



# GRADO EN MEDICINA

## TRABAJO FIN DE GRADO

El regulador transcripcional CTCF y su  
papel en cáncer

The transcriptional regulator CTCF and  
its role in cancer

**Autor/a:** Lucía Gómez González

**Director/es:** M. Dolores Delgado Villar  
Vanessa Junco Ruisánchez

**Santander, Mayo 2024**

## RESUMEN

CTCF es un importante regulador transcripcional y organizador de la estructura tridimensional del genoma, frecuentemente alterado en cáncer. La proteína CTCF está compuesta por tres dominios diferentes: un dominio N-terminal, uno C-terminal, y uno central con 11 dedos de zinc, que permiten su unión al ADN. Se encuentra expresada de forma ubicua y altamente conservada en los vertebrados, ya que es indispensable en el funcionamiento celular gracias a su actividad regulando la expresión génica. Entre sus múltiples y variadas funciones destaca su papel como regulador transcripcional, *insulator*, regulador epigenético, y regulador de la estructura tridimensional de la cromatina. Su implicación en el cáncer ha sido extensamente estudiada, desempeñando funciones clave en la proliferación celular, la apoptosis, la diferenciación y la invasión, entre otros procesos fundamentales en la carcinogénesis. Se ha descrito su alteración en numerosos tumores, incluyendo el cáncer de mama, endometrio y tumores hematológicos, en los que su pérdida o disfunción se asocia con la progresión tumoral, lo que subraya su importancia como posible biomarcador y diana terapéutica. La comprensión de los mecanismos moleculares subyacentes a la participación de CTCF en cáncer es fundamental, ya que está impulsando su caracterización como posible biomarcador, permitiendo un diagnóstico más temprano, un mejor pronóstico y el posible desarrollo de terapias dirigidas que lograrían un tratamiento más personalizado, más efectivo, y menos dañino para los pacientes.

**PALABRAS CLAVE:** CTCF, regulación génica, epigenética, cáncer, biomarcador.

## ABSTRACT

CTCF is an important transcriptional regulator and organizer of the three-dimensional genome structure, frequently altered in cancer. CTCF protein is composed of three different domains: a N-terminal domain, a C-terminal domain, and a central domain with 11 zinc fingers, which allow its binding to the DNA. It is ubiquitously expressed and highly conserved in vertebrates, as it is essential for cell functions due to its activity in the regulation of gene expression. Among its multiple and varied functions, its role as a transcriptional regulator, insulator, epigenetic regulator, and regulator of chromatin's three-dimensional structure stands out. Its involvement in cancer has been extensively studied, playing key roles in cell proliferation, apoptosis, differentiation, and invasion, among other fundamental processes in carcinogenesis. CTCF alterations have been described in several tumors, including breast cancer, endometrial cancer, and hematologic tumors, where its loss or dysfunction is associated with tumor progression, highlighting its importance as a potential biomarker and therapeutic target. Understanding the underlying molecular mechanisms of CTCF's involvement in cancer is crucial, as it is driving its characterization as a potential biomarker, enabling earlier diagnosis, better prognosis, and the potential development of targeted therapies that would achieve a more personalized, effective, and less harmful treatment for patients.

