



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS**

**INSTITUTO DE INVESTIGACIONES BIOMÉDICAS
BIOMEDICINA**

**REGULACIÓN TRANSCRIPCIONAL POR EL FACTOR NUCLEAR
CTCF EN LA LÍNEA CELULAR MCF-7**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRA EN CIENCIAS BIOLÓGICAS**

PRESENTA:

KARINA JÁCOME LÓPEZ

**TUTOR PRINCIPAL DE TESIS: DR. ALEJANDRO ZENTELLA DEHESA
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM
COMITÉ TUTOR: DR. ÁNGEL ALFONSO ZARAIN HERZBERG
FACULTAD DE MEDICINA, UNAM
DR. ERNESTO SOTO REYES SOLÍS
FACULTAD DE MEDICINA, UNAM**

CIUDAD UNIVERSITARIA, CD.MX. Enero, 2017



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.



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS**

**INSTITUTO DE INVESTIGACIONES BIOMÉDICAS
BIOMEDICINA**

**REGULACIÓN TRANSCRIPCIONAL POR EL FACTOR NUCLEAR
CTCF EN LA LÍNEA CELULAR MCF-7**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRA EN CIENCIAS BIOLÓGICAS**

PRESENTA:

KARINA JÁCOME LÓPEZ

**TUTOR PRINCIPAL DE TESIS: DR. ALEJANDRO ZENTELLA DEHESA
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM**

**COMITÉ TUTOR: DR. ÁNGEL ALFONSO ZARAIN HERZBERG
FACULTAD DE MEDICINA, UNAM
DR. ERNESTO SOTO REYES SOLÍS
FACULTAD DE MEDICINA, UNAM**

México, Cd. Mx. Enero, 2017

Lic. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 17 de octubre del 2016, se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** de la alumna **JÁCOME LÓPEZ KARINA** con número de cuenta **306188547** con la tesis titulada "**Regulación transcripcional por el factor nuclear CTCF en la línea celular MCF-7**", realizada bajo la dirección del **DR. ALEJANDRO ZENTELLA DEHESA**:

Presidente: Dr. Alfonso León del Río
Vocal: Dra. Sara Frías Vázquez
Secretaria: Dr. Ernesto Soto Reyes Solís
Suplente: Dra. Leda Carolina Torres Maldonado
Suplente: Dr. Angel Alfonso Zarain Herzberg

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

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, Cd. de Mx., a 25 de noviembre del 2016.

M. del Coro Arizmendi

DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



Agradecimientos

Agradezco al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México, por haberme permitido llevar a cabo mis estudios de maestría dentro de este programa.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo brindado en el programa de becarios (CVU 631104).

A mi tutor principal, el Doctor Alejandro Zentella Dehesa y a los miembros de mi comité tutor, el Dr. Ángel Alfonso Zaráin Herzberg y el Dr. Ernesto Soto Reyes, por los aportes realizados a mi trabajo de investigación.

Agradecimientos a título personal

A los miembros de la Unidad de Bioquímica del Instituto Nacional de Ciencias Médicas y Nutrición, en especial a José Luis Ventura Gallegos y a Eric Alejandro García López.

A los Doctores Víctor Manuel Dávila Borja y Sergio Juárez Méndez, del Laboratorio de Oncología Experimental en el Instituto Nacional de Pediatría, por el apoyo que me brindaron en el desarrollo de esta investigación.

A Jesús Rafael Rodríguez Aguilera por sus asesorías para la realización de los ensayos de inmunoprecipitación de cromatina.

Al Doctor Alfonso León del Río, a las Doctoras Sara Frías Vázquez y Leda Carolina Torres Maldonado, por el tiempo dedicado a la revisión de esta tesis y por sus valiosas aportaciones.

Índice

1. RESUMEN.....	7
2. ABSTRACT.....	8
3. INTRODUCCIÓN.....	9
Características estructurales de CTCF	9
Estructura genómica de CTCF.....	9
Estructura proteica de CTCF.....	9
Interacciones CTCF-DNA.....	11
Sitios de unión a CTCF en el DNA.....	11
Relación entre el estado de metilación del DNA con la unión de CTCF.....	12
Interacciones CTCF-PROTEINAS.....	13
Interacción CTCF-cohesina.....	14
Funciones de CTCF	16
Regulador de la expresión génica.....	16
Organizador de la cromatina.....	19
Alteraciones funcionales de CTCF	21
4. OBJETIVOS.....	24
5. METODOLOGÍA.....	26
Cultivo celular.....	26
Producción de partículas lentivirales y transducción.....	26
Extracción de proteínas totales	27
<i>Western Blot</i>	27
Microarreglos de expresión.....	28
Extracción de RNA.....	29
Transcripción reversa, PCR punto final y PCR cuantitativa.....	29
Análisis de metilación por MSP (<i>Methylation Specific PCR</i>)	30
Inmunoprecipitación de la cromatina (ChIP, <i>Chromatin Immunoprecipitation</i>)	30
6. RESULTADOS	32
Evaluación de la expresión proteica de CTCF en líneas celulares de cáncer de mama.....	32
Abatimiento de la expresión del factor nuclear CTCF en la línea celular MCF-7.....	33
Perfiles del transcriptoma en MCF-7 <i>mock</i> y <i>shCTCF</i>	34

Genes expresados diferencialmente MCF-7 <i>shCTCF</i> respecto a MCF-7 <i>mock</i>	35
Validación de microarreglo para <i>CTCF</i> y <i>EFEMP1</i>	37
Unión de CTCF a regiones regulatorias del gen <i>EFEMP1</i>	37
CTCF como barrera de metilación del promotor <i>EFEMP1</i>	39
CTCF, Rad21 y sitios de hipersensibilidad a DNasa I en CTSs de <i>EFEMP1</i>	41
7. DISCUSION.....	44
8. CONCLUSIONES.....	50
9. LITERATURA	51
10. ANEXO.....	56

1. RESUMEN

CTCF es un factor nuclear de 11 dedos de Zinc, que se une a diferentes secuencias de DNA, participando en diferentes mecanismos, como la activación de la transcripción, el silenciamiento, la inactivación del cromosoma X, el bloqueo de *enhancers*, impronta genómica y en la arquitectura nuclear, por lo que se le considera un factor maestro del genoma.

Reportes previos en cáncer demuestran que CTCF regula la actividad transcripcional de genes como *BRCA1*, *RASSF1A*, *CDH1*, *p16INK4a*, *BAX*, *RARRES* y *miR125b*, sin embargo no se han elucidado por completo los genes que son regulados por este factor. Con el objetivo de determinar genes blancos regulados por CTCF, empleamos un sistema lentiviral para abatir su expresión en la línea celular de adenocarcinoma mamario MCF-7 y mediante microarreglos de expresión de alta resolución (*Affymetrix GeneChip Human 1.0ST Array*).

Identificamos 15 genes expresados diferencialmente, 9 sobre expresados (*FXFD3*, *GDF15*, *RASGRP1*, *CMBL*, *PADI3*, *CEACAM5*, *MACF1*, *NBPF4*, *SEMA6A*) y 6 sub expresados (*EFEMP1*, *CTCF*, *GFRA1*, *DLX1*, *BAMBI*, *TGFB2*). De estos genes, fue de nuestro interés el gen *EFEMP1*, debido a que la disminución o ausencia en expresión se ha correlacionado con la progresión tumoral, en cáncer esporádico de mama, carcinoma hepatocelular, glioma, cáncer de próstata y carcinoma endometrial, por lo cual corroboramos perfil de expresión correspondiente a CTCF y *EFEMP1*, por RT-qPCR, posteriormente identificamos sitios de unión a CTCF (*CTCF Target Sites*, *CTSs*) en regiones regulatorias del gen *EFEMP1*. Encontramos dos *CTSs*, uno de ellos arriba del sitio de inicio a la transcripción y el segundo en el promotor del gen, este último incluye una porción del primer exón y esta embebido en una isla CpG. Asimismo, evaluamos la expresión del transcrito *EFEMP1* en las líneas celulares de cáncer de mama: MDA-MB-231, MDA-MB-453 y MDA-MB-468, encontramos que la línea celular MDA-MB-453 no presenta expresión, el promotor se encuentra metilado y CTCF no se une en los *CTSs* reportados para MCF-7, estos resultados indican que la expresión de este gen depende de la unión de CTCF a estas regiones, las cuales deben estar hipometiladas.

2. ABSTRACT

CTCF is a nuclear factor of 11 zinc fingers, highly versatile that bind to several DNA sequences and is involved in gene expression regulation through different epigenetic mechanisms, this factor is able to activate and repress transcription, also can work like insulator and intervenes in chromosome X inactivation, genomic imprinting and nuclear architecture.

In breast cancer, previous reports have shown that CTCF regulates the transcriptional activity of genes involved in carcinogenesis, such as *BRCA1*, *RASSF1A*, *CDH1*, *p16INK4a*, *BAX*, *RARRES* y *miR125b*, however have not been fully elucidated the genes that are regulated by this protein.

In order to determine genes regulated by CTCF, we used a lentiviral system to knock down CTCF mRNA by shRNA targeting in MCF-7 cells. Through Affymetrix GeneChip Human 1.0ST Array, we found 15 significantly and differentially expressed genes, 9 up regulated (*FXYD3*, *GDF15*, *RASGRP1*, *CMBL*, *PADI3*, *CEACAM5*, *MACF1*, *NBPF4*, *SEMA6A*) and 6 down regulated (*EFEMP1*, *CTCF*, *GFRA1*, *DLX1*, *BAMBI*, *TGFB2*).

CTCF and EFEMP1 expression profile were verify by RT-qPCR, next two CTCF target sites were identified in EFEMP1 gene sequence, one of them was located upstream from transcription start site, and the second is the promoter that comprise an exon portion and is inside of a CpG island. EFEMP1 expression was evaluate in MDA-MB-231, MDA-MB-453 and MDA-MB-468 cell lines, we found that MDA-MB-453 is negative to *EFEMP1*, also the promoter is methylate and CTCF cannot bind to target sites reported for MCF-7. The results show that CTCF binding and promoter hypomethylation are required for *EFEMP1* expression.

3. INTRODUCCIÓN

CTCF es un factor nuclear multifuncional de 11 dedos de Zinc, que tiene la capacidad de unirse a diversas secuencias de DNA, por el uso combinado de sus dedos de Zinc; se encuentra altamente conservado en vertebrados y se ha reportado su presencia en algunos invertebrados como *Anopheles gambiae*, *Aedes aegypti* y *Drosophila melanogaster* (Klenova et al. 1993; Ohlsson et al. 2001; Ong & Corces 2009). CTCF puede funcionar como activador o represor de la transcripción, elemento barrera, bloqueador de *enhancers* y participa en procesos como la inactivación del cromosoma X, impronta genómica y en la arquitectura nuclear (Ohlsson et al. 2001; Recillas-Targa 2006; Ong & Corces 2009; Fiorentino & Giordano 2012; Moon et al. 2005; Leers & Renkawitz 2005; Phillips & Corces 2009).

CARACTERÍSTICAS ESTRUCTURALES DE CTCF

Estructura genómica de CTCF

El gen CTCF se localiza en un segmento del cromosoma 16q22.1 y tiene un tamaño de 76.7 Kb, su promotor está embebido en una isla CpG, carece de caja CAAT y TATA, posee múltiples sitios de unión para el factor transcripcional Sp1 y el núcleo del promotor presenta una secuencia que es reconocida por el factor YY1 (Filippova et al. 1998a; Klenova et al. 1998; Pugacheva et al. 2006). El gen tiene 12 exones denominados: E₀, E_{0new}, E₁-E₁₀; los primeros dos no codifican, el exon E₁ codifica a la región N-terminal, los exones E₂-E₈ codifican a los 11 dedos de Zinc (ZFs, *Zinc Fingers*), el exon E₉ tiene una región rica en Adeninas-Timinas, el exon E₁₀ contiene el codón de paro a la transcripción y una región larga 3'UTR (*Untranslated Region*), en conjunto estos dos últimos exones dan origen a la región C-terminal (Figura 1) (Filippova et al. 2002).

Estructura proteica de CTCF

La proteína CTCF se compone de 727 residuos de aminoácidos y tiene un peso de 130 KDa, se divide en tres regiones: N-terminal, 11 dedos de Zinc y C-terminal (Figura 1) (Martinez & Miranda 2010).

La región N-terminal abarca los aminoácidos 1-273 y puede presentar dos tipos de modificaciones post-traduccionales: sumoilación en la Lisina 74 y poli(ADP)ribosilación en los aminoácidos de ácido glutámico presentes en esta región proteica (Farrar et al. 2010; Zlatanova & Caiafa 2009). Los dedos de Zinc 1-10 son del tipo C_2H_2 y el dedo de Zinc 11 es del tipo C_2HC , esto significa que los residuos de aminoácidos Cisteína (C) e Histidina (H) interactúan con el ion Zn^{2+} para formar los dedos, en conjunto se insertan en el surco mayor del DNA, generando contacto específico entre los aminoácidos de los dedos de Zinc con nucleótidos (Kaptein 1991; Filippova et al. 2002).

La región C-terminal se caracteriza por 3 motivos altamente conservados:

1. El motivo *SKKEDSSDSE*, susceptible de fosforilación por CKII (*Casein Kinase II*),
2. Dos secuencias de residuos de aminoácidos repetidas *PXXP*, característico de las proteínas de unión al dominio SH3,
3. El motivo *KRRGRP-type AT-hook* que está involucrado en la unión a DNA y proteínas (Ohlsson et al. 2001; Martínez & Miranda 2010; Filippova et al. 2002). Esta región, además de ser fosforilada en las Serinas del motivo *SKKEDSSDSE*, también puede sumoilarse en la Lisina 689 (Wang et al. 2012).

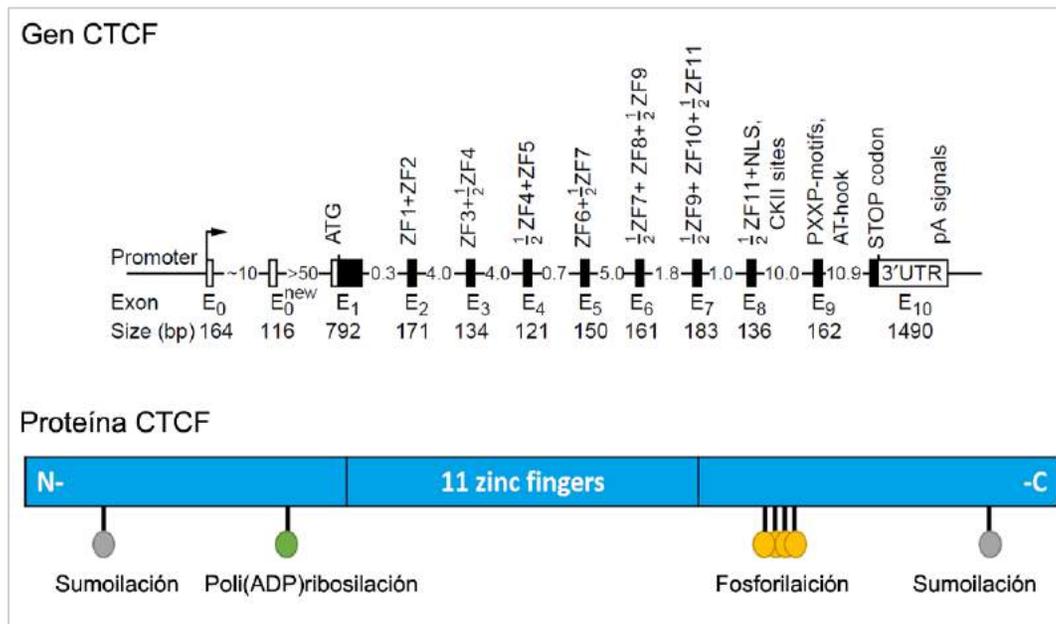


Figura 1. Estructura del gen y la proteína CTCF

El gen CTCF posee 12 exones, los dos primeros no codifican; la proteína puede ser modificada post-traduccionamente en las regiones N- y C-terminal, los 11 dedos de Zinc se unen al DNA (Tomada y modificada de Filippova et al. 2002).

INTERACCIONES CTCF-DNA

Sitios de unión a CTCF en el DNA

Los sitios de unión de CTCF al DNA se han denominado *CTSs* (*CTCF Target Sites*), debido al uso combinatorio de los dedos de Zinc para unirse a DNA se han reportado numerosas secuencias de nucleótidos que pueden ser reconocidas por este factor, no obstante con ensayos ChIP-exo (*Chromatin Immunoprecipitation-exonuclease*) se ha identificado un motivo “consenso” que consiste en una secuencia de 52pb (Figura 2), organizada en cuatro módulos (Rhee & Pugh 2011). Los módulos 2 y 3 constituyen el núcleo del motivo y son reconocidos por los dedos de Zinc 4-7, por su parte los dedos de Zinc 1-3 se unen al módulo 4 y los dedos 8-11 al módulo 1 (Nakahashi et al. 2013). CTCF puede unirse a los cuatro módulos, a combinatorias de módulos o reconocer únicamente al módulo 3 (Rhee & Pugh 2011).

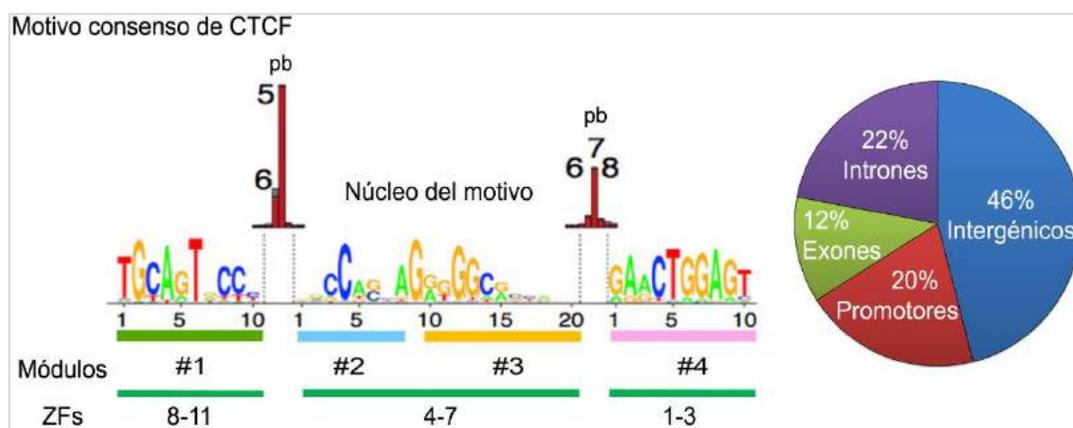


Figura 2. Motivo consenso de CTCF

El motivo de unión a CTCF se divide en cuatro módulos, los módulos 2-3 constituyen el núcleo del motivo, el cual es reconocido por los dedos de Zinc 4-7, siendo estos los más abundantes en el genoma. Este motivo se ha reportado en secuencias de promotores y en regiones inter- e intra- génicas (Tomada y modificada de Nakahashi et al. 2013; Rhee & Pugh 2011).

En lo que respecta a la localización genómica de los *CTSs*, se ha reportado que alrededor del 46% se ubican en regiones intergénicas, 20% en promotores y 34% en regiones intragénicas (exones e intrones) (Kim et al. 2007; Chen et al. 2012; Ong & Corces 2014). Esta localización genómica, provee información sobre la posible función de CTCF en un sitio específico, pues se pueden

correlacionar con la expresión génica, el estado de metilación del DNA, el estado de la cromatina, modificaciones post-traduccionales de histonas, sitios de hipersensibilidad a DNasa I o con otros elementos regulatorios del DNA (Chen et al. 2012).

En mamíferos, se estiman 55,000-65,000 *CTSs*, aproximadamente 5,000 son de alta afinidad y están ultra-conservados entre especies y tejidos (Ong & Corces 2014; Nakahashi et al. 2013).

