTC-Binding Factor Controls Its Own Nuclear Transport via Regulating the Expression of Importin 13. *Mol Cells*

2013;35:388–95. <https://doi.org/10.1007/s10059-013-2283-z>.

[58] Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell*

2009;34:271–84. <https://doi.org/10.1016/j.molcel.2009.04.001>.

[59] Zhao H, Sifakis EG, Sumida N, Millán-Ariño L, Scholz BA, Svensson JP, et al. PARP1- and CTCF-Mediated Interactions between Active and Repressed Chromatin at the Lamina Promote Oscillating Transcription. *Molecular Cell* 2015;59:984–97.

<https://doi.org/10.1016/j.molcel.2015.07.019>.

[60] Lutz M, Burke LJ, Barreto G, Goeman F, Greb H, Arnold R, et al.

Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Res* 2000;28:1707–13.

[61] Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife* n.d.;6.

<https://doi.org/10.7554/eLife.25776>.

[62] Damaschke NA, Yang B, Blute ML, Lin CP, Huang W, Jarrard DF. Frequent disruption of chromodomain helicase DNA-binding protein 8 (CHD8) and functionally associated chromatin regulators in prostate cancer. *Neoplasia* 2014;16:1018–27.

<https://doi.org/10.1016/j.neo.2014.10.003>.

[63] Rodenberg JM, Hoggatt AM, Chen M, Touw K, Jones R, Herring BP.

Regulation of serum response factor activity and smooth muscle cell apoptosis by chromodomain helicase DNA-binding protein 8. *Am J Physiol Cell Physiol*

2010;299:C1058–67. <https://doi.org/10.1152/ajpcell.00080.2010>.

[64] Li T, Hu J-F, Qiu X, Ling J, Chen H, Wang S, et al. CTCF Regulates Allelic

Expression of Igf2 by Orchestrating a Promoter-Polycomb Repressive Complex 2

Intrachromosomal Loop. *Mol Cell Biol* 2008;28:6473–82.

<https://doi.org/10.1128/MCB.00204-08>.

[65] Zhang H, Niu B, Hu J-F, Ge S, Wang H, Li T, et al. Interruption of

intrachromosomal looping by CCCTC binding factor decoy proteins abrogates genomic imprinting of human insulin-like growth factor II. *J Cell Biol* 2011;193:475–87.

<https://doi.org/10.1083/jcb.201101021>.

[66] Chernukhin IV, Shamsuddin S, Robinson AF, Carne AF, Paul A, El-Kady AI, et

al. Physical and functional interaction between two pluripotent proteins, the Y-box

DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *J Biol*

*Chem* 2000;275:29915–21. <https://doi.org/10.1074/jbc.M001538200>.

[67] Majumder P, Gomez JA, Chadwick BP, Boss JM. The insulator factor CTCF

controls MHC class II gene expression and is required for the formation of long-

distance chromatin interactions. *J Exp Med* 2008;205:785–98.

<https://doi.org/10.1084/jem.20071843>.

[68] Ku W-C, Chiu S-K, Chen Y-J, Huang H-H, Wu W-G, Chen Y-J.

Complementary Quantitative Proteomics Reveals that Transcription Factor AP-4

Mediates E-box-dependent Complex Formation for Transcriptional Repression of

HDM2. *Mol Cell Proteomics* 2009;8:2034–50. <https://doi.org/10.1074/mcp.M900013->

MCP200.

[69] Muck JS, Kandasamy K, Englmann A, Günther M, Zink D. Perinuclear positioning of the inactive human cystic fibrosis gene depends on CTCF, A-type lamins and an active histone deacetylase. *J Cell Biochem* 2012;113:2607–21.

<https://doi.org/10.1002/jcb.24136>.

[70] Lang F, Li X, Zheng W, Li Z, Lu D, Chen G, et al. CTCF prevents genomic instability by promoting homologous recombination-directed DNA double-strand break repair. *Proc Natl Acad Sci USA* 2017;114:10912–7.

<https://doi.org/10.1073/pnas.1704076114>.

[71] Mohan M, Bartkuhn M, Herold M, Philippen A, Heidl N, Bardenhagen I, et al. The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J* 2007;26:4203–14. <https://doi.org/10.1038/sj.emboj.7601851>.