**KEY WORDS:** CTCF, gene regulation, epigenetics, cancer, biomarker.

## ÍNDICE

|                                                    |    |
|----------------------------------------------------|----|
| <b>1. ABREVIATURAS</b>                             | 4  |
| <b>2. INTRODUCCIÓN</b>                             | 5  |
| <b>3. MÉTODOS</b>                                  | 6  |
| 3.1. Revisión bibliográfica                        | 6  |
| 3.2. ENCODE                                        | 6  |
| 3.3. PYMOL                                         | 6  |
| 3.4. CBIOPORTAL                                    | 6  |
| 3.5. GEPIA                                         | 7  |
| 3.6. CanProVar                                     | 7  |
| <b>4. ESTRUCTURA DE CTCF</b>                       | 7  |
| <b>5. UNIÓN DE CTCF AL ADN</b>                     | 9  |
| <b>6. MODIFICACIONES POSTRADUCCIONALES</b>         | 10 |
| <b>7. INTERACCIÓN DE CTCF CON OTRAS PROTEÍNAS</b>  | 12 |
| 7.1. Proteínas de unión a ADN                      | 13 |
| 7.2. Proteínas arquitectónicas                     | 14 |
| 7.3. Modificadores epigenéticos                    | 15 |
| 7.4. Proteínas de transcripción basal y otras      | 16 |
| <b>8. FUNCIONES MOLECULARES DE CTCF</b>            | 17 |
| 8.1. Regulador transcripcional                     | 18 |
| 8.2. Proteína de unión a <i>insulators</i>         | 18 |
| 8.3. Regulador epigenético                         | 19 |
| 8.4. Organización tridimensional de la cromatina   | 20 |
| 8.5. Unión al ARN                                  | 21 |
| <b>9. DESREGULACIÓN DE CTCF EN CÁNCER</b>          | 22 |
| 9.1. Variaciones en los niveles de CTCF            | 23 |
| 9.2. Alteraciones en la estructura de CTCF         | 24 |
| 9.3. Cambios en la unión de CTCF al genoma         | 25 |
| 9.3.1. Cambios en la secuencia del ADN             | 26 |
| 9.3.2. Cambios en la metilación del ADN            | 26 |
| <b>10. ALTERACIONES DE CTCF EN ALGUNOS TUMORES</b> | 27 |
| 10.1. Cáncer de mama                               | 27 |
| 10.2. Cáncer de endometrio                         | 28 |
| 10.3. Tumores hematológicos                        | 29 |
| <b>11. CTCF Y BORIS COMO BIOMARCADORES</b>         | 30 |
| <b>12. CONCLUSIONES</b>                            | 32 |
| <b>13. AGRADECIMIENTOS</b>                         | 32 |

|                                                                           |           |
|---------------------------------------------------------------------------|-----------|
| <b>14. BIBLIOGRAFÍA .....</b>                                             | <b>33</b> |
| <b>15. ANEXOS .....</b>                                                   | <b>41</b> |
| Anexo 1: Representación del gen <i>CTCF</i> con UCSC Genome Browser. .... | 41        |
| Anexo 2: Análisis con cBioPortal .....                                    | 42        |
| Anexo 3: Análisis con GEPIA .....                                         | 48        |

## 1. ABREVIATURAS

5-meC: 5-metilcitosina

BRCA1: *breast cancer type 1 susceptibility protein*

CpG: Citosina-Fosfato-Guanina

CTCF: factor de unión a CCCTC

CTS: sitio de unión a CTCF

ENCODE: *Encyclopedia of DNA Elements*

GIST: tumor del estroma gastrointestinal

ICGC: Consorcio Internacional del Genoma de Cáncer

IDH: isocitrato deshidrogenasa

IGF2: factor de crecimiento insulínico tipo 2

ILC: carcinoma lobulillar invasivo

kDa: Kilo daltons

LLA: leucemia linfoblástica aguda

LLC: leucemia linfocítica crónica

LMA: leucemia mieloide aguda

lncRNAs: ARN no codificantes de cadena larga

ncRNA: ARN no codificante

LOH: pérdida de heterocigosidad

PARP1: poli-(ADP-ribosa)–polimerasa 1

PARilación: poli-ADP-ribosilación

pb: pares de bases

CK2: caseína quinasa 2

PLK1: proteína poloquinasa 1

RBD: dominio de unión al receptor

SMC: proteínas de mantenimiento estructural de los cromosomas

SUMO: *small-like ubiquitin modifiers*

TAD: dominio topológicamente activado

TCGA: *The Cancer Genome Atlas*

TERT: telomerasa transcriptasa inversa

TGFB1: factor de crecimiento transformante beta 1

TSS: sitio de inicio de la transcripción

UTR: región no traducida

ZF: dedos de zinc

## 2. INTRODUCCIÓN

CTCF es un importante regulador transcripcional y organizador de la estructura tridimensional del genoma, frecuentemente alterado en cáncer. El gen *CTCF*, localizado en el cromosoma 16, contiene 12 exones, de los cuales 10 codifican la proteína CTCF, compuesta por 727 aminoácidos, que se encuentran distribuidos en tres dominios: un dominio N-terminal, uno C-terminal, y uno central compuesto por 11 dedos de zinc, encargados de su unión al ADN. Los estudios sobre CTCF han aumentado exponencialmente a lo largo de los años, y su importancia se pone de manifiesto al estar muy bien conservado durante la evolución en los vertebrados. Las funciones que desempeña CTCF dependen en gran medida de la localización de sus secuencias diana, sus modificaciones postraduccionales, y las proteínas con las que interactúa. En sus orígenes se describió como factor de transcripción, posteriormente se supo que desempeñaba también otras funciones como son la unión a secuencias *insulators*, regulación epigenética o la organización tridimensional de la cromatina, y más recientemente se ha descubierto su capacidad de unirse al ARN.