Además de las interacciones CTCF-DNA, esta proteína tiene una región de unión a RNA (RBR, *RNA-Binding Region*) que comprende a los ZFs 10-11 y la región C-terminal; mediante experimentos PAR-CLIP (*Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation*) se identificaron los RNAs que interactúan con CTCF en la línea celular de osteosarcoma humano U2OS, donde 71% corresponden a transcritos de genes codificantes, 12% transcritos de pseudogenes, 8.3% RNAs antisentido, 5% RNAs intergénicos largos no codificantes (lincRNAs) y 3.7% a RNAs de otros tipos; estas interacciones son relevantes para la regulación de la expresión génica, un ejemplo de esta función es el transcrito antisentido *Wrap53*, sin embargo también se ha propuesto que estas interacciones participan en la multimerización de CTCF (Saldaña-Meyer et al. 2014).

Relación entre el estado de metilación del DNA con la unión de CTCF

En mamíferos, la metilación del DNA consiste en la adición covalente de un grupo metilo (-CH₃) en el carbono 5 de Citosinas (C) contiguas a Guaninas (G) -dinucleótidos CpG- produciendo a la 5-metilCitosina (5mC), esta reacción es catalizada por las enzimas DNMTs (*DNA methyl transferases*). La metilación del DNA puede regular la actividad transcripcional de los genes mediante varios mecanismos: 1. Impidiendo el acceso de factores de transcripción a los promotores por impedimento estérico; 2. Por reclutamiento de enzimas, a través de proteínas con dominios de unión a DNA metilado, que des-acetilan las colas de histonas (HDACs, *Histone Deacetylases*) e inducen la compactación de la cromatina, haciéndola inaccesible para los factores de transcripción; 3. Las DNMTs pueden reclutar Complejos Proteicos Remodeladores de la Cromatina produciendo el silenciamiento génico (Kulis & Esteller 2010).

Con la caracterización del motivo “consenso” a CTCF, también se identificaron dinucleótidos CpG en dicha secuencia, lo cual llevó a la siguiente pregunta: ¿Qué efecto tendrá la presencia de CpGs en el motivo de CTCF?

Con análisis masivos del genoma y en diferentes líneas celulares de humano, se han identificado *CTSs* constitutivos y *CTSs* variables, los primeros se refieren a sitios de unión a CTCF con la misma localización en todas las líneas celulares, los cuales generalmente no presentan metilación del DNA, en contra parte los *CTSs* variables son específicos de cada tipo celular y muestran patrones diferenciales de metilación; ésta información indica que la metilación de los *CTSs* se asocia negativamente con la unión de CTCF (Figura 3) (Chen et al. 2012; Ong & Corces 2014).

En el proceso de desmetilación de la 5mC, se generan otras formas de citosina modificada, éstas son 5hmC (5-hidroximetilcitosina), 5fC (5-formilcitosina) y 5caC (5-carboxil citosina) (Guo et al. 2011). En particular, la 5hmC es una modificación que se asocia fuertemente a elementos como promotores, TSSs (*Transcription Start Sites*), exones y *enhancers*; asimismo se ha reportado que los *CTSs* están enriquecidos con 5hmC y en menor proporción con 5fC. Esto llama la atención, pues se ha reportado que CTCF interactúa con las proteínas TET, enzimas oxidan el grupo metilo de la 5mC, y generando la hidroximetilación del DNA y consecuentemente la activación de la transcripción (Yu et al. 2012; Song et al. 2013; Dubois-Chevalier et al. 2014).

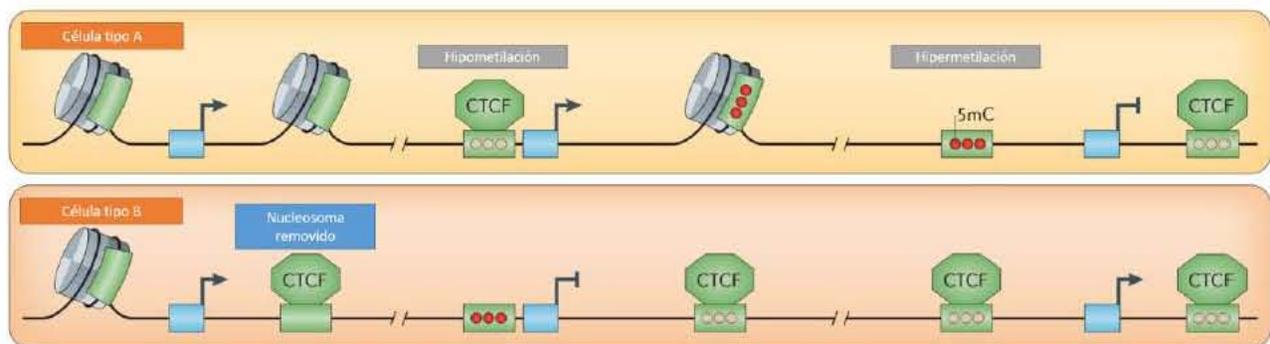


Figura 3. CTCF y la metilación del DNA

Los tipos celulares presentan patrones de metilación del DNA diferenciales, CTCF generalmente se asocia a *CTSs* hipometilados y a su vez se promueve la accesibilidad al DNA y la transcripción (Modificada de Ong & Corces 2014).

INTERACCIONES CTCF-PROTEINAS

Otros factores que contribuyen a la versatilidad funcional de CTCF, es su participación en complejos proteicos y su interacción con diversas proteínas. Se han reportado muchas proteínas que interactúan con CTCF y se han clasificado en los siguientes grupos (Figura 4): 1. Factores de transcripción, 2. Enzimas, 3. Proteínas asociadas a la cromatina, 4. Proteínas multifuncionales (Zlatanova & Caiafa 2009; Holwerda & de Laat 2013). De las proteínas mencionadas, llama la atención el complejo CTCF-cohesina por su papel biológico, el cual se describe en los siguientes párrafos.

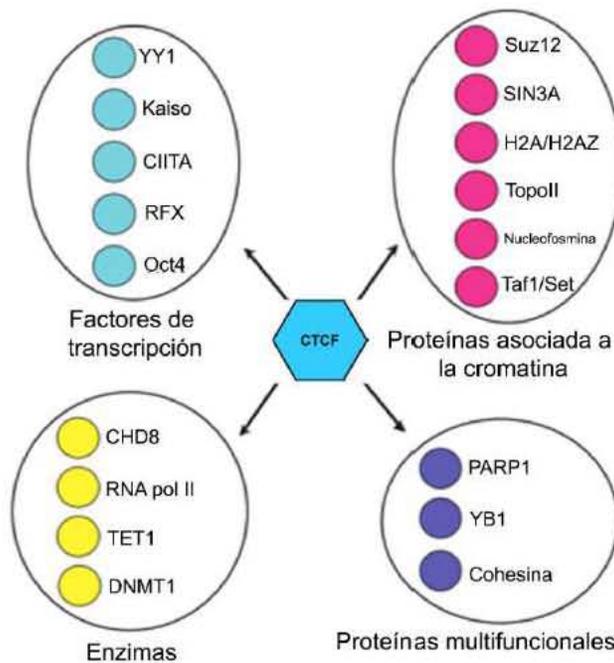


Figura 4. Proteínas que interactúan con CTCF
Las proteínas se han clasificado en cuatro grupos según su función, la imagen muestra solo algunas de las proteínas reportadas que interactúan con CTCF (Modificada de Zlatanova & Caiafa 2009).

Interacción CTCF-cohesina

El complejo CTCF-cohesina regula la expresión génica, mediante la formación de asas de cromatina y su interacción con otras proteínas, por lo que tiene un papel importante en la organización de la cromatina de cada tipo celular (Lee & Iyer 2012). La cohesina está compuesta por las subunidades SMC1, SMC3, SCC3/SA1/SA2 y SCC1/RAD21, quienes forman una estructura de anillo que “abrazo” a la cromatina, para mantener a las cromátidas hermanas juntas durante la mitosis y la reparación del DNA (Rubio et al. 2008). El anillo de cohesina no tiene un motivo de unión

específico al DNA, su localización depende de otras proteínas, por ejemplo CTCF sirve para posicionar al complejo cohesina sobre los cromosomas, en conjunto generan asas de cromatina y regulan la expresión génica (Figura 5), se ha reportado que el 50-80% de los *CTSs*, dependiendo del tipo celular, se asocian con la presencia del complejo cohesina (Lee & Iyer 2012; Merkenschlager & Odom 2013).

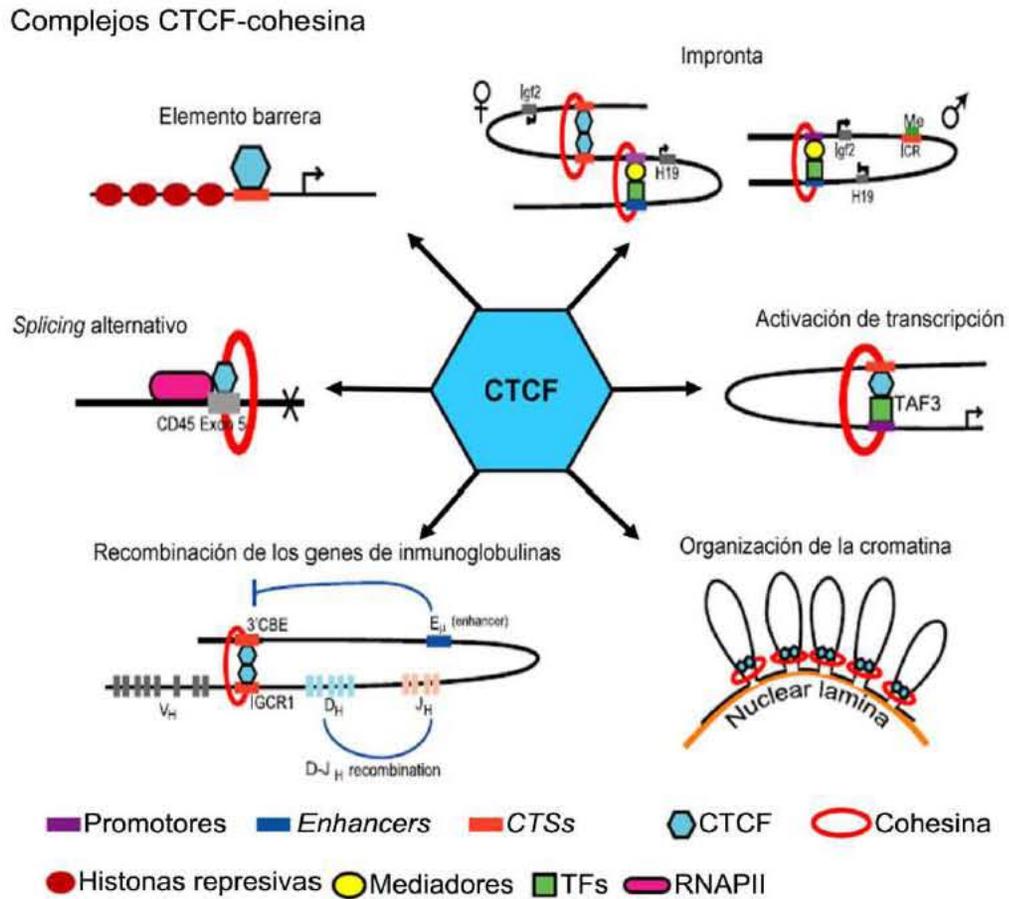


Figura 5. Complejo CTCF-cohesina

La cohesina colabora con CTCF para regular la expresión génica y la arquitectura genómica, por ejemplo, el caso de impronta; activación de genes a través de *enhancer*, en la formación de *asas*, e interacción con la lámina nuclear. Por otro lado, el complejo cohesina; independientemente de CTCF estabiliza complejos necesarios para la transcripción, por ejemplo en el caso de *ERα* o bien en el proceso de mitosis para mantener juntas a las cromátidas hermanas (Modificada de Lee & Iyer 2012).

FUNCIONES DE CTCF

Como se mencionó anteriormente CTCF es una proteína funcionalmente versátil, desde un punto de vista global podemos dividir sus funciones en dos: 1. Regulador de la expresión génica y 2. Organizador de la cromatina.

Regulador de la expresión génica

La regulación de la expresión génica por CTCF ocurre mediante diferentes mecanismos, para esto CTCF puede actuar como elemento barrera, bloquear o activar *enhancers*, unirse a promotores y a secuencias silenciadoras, o al cuerpo del gen. Existen diversos reportes de genes que son regulados por estos mecanismos, los cuales se describen brevemente a continuación:

1. CTCF como elemento barrera

Los elementos barrera, se definen como proteínas de unión a DNA, que evitan el esparcimiento de heterocromatina a zonas transcripcionalmente activas y pueden reclutar enzimas involucradas en la modificación de histonas para detener el esparcimiento de marcas de represión (Yang & Corces 2011). Se ha reportado que CTCF actúa como elemento barrera para los genes *p16INK4a*, *pRB*, *BRCA1*, *RASSF1A* (Figura 6), entre otros.

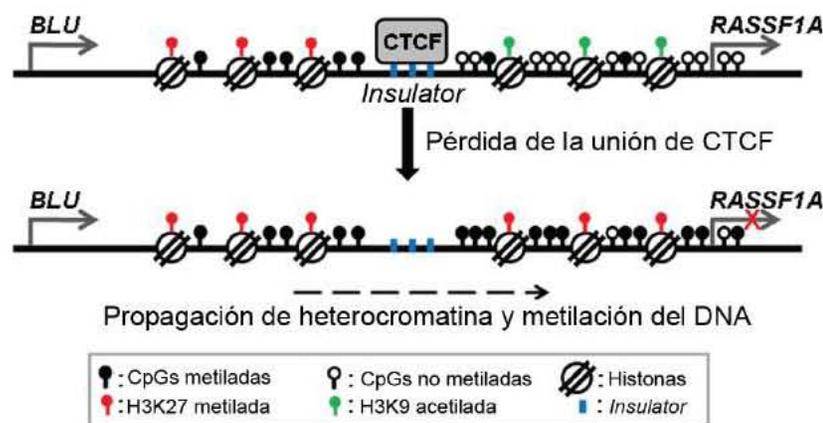


Figura 6. CTCF como elemento barrera para el gen *RASSF1A*

CTCF permite la transcripción del gen *RASSF1A* al evitar el esparcimiento de cromatina represiva, la cual se caracteriza por la marca de histona H3K27me₃, en este ejemplo CTCF también evita la propagación de la metilación del DNA (Modificada de Chang et al. 2010).

2. CTCF como bloqueador de *enhancers*

Los *enhancers* son reconocidos por proteínas mediadoras y factores de transcripción, forman asas de cromatina con la finalidad de actuar sobre los promotores y potencializar la transcripción (Gaszner & Felsenfeld 2006). Algunos genes regulados por CTCF mediante el bloqueo de *enhancers* son: los genes β -globina, *c-myc*, genes improntados (p. ej. *Igf2/H19*, Figura 7)(Bell et al. 1999; Phillips & Corces 2009).

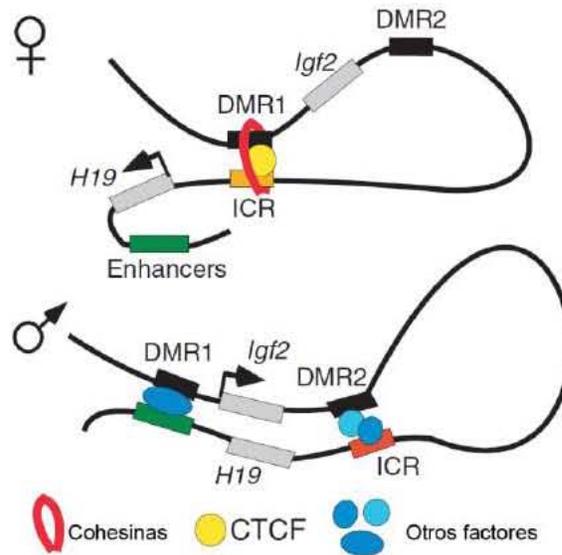


Figura 7. Bloqueo de *enhancer* en cromosoma materno (impronta).

En el cromosoma materno, CTCF identifica la región control de impronta no metilada y evita que un *enhancer* actúe sobre el gen *Igf2*, pero a su vez permite la expresión del gen *H19*. En el cromosoma paterno la región de control de impronta se encuentra metilada, esto evita su reconocimiento por CTCF de tal forma que el *enhancer* sí actúa en el gen *Igf2* y al no estar presente CTCF la metilación de la región control se esparce hacia el promotor de *H19* (Modificada de Ong & Corces 2009).

3. CTCF y su unión a promotores

La transcripción del gen *APP* (*amyloid β -protein precursor*), fue el primer ejemplo donde se observó que CTCF actúa como un activador de la transcripción, mediante su unión con el promotor proximal del gen, la depleción de esta región reduce la actividad transcripcional en un 70-90% (Vostrov & Quitschke 1997). Cabe mencionar que en muchos casos, CTCF facilita la interacción de *enhancers* con promotores (Figura 8), por lo que este factor tiene la función dual de bloquear o facilitar la interacción con estas secuencias regulatorias (Ong & Corces 2014).

4. CTCF y su unión a *silencers*

Los *silencers* son secuencias de DNA, que impiden la transcripción de un gen y en ocasiones conducen a la compactación de la cromatina (Gaszner & Helsenfeld 2006). CTCF reconoce estas secuencias y evita la expresión, por ejemplo, la transcripción del gen de la lisozima de pollo es silenciada por la unión de CTCF a una secuencia silenciadora río arriba del sitio de inicio de la transcripción (Baniahmad et al. 1990; Awad 1999).

5. CTCF y su unión en el cuerpo de genes

CTCF puede unirse en regiones intragénicas y de esta forma regular expresión o bien puede participar en *splicing*. Por ejemplo, CTCF se une a una secuencia localizada en los primeros dos exones del gen *hTERT* (Figura 9) y reprime la transcripción (Renaud et al. 2005; S Renaud et al. 2007). En lo que respecta a *splicing*, el ejemplo es el gen *CD45* (Figura 5), en este caso CTCF se une al exon 5 no metilado, promoviendo su inserción en el pre-mRNA y en el mRNA maduro (Shukla et al. 2011).

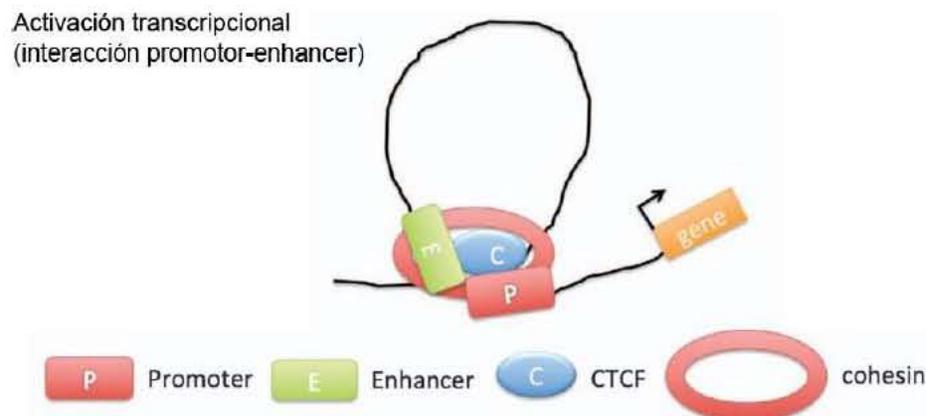


Figura 8. Interacciones promotor-*enhancer* mediadas por CTCF

CTCF puede promover la transcripción uniéndose al promotor y/o potencializando el efecto de *enhancers* sobre el promotor donde se encuentre, como se muestra en la figura (Modificado de Kim et al. 2015).