[72] Xiao T, Wongtrakoongate P, Trainor C, Felsenfeld G. CTCF Recruits Centromeric Protein CENP-E to the Pericentromeric/Centromeric Regions of Chromosomes through Unusual CTCF Binding Sites. *Cell Rep* 2015;12:1704–14.

<https://doi.org/10.1016/j.celrep.2015.08.005>.

[73] Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A, et al. Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol* 2016;17. <https://doi.org/10.1186/s13059-016-1043-8>.

[74] Liu Q, Yang B, Xie X, Wei L, Liu W, Yang W, et al. Vigilin interacts with CCCTC-binding factor (CTCF) and is involved in CTCF-dependent regulation of the imprinted genes *Igf2* and *H19*. *FEBS J* 2014;281:2713–25.

<https://doi.org/10.1111/febs.12816>.

[75] Ross-Innes CS, Brown GD, Carroll JS. A co-ordinated interaction between

CTCF and ER in breast cancer cells. *BMC Genomics* 2011;12:593.

<https://doi.org/10.1186/1471-2164-12-593>.

[76] Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, et al. Histone deimination antagonizes arginine methylation. *Cell* 2004;118:545–53. <https://doi.org/10.1016/j.cell.2004.08.020>.

[77] Deng X, Shao G, Zhang H-T, Li C, Zhang D, Cheng L, et al. Protein arginine methyltransferase 5 functions as an epigenetic activator of the androgen receptor to promote prostate cancer cell growth. *Oncogene* 2016.

<https://doi.org/10.1038/onc.2016.287>.

[78] Ramadoss S, Guo G, Wang C-Y. Lysine demethylase KDM3A regulates breast cancer cell invasion and apoptosis by targeting histone and the non-histone protein p53. *Oncogene* 2017;36:47–59. <https://doi.org/10.1038/onc.2016.174>.

[79] Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell* 2007;128:707–19. <https://doi.org/10.1016/j.cell.2007.01.015>.

[80] Guerra-Calderas L, González-Barrios R, Herrera LA, Cantú de León D, Soto-Reyes E. The role of the histone demethylase KDM4A in cancer. *Cancer Genet* 2015;208:215–24. <https://doi.org/10.1016/j.cancergen.2014.11.001>.

[81] Kim S, Benoiton L, Paik WK. EPSILON-ALKYLLYSINASE. PURIFICATION AND PROPERTIES OF THE ENZYME. *J Biol Chem* 1964;239:3790–6.

[82] Paik WK, Kim S. Enzymatic demethylation of calf thymus histones. *Biochem Biophys Res Commun* 1973;51:781–8.

[83] Paik WK, Kim S. Epsilon-alkyllysine. New assay method, purification, and biological significance. *Arch Biochem Biophys* 1974;165:369–78.

[84] Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, et al. New



nomenclature for chromatin-modifying enzymes. *Cell* 2007;131:633–6.

<https://doi.org/10.1016/j.cell.2007.10.039>.

[85] Metzger E, Wissmann M, Yin N, Müller JM, Schneider R, Peters AHFM, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 2005;437:436–9. <https://doi.org/10.1038/nature04020>.

[86] Huang J, Berger SL. The emerging field of dynamic lysine methylation of non-histone proteins. *Curr Opin Genet Dev* 2008;18:152–8.

<https://doi.org/10.1016/j.gde.2008.01.012>.

[87] Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, et al. p53 is regulated by the lysine demethylase LSD1. *Nature* 2007;449:105–8.

<https://doi.org/10.1038/nature06092>.

[88] Kontaki H, Talianidis I. Lysine methylation regulates E2F1-induced cell death. *Mol Cell* 2010;39:152–60. <https://doi.org/10.1016/j.molcel.2010.06.006>.

[89] Lee J-Y, Park J-H, Choi H-J, Won H-Y, Joo H-S, Shin D-H, et al. LSD1 demethylates HIF1 $\alpha$  to inhibit hydroxylation and ubiquitin-mediated degradation in tumor angiogenesis. *Oncogene* 2017;36:5512–21. <https://doi.org/10.1038/onc.2017.158>.