En este TFG se ha hecho una revisión bibliográfica de artículos desde el descubrimiento de CTCF hasta la actualidad, describiendo su estructura, unión al ADN, modificaciones postraduccionales, interacciones con otras proteínas, funciones, y especialmente su papel en cáncer. Además de la revisión bibliográfica, hemos utilizado diferentes plataformas y programas en línea para completar el análisis y descripción de las diferentes secciones del trabajo. En la parte de su estructura, por medio de los programas ENCODE y PYMOL hemos analizado la estructura del gen y de la proteína. El resto de los programas se han utilizado para los apartados de la implicación de CTCF en cáncer: CanProVar para estudiar aquellas proteínas con las que interacciona CTCF implicadas en procesos cancerígenos; GEPIA para determinar la expresión diferencial de CTCF en tejidos tumorales y normales, y cBioPortal para analizar tipos de mutaciones en cada tumor, rutas en las que CTCF está alterado, variantes estructurales, etc.

En el contexto del cáncer, se sabe que CTCF está implicado en la regulación de una amplia variedad de procesos alterados en células tumorales, como son la expresión génica, la metilación del ADN y la arquitectura cromosómica, entre otras. El cáncer representa una de las principales causas de morbilidad y mortalidad. Aunque la incidencia varía según la región y el tipo de cáncer, se observa año tras año un aumento en la prevalencia, impulsado por factores como el envejecimiento, la urbanización y los cambios en el estilo de vida. Se origina a partir de mutaciones en el ADN que afectan a genes involucrados en la regulación del crecimiento celular, la reparación del ADN, la apoptosis, etc. Las principales clases de genes implicados en el cáncer incluyen: los oncogenes, que promueven el crecimiento celular, y los genes supresores de tumores, que inhiben el crecimiento celular y la progresión tumoral. Las alteraciones cromosómicas, como translocaciones, amplificaciones y deleciones, son comunes en el cáncer. Los avances en tecnologías de secuenciación y análisis genómico han permitido identificar mutaciones específicas asociadas con diferentes tipos de cáncer, lo que está llevando al desarrollo de terapias dirigidas y la medicina de precisión. En el caso de CTCF, la comprensión de sus mecanismos moleculares alterados en los tumores, podrían impulsar su investigación como nuevo biomarcador, terapias innovadoras y enfoques de prevención y detección temprana basadas en esta proteína.

### 3. MÉTODOS

#### 3.1. Revisión bibliográfica

Se ha realizado una revisión bibliográfica sobre CTCF, su estructura, modificaciones, *partners* y funciones, y su implicación en cáncer. En la bibliografía se han incluido 83 referencias, principalmente extraídas de la base de datos PubMed (<http://pubmed.ncbi.nlm.nih.gov>). PubMed es un motor de búsqueda gratuito que accede principalmente a la base de datos MEDLINE, de acceso libre y especializada en ciencias de la salud, con más de 19 millones de referencias bibliográficas. Como gestor bibliográfico se ha empleado Mendeley Cite (<https://www.mendeley.com/reference-management/mendeley-cite>).

#### 3.2. ENCODE

La *Encyclopedia of DNA Elements* (ENCODE, UCSC GENOME BROWSER) es una herramienta en línea que permite explorar y analizar genomas completos de multitud de organismos. El sitio web (<https://genome.ucsc.edu/>) proporciona una interfaz intuitiva para analizar las secuencias de ADN, variantes genéticas, regiones reguladoras, etc. Además, permite comparar genomas de diferentes especies y realizar análisis avanzados.

Se ha utilizado la Versión Diciembre 2013. El término de búsqueda fue: *CTCF (Homo sapiens CCCTC-binding factor (CTCF), transcript variant 1, mRNA (from RefSeq NM\_006565))*. Se analizó el locus de CTCF, la distribución de intrones y exones, y la secuencia codificante. También se analizó la localización de los principales dominios estructurales en la secuencia de CTCF, extraído de Uniprot.

#### 3.3. PYMOL

PyMOL (<https://pymol.org/>) es un programa de visualización molecular de código abierto que permite visualizar y analizar las estructuras tridimensionales de proteínas, ácidos nucleicos y otras macromoléculas. Además de la visualización, también permite realizar mediciones, superposiciones de estructuras, animaciones y analizar interacciones moleculares. En este TFG se ha utilizado PyMOL para crear una figura de la estructura terciaria de CTCF (Figura 2). La estructura se descargó de Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB), que es un recurso en línea (<https://www.rcsb.org/>) que almacena y proporciona acceso a estructuras tridimensionales de proteínas, ácidos nucleicos y otras macromoléculas biológicas. Se editaron los colores de los aminoácidos para distinguir los dedos de zinc de CTCF.