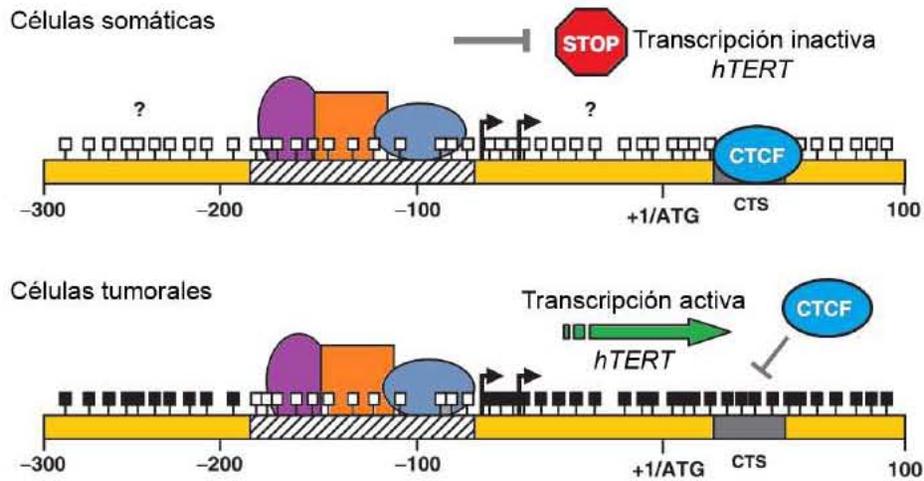


Figura 9. CTCF y su unión al cuerpo del gen *hTERT*

Cuando el gen *hTERT* se encuentra hipometilado, CTCF se une a los primeros exones y evita la expresión (Modificado de Stéphanie Renaud et al. 2007)

Organizador de la cromatina

Por otra parte, CTCF contribuye a la organización espacial del genoma, mediante la formación de asas de cromatina, de tal forma que regula la interacción a distancia dentro de los cromosomas y entre ellos (Ong & Corces 2014).

La organización estructural de la cromatina se refleja en los perfiles de expresión de cada tipo celular; durante la interfase celular los cromosomas se organizan en "territorios cromosómicos" y a su vez la cromatina se compone en eucromatina o en heterocromatina, esta última prepondera unida a la lámina nuclear y en la periferia del nucléolo (Cremer & Cremer 2001). La lámina nuclear funciona como un compartimento de represión transcripcional, donde proteínas de esta región se unen directamente a la cromatina y a reguladores transcripcionales, en conjunto originan "dominios asociados a la lámina" (*LADs, Lamin Associated Domains*), los cuales se caracterizan por una cromatina represiva. Dentro de los territorios cromosómicos, segmentos de los cromosomas pueden adoptar una organización, en forma de asas, la cuales representan un elemento sub estructural de la organización espacial del genoma (Guelen et al. 2008; Zhao et al. 2009).

CTCF se asocia con proteínas de la lámina nuclear promoviendo la formación de asas, por ejemplo los bordes de los *LADs* están delimitados por CTCF y se caracterizan por tener la marca de represión

H3K27me3 (Guelen et al. 2008). Bing Ren y colaboradores (2012), propusieron que dentro de cada territorio cromosomal se forman asas que regulan la interacción a larga distancia, a los que llamaron “dominios topológicos” (*TADs*, *Topologic Associated Domains*), los cuales pueden contener *CTSs*, que al ser reconocidos por CTCF logran establecer otras asas mediadas por la dimerización de esta proteína, a estas estructuras se les ha denominado “subdominios ad hoc” (Figura 10) (Dixon et al. 2012; Li et al. 2012).

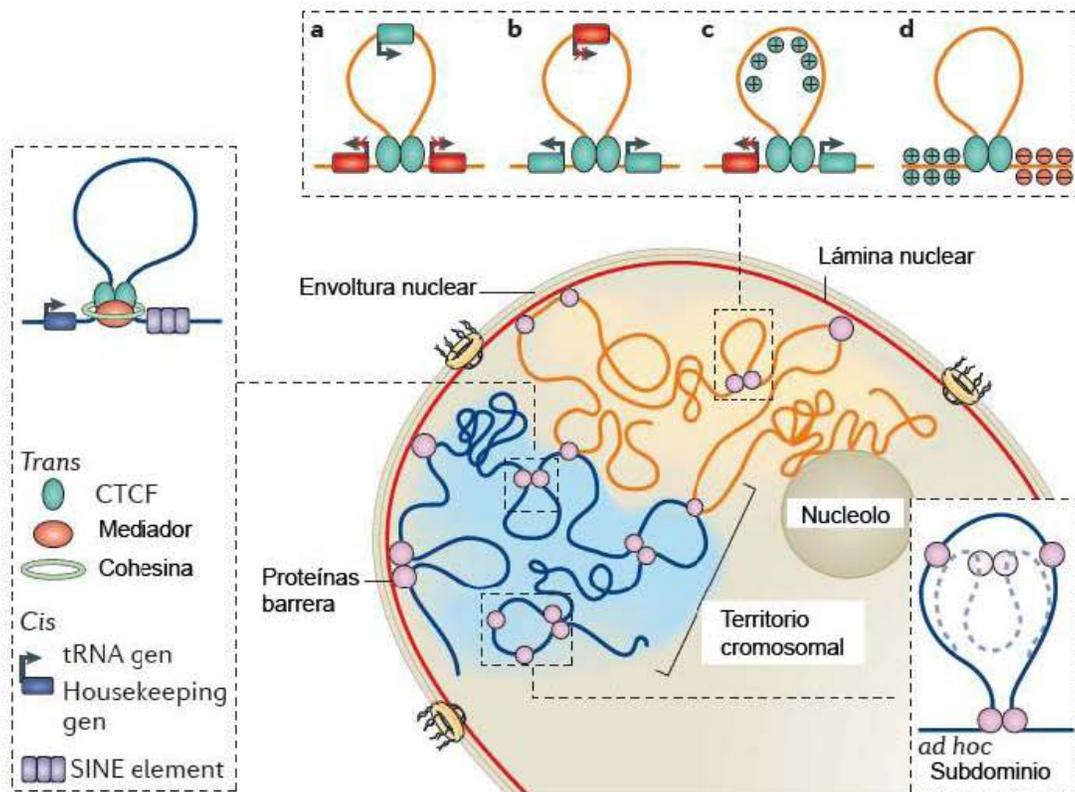


Figura 10. CTCF y la organización de la cromatina

Modelo de los dominios topológicos en el genoma mediados por CTCF, a la izquierda se esquematizan los factores proteicos que están implicados en la formación de dominios (CTCF, Mediator, Cohesina), así mismo se ha descrito que en la base del dominio las secuencias de DNA generalmente encontradas corresponden a genes constitutivos y secuencias SINE. En la parte superior se representan los tipos de dominios que pueden generarse, en a) al interior del asa hay actividad transcripcional contrario a lo observado al exterior del asa; b) caso contrario en el representado en a); c) el interior del asa está enriquecido con marcas de histonas activas; d) el asa separa una región cromatínica activa de una inactiva (Li et al. 2012).

ALTERACIONES FUNCIONALES DE CTCF

La versatilidad funcional de CTCF, hace de éste un factor maestro del genoma, pues regula el arreglo espacial de los cromosomas en cada tipo celular y la expresión de genes involucrados en distintos procesos celulares. La desregulación de estas funciones conduce al desarrollo de enfermedades, como el cáncer, en esta patología se han reportado varias alteraciones en CTCF que contribuyen al desarrollo de la tumorigénesis, las cuales pueden ser: 1. Mutaciones en el gen, 2. Modificaciones en los CTSs, 3. Modificaciones post-traduccionales, 4. Localización celular y niveles de expresión; estos aspectos se describen a continuación.

1. Mutaciones en el gen CTCF

En cáncer esporádico de mama y próstata, se han reportado mutaciones del tipo: puntuales, sin sentido e inserciones en el cuerpo del gen, asimismo el locus del gen *CTCF* frecuentemente se encuentra deletada (Filippova et al. 1998a; Aulmann et al. 2003; Kemp et al. 2014). En tumores de Wilms, mama y próstata se han reportado mutaciones en los ZFs 3 y 7 (Figura 11), las cuales impiden la unión de CTCF a secuencias reguladoras de genes como *MYC*, *Igf2* y *p19ARF* (Filippova et al. 2002).

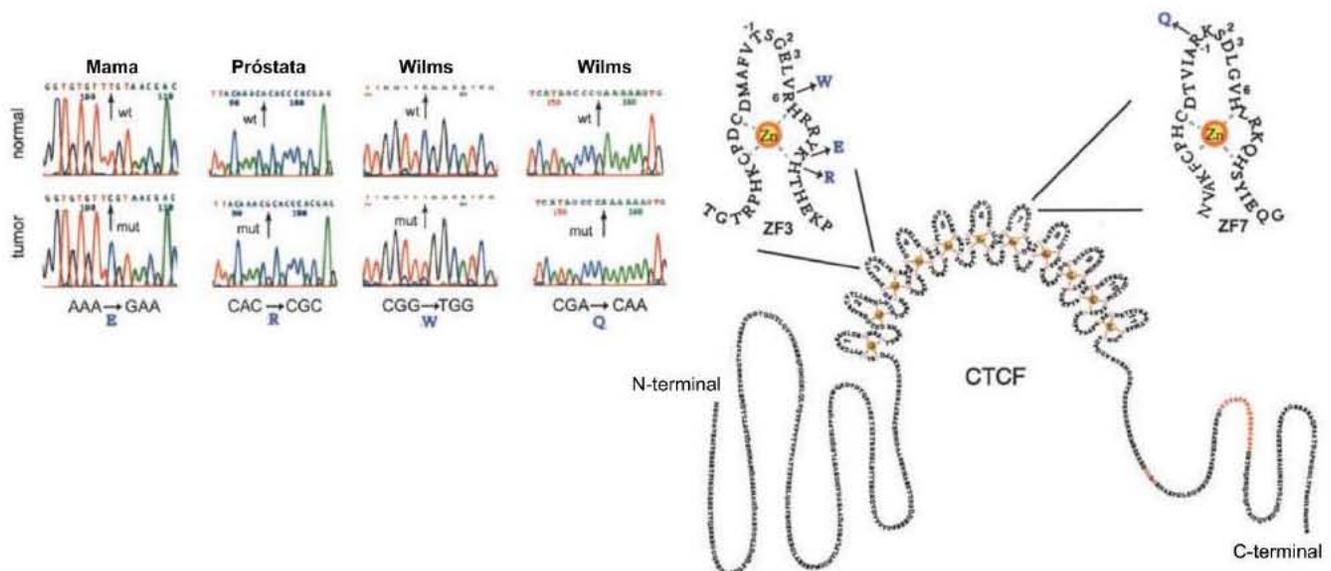


Figura 11. Mutaciones en los dedos de Zinc de CTCF

La imagen de la izquierda muestra las mutaciones que se han encontrado en tumores de mama, próstata y Wilms, las cuales resultan en sustituciones de aminoácidos en los dedos de Zinc 3 y 7, como se muestra en la imagen derecha, estas posiciones son críticas para el reconocimiento del DNA y para la formación de los dedos de Zinc (Modificado de Filippova et al. 2002).

2. Modificaciones en los CTSs

La unión de CTCF a sus CTSs puede ser alterada por mutaciones o metilación del DNA en estas secuencias blanco, los genes reportados cuya expresión se desregula por modificaciones en los CTSs son numerosos. En el caso de mutaciones, éstas pueden ocurrir en promotores o regiones de control de impronta, como se ha reportado para los genes *Xist* y *H19* respectivamente; además se ha reportado que las secuencias blanco del complejo CTCF-cohesina presentan alto porcentaje de mutaciones (Katainen et al. 2015; Pugacheva et al. 2005; Pant et al. 2004). En cuanto a metilación, la ganancia de metilación en secuencias regulatorias de genes como *hTERT*, *p16INK4a*, *BRCA1*, *Rb*, *RASSF1A*, *RARRES1*, *miR-125b1* (Figura 12), entre otros, impide la unión de CTCF (S Renaud et al. 2007; Witcher & Emerson 2009; Butcher & Rodenhiser 2007; De La Rosa-Velázquez et al. 2007; Chang et al. 2010; Peng et al. 2012; Saito & Saito 2012).

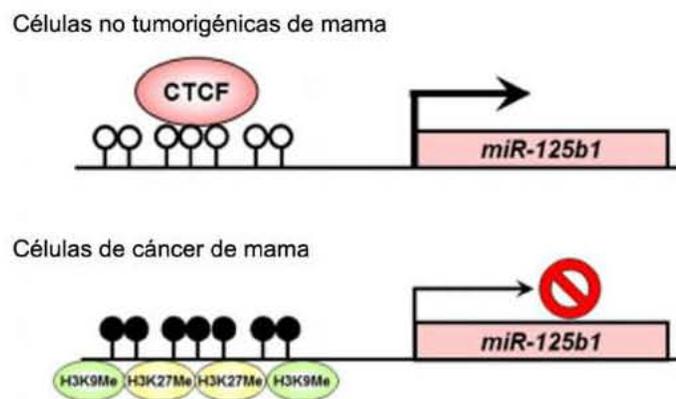


Figura 12. Modificación de CTSs por metilación del DNA

El CTSs del promotor del gen *miR-1251*, al presentar metilación del DNA y en este caso marcas de histonas represivas, evita que CTCF se una a esta región y se inactiva la expresión, esto se ha reportado en cáncer de mama (Saito & Saito 2012)

3. Modificaciones post-traduccionales

Como se mencionó en párrafos anteriores, CTCF puede presentar tres tipos de modificación post-traduccionales: fosforilación, poli(ADP)ribosilación y sumoilación, lo cual le permite interactuar con otras proteínas y le confiere mayor versatilidad funcional. Diversos reportes demuestran la importancia de dichas modificaciones en CTCF, las cuales le permiten regular la expresión en

tiempo y espacio, cuando estas señales se alteran y cambian los perfiles de expresión se puede promover la proliferación celular descontrolada.

La fosforilación de CTCF es dinámica durante el ciclo celular, regula la transcripción positiva y negativamente, un ejemplo ampliamente reportado es *c-myc*, cuya expresión ocurre solamente cuando CTCF fosforilada se une al promotor de dicho gen (El-Kady & Klenova 2005; Klenova et al. 2001; Reimand et al. 2013). CTCF poli(ADP)ribosilada lleva a cabo diversas funciones, como: *insulator*, elemento barrera, protege al DNA de metilación, inhibe transcripción de genes ribosomales y su ausencia o bajos niveles se han asociado con aumento en la tasa de proliferación celular (Yu et al. 2004; Witcher & Emerson 2009; Torrano et al. 2006; Docquier et al. 2009; Zampieri et al. 2012). Por otra parte, CTCF sumoileda participa en el bloqueo de *enhancers* y en la represión transcripcional, la hipoxia y el estrés oxidativo provocan una disminución de CTCF sumoileda, lo cual se ha relacionado con la activación transcripcional de genes implicados en el desarrollo del cáncer (Kitchen & Schoenherr 2010; Wang et al. 2012).

4. Localización celular y niveles de expresión

CTCF es una proteína nuclear con distribución dinámica durante el ciclo celular, sin embargo en células de cáncer, la localización y los niveles de expresión parecen estar relacionados con la progresión tumorigénica (Zhang et al. 2004). En muestras de carcinoma mamario, se observó que CTCF tiene una localización nuclear y citoplásmica, donde la expresión nuclear correlacionó con tumores de bajo grado y una actividad proliferativa baja, mientras que la baja o ausencia de expresión nuclear se asoció a tumores de alto grado, mientras que la expresión citoplásmica se correlacionó con tumores pequeños e invasión vascular (Rakha et al. 2004).

Estudios en líneas celulares de cáncer de mama han mostrado que la sobre expresión de CTCF inhibe la proliferación celular sin inducción de apoptosis, lo que sugiere que la ausencia del factor CTCF puede estar asociada a la actividad transcripcional de genes que promueven la proliferación celular y por tanto la tumorigénesis (Rasko et al. 2001; Tiffen et al. 2013). Otro estudio, en el que se observó que los niveles de CTCF son variables a nivel de proteína y mRNA en diferentes líneas celulares de mama, reportó que la ausencia de CTCF correlaciona con la apoptosis (Docquier et al. 2005). La contraposición de ambos estudios enfatizan la acción dual de CTCF en cáncer, la cual

depende de sus modificaciones post-traduccionales, proteínas con las que interactúa, metilación de los *CTS*s, el contexto celular, entre otros factores, no obstante las redes de regulación en el genoma llevadas a cabo por CTCF se encuentran alteradas (Tiffen et al. 2013).

CTCF tiene una expresión diferencial entre varias líneas celulares tumorogénicas de mama, lo cual resulta interesante ya que nos permite suponer un papel importante de CTCF en la carcinogénesis. En este sentido, en cáncer de mama se ha reportado una desregulación en la expresión génica a causa de CTCF en los genes *BRCA1*, *RASSF1A*, *CDH1*, *p16INK4a*, *BAX*, *RARRES* y *miR125b* (Butcher & Rodenhiser 2007; Chang et al. 2010; Witcher & Emerson 2009; Méndez-Catalá et al. 2013; Peng et al. 2012; Soto-Reyes et al. 2012). Además de la función de CTCF en la regulación de la expresión, la haploinsuficiencia de este factor, se ha relacionado con una mayor tasa de proliferación celular y con mayor capacidad de desarrollar tumores (Kemp et al. 2014).

A pesar de las evidencias, que demuestran la desregulación de las funciones llevadas a cabo por CTCF en cáncer, actualmente los genes blancos de CTCF no son bien conocidos. Por lo anterior, estamos interesados en investigar, en un modelo *in vitro* de cáncer de glándula mamaria: 1. Genes blanco de CTCF, 2. Predecir los posibles mecanismos de regulación de CTCF sobre sus genes blanco y 3. Correlacionar genes blancos de CTCF con eventos celulares específicos, con la finalidad de entender mejor la participación de CTCF sobre la biología de esta enfermedad.

4. OBJETIVOS

Objetivo general

Determinar los genes regulados transcripcionalmente por CTCF en la línea celular de cáncer de glándula mamaria MCF-7.

Objetivos particulares

1. Evaluar la expresión del factor CTCF en la línea celular de cáncer de glándula mamaria MCF-7.
2. Abatir la expresión de CTCF en una línea celular
3. Identificar genes diferencialmente expresados en ausencia de CTCF
4. Evaluar la unión de CTCF en regiones regulatorias de los genes encontrados y su relación con metilación y marcas de histonas.

5. METODOLOGÍA

Cultivo celular

Las líneas celulares MCF-7, MDA-MB-231 se crecieron en medio RPMI-1640 (GIBCO), mientras que las líneas MDA-MB-453 y MDA-MB-468 se cultivaron en medio DMEM; los medios se suplementaron con suero fetal bovino inactivado (SFB, GIBCO) al 10%, antibióticos (penicilina 100U/ml y estreptomina 100µg/ml), antimicótico (anfotericina B 0.25 µg/ml) y L-glutamina 2mM (GIBCO); los cultivos se mantuvieron a 37°C en atmósfera del 5% de CO₂ y una humedad relativa del 100%. Todas las líneas se obtuvieron de la *American Type Culture Collection* (ATCC). Para el abatimiento del factor CTCF se transdujeron cultivos celulares de MCF-7, con RNA de interferencia, en un sistema lentiviral con un *cassette* de resistencia a puromicina. Los cultivos controles se designaron como MCF-7 *mock* y el experimental como MCF-7 *shCTCF*; se mantuvieron con medio RPMI-1640 (GIBCO), suplementado con suero fetal bovino inactivado (SFB, GIBCO) al 10%, antibióticos (penicilina 100U/ml y estreptomina 100µg/ml), antimicótico (anfotericina B 0.25 µg/ml), L-glutamina 2mM (GIBCO) y 5µg/ml de puromicina; a 37°C en atmósfera del 5% de CO₂ y una humedad relativa del 100%.