[90] Sheng W, LaFleur MW, Nguyen TH, Chen S, Chakravarthy A, Conway JR, et al. LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. *Cell* 2018;174:549-563.e19. <https://doi.org/10.1016/j.cell.2018.05.052>.

[91] Majello B, Gorini F, Saccà CD, Amente S. Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers (Basel)* 2019;11.

<https://doi.org/10.3390/cancers11030324>.

[92] Burg JM, Link JE, Morgan BS, Heller FJ, Hargrove AE, McCafferty DG.

KDM1 class flavin-dependent protein lysine demethylases. *Biopolymers* 2015;104:213–46. <https://doi.org/10.1002/bip.22643>.

- [93] Yang Y, Yin X, Yang H, Xu Y. Histone Demethylase LSD2 Acts as an E3 Ubiquitin Ligase and Inhibits Cancer Cell Growth through Promoting Proteasomal Degradation of OGT. *Molecular Cell* 2015;58:47–59.  
<https://doi.org/10.1016/j.molcel.2015.01.038>.
- [94] Anand R, Marmorstein R. Structure and mechanism of lysine-specific demethylase enzymes. *J Biol Chem* 2007;282:35425–9.  
<https://doi.org/10.1074/jbc.R700027200>.
- [95] Spannhoff A, Hauser A-T, Heinke R, Sippl W, Jung M. The emerging therapeutic potential of histone methyltransferase and demethylase inhibitors. *ChemMedChem* 2009;4:1568–82. <https://doi.org/10.1002/cmdc.200900301>.
- [96] Yang M, Gocke CB, Luo X, Borek D, Tomchick DR, Machius M, et al. Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Mol Cell* 2006;23:377–87.  
<https://doi.org/10.1016/j.molcel.2006.07.012>.
- [97] Lu T, Jackson MW, Wang B, Yang M, Chance MR, Miyagi M, et al. Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. *Proc Natl Acad Sci USA* 2010;107:46–51. <https://doi.org/10.1073/pnas.0912493107>.
- [98] Ponnaluri VKC, Vavilala DT, Putty S, Gutheil WG, Mukherji M. Identification of non-histone substrates for JMJD2A-C histone demethylases. *Biochem Biophys Res Commun* 2009;390:280–4. <https://doi.org/10.1016/j.bbrc.2009.09.107>.
- [99] Arrowsmith CH, Bountra C, Fish PV, Lee K, Schapira M. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov* 2012;11:384–400.  
<https://doi.org/10.1038/nrd3674>.
- [100] Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, et al. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell*

2006;125:467–81. <https://doi.org/10.1016/j.cell.2006.03.028>.

[101] Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* 2007;25:15–30. <https://doi.org/10.1016/j.molcel.2006.12.014>.

[102] De Santa F, Totaro MG, Prosperini E, Notarbartolo S, Testa G, Natoli G. The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* 2007;130:1083–94. <https://doi.org/10.1016/j.cell.2007.08.019>.

[103] Hong Z, Li H, Li L, Wang W, Xu T. Different expression patterns of histone H3K27 demethylases in renal cell carcinoma and bladder cancer. *Cancer Biomark* 2017;18:125–31. <https://doi.org/10.3233/CBM-160003>.

[104] Lan F, Zaratiegui M, Villén J, Vaughn MW, Verdel A, Huarte M, et al. S. pombe LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Mol Cell* 2007;26:89–101. <https://doi.org/10.1016/j.molcel.2007.02.023>.

[105] Rudolph T, Yonezawa M, Lein S, Heidrich K, Kubicek S, Schäfer C, et al. Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol Cell* 2007;26:103–15. <https://doi.org/10.1016/j.molcel.2007.02.025>.

[106] Shiao C, Trnka MJ, Bozicevic A, Ortiz Torres I, Al-Sady B, Burlingame AL, et al. Reconstitution of nucleosome demethylation and catalytic properties of a Jumonji histone demethylase. *Chem Biol* 2013;20:494–9. <https://doi.org/10.1016/j.chembiol.2013.03.008>.

[107] Garcia-Bassets I, Kwon Y-S, Telese F, Prefontaine GG, Hutt KR, Cheng CS, et al. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell* 2007;128:505–18.