#### 3.4. CBIOPORTAL

cBioPortal for Cancer Genomics (<https://www.cbioportal.org/>) es una plataforma en línea de acceso abierto para la exploración interactiva de multitud de datos sobre genómica, como perfiles de expresión génica, mutaciones genéticas y alteraciones cromosómicas, en el contexto de estudios sobre el cáncer, permitiendo entender mejor la biología subyacente de la enfermedad y descubrir posibles objetivos terapéuticos.

Se ha utilizado para analizar la implicación de CTCF en el cáncer, a partir de un conjunto de 32 estudios de The Cancer Genome Atlas (TCGA). La plataforma nos permite analizar diferentes aspectos como los tipos de mutaciones y su frecuencia en cada tipo de cáncer (*oncoprint* y *cancer types summary*), las mutaciones y modificaciones postraduccionales (*mutations*), las rutas en las que CTCF está alterado en cáncer (*pathways*), y una nueva función que son las variantes estructurales (*structural variants*), entre otras.

### 3.5. GEPIA

GEPIA (Gene Expresión Profiling Interactive Analysis) (<http://gepia.cancer-pku.cn/index.html>) es una herramienta que permite a los usuarios explorar y analizar datos de expresión génica en el contexto del cáncer. Actúa como mediador entre los datos sobre genómica del cáncer y la información integrada a los usuarios, facilitando la extracción de datos. Utiliza datos del TCGA y del Consorcio Internacional del Genoma de Cáncer (ICGC) para proporcionar información sobre la expresión diferencial de genes en distintos tipos de tumores, que es para lo que se ha utilizado en este TFG.

### 3.6. CanProVar

CanProVar (Human Cancer Proteome Variation Database) (<http://canprovar2.zhanglab.org/>) es una base de datos que muestra información detallada sobre las variantes genéticas identificadas en diferentes tipos de cáncer, incluidas mutaciones somáticas, variantes de genes relacionados con el cáncer, y su relevancia clínica y funcional. La búsqueda puede hacerse no solo mediante los nombres/ID de las proteínas o el tipo de cáncer, sino también la lista de proteínas, la ubicación o la ruta de los cromosomas. En este caso se ha utilizado para obtener la lista de proteínas que interaccionan con CTCF alteradas en cáncer (Tabla 1).

## 4. ESTRUCTURA DE CTCF

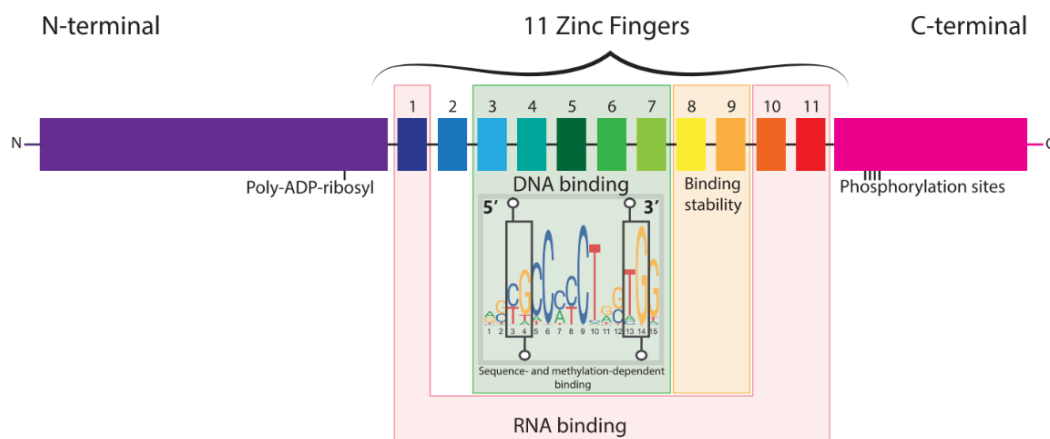
CTCF se describió por primera vez como un represor de la transcripción, que se unía al promotor de *c-myc* de pollo. Su nombre proviene de su sitio de unión al ADN, que posee tres repeticiones espaciadas de la secuencia CCCTC, por lo que se denominó factor de unión a CCCTC o CTCF (Klenova et al., 1993; Lobanenko et al., 1990). Posteriormente, se descubrió que también actúa como activador transcripcional (Vostrov & Quitschke, 1997) además de tener muchas otras funciones, ya que se trata de una proteína expresada de forma ubicua y altamente conservada en los vertebrados, especialmente en mamíferos (revisado recientemente en Del Moral-Morales et al., 2023).