Producción de partículas lentivirales y transducción

Las partículas lentivirales se produjeron en la línea celular HEK293ft, con los vectores pVSVG, pRSV/REV, pMDLg/RRE, y con el vector de expresión pLL3.7, los cuales codifican para las diferentes partículas de ensamble del virus, el vector de expresión contiene la secuencia *short hairpin* (*shRNA*) contra CTCF. La línea celular HEK293ft se cultivó en medio DMEM, SFBi al 10%, antibiótico estreptomina/ampicilina (*Invitrogen*), L-glutamina 200mM (GIBCO) y piruvato de sodio. Los vectores se agregaron al cultivo de HEK293ft en una solución de CaCl₂, 48 horas después se colectó el sobrenadante y se centrifugó a 2,000rpm por 7 minutos a 4°C, se recuperó el sobrenadante y se centrifugó por 90 minutos a 25,000rpm, se eliminó el sobrenadante, se dejaron secando por 10 minutos, posteriormente se agregaron 50µl de PBS frío pH 7.4 y se incubó a 4°C

por una noche, las partículas virales se resuspendieron en el PBS y se alicuotaron en 5 tubos, cada uno con 10 μ l, finalmente se guardaron a -80°C .

Para la transducción de las células MCF-7 se utilizaron dos cultivos con 2×10^5 células en botellas de cultivo rectangulares Corning, con área de crecimiento de 25 cm^2 , se preparó medio RPMI-1640 suplementado con SFBi al 10%, polibreno $8 \mu\text{g}/\text{ml}$ y una alícuota viral (sin antibióticos), se incubó a 37°C en atmósfera del 5% de CO_2 y una humedad relativa del 100%, al día siguiente se adicionó medio de cultivo nuevo y se mantuvieron en incubación durante 2 días, después del periodo de transducción las células se mantuvieron con RPMI-1640 suplementado con SFB al 10% y puromicina ($5 \mu\text{g}/\text{ml}$), los cultivos controles se designaron como MCF-7 *mock* y el experimental como MCF-7 *shCTCF*.

Extracción de proteínas totales

Para la extracción de proteínas se usaron cultivos celulares con una confluencia mínima del 70%. Los cultivos se lavaron dos veces con 10ml de PBS frío pH 7.4 y se adicionaron 150-200 μ l de buffer de lisis (1ml RIPA, 40 μ l Complete EDTA Free 25x y 10 μ l PMSF 100mM). Se incubó por 5 minutos en hielo, posteriormente se recuperó el lisado con un gendarme en un tubo eppendorf y se centrifugó a 14,000 rpm por 15 minutos a 4°C . Se recuperó el sobrenadante y se determinó la concentración de proteínas por el método de Bradford. Finalmente las proteínas se conservaron a -20°C en buffer Laemmli al 1X.

Western Blot

Para el ensayo de *Western Blot*, la electroforesis de las proteínas se realizó en un gel separador SDS-PAGE al 7% y un gel concentrador al 4%, se usaron 40 μ g de proteínas totales y se corrieron en la cámara de electroforesis con buffer de corrida 1X pH 8.3 (buffer de corrida 1 litro: 3.02g Tris Base, 14.4g Glicina, 1g SDS), como marcador de peso molecular se utilizó *Precision Plus Protein Prestained Standards Dual Xtra* de Bio-Rad. Las proteínas del gel SDS-PAGE se transfirieron a una membrana de PVDF (*Polyvinylidene Fluoride*) con el equipo *Trans-Blot Description & Assembly of Parts* de Bio-Rad, que se llevó a cabo a 10°C con buffer de transferencia (buffer de transferencia 2 litros: 6g Tris-base, 29g glicina, 200ml metanol) y 30-62 volts durante toda la noche. La membrana se cortó justo

por arriba de la banda de 75kDa, la parte superior se utilizó para identificar la proteína problema, CTCF (130kDa), y la parte inferior para identificar a la proteína constitutiva α -tubulina (55kDa). El bloqueo de la membrana se realizó con TBS1X-Tween 0.1%-Leche Svelty 5% por una hora a temperatura ambiente y en agitación, después del bloqueo se realizaron 3 lavados de 5 minutos con TBS1X-Tween 0.1%. El anticuerpo primario se preparó en TBS 1X-Tween 0.1%-BSA 1%, para el caso de CTCF se utilizó una dilución 1:2,000 (Anti-CTCF *Polyclonal Antibody Millipore* Cat. # 07-729) y para α -tubulina la dilución fue de 1:15,000 (Santa Cruz), los anticuerpos primarios se incubaron a 4°C toda la noche en agitación. Los anticuerpos secundarios se usaron en dilución 1:10,000 (anti-conejo para CTCF y anti-ratón para α -tubulina, Santa Cruz), se incubó por 45 minutos a temperatura ambiente, posteriormente se hicieron 3 lavados de 5 minutos, dos de ellos con TBS 1X-Tween 0.1% y un último lavado con TBS 1X. Las proteínas se detectaron por quimioluminiscencia con *SuperSignal West Pico Chemiluminescent Substrate™* de Thermo Scientific.

Microarreglos de expresión

Las líneas celulares se disgregaron con el *TissueLiser™* (Qiagen Valencia, CA, USA) por 60 segundos a 30Hz, posteriormente se purificaron con el kit *RNeasy mini Kit™* (Qiagen Valencia, CA, USA) bajo el protocolo de la casa comercial. La concentración y calidad del RNA purificado, se analizó con el NanoDrop ND-1000 espectrofotómetro y con el Bioanalizador 2100 de *Agilent (Agilent Technologies, Santa Clara, CA, USA)* respectivamente.

El microarreglo se realizó por triplicado para MCF-7 *mock* y MCF-7 *shCTCF*, con el sistema *GeneChip® Human Gene 1.0 ST Array de Affymetrix*, el cual está diseñado con 764,885 sondas que reconocen los exones de 28,869 genes, se utilizó el kit *Ambion® WT Expression* para generar cDNA a partir de 100ng de RNA total, este cDNA se transcribió *in vitro* para producir cRNA (la hebra antisentido) y a partir de este se obtuvo nuevamente cDNA, es decir, se genera la cadena complementaria al cRNA antisentido. El siguiente paso consistió en el uso del kit *Affymetrix® Reagent* para generar fragmentos de 40-70nt, con las enzimas APE1 y UDG, los cuales se marcaron con biotina. El cDNA marcado y fragmentado se hibrida en los microarreglos, de acuerdo a los parámetros establecidos en el manual *GeneChip® WT, Terminal Labeling and Hybridization* del kit

Ambion[®] *WT Expression*. Se escanearon los microarreglos con el escáner *GeneChip*[®] *Scanner 3000* 7G bajo el protocolo establecido *GeneChip*[®] *Expression Wash, Stain and Scan User Manual*. Finalmente, el análisis de los microarreglos se realizó con el software *Partek Genomics Suite 6.5v*.

Extracción de RNA

Para la extracción de RNA se usaron cultivos celulares con una confluencia mínima del 70%, los cultivos se lavaron 3 veces con PBS pH 7.4 y se agregaron 150-300µl de Trizol[®] *RNA Isolation Reagent*, los lisados se depositaron en tubos eppendorf y se les adicionó 200µl de cloroformo por cada mililitro de Trizol usado, se agitaron 15 segundos en vórtex y se incubaron por 3 minutos a temperatura ambiente, transcurrido el tiempo se centrifugaron a 12,000g por 15 minutos a 4°C. Se recuperó la fase acuosa en tubos eppendorf nuevos a los cuales se les agregaron 250µl de isopropanol, 1µl de acetato de sodio 3M por cada 10µl de muestra, se agitó en vórtex 15 segundos dejando reposar 10 minutos a temperatura ambiente, se centrifugó a 12,000g por 10 minutos a 4°C y se eliminó el sobrenadante, posteriormente se adicionó 1ml etanol al 75% frío, se agitó en vórtex y se centrifugó a 7,500g por 5 minutos a 4°C, se eliminó el etanol, las muestras se dejaron secar y posteriormente se resuspendieron en 22µl de agua libre de RNasas.

Transcripción reversa, PCR punto final y PCR cuantitativa

Para la síntesis de cDNA se usaron 1000ng de RNA total, a los cuales se les adicionó 1U de DNasa I, 1µl de buffer DNasa (10X), se llevó a un volumen final de 10µl y se incubó por 30 minutos a 37°C, posteriormente se agregó 1µl EDTA 25Mm y se incubó 10 minutos a 65°C. Finalmente se adicionó el mix *SuperScript*[™] *II Reverse Transcriptase* de *Invitrogen, Life Technologies*: 1µl Random primers (50ng/µl), 1µl dNTPs (10Mm), 2µl DTT (0.1M), 1µl inhibidor RNasa (40U/µl), 4µl Buffer M-MLV (5X) y 1µl Enzima M-MLV (200U/µl), y se incubó a 37°C-15min, 25°C-10min, 37°C-60min, 70°C-10min y 4°C-∞.

Las reacciones de PCR punto final se realizaron con *KAPA2G Fast HotStart ReadyMix*, cada reacción se preparó con 25ng de RNA, 1µl de *primers* 10µM y ajustando a un volumen final de 15 µl; se incubó a 95°C-3min, (95°C-15s, 60°C-15s, 72°C-5s) (35 ciclos), 72°C-1min y 4°C-∞; los productos se visualizaron en geles de agarosa 1%.

Las PCR cuantitativas se realizaron con *Maxima SYBR Green/ROX qPCR Master Mix (2X)*, ajustando las reacciones a un volumen final de 10µl, con 10ng de templado, los *primers* se utilizaron a una concentración de 0.5µM (Anexo, Tabla 1), el programa de amplificación consistió en: 95°C-10min, (95°C-15s, 60°C-50s) (33ciclos), las cuales se realizaron en el sistema *StepOne™ Real-Time PCR System de Applied Biosystems*. Para todos los *primers* empleados se realizaron las curvas estándar y curvas *melting* (Anexo, Figura 25), como *housekeeping* se utilizó el gen ribosomal *OAZ1*. La expresión de los genes problema se determinó mediante cuantificación relativa (Pfaffl 2004).

Análisis de metilación por MSP (Methylation Specific PCR)

El DNA empleado se obtuvo a partir de cultivos celulares con el kit *Quick-gDNA™ MiniPrep de Zymo Research*, los productos se cuantificaron y se evaluó su integridad mediante geles de agarosa. La modificación con bisulfito sodio se realizó con el kit *EZ-DNA Methylation-Gold™ de Zymo Research*, se utilizaron 500ng de DNA y se siguió el protocolo de la casa comercial. Las reacciones de PCR se realizaron con *KAPA2G Fast HotStart ReadyMix*, con 1µl de eluido, 1µl de *primers* 10µM y ajustando a un volumen final de 15 µl; se incubó a 95°C-3min, (95°C-15s, 60°C-15s, 72°C-5s) (35 ciclos), 72°C-1min y 4°C-∞; los productos se visualizaron en geles de agarosa 2%. Los primers para la evaluación de metilación del promotor de *EFEMP1* se muestran en el Anexo, Tabla 2 (Yang et al. 2013; Nomoto et al. 2010); para la identificación de islas CpG “*MethPrimer CpG island prediction*” y *Methyl Primer Express v1.0 de Applied Biosystems* (<http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi>).

Inmunoprecipitación de la cromatina (ChIP, *Chromatin Immunoprecipitation*)

El ensayo de ChIP se realizó con aproximadamente 1x10⁷ células MCF-7, para el *crosslinking* se incubó por 10 minutos con formaldehído al 1% y se neutralizó la reacción con 1ml de glicina 2.5M. Todos los lavados se hicieron con PBS suplementado con PMSF, asimismo el buffer de lisis se suplementó con PMSF y Complete 25x. La cromatina se sonicó con 9 pulsos, durante 0.5 segundos *on/off*; con una amplitud del 35%, con la finalidad de obtener fragmentos de 300-600pb, posteriormente se evaluó la integridad y tamaño de los fragmentos de cromatina mediante geles de agarosa y SDS-PAGE. La inmunoprecipitación se realizó con los anticuerpos *Biotinylated Goat Anti-*

Rabbit IgG Ab (Vector Laboratories #BA-1000), Anti-CTCF Polyclonal Antibody Millipore (Cat. # 07-729), Anti-H4K16ac Polyclonal Antibody Millipore (Cat. #07-329), Anti-H3K27me3 Polyclonal Antibody Millipore (Cat. #07-449) y se precipitaron con *Magna ChIPTM Protein A+G Magnetic Beads* (Cat. #16-663); antes de cada inmunoprecipitación, los fragmentos de cromatina se incubaron con perlas magnéticas durante 2 horas a 4°C. Los inmunoprecipitados fueron lavados y se sometieron a *descrosslinking* e incubaron con proteinasa K y RNasa, finalmente el DNA se obtuvo con fenol-cloroformo-alcohol isoamílico (25:24:1) y se resuspendieron en 50µl. Las reacciones de PCR se realizaron con *KAPA2G Fast HotStart ReadyMix*, con 1µl de templado, 1µl de *primers* 10µM y ajustando a un volumen final de 15 µl; se incubó a 95°C-3min, (95°C-15s, 60°C-15s, 72°C-5s) (35 ciclos), 72°C-1min y 4°C-∞; los productos se visualizaron en geles de agarosa 1%. Para la identificación de sitios de unión a CTCF se emplearon las bases de datos *Encyclopedia of DNA Elements* (ENCODE) y *Roadmap Epigenome Browser V1.19*. Los *primers* utilizados se muestran en Anexo, Tabla 3.

Análisis bioinformático

La identificación de secuencias promotoras se realizó con el programa “*The Eukaryotic Promoter Database*” (<http://epd.vital-it.ch/>). Usamos el programa “*MethPrimer CpG island prediction*” (<http://www.urogene.org/methprimer/>) para evaluar si los promotores son de tipo HCP (*High CpG Promoters*), bajo los parámetros: tamaño ≥100pb, 50% de contenido GC y frecuencia observada/esperada de 0.6. Posteriormente, buscamos sitios de unión a CTCF en los promotores y hasta -15,000 pb río arriba del sitio de inicio a la transcripción, con el programa “*CTCFBSDB 2.0*” de Ziebarth et al (2013), en la línea celular MCF-7 (Human, hg19) y *Encyclopedia of DNA Elements* (ENCODE) y *Roadmap Epigenome Browser V1.19*. Para la búsqueda de *enhancers* se usó la herramienta “*VISTA Enhancer Browser*” (<http://enhancer.lbl.gov/>).

6. RESULTADOS

Evaluación de la expresión proteica de CTCF en líneas celulares de cáncer de mama

Ensayos previos de *Western Blot* en nuestro laboratorio nos permitieron observar que la línea celular de cáncer de glándula mamaria MCF-7, comparada con otras líneas celulares de mama, muestra mayor cantidad de CTCF, tanto en extractos totales como en extractos nucleares y citoplásmicos (datos no publicados). Debido a lo anterior, verificamos los niveles proteicos de CTCF en las líneas celulares de cáncer de mama MCF-7, MDA-MB-231, MDA-MB-453 y MDA-MB-468, a partir de extractos totales de proteína, realizamos un análisis densitométrico del *Western Blot* con *ImageJ* (Figura 13). Los resultados corroboran que MCF-7 tiene mayor abundancia de CTCF, con base en este resultado, decidimos emplear la línea celular MCF-7 para determinar genes blancos de CTCF.

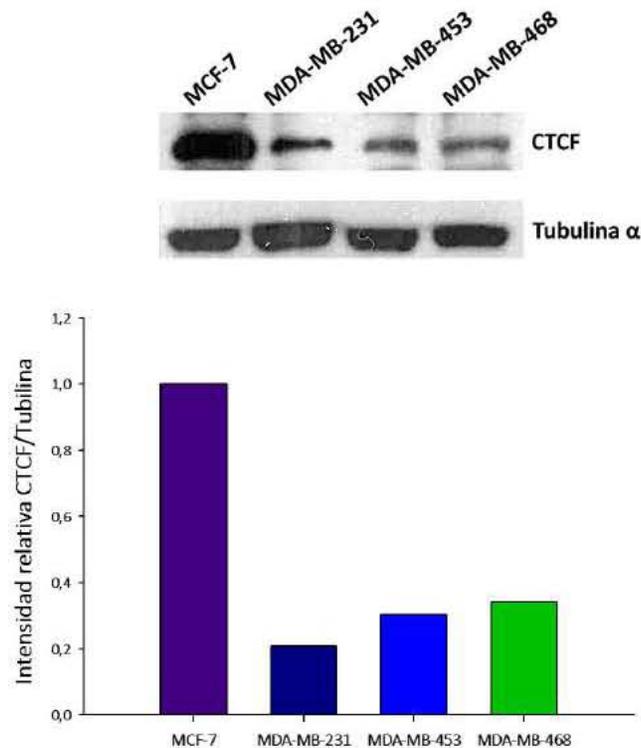


Figura 13. Expresión de CTCF en líneas celulares de cáncer de mama

Western Blot a partir de extractos totales de proteína de las líneas celulares MCF-7, MDA-MB-231, MDA-MB-453 y MDA-MB-468, la película del *Western Blot* se escaneó y cuantificó con *Image J*, para comparar la abundancia de CTCF en las diferentes líneas celulares, se determinaron los radios de intensidad de la señal de CTCF en MCF-7, respecto a la intensidad de la señal de tubulina α .

Abatimiento de la expresión del factor nuclear CTCF en la línea celular MCF-7

Después de verificar los altos niveles proteicos de CTCF en MCF-7, abatimos su expresión mediante un sistema lentiviral, con el objetivo de determinar los posibles cambios en la actividad transcripcional de genes, sugerentes de la participación de CTCF sobre su regulación.

Para dicho abatimiento, se utilizó el vector pLL3.7 para clonar la secuencia interferente (*shRNA*, *short hairpin*) contra el mRNA de CTCF, las células se denominaron MCF-7 *shCTCF* y MCF-7 *mock* (control). Después de la transducción viral, los cultivos celulares se mantuvieron en selección con puromicina, cabe mencionar que originalmente el vector pLL3.7 tiene un *cassette* que codifica a GFP (*Green Fluorescent Protein*), el cual fue reemplazado con un *cassette* de resistencia a puromicina. Para corroborar el abatimiento de la proteína CTCF, se realizó un ensayo de *Western Blot* con extractos de proteínas totales, los resultados permitieron confirmar el abatimiento de CTCF (Figura 14).

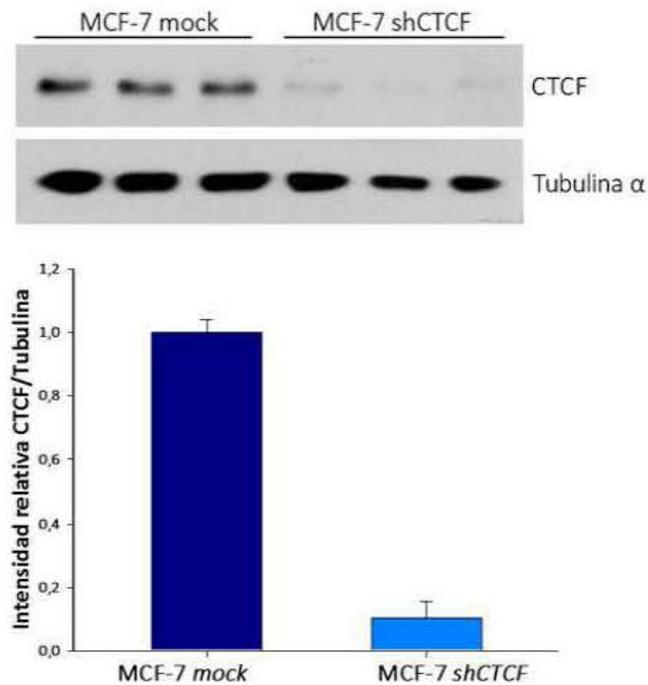


Figura 14. Abatimiento de CTCF en la línea celular MCF-7 (*shCTCF*)

Arriba se muestra la película del *Western Blot*, donde se observa el abatimiento de la proteína como efecto del *shCTCF*. Abajo se presenta la gráfica correspondiente al análisis densitométrico, donde las barras corresponden a la desviación estándar.

Perfiles del transcriptoma en MCF-7 *mock* y *shCTCF*

Con el objetivo de determinar genes blanco de CTCF en cáncer de mama, se realizaron microarreglos de expresión, con el sistema *GeneChip® Human Gene 1.0 ST Array de Affymetrix*. Los datos obtenidos se agruparon y analizaron mediante el sistema denominado “Análisis de Componentes Principales” (PCA, *Principal Component Analysis*), con el *software Partek Genomics Suite 6.5v*, el cual mostró diferencias globales del transcriptoma entre MCF-7 *mock* y MCF-7 *shCTCF* (Figura 15).