<https://doi.org/10.1016/j.cell.2006.12.038>.

[108] Secombe J, Li L, Carlos L, Eisenman RN. The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Genes Dev* 2007;21:537–51. <https://doi.org/10.1101/gad.1523007>.

[109] Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, et al. Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature* 2007;446:882–7. <https://doi.org/10.1038/nature05671>.

[110] Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villén J, Li J, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci USA* 2004;101:12130–5. <https://doi.org/10.1073/pnas.0404720101>.

[111] Franci G, Sarno F, Nebbioso A, Altucci L. Identification and characterization of PKF118-310 as a KDM4A inhibitor. *Epigenetics* 2017;12:198–205. <https://doi.org/10.1080/15592294.2016.1249089>.

[112] Guillade L, Sarno F, Tarhonskaya H, Nebbioso A, Alvarez S, Kawamura A, et al. Synthesis and Biological Evaluation of Tripartin, a Putative KDM4 Natural Product Inhibitor, and 1-Dichloromethylinden-1-ol Analogues. *ChemMedChem* 2018;13:1949–56. <https://doi.org/10.1002/cmdc.201800377>.

[113] Hamada S, Kim T-D, Suzuki T, Itoh Y, Tsumoto H, Nakagawa H, et al. Synthesis and activity of N-oxalylglycine and its derivatives as Jumonji C-domain-containing histone lysine demethylase inhibitors. *Bioorg Med Chem Lett* 2009;19:2852–5. <https://doi.org/10.1016/j.bmcl.2009.03.098>.

[114] Laukka T, Myllykoski M, Looper RE, Koivunen P. Cancer-associated 2-oxoglutarate analogues modify histone methylation by inhibiting histone lysine demethylases. *J Mol Biol* 2018;430:3081–92. <https://doi.org/10.1016/j.jmb.2018.06.048>.

- [115] Metzger E, Stepputtis SS, Strietz J, Preca B-T, Urban S, Willmann D, et al. KDM4 Inhibition Targets Breast Cancer Stem-like Cells. *Cancer Res* 2017;77:5900–12. <https://doi.org/10.1158/0008-5472.CAN-17-1754>.
- [116] Kristensen LH, Nielsen AL, Helgstrand C, Lees M, Cloos P, Kastrup JS, et al. Studies of H3K4me3 demethylation by KDM5B/Jarid1B/PLU1 reveals strong substrate recognition in vitro and identifies 2,4-pyridine-dicarboxylic acid as an in vitro and in cell inhibitor. *FEBS J* 2012;279:1905–14. <https://doi.org/10.1111/j.1742-4658.2012.08567.x>.
- [117] Jeong YS, Park JS, Ko Y, Kang Y-K. JHDM3A module as an effector molecule in guide-directed modification of target chromatin. *J Biol Chem* 2011;286:4461–70. <https://doi.org/10.1074/jbc.M110.176040>.
- [118] Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science* 1975;187:226–32.
- [119] Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;16:6–21. <https://doi.org/10.1101/gad.947102>.
- [120] Razin A, Riggs AD. DNA methylation and gene function. *Science* 1980;210:604–10.
- [121] Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 1992;69:915–26.
- [122] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 1999;99:247–57.
- [123] Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 2006;38:1378–85. <https://doi.org/10.1038/ng1909>.

- [124] Illingworth R, Kerr A, DeSousa D, Jørgensen H, Ellis P, Stalker J, et al. A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. *PLoS Biol* 2008;6. <https://doi.org/10.1371/journal.pbio.0060022>.
- [125] Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 2017;18:517–34. <https://doi.org/10.1038/nrg.2017.33>.
- [126] Pfaffeneder T, Hackner B, Truss M, Münzel M, Müller M, Deiml CA, et al. The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew Chem Int Ed Engl* 2011;50:7008–12. <https://doi.org/10.1002/anie.201103899>.
- [127] He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011;333:1303–7. <https://doi.org/10.1126/science.1210944>.
- [128] Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 2011;333:1300–3. <https://doi.org/10.1126/science.1210597>.
- [129] Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* 2011;286:35334–8. <https://doi.org/10.1074/jbc.C111.284620>.
- [130] Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu G-L, et al. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol* 2012;8:328–30. <https://doi.org/10.1038/nchembio.914>.
- [131] Guo JU, Su Y, Zhong C, Ming G, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 2011;145:423–34. <https://doi.org/10.1016/j.cell.2011.03.022>.
- [132] Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, et al.

- AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol* 2012;8:751–8. <https://doi.org/10.1038/nchembio.1042>.
- [133] Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 2008;135:1201–12. <https://doi.org/10.1016/j.cell.2008.11.042>.
- [134] Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, et al. DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell* 2010;142:930–42. <https://doi.org/10.1016/j.cell.2010.08.030>.
- [135] Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 2010;463:1101–5. <https://doi.org/10.1038/nature08829>.
- [136] Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 2010;463:1042–7. <https://doi.org/10.1038/nature08752>.
- [137] Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 2011;146:67–79. <https://doi.org/10.1016/j.cell.2011.06.020>.
- [138] Wiehle L, Thorn GJ, Raddatz G, Clarkson CT, Rippe K, Lyko F, et al. DNA (de)methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. *Genome Res* 2019;29:750–61. <https://doi.org/10.1101/gr.239707.118>.
- [139] Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, et al. Cohesin Loss Eliminates All Loop Domains. *Cell* 2017;171:305–320.e24. <https://doi.org/10.1016/j.cell.2017.09.026>.
- [140] Yu Z, Genest P-A, ter Riet B, Sweeney K, DiPaolo C, Kieft R, et al. The protein

- that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase. *Nucleic Acids Res* 2007;35:2107–15. <https://doi.org/10.1093/nar/gkm049>.
- [141] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009;324:929–30. <https://doi.org/10.1126/science.1169786>.
- [142] Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* 2013;152:1146–59. <https://doi.org/10.1016/j.cell.2013.02.004>.
- [143] Szulwach KE, Li X, Li Y, Song C-X, Han JW, Kim S, et al. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet* 2011;7:e1002154. <https://doi.org/10.1371/journal.pgen.1002154>.
- [144] Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res* 2002;62:4075–80.
- [145] Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010;466:1129–33. <https://doi.org/10.1038/nature09303>.
- [146] Li D, Guo B, Wu H, Tan L, Lu Q. TET Family of Dioxygenases: Crucial Roles and Underlying Mechanisms. *Cytogenet Genome Res* 2015;146:171–80. <https://doi.org/10.1159/000438853>.
- [147] Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep* 2011;13:28–35. <https://doi.org/10.1038/embor.2011.233>.
- [148] Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic



- acids. *Cell Cycle* 2009;8:1698–710. <https://doi.org/10.4161/cc.8.11.8580>.
- [149] Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, et al. Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* 2013;155:1545–55. <https://doi.org/10.1016/j.cell.2013.11.020>.
- [150] Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 2011;8:200–13. <https://doi.org/10.1016/j.stem.2011.01.008>.
- [151] Xiao M, Yang H, Xu W, Ma S, Lin H, Zhu H, et al. Inhibition of  $\alpha$ -KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev* 2012;26:1326–38. <https://doi.org/10.1101/gad.191056.112>.
- [152] Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim S-H, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of  $\alpha$ -ketoglutarate-dependent dioxygenases. *Cancer Cell* 2011;19:17–30. <https://doi.org/10.1016/j.ccr.2010.12.014>.
- [153] Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem* 2014;6:1049–55. <https://doi.org/10.1038/nchem.2064>.
- [154] Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013;500:222–6. <https://doi.org/10.1038/nature12362>.
- [155] Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem* 2013;288:13669–74. <https://doi.org/10.1074/jbc.C113.464800>.

- [156] Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc* 2013;135:10396–403. <https://doi.org/10.1021/ja4028346>.
- [157] Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev* 2016;30:733–50. <https://doi.org/10.1101/gad.276568.115>.
- [158] Wang Y, Zhang Y. Regulation of TET protein stability by calpains. *Cell Rep* 2014;6:278–84. <https://doi.org/10.1016/j.celrep.2013.12.031>.
- [159] Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 2011;473:389–93. <https://doi.org/10.1038/nature09934>.
- [160] Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, et al. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* 2011;42:451–64. <https://doi.org/10.1016/j.molcel.2011.04.005>.
- [161] Xu Y, Xu C, Kato A, Tempel W, Abreu JG, Bian C, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* 2012;151:1200–13. <https://doi.org/10.1016/j.cell.2012.11.014>.
- [162] Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* 2010;20:1390–3. <https://doi.org/10.1038/cr.2010.156>.
- [163] Putiri EL, Tiedemann RL, Thompson JJ, Liu C, Ho T, Choi J-H, et al. Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. *Genome Biol* 2014;15:R81. <https://doi.org/10.1186/gb-2014-15-6-r81>.