El gen *CTCF* está localizado en el cromosoma 16q22.1, una región frecuentemente delecionada en diversas enfermedades malignas humanas. Utilizando la plataforma ENCODE (UCSC GENOME BROWSER), en el anexo 1 (Figura A1) se muestra la estructura genómica de *CTCF*, la secuencia codificante y los dominios que se detallan a continuación.

El gen *CTCF* contiene 12 exones. Los exones 3-12 codifican una proteína de 82 kDa, compuesta por 727 aminoácidos, que se encuentran distribuidos en tres dominios



diferentes (Figura 1): un dominio N-terminal, uno C-terminal, y uno central, que es el dominio de unión al ADN (Del Moral-Morales et al., 2023).



**Figura 1: Organización de los dominios de la proteína CTCF:** organización lineal de CTCF con sus tres dominios funcionales y la modificación postraduccional más frecuente en cada uno de ellos. Dominio central: los dedos de zinc (ZF) del 3 al 7 (cuadrado verde) se unen al ADN. Los ZF 1, 10 y 11 son los encargados de unirse al ARN (cuadrado rojo). Los ZF 8 y 9 son los que proporcionan estabilidad a estas uniones (cuadrado amarillo). Los recuadros negros resaltan las posiciones donde la presencia de un dinucleótido CpG metilado en su secuencia de unión puede prevenir la unión de CTCF al ADN (Segueni & Noordermeer, 2022).

El dominio central está a su vez compuesto por 11 dominios de dedos de zinc (ZF), especialmente bien conservados durante la evolución en los vertebrados, pues son fundamentales para el funcionamiento de CTCF (Dehingia et al., 2022). CTCF utiliza combinaciones diferentes de estos dedos de zinc para ser capaz de reconocer y unirse a gran variedad de secuencias de ADN, fundamentalmente en áreas intergénicas, pero también intragénicas y promotoras, pudiendo así modificar la arquitectura de los cromosomas de múltiples formas. Además, puede unirse a la matriz nuclear y estabilizar su arquitectura. Es por esto que, a día de hoy, se le considera una proteína multivalente (Li et al., 2020). A pesar de poderse unir a regiones del ADN muy dispares, aproximadamente el 80% de las secuencias a las que se une contienen la secuencia consenso de 12 a 15 pb (Figura 1) 5'-NCA-NNA-G(G/A)N-GGC-(G/A)(C/G)(T/C)-3' (Hashimoto et al., 2017), que es reconocida por los dedos del 3 al 7. Por su parte, las secuencias flanqueantes no conservadas son reconocidas por los dedos 1 y 2 o del 8 al 11, que ayudan a estabilizar el complejo CTCF-ADN. Además de al ADN, CTCF también tiene la peculiaridad de ser capaz de unirse al ARN, gracias a los ZF1, ZF10 y ZF11, que contienen un dominio de unión al ARN (RBD) proporcionando a la proteína puntos de anclaje adicionales (Debaugny & Skok, 2020).

Como vemos, las acciones y funciones de CTCF dependen en gran medida de la localización de sus secuencias diana, además de las modificaciones postraduccionales que pueden sufrir cada uno de los 3 dominios.

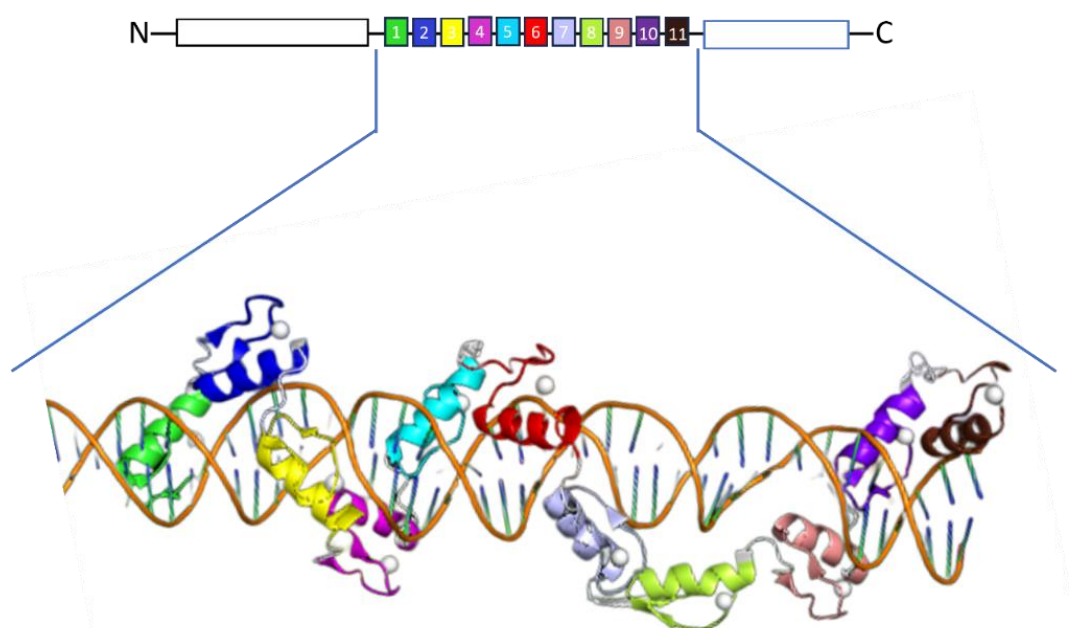
CTCFL (*CTCF like*), también conocido como BORIS, es el parólogo de CTCF. Su origen se encuentra en la duplicación genética de CTCF durante la evolución. A diferencia de CTCF, que se encuentra expresada de forma constitutiva y ubicua, CTCFL solo se expresa de forma transitoria y únicamente en células germinales masculinas premeióticas (Nishana