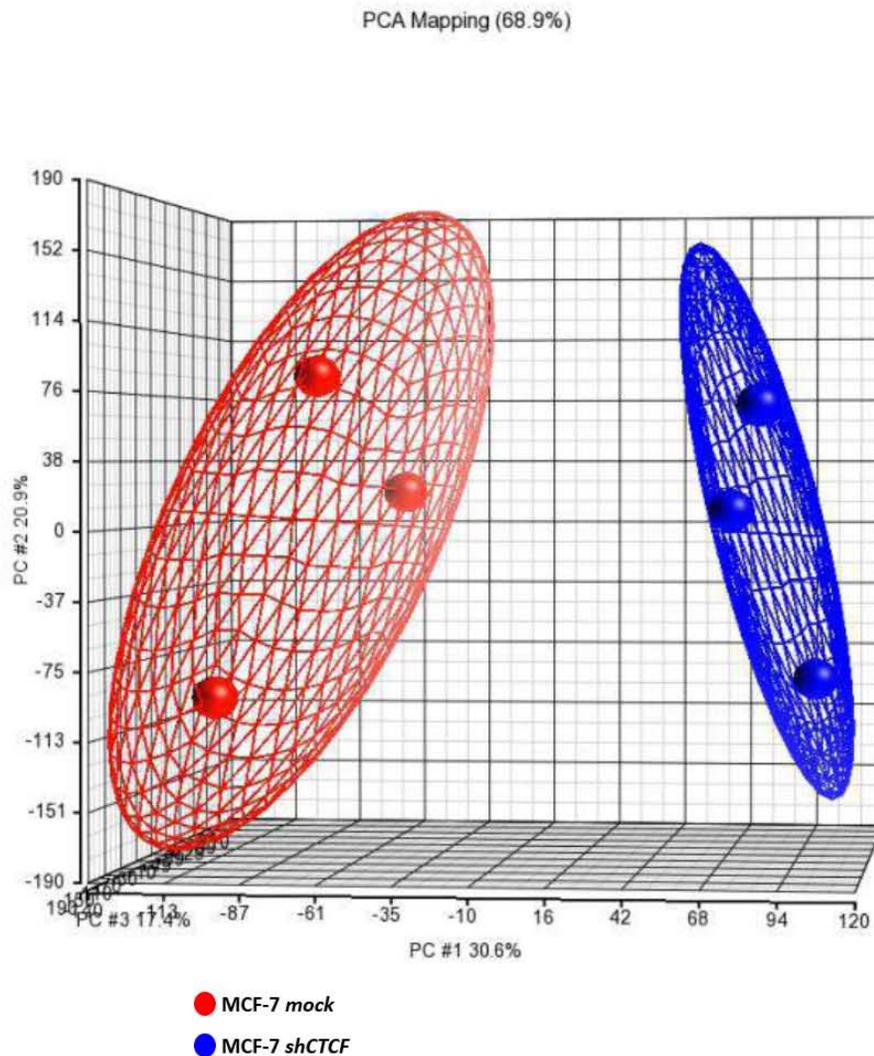


Figura 15. Análisis de componentes principales

El PCA muestra que existen diferencias a nivel global del transcriptoma entre MCF-7 *mock* (grupo rojo) y MCF-7 *shCTCF* (grupo azul).

Genes expresados diferencialmente MCF-7 *shCTCF* respecto a MCF-7 *mock*.

Para determinar genes expresados diferencialmente y correlacionarlos con la acción de CTCF, se analizaron los datos mediante una ANOVA con el programa *Partek Genomics Suite 6.5v*, utilizando el valor $p < 0.05$ y un *fold change* de 1.2 (cambio de expresión $\geq \pm 120\%$). De este análisis obtuvimos un *cluster* jerárquico, el cual muestra 15 genes que se encuentran diferencialmente expresados, 9 sobre expresados (*FXYD3*, *GDF15*, *RASGRP1*, *CMBL*, *PADI3*, *CEACAM5*, *MACF1*, *NBPF4*, *SEMA6A*) y 6 sub expresados (*EFEMP1*, *CTCF*, *GFRA1*, *DLX1*, *BAMBI*, *TGFB2*) (Figura 16). Dentro de los genes reprimidos, en MCF-7 *shCTCF* se observó al gen CTCF, lo cual permitió corroborar el efecto del *short hairpin* (Figura 17, anexo tabla 4).

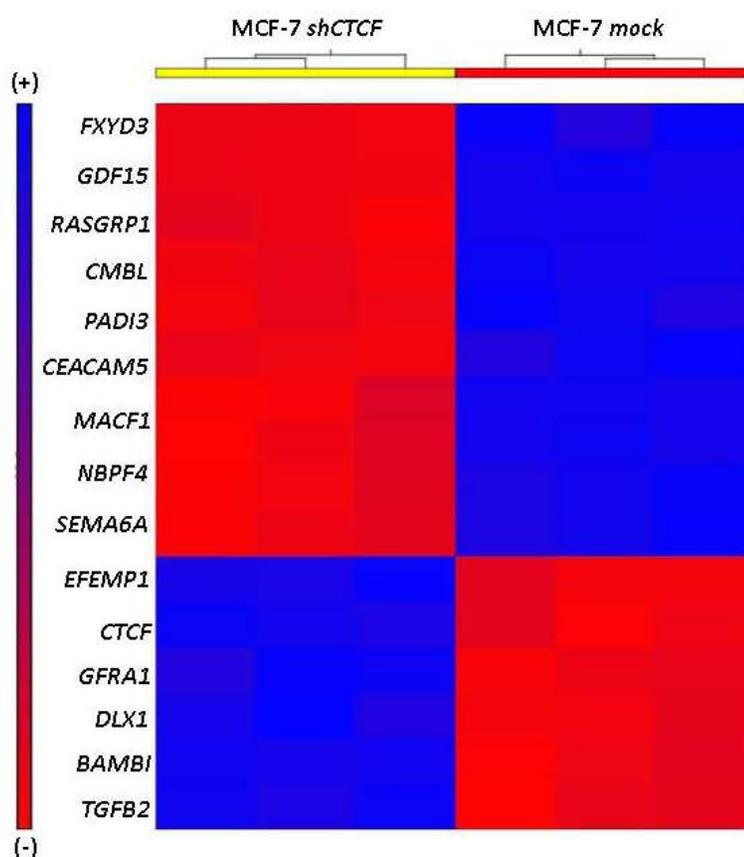


Figura 16. *Cluster* jerárquico de los genes diferencialmente expresados en MCF-7 *shCTCF*.

En MCF-7 *shCTCF* se sobre expresan los genes *FXYD3*, *GDF15*, *RASGRP1*, *CMBL*, *PADI3*, *CEACAM5*, *MACF1*, *NBPF4* y *SEMA6A*; se sub expresan los genes *EFEMP1*, *CTCF*, *GFRA1*, *DLX1*, *BAMBI* y *TGFB2*, ($p < 0.05$ y un *fold change* de 1.2).

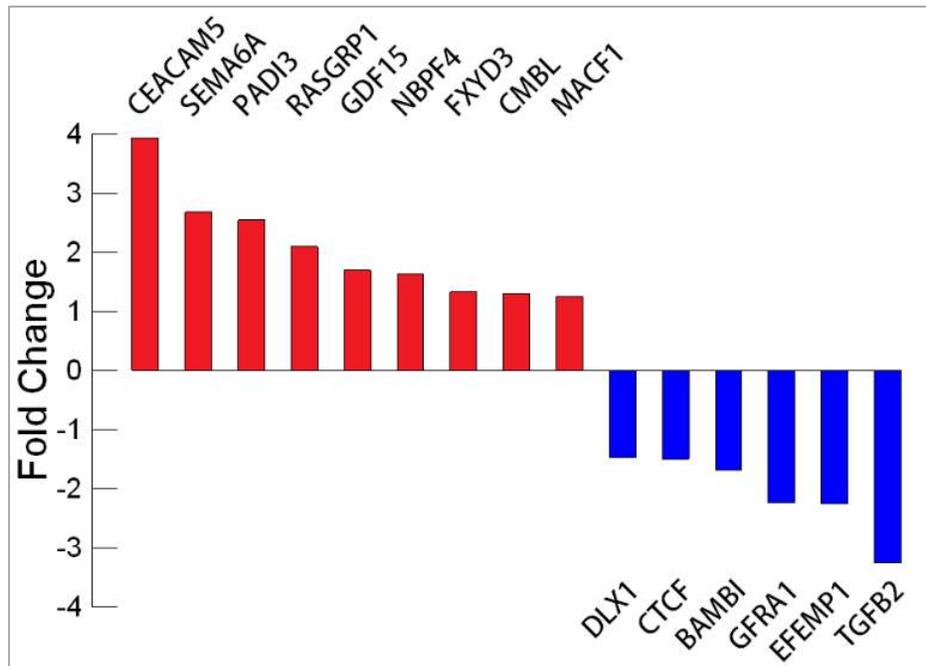


Figura 17. Cambio de expresión génica en MCF-7 *shCTCF* respecto a MCF-7 *mock*.

Los genes sobre expresados están dentro de un rango de *fold change* de 1.2 a 4, es decir se expresan de 120% a 400% más, respecto a MCF-7 *mock*. Los genes sub expresados tienen un rango del *fold change* de -1.2 a -3.3, esto significa que su expresión es de 120% a 330% menos que en MCF-7 *mock*.

Validación de microarreglo para *CTCF* y *EFEMP1*

Con la finalidad de validar los datos obtenidos del microarreglo, cuantificamos la expresión de los genes *CTCF* y *EFEMP1*, por RT-qPCR. Los resultados confirman, que cuando *CTCF* está abatido, *EFEMP1* se encuentra sub expresado, asimismo se verificó el abatimiento del mRNA de *CTCF* (Figura 18). Estos datos, sugieren que *CTCF* podría tener un papel positivo en la expresión de *EFEMP1*, por lo que decidimos buscar sitios de unión a *CTCF* en regiones regulatorias de dicho gen.

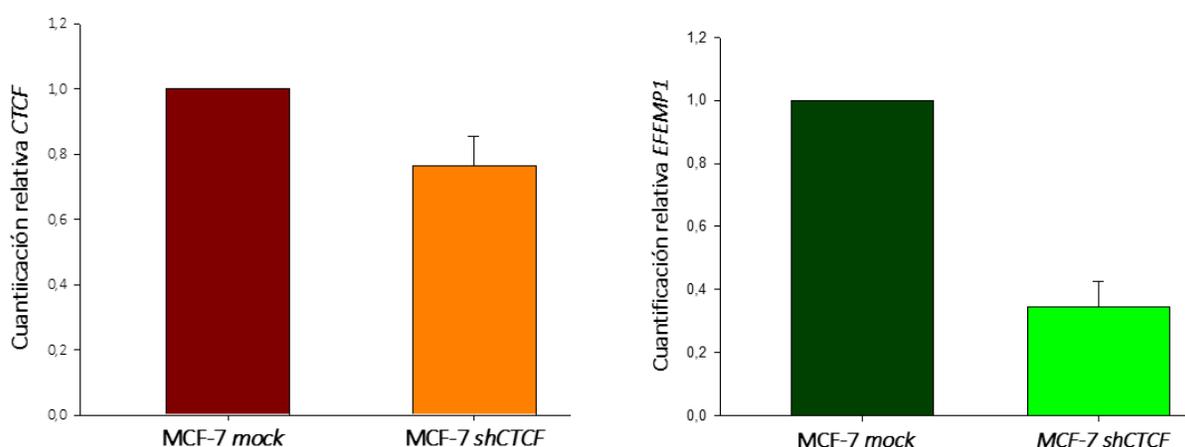


Figura 18. Niveles de expresión de *CTCF* y *EFEMP1* en MCF-7 *shCTCF*

A la izquierda se muestra la gráfica que representa los niveles de expresión de *CTCF* ($p < 0,1$), a la derecha se muestra la gráfica de cuantificación para *EFEMP1* ($p < 0,05$), ambos genes se encuentran sub expresados en MCF-7 *shCTCF*. Se gráfica desviación estándar en ambos casos de 3 experimentos.

Unión de *CTCF* a regiones regulatorias del gen *EFEMP1*

Para determinar si *CTCF* se une a regiones regulatorias del gen *EFEMP1*, usamos las bases de datos *Encyclopedia of DNA Elements* (ENCODE) y *Roadmap Epigenome Browser V1.19*, encontramos que hay dos posibles *CTCF target sites* (*CTSs*), uno de ellos se ubica a 4,167pb río arriba del promotor y otro sitio potencial se encuentra en el promotor del gen, el cual abarca una porción del primer exón, a estos sitios los nombramos *CTS 1* y *CTS 2*, respectivamente (Figura 19). Para determinar si estas regiones son reconocidas por *CTCF*, realizamos un ensayo de ChIP, para esto obtuvimos la secuencia de estos sitios, diseñamos oligonucleótidos para amplificar dichas regiones y usando DNA genómico de MCF-7 evaluamos que el amplicon correspondiera con el diseño y a la secuencia

(Anexo, Figura 26). Los resultados muestran que CTCF se une a las secuencias reportadas en la base de datos (Figura 20), por lo que nos permitimos suponer que esta proteína tiene efectos sobre la regulación en la transcripción de *EFEMP1*.

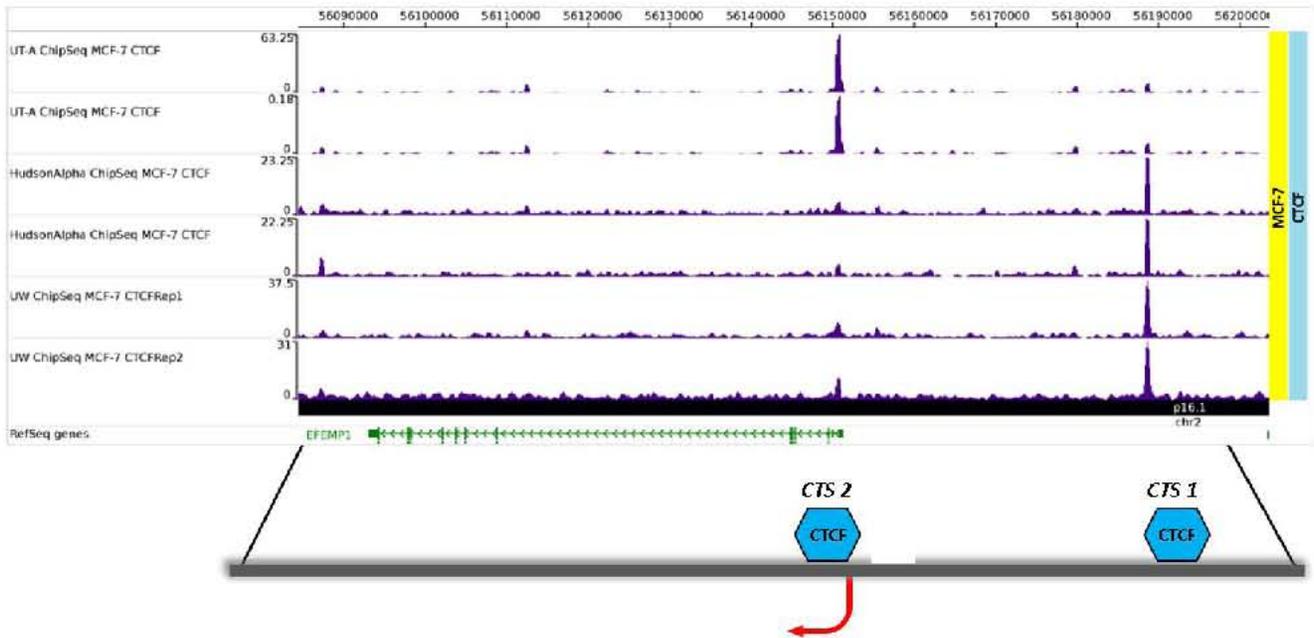


Figura 19. Sitios de enriquecimiento CTCF

Para el gen *EFEMP1* en la línea celular MCF-7, existe información de diferentes ChIP-seq, dos de ellos muestran un sitio de enriquecimiento para CTCF en la región promotora del gen (*CTS 2*) y los restantes muestran un sitio río arriba del promotor (*CTS 1*). El gen *EFEMP1* se ubica en el cromosoma 2p16.1 y en la cadena antisentido.

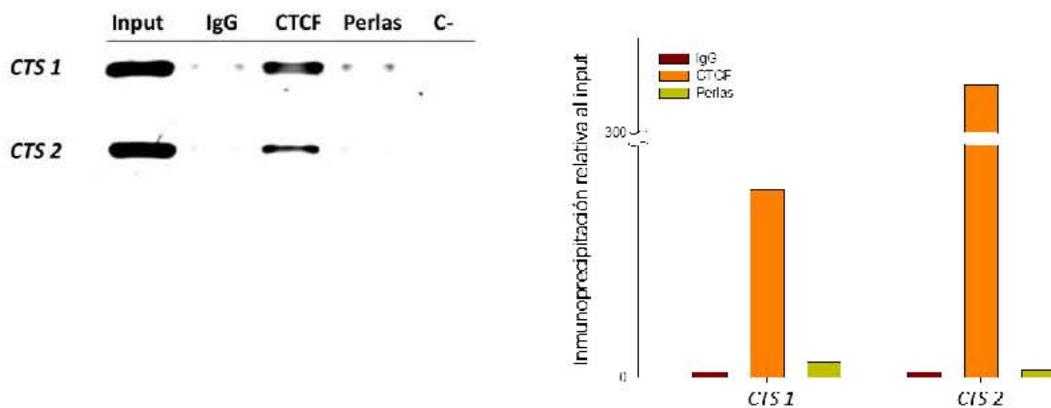


Figura 20. CTCF se une a los *CTS*s reportados para el gen *EFEMP1*

La unión de CTCF al *CTS 1* y *2* se evaluó por ChIP en la línea celular MCF-7, la imagen superior muestra los niveles de amplificación por PCR de los fragmentos inmunoprecipitados para estas regiones. La gráfica representa la densitometría del gel, realizada con *ImageJ*, donde se visualiza un enriquecimiento de CTCF en ambas regiones de DNA, en MCF-7.

CTCF como barrera de metilación del promotor *EFEMP1*

Una vez identificados los *CTS*s en el gen *EFEMP1* y teniendo como antecedente que la unión de CTCF protege de metilación en el DNA y promueve la expresión, decidimos evaluar si estos sitios se encuentran en una isla CpG, para esto empleamos los programas *MethPrimer* y *Methyl Primer Express*, bajo los parámetros: tamaño ≥ 200 pb, 50% de contenido GC y frecuencia observada/esperada de 0.6. Encontramos que el *CTS 2* se encuentra embebido en una isla CpG (Figura 21), estos resultados nos sugirieron que la expresión del gen *EFEMP1*, podría depender de la unión de CTCF al *CTS 2* y que esta secuencia debe encontrarse libre de metilación.