- [164] Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 2011;480:490–5. <https://doi.org/10.1038/nature10716>.
- [165] Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* 2013;497:122–6. <https://doi.org/10.1038/nature12052>.
- [166] Zhang W, Xia W, Wang Q, Towers AJ, Chen J, Gao R, et al. Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development. *Molecular Cell* 2016;64:1062–73. <https://doi.org/10.1016/j.molcel.2016.10.030>.
- [167] Jain D, Baldi S, Zabel A, Straub T, Becker PB. Active promoters give rise to false positive “Phantom Peaks” in ChIP-seq experiments. *Nucleic Acids Res* 2015;43:6959–68. <https://doi.org/10.1093/nar/gkv637>.
- [168] Guilhamon P, Eskandarpour M, Halai D, Wilson GA, Feber A, Teschendorff AE, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. *Nat Commun* 2013;4:2166. <https://doi.org/10.1038/ncomms3166>.
- [169] Rampal R, Alkalin A, Madzo J, Vasanthakumar A, Pronier E, Patel J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep* 2014;9:1841–55. <https://doi.org/10.1016/j.celrep.2014.11.004>.
- [170] Wang Y, Xiao M, Chen X, Chen L, Xu Y, Lv L, et al. WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol Cell* 2015;57:662–73. <https://doi.org/10.1016/j.molcel.2014.12.023>.
- [171] Feldmann A, Ivanek R, Murr R, Gaidatzis D, Burger L, Schübeler D. Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLoS Genet* 2013;9:e1003994.

<https://doi.org/10.1371/journal.pgen.1003994>.

[172] Sun Z, Terragni J, Jolyon T, Borgaro JG, Liu Y, Yu L, et al. High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell Rep* 2013;3:567–76. <https://doi.org/10.1016/j.celrep.2013.01.001>.

[173] Teif VB, Beshnova DA, Vainshtein Y, Marth C, Malm J-P, Höfer T, et al. Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development. *Genome Res* 2014;24:1285–95.

<https://doi.org/10.1101/gr.164418.113>.

[174] Nanan KK, Sturgill DM, Prigge MF, Thenoz M, Dillman AA, Mandler MD, et al. TET-Catalyzed 5-Carboxylcytosine Promotes CTCF Binding to Suboptimal Sequences Genome-wide. *IScience* 2019;19:326–39.

<https://doi.org/10.1016/j.isci.2019.07.041>.

## **FIGURE LEGENDS.**

**Figure 1:** The participation of CTCF, CTCF-s and BORIS in the chromatin loop formation. (A) CTCF physically binds to itself to form homodimers which promotes chromatin loop formation. (B) CTCF-s (CTCF-s) competes with CTCF to alter chromatin architecture and loop formation. (C) BORIS can alter chromatin loops by either a cooperative or competitive mechanism with CTCF.

**Figure 2:** CTCF and its partners. (A) CTCF interacts with cohesin allowing loop formation. CTCF also interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites. CTCF regulates the p53 gene through direct interaction with Wrap53, ncRNA. CTCF interacts with IPO13, and this interaction mediates the nuclear import of CTCF. (B) A novel interaction of CTCF is with the histone demethylases CTCF interacts with KDM5B and regulates basal/stem cell genes in luminal breast cancer lines. Finally, CTCFs interaction with KDM4A is involved in the downregulation of the CHD5 gene expression in MCF7 cells. (C) The other new partners are the DNA demethylases, CTCF interaction with TET 1 and TET2 proteins, this association is involved in enhancer activation. This nuclear factor can also interact with 5caC and is involved in RNA pol II pausing and alternative exon inclusion of the CD45+ gene.

Figure 1