et al., 2020). Sin embargo, se expresa en células tumorales y tiene un importante papel en cáncer, como se describirá más adelante. El dominio de 11 dedos de zinc de BORIS presenta alta homología con el de CTCF. Sin embargo, los dominios N y C terminales son diferentes, lo que sugiere funciones diferentes de estos dominios en las dos proteínas (Ohlsson et al., 2010).

## 5. UNIÓN DE CTCF AL ADN

Tras la realización de múltiples experimentos para mapear los sitios de unión de CTCF en todo el genoma, se ha demostrado que en él existen entre 40000 y 80000 sitios de unión a CTCF, dependiendo del contexto celular (Bastiaan Holwerda & de Laat, 2013).

A diferencia de la mayoría de factores de transcripción, CTCF se une fundamentalmente a secuencias intergénicas, y normalmente a una determinada distancia del sitio de inicio de la transcripción (TSS). CTCF fue una de las primeras proteínas que demostró mediar el bucle de cromatina entre sus sitios de unión (Debaugny & Skok, 2020). Como se ha comentado previamente, la mayoría de las secuencias de ADN a las que se une CTCF contienen la secuencia 5'-NCA-NNA-G(G/A)N-GGC-(G/A)(C/G)(T/C)-3' (Hashimoto et al., 2017). A estas secuencias consenso se les denomina CTS (Ohlsson et al., 2010). De todos los CTS descritos, un 46% se encuentran en regiones intergénicas. El resto se reparten entre regiones promotoras (20%), exones (12%) e intrones (22%).



**Figura 2: Unión de CTCF al ADN.** Los ZF del 3 al 7 contactan cada uno con 3 bases de la secuencia consenso constituida por 15 pares de bases. El código de colores de los ZF unidos al ADN corresponde a los cuadrados numerados en el esquema superior. Esta figura ha sido generada utilizando el programa PYMOL.

Las diferentes mutaciones que pueden ocurrir en los CTS pueden cambiar la especificidad de unión de CTCF al ADN. Por ejemplo, las mutaciones C por A o C por G en un CTS en la posición 43 están relacionadas con el sitio de inicio de la transcripción

del gen *XIST*. La primera de ellas lleva a la inactivación de *XIST*, y la segunda a su activación (Ohlsson et al., 2010).

Cada uno de los once dedos de zinc desempeña una función específica, como se ha mencionado en el apartado anterior (Figura 2). Los dedos del 3 al 7 establecen el 80% de los contactos específicos de secuencia, ZF 2 también es capaz de hacerlo, pero se encuentra demasiado lejos como para formar enlaces específicos. Por su parte, los dedos 2, 8 y 9 son capaces de estabilizar dichas uniones, aunque cabe resaltar que ZF 8 es capaz de hacerlo por sí solo (Nishana et al., 2020). Existe cierta controversia sobre la unión de los ZF 10 y 11 al ADN, aunque sí que se sabe que se pueden unir al ARN junto con ZF 1 (Hashimoto et al., 2017). Las secuencias de ADN que flanquean la secuencia central pueden modular la unión de CTCF, cambiando su afinidad, tanto incrementándola (*upstream*) como disminuyéndola (*downstream*). La secuencia conservada *upstream* se une a ZF 9-11 y la secuencia *downstream* a ZF 1 y 2 (Hashimoto et al., 2017).