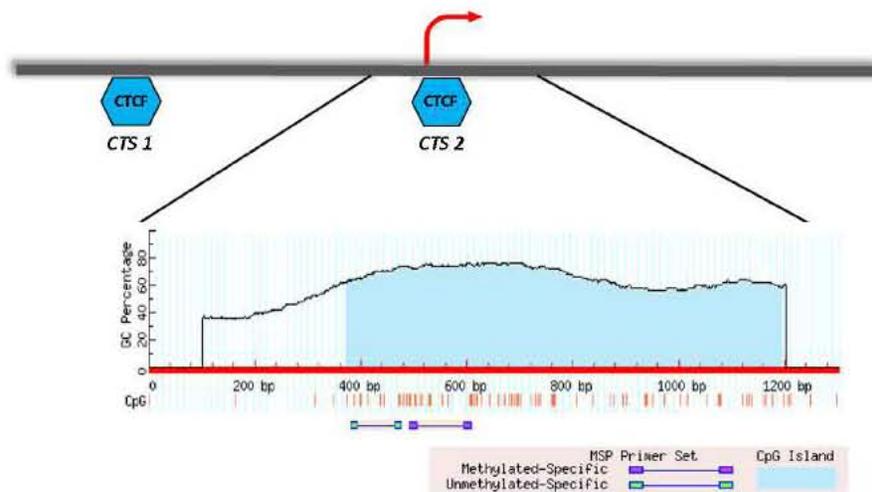


Figura 21. Promotor *EFEMP1* embebido en una isla CpG

La secuencia analizada para la identificación de isla CpG, parte de -300 a +1000 respecto al TSS del gen; el promotor y el primer exón del cuerpo del gen de *EFEMP1*, se encuentran embebidos en una isla CpG de alta densidad. La isla se muestra en color azul, representa el porcentaje de GC, bajo el eje X se muestran con líneas verticales rojas los dinucleótidos CG, asimismo se observan los sitios amplificados por los juegos de *primers* usados para la prueba MSP (Figura 23)

Para abordar esta idea, decidimos evaluar la expresión de *EFEMP1* en las líneas celulares MCF-7, MDA-MB-231, MDA-MB-453 y MDA-MB-468, con la finalidad de encontrar un modelo celular que no exprese el gen (Figura 22). Los resultados mostraron que la línea celular MDA-MB-453 es negativa a *EFEMP1*, por lo que si nuestra hipótesis se confirma, en este modelo celular el *CTS 2* debería presentar metilación, lo cual no ocurriría en los modelos celulares que sí tienen expresión de *EFEMP1*; para explorar esta idea realizamos ensayos *Methyl Specific Primers* (MSP), cabe

mencionar que los *primers* utilizados amplifican dentro de la secuencia reportada como *CTS 2* (Anexo, Figura 27). Encontramos que las líneas celulares MCF-7 y MDA-MB-231 no presentan metilación, la línea MDA-MB-468 tiene una metilación parcial que no impide la expresión de *EFEMP1* y la línea celular MDA-MB-453 presenta metilación (Figura 23), estos resultados sustentan nuestra hipótesis.

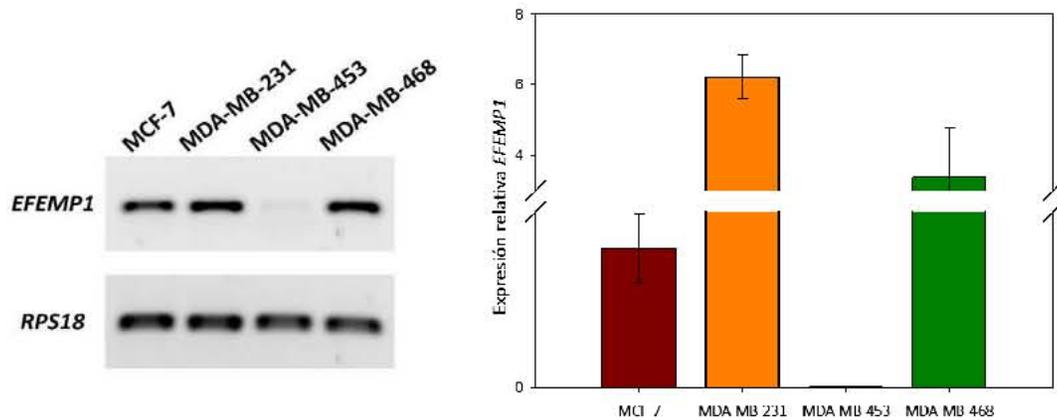


Figura 22. Expresión de *EFEMP1* en líneas celulares de mama

A la izquierda se muestran los productos de PCR punto final en un gel de agarosa al 1%, a la derecha se encuentra la gráfica obtenida de la cuantificación relativa por qPCR, en ambos ensayos se encontró que la línea celular MDA-MB-453 es negativa a *EFEMP1*.

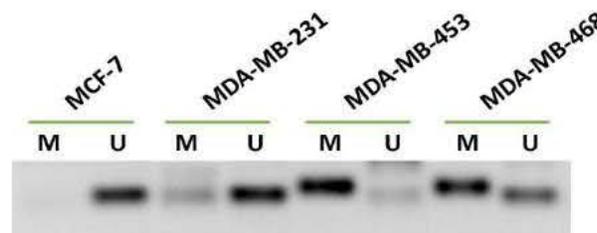


Figura 23. MSP *EFEMP1*

La línea celular MDA-MB-453 es positiva para metilación de la secuencia analizada, las líneas MCF-7 y MDA-MB-231 son negativas y la línea celular MDA-MB-468 presenta una metilación intermedia, esto se debe a que ambos juegos de primer amplifican (ver figura 21).

CTCF, Rad21 y sitios de hipersensibilidad a DNasa I en *CTS*s de *EFEMP1*

Una vez determinado que la línea celular MCF-7 presenta expresión de *EFEMP1* y que CTCF se une a los *CTS*s de dicho gen, donde el *CTS 2* se encuentra libre de metilación, decidimos evaluar en la línea celular MDA-MB-453 la unión de CTCF a estas regiones; además con el objetivo de evaluar si en esas regiones hay marcas de histonas de cromatina abierta o cerrada, incluimos los anticuerpos para H4K16ac y H3K27me3.

Los resultados nos confirman que la expresión de *EFEMP1* depende de la unión de CTCF a los *CTS*s, donde la metilación del *CTS 2* evita la unión de CTCF, de tal forma que no hay expresión de *EFEMP1*. En lo que respecta a las marcas de cromatina evaluadas, la línea celular MDA-MB-453 se caracteriza por mayor presencia de H3K27me3 en los *CTS*s en comparación con MCF-7, mientras que no hay cambios significativos en lo que respecta a la marca de cromatina abierta H4K16ac (Figura 24).

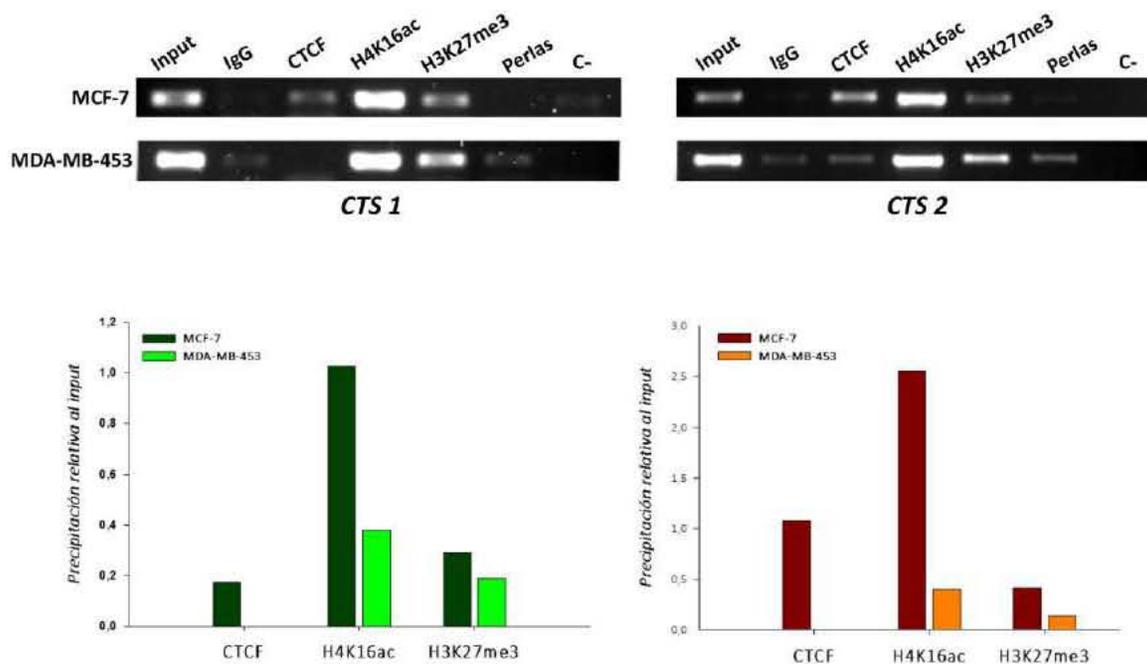


Figura 24. ChIP CTCF y marcas de histonas en MCF-7 y MDA-MB-453
Las gráficas representan el análisis densitométrico realizado en *ImageJ*.

Con la finalidad de complementar el modelo propuesto para regulación transcripcional de *EFEMP1*, decidimos buscar otras proteínas que pudieran unirse a los *CTSs* o a regiones aledañas a estos y que tal vez jueguen un papel importante en la regulación. Para esto, hicimos uso nuevamente del *Roadmap Epigenome Browser V1.19*, encontramos sitios de enriquecimiento para Rad21 y sitios de hipersensibilidad a DNasa I, en MCF-7.

Rad21 es una subunidad del complejo cohesina, ambos *CTSs* presentaron sitios de enriquecimiento a esta proteína, además parece que se forma un asa de cromatina, en donde los dos *CTSs* interactúan, por lo que el *CTS 1* que se encuentra río arriba del sitio de inicio a la transcripción, pudiera actuar como un enhancer (Figura 25).

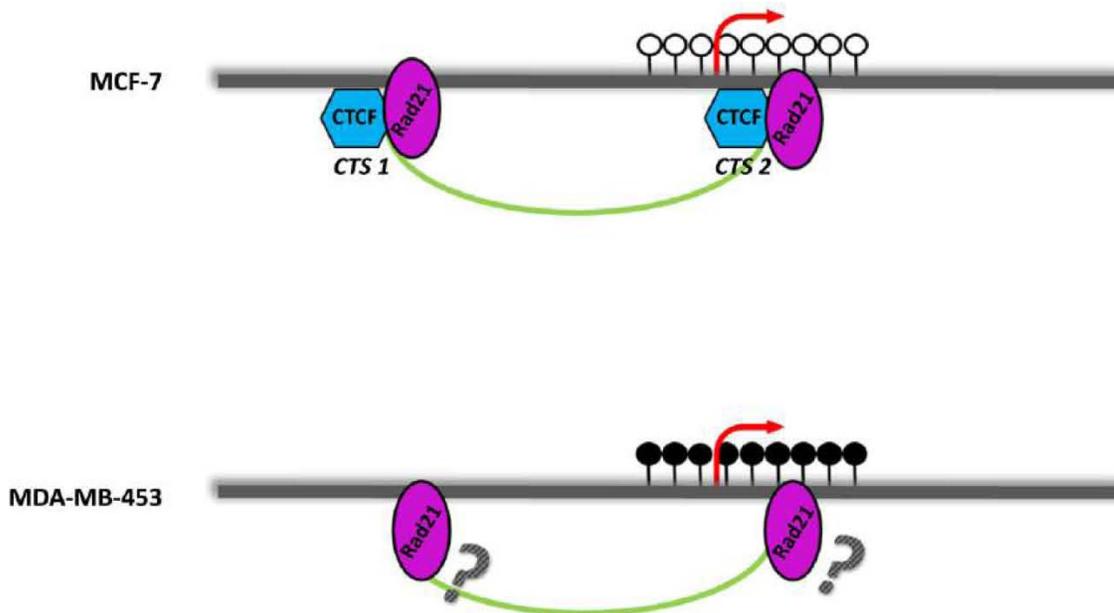
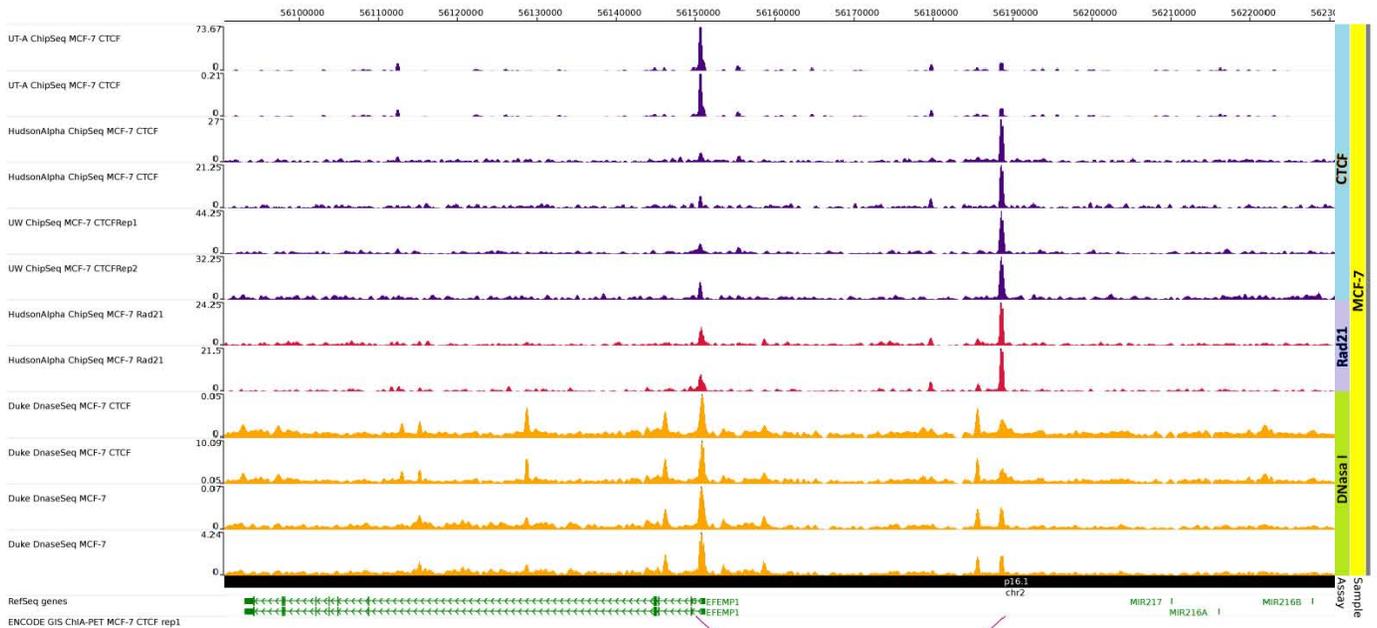


Figura 25. Modelo de regulación transcripcional para *EFEMP1*

La imagen superior muestra el análisis obtenido del *Roadmap Epigenomics*, donde se observan sitios de enriquecimiento en los *CTS*s para Rad21, los cuales coinciden con hipersensibilidad a DNasaI, además se observa una interacción entre los *CTS*s. La imagen inferior refleja nuestra propuesta de modelo transcripcional, donde en líneas celulares positivas a *EFEMP1*, como MCF-7, la unión de CTCF y Rad21 en los *CTS*s, además de un promotor hipometilado son requisitos para la expresión del gen. Cuando el promotor se hipermetila, CTCF no se encuentra unido al *CTS 2* y no hay expresión, como se determinó para MDA-MB-453.

7. DISCUSION

Este estudio se basó inicialmente en reportes previos de CTCF en cáncer de mama, concluyendo que la regulación en la expresión génica llevadas a cabo por CTCF se encuentran desreguladas, dentro de estos estudios se han reportado genes que son regulados directamente por CTCF en diferentes líneas celulares y que están involucrados en el desarrollo de la tumorigénesis, sin embargo no se conocen todos los genes regulados por CTCF, por lo que en el presente estudio decidimos buscar genes blancos de CTCF en la línea celular cancerosa de glándula mamaria MCF-7, caracterizada por altos contenidos de CTCF, así como identificar mecanismos de regulación en los mismos.

Niveles de CTCF en diferentes líneas celulares de cáncer de mama

La evaluación de los niveles CTCF a nivel de proteína en las líneas celulares MCF-7, MDA-MB-231, MDA-MB-453 y MDA-MB-468, corroboran que MCF-7 se caracteriza por tener altas cantidades de CTCF, lo cual correlaciona con reportes previos (Docquier et al. 2005). Se ha reportado que las células cancerígenas tienen altos niveles de CTCF no poli-ADP-ribosilada (CTCF-130kDa), en cambio las células no cancerígenas y tejidos sanos presentan pocas cantidades de CTCF-130kDa y mayores niveles de CTCF poli-ADP-ribosilada (CTCF-180kDa), en este contexto la identificación de CTCF-130kDa podría funcionar como un marcador para el diagnóstico de la enfermedad (Docquier et al. 2009). Asimismo, reportes previos demuestran que los niveles proteicos de CTCF-130kDa en diferentes líneas celulares de mama son altamente variables y no correlacionan con la presencia tanto proteica como de mRNA de *ER* (*Estrogen Receptor*), *p53*, *pRB*, *c-MYC* y *BRCA1* (Docquier et al. 2009). Se ha sugerido que la razón por la cual las células cancerígenas no presenten la forma CTCF-180kDa es que el mecanismo de poli(ADP)ribosilación se encuentra alterado, lo cual implicaría que a nivel global otras proteínas no presentarían poli(ADP)ribosilación, produciendo alteraciones en las funciones celulares reguladas a través de esta modificación post-traducciona. Witcher y Emerson (2009) analizaron la poli(ADP)ribosilación, en línea celular de cáncer de mama T47D, y observaron que CTCF poli(ADP)ribosilada es capaz de unirse a la región regulatoria del gen *p16^{INK4a}* permitiendo su

transcripción y que la falta de poli(ADP)-ribosilación deriva en el silenciamiento de *p16^{INK4a}*; el grupo consideró una posible alteración en los mecanismos de poli(ADP)ribosilación, sin embargo, al evaluar la poli(ADP)ribosilación global concluyeron que esta modificación post-traducciona no está alterada por lo que es necesario indagar con más detalle los mecanismos implicados en la poli(ADP)ribosilación específica de CTCF en cáncer.

Un punto interesante para discutir es el papel de CTCF como un supresor de tumores, ya que inicialmente se le atribuyó esta función, debido a que el gen CTCF se localiza en una región frecuentemente deletada en cáncer de mama y próstata, además se reportaron mutaciones en el dominio ZFs y que la sobre expresión de CTCF inhibe el crecimiento celular (Filippova et al. 1998b; Filippova et al. 2002); sin embargo se ha observado que los niveles de CTCF en tejidos de cáncer de mama aumentan en comparación con tejidos sanos (Rakha et al. 2004). Cabe recordar que generalmente en el proceso tumorigénico, los niveles de expresión de los oncogenes aumentan y los genes supresores disminuyen, no obstante, los niveles de CTCF en cáncer de mama se incrementan, por lo que se podría más como un oncogen que un supresor de tumores (Rakha et al. 2004). Nuestro grupo de trabajo ha considerado que CTCF, al igual que otros factores nucleares, puede participar en varias etapas de la progresión tumoral, en las que al inicio podría regular genes implicados en proliferación celular y evasión de la apoptosis, posteriormente en la migración celular y por último en la metástasis de las células tumorales. Dicha premisa se ha fundamentado en las observaciones de la expresión diferencial de CTCF en varias líneas celulares de mama con distinta susceptibilidad a la quimioterapia, tasa de proliferación, migración, capacidad tumorigénica y metastásica, entre otras (Docquier et al. 2005; Rasko et al. 2001; Docquier et al. 2009).

Cabe resaltar que el efecto observado por la manipulación de los niveles de CTCF en varios contextos celulares ha llevado a proponer que CTCF tiene un papel de señalización dual (*CTCF Dual Signal Model*), por ejemplo en MCF-7 se ha reportado que la reducción en los niveles de CTCF conducen a apoptosis, mientras que la sobre expresión protege parcialmente de la apoptosis (Docquier et al. 2005); en otras líneas celulares se ha documentado que la ausencia de CTCF promueve el crecimiento celular y favorece la tumorigénesis, sugiriendo que CTCF también tiene propiedades anti-proliferativas (Rasko et al. 2001; Rakha et al. 2004; Qi et al. 2003). De igual

forma, recientemente se observó que la sobre expresión de CTCF en MCF-7 produce una disminución en la proliferación y en la capacidad clonogénica; conjuntamente se demostró, en un modelo *in vivo*, que la sobre expresión de CTCF tiene efectos inhibitorios sobre el crecimiento celular, sin embargo cuando las células que sobre expresan CTCF son expuestas a condiciones de estrés se observa una resistencia a la apoptosis (Tiffen et al. 2013). Además se ha reportado que bajo condiciones de estrés ocurre la de-sumoilación de CTCF, sin embargo no existen evidencias de cómo esta de-sumoilación podría afectar la actividad de CTCF en la regulación génica (Wang et al. 2012).

Genes diferencialmente expresados

La transcripción por CTCF en cáncer de mama ha sido poco estudiada, sin embargo nuestros resultados del microarreglo de expresión muestran que cuando el factor nuclear CTCF es abatido, los genes *CEACAM5*, *SEMA6A*, *PADI3*, *RASGRP1*, *GDF15*, *NBPF4*, *FXVD3*, *CMBL* y *MACF1* se sobre expresan, mientras que los genes *DLX1*, *BAMBI*, *GFRA1*, *EFEMP1* y *TGFB2* se sub expresan. Cabe mencionar que para la obtención de estos datos se realizó un análisis estadístico muy restrictivo con el propósito de seleccionar los genes con la mayor diferencia de expresión, lo cual pudiera ser sugerente de un efecto evidente de CTCF sobre su actividad transcripcional.

Como esperábamos nuestros resultados muestran que CTCF regula la actividad transcripcional de genes implicados en la proliferación celular y tumorigénesis (Kim et al. 2011; Blumenthal et al. 2007; Urbich et al. 2012; Stone 2011; Yamamoto et al. 2009; Pils et al. 2010; Wu et al. 2013; Barcellos-Hoff & Akhurst 2009), sin embargo debido a las múltiples funciones de regulación descritas para CTCF, esperábamos que el número de genes expresados diferencialmente fuese mayor a los obtenidos en nuestro análisis o bien, que los genes modificados transcripcionalmente correlacionaran con genes previamente reportados como *mir125b*, *BRCA1*, *RASSF1A*, *CDH1*, *p16^{INK4a}*, *BAX* y *RARRES*. En estudios previos de Wendt et al., (2008), quienes realizaron un análisis de transcritos por microarreglo en células HeLa CTCF *knockdown* en fase G2, reportaron un cambio de expresión en un rango de *fold change* de 2 a 4, en el que observaron 127 genes sobre expresados y 77 sub expresados. Asimismo, en las células HeLa CTCF *knockdown* se han observado cambios

en la arquitectura nucleolar y una sobre regulación en la transcripción de genes ribosomales (Hernández-Hernández et al. 2012).

A pesar de que con nuestro análisis por microarreglo no es posible saber si CTCF regula la transcripción génica de manera directa para todos los genes, si nos permite considerar genes blancos de CTCF, así como sugerir sus posibles mecanismos de regulación génica tejido-específica.

A partir del análisis de los genes expresados diferencialmente, decidimos explorar la regulación transcripcional por CTCF en solo uno de los genes que identificamos en el microarreglo, para esto previamente realizamos una búsqueda de sitios de enriquecimiento para CTCF en regiones río arriba del TSS en todos los genes, encontramos que los genes sub expresados se caracterizan por la unión de CTCF en promotores, por el contrario los genes sobre expresados presentan CTCF río arriba del promotor, con excepción de RASGRP1. Con estos datos decidimos explorar la participación de CTCF en el promotor de un gen, *EFEMP1*, debido a que consideramos que la unión de CTCF en regiones inter génicas tienen mayor relevancia en la formación de dominios topológicamente asociados.

***EFEMP1* y su regulación transcripcional**

EFEMP1 o Fibulina-3, es una glucoproteína extracelular que se expresa durante el desarrollo y en tejidos de adulto, en este ultimo la expresión corresponde a células epiteliales y endoteliales, además de encontrarse en matriz extracelular. A EFEMP1 se le ha atribuido una función anti-angiogénica, esto se debe a que promueve la expresión de proteínas inhibitoras de metaloproteinasas (TIMP-1 y TIMP-3), a su vez inhibe la expresión de las metaloproteinasas MMP-2, MMP-3 y MMP-9 (Zhang & Marmorstein 2010; Sadr-Nabavi et al. 2009).

Estudios previos indican que *EFEMP1* es un gen supresor de tumores (TSG, *tumor supresor gene*), debido a que la disminución o ausencia en expresión, correlaciona con la progresión tumoral, estos reportes coinciden para cáncer esporádico de mama, carcinoma hepatocelular, glioma, cáncer de próstata y carcinoma endometrial (Sadr-Nabavi et al. 2009; Nomoto et al. 2010; Hu et al. 2011; Kim et al. 2011; Yang et al. 2013). Uno de los factores que se ha relacionado directamente con la disminución o eliminación en la expresión de *EFEMP1*, corresponde a la metilación del promotor,

por ejemplo en cáncer de mama se evaluó la expresión del transcrito y la proteína en muestras de pacientes, se reportó que la ausencia de *EFEMP1* correlaciona con un patrón de metilación en el promotor de al menos 25% en 60% del total de las muestras, además esta relación metilación-promotor se exploró en líneas celulares de mama, entre ellas MCF-7 (Sadr-Nabavi et al. 2009). En otro estudio, se corroboró que la hipermetilación del promotor evita la expresión del gen, en líneas celulares de carcinoma endometrial, a las cuales se expuso con 5-aza-2'-deoxicitidina y con tricostatina A, con la finalidad de revertir la metilación y recuperar la expresión. Con base en estos antecedentes, decidimos explorar la posibilidad de que CTCF se relacionara con un promotor hipometilado y que su ausencia condujera al efecto contrario, a pesar de que nuestros resultados muestran una correlación entre la presencia de CTCF con un promotor hipometilado, el mecanismo por el cual la isla CpG evita ser metilado es desconocido. No obstante se ha reportado que CTCF, PARP-1 y la DNMT, forman un complejo que es capaz de mantener libre de metilación a islas CpG, por lo que sería interesante explorar si este complejo proteico se encuentra en el promotor, como se ha descrito para otros genes (Zampieri et al. 2012).

Los sitios de enriquecimiento para Rad21 y sitios de hipersensibilidad a DNasa I, en la región correspondiente al sitio de unión para CTCF en el promotor de *EFEMP1*, más la interacción entre sitios de unión a CTCF, sugieren que la expresión del gen no solo depende de una proteína, sino de un conjunto de elementos (proteínas, RNAs, histonas, metilación del DNA) que son distintivos en todo los sitios en donde se encuentra unido CTCF, de tal manera que la regulación es precisa y específica para cada gen. En particular, la actividad del complejo CTCF-cohesina llama la atención, pues ambas proteínas de manera individual llevan a cabo diversas funciones, en conjunto parecen ser altamente relevantes para la formación de dominios topológicamente asociados (TADs), sin embargo aún no queda claro cómo es que la formación de los dominios repercute en la regulación de la expresión génica, o cómo se forman los complejos que delimitan a los dominios, los reportes indican que primero se une CTCF seguido de la cohesina, específicamente la subunidad SCC3/SA1/SA2 se une al extremo C-terminal de CTCF y las subunidades restantes de la cohesina se reclutan a través de la interacción con SCC3 (Xiao et al. 2011). Esto resulta muy importante pues no en todos los sitios de unión a CTCF se encuentra colocalizado el complejo cohesina, por lo que

quizá existan otros factores que otorguen esta localización tan particular y específica en cada tipo celular.

A pesar de la información que se ha reportado sobre la funcionalidad de CTCF, aún hay muchas interrogantes sobre cómo regula la actividad transcripcional y la arquitectura del genoma, esto se debe a las múltiples características de la proteína: su unión a diversas secuencias de DNA, interacción con otras proteínas, su participación en la formación de asas de cromatina, su capacidad para formar dímeros, además de estas propiedades también se ha reportado que CTCF tiene un dominio de unión a RNAs, lo cual le otorga mayor complejidad funcional (Saldaña-Meyer et al. 2014).

Con base en esta información podemos concluir que la regulación génica llevada a cabo por CTCF ocurre de manera específica en cada tipo celular. CTCF puede promover la diferenciación celular, senescencia o arresto celular dependiendo de las características de cada tipo celular y en coordinación con los patrones de metilación de los CTSs, proteínas y RNAs asociadas, modificaciones post-traduccionales, de tal forma que CTCF actúa principalmente como un regulador de la expresión génica, en conjunto con la arquitectura genómica, y no como un gen supresor de tumores u oncogen en un sentido estricto, como se sugirió en reportes previos.

8. CONCLUSIONES

1. MCF-7 presenta altos niveles proteicos de CTCF en comparación con las líneas MDA-MB-231/453/468
2. El abatimiento de CTCF ocasiona cambios globales en el transcriptoma
3. La expresión del gen *EFEMP1* disminuye cuando CTCF es abatido
4. *EFEMP1* no se expresa en la línea celular MDA-MB-453 pero sí en MCF-7
5. En MCF-7, CTCF se une a dos regiones regulatorias del gen *EFEMP1* (*CTS 1* y *CTS 2*)
6. El *CTS 2* del gen *EFEMP1* se encuentra dentro de una isla CpG
7. En MDA-MB-453 el *CTS 2* se encuentra metilado, mientras que en MCF-7 no presenta metilación
8. Para la expresión de *EFEMP1* se requiere la unión de CTCF y la ausencia de metilación del DNA
9. En MCF-7, los *CTSs 1* y *2* de *EFEMP1* son sitios de unión para CTCF, Rad21 y son regiones hipersensibles a Dnasa I.
10. El *CTS 1* podría actuar como enhancer.

PERSPECTIVAS

1. Someter a la línea celular MDA-MB-453 a tratamiento con 5-azacitidina y evaluar:
 - a) La expresión de *EFEMP1*
 - b) El estado de metilación del *CTS 2*
 - c) La unión de CTCF
2. Evaluar la unión de Rad21 en los sitios de unión a CTCF en *EFEMP1*

9. LITERATURA

- Aulmann, S. et al., 2003. CTCF gene mutations in invasive ductal breast cancer. *Breast cancer research and treatment*, 80(3), pp.347–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14503807>.
- Awad, T. a., 1999. Negative Transcriptional Regulation Mediated by Thyroid Hormone Response Element 144 Requires Binding of the Multivalent Factor CTCF to a Novel Target DNA Sequence. *Journal of Biological Chemistry*, 274(38), pp.27092–27098. Available at: <http://www.jbc.org/cgi/doi/10.1074/jbc.274.38.27092> [Accessed September 7, 2013].
- Baniahmad, A. et al., 1990. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell*, 61(3), pp.505–514. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2159385>.
- Barcellos-Hoff, M.H. & Akhurst, R.J., 2009. Transforming growth factor-beta in breast cancer: too much, too late. *Breast cancer research: BCR*, 11(1), p.202. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2687712&tool=pmcentrez&rendertype=abstract> [Accessed January 22, 2014].
- Bell, a C., West, a G. & Felsenfeld, G., 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98(3), pp.387–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10458613>.
- Blumenthal, R.D. et al., 2007. Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. *BMC cancer*, 7, p.2. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1769503&tool=pmcentrez&rendertype=abstract> [Accessed February 3, 2014].
- Butcher, D.T. & Rodenhiser, D.I., 2007. Epigenetic inactivation of BRCA1 is associated with aberrant expression of CTCF and DNA methyltransferase (DNMT3B) in some sporadic breast tumours. *European journal of cancer (Oxford, England: 1990)*, 43(1), pp.210–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17071074> [Accessed July 31, 2013].
- Chang, J.-W. et al., 2010. Distinct epigenetic domains separated by a CTCF bound insulator between the tandem genes, BLU and RASSF1A. *PLoS one*, 5(9), p.e12847. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2942851&tool=pmcentrez&rendertype=abstract> [Accessed August 1, 2013].
- Chen, H. et al., 2012. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS one*, 7(7), p.e41374. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3400636&tool=pmcentrez&rendertype=abstract> [Accessed August 1, 2013].
- Cremer, T. & Cremer, C., 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews. Genetics*, 2(4), pp.292–301. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11283701>.
- Dixon, J.R. et al., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398), pp.376–380. Available at: <http://www.nature.com/doi/10.1038/nature11082>.
- Docquier, F. et al., 2009. Decreased poly(ADP-ribosyl)ation of CTCF, a transcription factor, is associated with breast cancer phenotype and cell proliferation. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 15(18), pp.5762–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19737964> [Accessed January 10, 2014].
- Docquier, F. et al., 2005. Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. *Cancer research*, 65(12), pp.5112–22. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15958555> [Accessed July 31, 2013].
- Dubois-Chevalier, J. et al., 2014. A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. *Nucleic acids research*, 42(17), pp.10943–10959.
- El-Kady, A. & Klenova, E., 2005. Regulation of the transcription factor, CTCF, by phosphorylation with protein

- kinase CK2. *FEBS Letters*, 579(6), pp.1424–1434.
- Farrar, D. et al., 2010. Mutational analysis of the poly(ADP-ribosyl)ation sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-ribosyl)ation. *Molecular and cellular biology*, 30(5), pp.1199–1216.
- Filippova, G.N. et al., 1998a. A Widely Expressed Transcription Factor With Multiple DNA Sequence Specificity , CTCF , Is Localized at Chromosome Segment 16q22 . 1 Within One of the Smallest Regions of Overlap for Common Deletions in Breast and Prostate Cancers. , 36(April 1997), pp.26–36.
- Filippova, G.N. et al., 2002. Tumor-associated Zinc Finger Mutations in the CTCF Transcription Factor Selectively Alter Its DNA-binding Specificity Advances in Brief Tumor-associated Zinc Finger Mutations in the CTCF Transcription Factor. , pp.48–52.
- Filippova, G.N. et al., 1998b. With Multiple DNA Sequence Specificity , CTCF , Is Localized at Chromosome Segment 16q22 . 1 Within One of the Smallest Regions of Overlap for Common Deletions in Breast and Prostate Cancers. , 36(April 1997), pp.26–36.
- Fiorrentino, F.P. & Giordano, A., 2012. The tumor suppressor role of CTCF. *Journal of cellular physiology*, 227(2), pp.479–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21465478> [Accessed July 31, 2013].
- Gaszner, M. & Felsenfeld, G., 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nature reviews. Genetics*, 7(9), pp.703–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16909129> [Accessed July 31, 2013].
- Guelen, L. et al., 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, 453(7197), pp.948–51. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18463634> [Accessed August 7, 2013].
- Guo, J.U. et al., 2011. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle*, 10(16), pp.2662–2668.
- Hernández-Hernández, a et al., 2012. Changes of the nucleolus architecture in absence of the nuclear factor CTCF. *Cytogenetic and genome research*, 136(2), pp.89–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22286186> [Accessed October 22, 2013].
- Holwerda, S.J.B. & de Laat, W., 2013. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1620), p.20120369. Available at: <http://rsta.royalsocietypublishing.org/content/368/1620/20120369>.
- Hu, Y. et al., 2011. EFEMP1 suppresses malignant glioma growth and exerts its action within the tumor extracellular compartment. *Molecular cancer*, 10(1), p.123. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3204287&tool=pmcentrez&rendertype=abstract>.
- Kaptein, R., 1991. Zinc fingers. *Current Opinion in Structural Biology*, 1(1), pp.63–70. Available at: <http://linkinghub.elsevier.com/retrieve/pii/0959440X9190013>.
- Katainen, R. et al., 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nature Genetics*, advance on(7), pp.818–21. Available at: <http://dx.doi.org/10.1038/ng.3335> <http://www.ncbi.nlm.nih.gov/pubmed/26053496>.
- Kemp, C.J. et al., 2014. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Reports*, 7(4), pp.1020–1029. Available at: <http://dx.doi.org/10.1016/j.celrep.2014.04.004>.
- Kim, S., Yu, N.-K. & Kaang, B.-K., 2015. CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental & molecular medicine*, 47(6), p.e166. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4491725&tool=pmcentrez&rendertype=abstract>.
- Kim, T.H. et al., 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*, 128(6), pp.1231–45. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2572726&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Kim, Y.-J. et al., 2011. EFEMP1 as a novel DNA methylation marker for prostate cancer: array-based DNA methylation and expression profiling. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 17(13), pp.4523–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21571867>

- [Accessed February 14, 2014].
- Kitchen, N.S. & Schoenherr, C.J., 2010. Sumoylation modulates a domain in CTCF that activates transcription and decondenses chromatin. *Journal of cellular biochemistry*, 111(3), pp.665–75. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20589758> [Accessed July 31, 2013].
- Klenova, E.M. et al., 1998. Characterization of the chicken CTCF genomic locus, and initial study of the cell cycle-regulated promoter of the gene. *The Journal of biological chemistry*, 273(41), pp.26571–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9756895>.
- Klenova, E.M. et al., 1993. 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.
- Klenova, E.M. et al., 2001. Functional Phosphorylation Sites in the C-Terminal Region of the Multivalent Multifunctional Transcriptional Factor CTCF. *Molecular and cellular biology*, 21(6), pp.2221–2234.
- Kulis, M. & Esteller, M., 2010. DNA methylation and cancer. *Advances in genetics*, 70(10), pp.27–56. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20920744> [Accessed July 31, 2013].
- De La Rosa-Velázquez, I. a et al., 2007. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer research*, 67(6), pp.2577–85. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17363576> [Accessed July 31, 2013].
- Lee, B.-K. & Iyer, V.R., 2012. Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation. *The Journal of biological chemistry*, 287(37), pp.30906–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22952237> [Accessed July 31, 2013].
- Leers, J. & Renkawitz, R., 2005. The DNA-binding factor CTCF co-ordinates gene expression. *B.I.F. FUTURA*, 20, pp.92–96.
- Li, M., Liu, G.-H. & Izpisua Belmonte, J.C., 2012. Navigating the epigenetic landscape of pluripotent stem cells. *Nature reviews. Molecular cell biology*, 13(8), pp.524–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22820889> [Accessed January 21, 2014].
- Martinez, S.R. & Miranda, J.L., 2010. CTCF terminal segments are unstructured. *Protein science: a publication of the Protein Society*, 19(5), pp.1110–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2868253&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Méndez-Catalá, C.F. et al., 2013. A Novel Mechanism for CTCF in the Epigenetic Regulation of Bax in Breast Cancer Cells. , 15(8), pp.898–912.
- Merkenschlager, M. & Odom, D.T., 2013. CTCF and cohesin: Linking gene regulatory elements with their targets. *Cell*, 152(6), pp.1285–1297. Available at: <http://dx.doi.org/10.1016/j.cell.2013.02.029>.
- Moon, H. et al., 2005. CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO reports*, 6(2), pp.165–70. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1299244&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Nakahashi, H. et al., 2013. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell reports*, 3(5), pp.1678–89. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23707059> [Accessed August 13, 2013].
- Nomoto, S. et al., 2010. Epidermal growth factor-containing fibulin-like extracellular matrix protein 1, EFEMP1, a novel tumor-suppressor gene detected in hepatocellular carcinoma using double combination array analysis. *Annals of surgical oncology*, 17(3), pp.923–932.
- Ohlsson, R., Renkawitz, R. & Lobanenkova, V., 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in genetics: TIG*, 17(9), pp.520–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11525835>.
- Ong, C. & Corces, V.G., 2009. Insulator as mediator of intra- and inter-chromosomal interactions: a common evolutionary theme. *Journal of Biology*, 8(73), p.73.1-73.4.
- Ong, C.-T. & Corces, V.G., 2014. CTCF: an architectural protein bridging genome topology and function. *Nature Reviews Genetics*, (March), pp.1–13. Available at: <http://www.nature.com/doi/10.1038/nrg3663> [Accessed March 11, 2014].
- Pant, V. et al., 2004. Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Molecular and*