Existe un mecanismo capaz de mediar la interacción entre regiones distantes del ADN, a través de la dimerización de CTCF. Consiste en la interacción de una molécula de CTCF con otra mediante los dedos de zinc y el dominio C-terminal, aunque no se descarta que intervenga también el dominio N-terminal. La superficie para la dimerización va variando en función del CTS (Ohlsson et al., 2010).

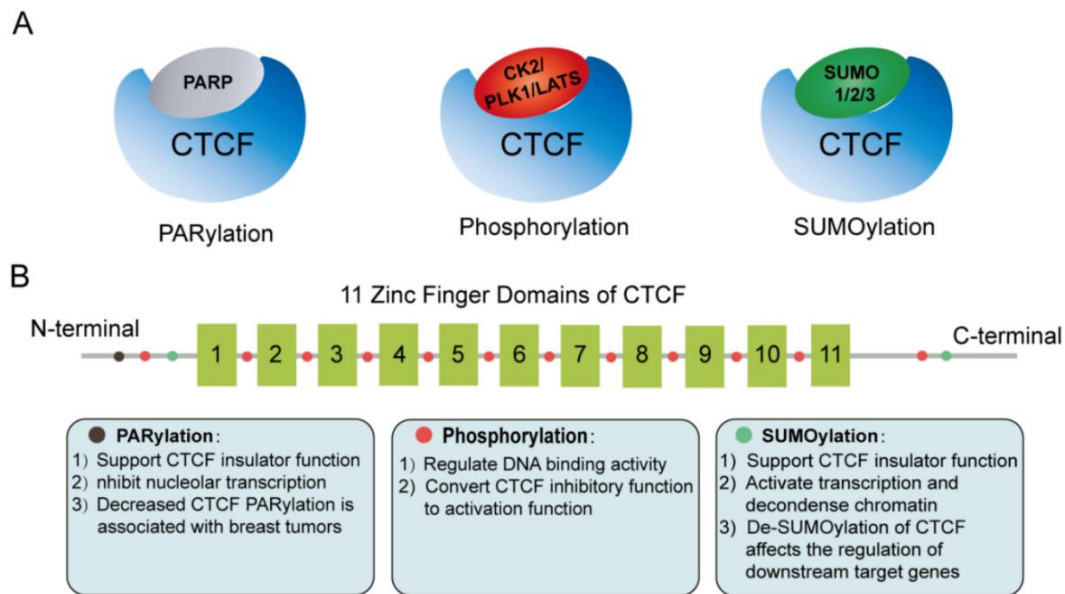
Debido a la similitud en sus regiones de unión al ADN, CTCFL (BORIS) tiene la capacidad de unirse y competir con CTCF en algunos de sus sitios de unión (Debaugny & Skok, 2020). Los dedos de zinc de CTCFL tienen un 74,1% de homología con los de CTCF; sin embargo, difieren mucho en los otros dos dominios (N-terminal y C-terminal), ya que solo comparten un 15% de las secuencias, otorgándoles esto funciones únicas y diferentes (Bailey et al., 2021).

El ARN es capaz de regular la unión de CTCF al ADN tanto de forma directa como indirecta. Por un lado, es capaz de estabilizar la unión del ADN y la cromatina, y por otro puede competir con el ADN y bloquear dicha unión (Sun et al., 2022).

## 6. MODIFICACIONES POSTRADUCCIONALES

Las modificaciones postraduccionales de las proteínas se definen como aquellos cambios funcionales y estructurales que ocurren después de su síntesis por los ribosomas. El objetivo de estos cambios es, por un lado, permitir que algunas proteínas ejerzan sus funciones, y por otro regular su actividad biológica (Lee et al., 2023).

En el caso de CTCF, fundamentalmente son tres (Figura 3): poli(ADP-ribosil)ación (PARilación), SUMOilación y fosforilación.



**Figura 3: Modificaciones postraduccionales de CTCF** (A) CTCF sufre tres modificaciones: PARilación, fosforilación y SUMOilación. (B) Sitios de las modificaciones postraduccionales y su papel regulador de la función de CTCF (Sun et al., 2022).

**PARilación.** La poli (ADP-ribosil)ación está mediada por la poli(ADP-ribosa) polimerasa (PARP), y está directamente implicada en la regulación de la función de CTCF. El dominio más frecuentemente afectado por esta modificación es el dominio N-terminal, formándose un complejo CTCF/PARP1 que modifica de forma reversible dicha modificación para poder regular la función de CTCF (Yu et al., 2004).