- cellular biology*, 24(8), pp.3497–3504.
- Peng, Z. et al., 2012. Epigenetic repression of RARRES1 is mediated by methylation of a proximal promoter and a loss of CTCF binding. *PloS one*, 7(5), p.e36891. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3355180&tool=pmcentrez&rendertype=abstract> [Accessed August 1, 2013].
- Pfaffl, M.W., 2004. Relative quantification. , pp.63–82.
- Phillips, J.E. & Corces, V.G., 2009. CTCF: master weaver of the genome. *Cell*, 137(7), pp.1194–211. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3040116&tool=pmcentrez&rendertype=abstract> [Accessed October 18, 2013].
- Pils, D. et al., 2010. BAMBI is overexpressed in ovarian cancer and co-translocates with Smads into the nucleus upon TGF-beta treatment. *Gynecologic oncology*, 117(2), pp.189–97. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20189233> [Accessed February 14, 2014].
- Pugacheva, E.M. et al., 2006. Cloning and characterization of zebrafish CTCF: Developmental expression patterns, regulation of the promoter region, and evolutionary aspects of gene organization. *Gene*, 375, pp.26–36. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16647825> [Accessed July 31, 2013].
- Pugacheva, E.M. et al., 2005. Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Human molecular genetics*, 14(7), pp.953–65. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15731119> [Accessed September 8, 2013].
- Qi, C.-F. et al., 2003. CTCF functions as a critical regulator of cell-cycle arrest and death after ligation of the B cell receptor on immature B cells. *Proceedings of the National Academy of Sciences of the United States of America*, 100(2), pp.633–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=141048&tool=pmcentrez&rendertype=abstract>.
- Rakha, E. a et al., 2004. Expression of the transcription factor CTCF in invasive breast cancer: a candidate gene located at 16q22.1. *British journal of cancer*, 91(8), pp.1591–6. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2409935&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Rasko, J.E.J. et al., 2001. Cell Growth Inhibition by the Multifunctional Multivalent Zinc-Finger Factor CTCF Advances in Brief Cell Growth Inhibition by the Multifunctional Multivalent Zinc-Finger. , pp.6002–6007.
- Recillas-Targa, F., 2006. Epigenetic boundaries of tumour suppressor gene promoters: the CTCF connection and its role in carcinogenesis. *Journal of Cellular and Molecular Medicine*, 10(3). Available at: <http://doi.wiley.com/10.2755/jcmm010.003.02> [Accessed July 31, 2013].
- Reimand, J., Wagih, O. & Bader, G.D., 2013. The mutational landscape of phosphorylation signaling in cancer. *Scientific reports*, 3, p.2651. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3788619&tool=pmcentrez&rendertype=abstract>.
- Renaud, S. et al., 2005. CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic acids research*, 33(21), pp.6850–60. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1301598&tool=pmcentrez&rendertype=abstract> [Accessed August 2, 2013].
- Renaud, S. et al., 2007. Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. *Nucleic acids research*, 35(4), pp.1245–56. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1851636&tool=pmcentrez&rendertype=abstract> [Accessed September 8, 2013].
- Renaud, S. et al., 2007. Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic acids research*, 35(21), pp.7372–88. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2175345&tool=pmcentrez&rendertype=abstract> [Accessed August 15, 2013].
- Rhee, H.S. & Pugh, B.F., 2011. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell*, 147(6), pp.1408–1419. Available at:

- <http://dx.doi.org/10.1016/j.cell.2011.11.013>.
- Rubio, E.D. et al., 2008. CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, 105(24), pp.8309–14. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2448833&tool=pmcentrez&rendertype=abstract>.
- Sadr-Nabavi, A. et al., 2009. Decreased expression of angiogenesis antagonist EFEMP1 in sporadic breast cancer is caused by aberrant promoter methylation and points to an impact of EFEMP1 as molecular biomarker. *International journal of cancer. Journal international du cancer*, 124(7), pp.1727–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19115204> [Accessed February 14, 2014].
- Saito, Y. & Saito, H., 2012. Role of CTCF in the regulation of microRNA expression. *Frontiers in genetics*, 3(September), p.186. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3457075&tool=pmcentrez&rendertype=abstract> [Accessed August 20, 2013].
- Saldaña-Meyer, R. et al., 2014. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes & development*, 28(7), pp.723–34. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4015496&tool=pmcentrez&rendertype=abstract>.
- Shukla, S. et al., 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, 479(7371), pp.74–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21964334> [Accessed July 31, 2013].
- Song, C.-X. et al., 2013. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 19(2), pp.161–169.
- Soto-Reyes, E. et al., 2012. Disruption of CTCF at the miR-125b1 locus in gynecological cancers. *BMC cancer*, 12, p.40. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3297514&tool=pmcentrez&rendertype=abstract>.
- Stone, J.C., 2011. Regulation and Function of the RasGRP Family of Ras Activators in Blood Cells. *Genes & cancer*, 2(3), pp.320–34. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3128638&tool=pmcentrez&rendertype=abstract> [Accessed January 20, 2014].
- Tiffen, J.C. et al., 2013. The cancer-testis antigen BORIS phenocopies the tumor suppressor CTCF in normal and neoplastic cells. *International journal of cancer. Journal international du cancer*, 133(7), pp.1603–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23553099> [Accessed January 2, 2014].
- Torrano, V. et al., 2006. Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosyl)ation-dependent mechanism. *Journal of cell science*, 119(Pt 9), pp.1746–59. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16595548> [Accessed October 19, 2013].
- Urbich, C. et al., 2012. MicroRNA-27a/b controls endothelial cell repulsion and angiogenesis by targeting semaphorin 6A. *Blood*, 119(6), pp.1607–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22184411> [Accessed February 6, 2014].
- Vostrov, A.A. & Quitschke, W.W., 1997. The Zinc Finger Protein CTCF Binds to the APB⁺ Domain of the Amyloid⁺-Protein Precursor Promoter. , 272(52), pp.33353–33359.
- Wang, J., Wang, Y. & Lu, L., 2012. De-SUMOylation of CCCTC binding factor (CTCF) in hypoxic stress-induced human corneal epithelial cells. *The Journal of biological chemistry*, 287(15), pp.12469–79. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3320996&tool=pmcentrez&rendertype=abstract> [Accessed January 13, 2014].
- Wendt, K.S. et al., 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, 451(7180), pp.796–801. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18235444> [Accessed July 31, 2013].
- Witcher, M. & Emerson, B.M., 2009. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with

- loss of CTCF binding and a chromatin boundary. *Molecular cell*, 34(3), pp.271–84. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2723750&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Wu, Z.-S. et al., 2013. Prognostic significance of the expression of GFR α 1, GFR α 3 and syndecan-3, proteins binding ARTEMIN, in mammary carcinoma. *BMC cancer*, 13(1), p.34. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3562211&tool=pmcentrez&rendertype=abstract> [Accessed February 14, 2014].
- Xiao, T., Wallace, J. & Felsenfeld, G., 2011. 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. *Molecular and cellular biology*, 31(11), pp.2174–83. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3133248&tool=pmcentrez&rendertype=abstract> [Accessed September 7, 2013].
- Yamamoto, H. et al., 2009. FXYD3 protein involved in tumor cell proliferation is overproduced in human breast cancer tissues. *Biological & pharmaceutical bulletin*, 32(7), pp.1148–54. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19571376>.
- Yang, J. & Corces, V.G., 2011. *Chromatin insulators: a role in nuclear organization and gene expression.*, Elsevier Inc. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3175007&tool=pmcentrez&rendertype=abstract> [Accessed July 31, 2013].
- Yang, T. et al., 2013. Epigenetic inactivation of EFEMP1 is associated with tumor suppressive function in endometrial carcinoma. *PLoS one*, 8(6), p.e67458. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3696089&tool=pmcentrez&rendertype=abstract> [Accessed February 14, 2014].
- Yu, M. et al., 2012. Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. *Cell*, 149(6), pp.1368–1380.
- Yu, W. et al., 2004. Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature genetics*, 36(10), pp.1105–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15361875> [Accessed July 31, 2013].
- Zampieri, M. et al., 2012. ADP-ribose polymers localized on Ctfp-Parp1-Dnmt1 complex prevent methylation of Ctfp target sites. *The Biochemical journal*, 441(2), pp.645–52. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3258657&tool=pmcentrez&rendertype=abstract> [Accessed January 11, 2014].
- Zhang, R. et al., 2004. Dynamic association of the mammalian insulator protein CTCF with centrosomes and the midbody. *Experimental cell research*, 294(1), pp.86–93. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14980504> [Accessed September 8, 2013].
- Zhang, Y. & Marmorstein, L.Y., 2010. Focus on molecules: Fibulin-3 (EFEMP1). *Experimental Eye Research*, 90(3), pp.374–375. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0014483509002863>.
- Zhao, R., Bodnar, M.S. & Spector, D.L., 2009. Nuclear Neighborhoods and Gene Expression. *Current opinion in genetics & development*, 19(2), pp.172–179.
- Zlatanova, J. & Caiafa, P., 2009. CTCF and its protein partners: divide and rule? *Journal of cell science*, 122(Pt 9), pp.1275–84. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19386894> [Accessed July 31, 2013].
92. Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics*. 2002 Nov;18(11):1427-31. [PMID: 12424112](https://pubmed.ncbi.nlm.nih.gov/12424112/)

10. ANEXO

Tabla 1. *Primers* de expresión

Nombre	Forward	Reverse	Amplicon
OAZ1	5'-ACG TCC AAC GAC AAG ACG AGG ATT-3'	5'-TCA GCA CTG TTC GCC AGT TAA TGC-3'	82pb
RPS18	5'-CAG CCA GGT CCT AGC CAA TG -3'	5'- CCC TCT ATG GGC CCG AAT CT-3'	82pb
CTCF	5'-TGA CAC AGT CAT AGC CCG AAA A-3'	5'-TGC CTT GCT CAA TAT AGG AAT GC-3'	74pb
EFEMP1	5'-TGT ATG TGC CCC CAG GGA TA-3'	5'-AAC GGA AGC CGC CAT GAT AA-3'	124pb

Tabla 2. *Primers* usados para el análisis de metilación (M, *Methylation*/UM, *UnMethylation*) (Yang et al. 2013; Nomoto et al. 2010).

Nombre	Forward	Reverse	Amplicon
M	5'-CGT CGG GTT CGT AAC GTT GG-3'	5'-GAC AAC GAC CGC GAC GAC AA-3'	120pb
UM	5'-GGT GTG GGG TTT TTG TGA GTT TG-3'	5'-CAC AAC AAA ATC CAC ACC CCC TA-3'	103pb

Tabla 3. *Primers* ChIP, (CTS 1: CTCF Target Site 1, CTS 2: CTCF Target Site 2)

Nombre	Forward	Reverse	Amplicon
EFEMP1 CTS 1	5'-CAGAGCAATTTTAGGAAGGTGAA-3'	5'-ACTTTATGGATTCAAACCTGGATGC-3'	84pb
EFEMP1 CTS 2	5'-ACTTCGCCAGATCAGACCCA-3'	5'-TCGGGTACTCAACACCCCTC-3'	187pb

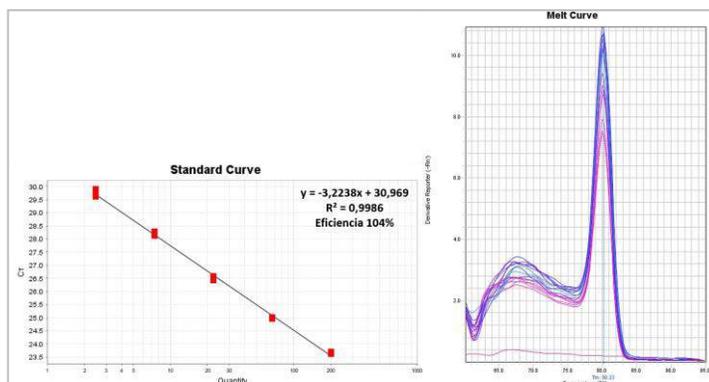
	Gen	<i>p-value</i> (<i>shCTCF</i> vs. control)	<i>Fold-Change</i> (<i>shCTCF</i> vs. control)
Genes sobre expresados	<i>CEACAM5</i>	1,55E-05	3,9345
	<i>SEMA6A</i>	1,43E-05	2,67432
	<i>PADI3</i>	8,73E-06	2,53599
	<i>RASGRP1</i>	3,10E-06	2,08881
	<i>GDF15</i>	1,37E-07	1,68717
	<i>NBPF4</i>	1,90E-05	1,62437
	<i>FXYD3</i>	1,99E-05	1,32002
	<i>CMBL</i>	1,23E-06	1,29157
	<i>MACF1</i>	2,24E-05	1,2423
	Genes sub expresados	<i>DLX1</i>	2,58E-05
<i>CTCF</i>		1,39E-05	-1,49904
<i>BAMBI</i>		7,20E-06	-1,69715
<i>GFRA1</i>		2,24E-05	-2,24356
<i>EFEMP1</i>		1,79E-05	-2,25226
<i>TGFB2</i>		2,02E-05	-3,25325

Tabla 4. Cambio de expresión génica en MCF-7 *shCTCF* respecto a MCF-7 *mock*.

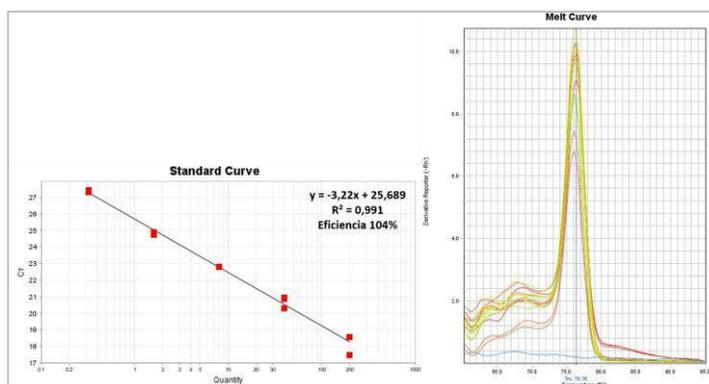
Los genes sobre expresados están dentro de un rango de *fold change* de 1.2 a 4, es decir se expresan de 120% a 400% más, respecto a MCF-7 *mock*. Los genes sub expresados tienen un rango del *fold change* de -1.2 a -3.3, esto significa que su expresión es de 120% a 330% menos que en MCF-7 *mock*

Figura 25. Curvas estándar y melting

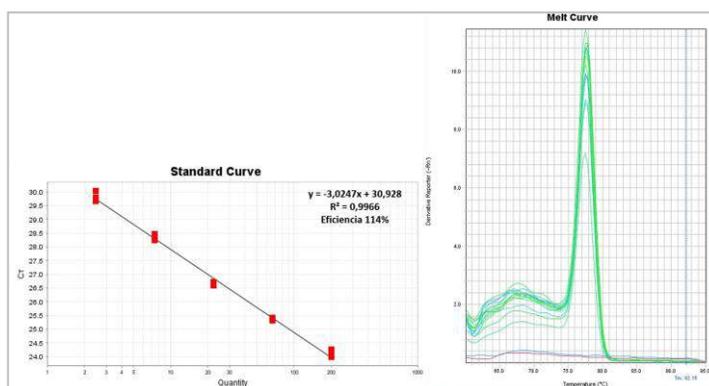
OAZ1



CTCF



EFEMP1



Secuencia CTCF Target Site 1

5' ACCGTGCCAGCCAAATCTTTAAATCTCAGATTCCTCTGAGCT **CAGAGCAATTTAGGAAGGTGAA**TTTACTGCAATT
TCCACTTGCATTCTGCTAAATCCAT **GCATCCAGTTTGAATCCATAAAGT**TGGACTCCATTAATAAATCAT 3'

Promotor EFEMP1 -300 a 1000 del TSS

Secuencia CTCF Target Site 2

5' AGACCATA~~AAAA~~CA~~AAAC~~CC~~CC~~AC~~CT~~GC~~AA~~AT~~TTT~~GT~~T~~AGT~~G~~CA~~AG~~TT~~CT~~AT~~TT~~TGG~~CC~~TA~~TT~~CC~~CA~~AC~~AC~~TA~~AG~~CAC~~AC~~AGT~~TA~~
GGTGGCA~~AA~~TAT~~TT~~TGTTGA~~AT~~GGATGA~~AT~~TCTC~~AT~~TTT~~TT~~TCC~~AT~~TTGTGGA~~AG~~CAGAC~~AC~~CTC~~CT~~CA~~AAAA~~AAAA~~AAAA~~
AATCCGAA~~AG~~AAAAGGTAA~~AG~~GAACAGTTCTATGA~~AT~~GGGGT~~TA~~ATTTACTCTCCTAGCATTCACTTATGGGAAAA
AACTAACACCAGTTTCTACA~~ACC~~CTCTGCTCTC~~TA~~CCGACATCCTCTCT~~CC~~CTCC~~CT~~CTCAGACA~~ACA~~AAAA~~CC~~CACT
AGAGACTGGAGCATCC~~CA~~AG~~CT~~GA~~AG~~CC~~CA~~TCCAGGACATCACA~~TT~~CA~~AC~~TTCT~~CC~~CTCTGGCGTGGGGTCC~~CC~~CG
CGAGTCTGGGAAACGAGGAGCTGGCA~~AG~~GGGAGGGCA~~AG~~CGGGGGCGGCA~~AG~~AGAGGGCCCTGGAGCAGGGGGCGCGGACC
TCGCTGCGCTGCGCGCTCTACC~~CC~~GGGCTCGCA~~AC~~CGTGGGCTCAGCGCTCGCGCTCCCTCAGCTCTCTCTCCGCCC
CCCTTCGCCCCTCC~~CC~~CTTCCCTCCCTTTCTCTCCTCCTCCTG~~CC~~CGCGCGCGCGCTG~~CC~~CGGACTTCGCCAGATCAGAC
CCACGGGGCTGCCCTCC~~CT~~TGCCCACTCC~~CT~~CGTGC~~CC~~GGGCGCGGAGCGCAGCGCGCGCCACAGGTAGGAGCCCC
AACGGGGGTCGCCCCGAGCGAGCCCACTACTCACTGGGCGCGCGGGGAGAGTGAGGTGGGGTGAGCT**GAGGGGTGTT**
GAGTACCCGAGAACTGCCTGGACTGTGCAGGGTGTCTCGGGCAGAGAGGGAGGTGGAGGTGTGCAGCCTCGGAATGCGA
TCTGTCAAGCAAGTGGAGTACTCGGGTGCATCTGTTGGGATGGGCATGTGTGTGCGCGCCGACAACAGCCGGCAGGGC
TTTCCCATAAACGTGGTTTTTGA~~AG~~TTTAGGGGACCAGTGTGCGAGATTTAAGCCGCACCTGGATTCCATAGGAGCTGG
TTAGAAGCTGGGACGCTGAGCAGCTCCAGGGGACCGCCGCTTAGCTTGCTGTTAAGAAAGGGACCTCATCTCCCTGCC
GGGCCAGGCCCGCCCGCAA~~ACT~~TGGTACCTTGGGCTGCGGTGCGATCCCTGGTTCCGGTCTTAGGCAGCCTGAAACCGA
AGGTAGCGTGTGGGGACCCAGACTGATAAGACAAAAGAGAATCAGTTCGCTTTGGGCTGCCCTCCACACA~~ACT~~GGGAC
TTTTAAACAAAGCTGTGCGCA_3'

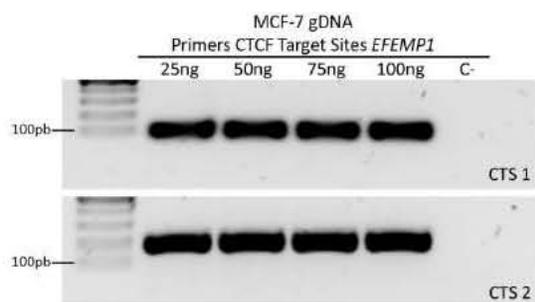


Figura 26. Secuencias CTS:

Secuencias correspondientes a los CTS 1 y 2, el CTS 2 se muestra en letra negra dentro de la secuencia del promotor de EFEMP1 (-300 a 1000 del TSS, marcado con flecha verde). En rojo se muestran los sitios donde se anclan los primers diseñados para el ensayo de ChIP, los cuales se probaron en DNA genómico (gDNA) de MCF-7 como se muestra en el gel de agarosa.