Tras la PARilación, se forma la isoforma CTCF PARilada (180 kDa), que sobre todo se encuentra en el nucleolo, y la PARilación de CTCF reprime la transcripción nucleolar, tal como describió nuestro grupo (Torrano et al., 2006). Además, la PARilación de CTCF afecta a la unión de este a la cromatina, relacionándose su disminución con procesos de proliferación celular y tumorigénesis (Sun et al., 2022).

**SUMOilación.** CTCF puede ser modificado por unión de pequeñas proteínas similares a la ubiquitina, que se denominan SUMO. Se incluyen en tres grupos, SUMO 1,2 y 3, y se sabe que todas ellas pueden modificar a CTCF (MacPherson et al., 2009). La vía bioquímica para la modificación de proteínas por las proteínas SUMO es mecánicamente similar, pero con funciones distintas a la de la ubiquitina. A diferencia de la poliubiquitina, la modificación SUMO de una proteína no la lleva a la degradación, sino que desempeña papeles clave en la regulación transcripcional, estructura de la cromatina y reparación del ADN. Actualmente se conocen dos sitios fundamentales de SUMOilación, ubicados en los dominios N y C terminales. Especialmente, se identificó un dominio de 107 aminoácidos en la región N-terminal de CTCF que activa la transcripción, y la SUMOilación completa de este dominio suprime la actividad transcripcional de CTCF y previene la apertura de la cromatina (Sun et al., 2022). Se sabe que la SUMOilación es funcionalmente importante para la acción represiva de CTCF en el Promotor P2 de c-MYC (Sun et al., 2022).

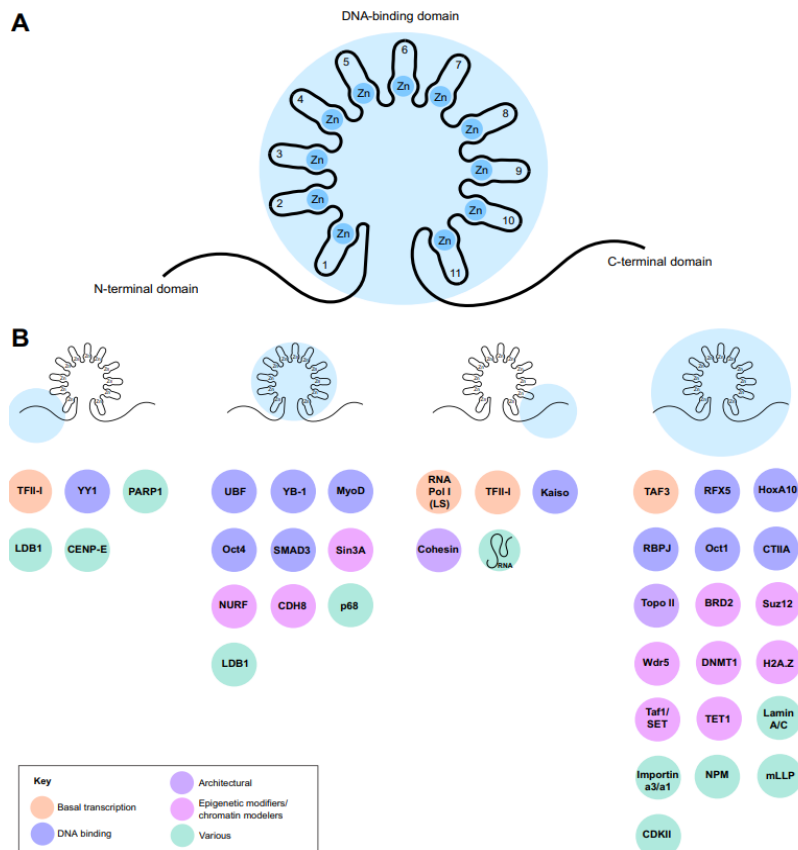
**Fosforilación.** Por último, la fosforilación es una de las modificaciones clásicamente más estudiadas. La fosforilación de CTCF es dinámica durante el desarrollo y diferenciación celular (Delgado et al., 1999; Klenova et al., 2001). En este caso, es el dominio C-terminal el que sufre mayoritariamente la fosforilación, sobre todo mediada por la caseína quinasa 2 (CK2). CK2 es capaz de convertir la función inhibidora de CTCF en una función activadora. Además, los residuos de aminoácidos Thr289, Thr317, Thr346, Thr374, Ser402, Ser461 y Thr518 de CTCF se fosforilan durante la mitosis para regular su actividad de unión al ADN (Sun et al., 2022).

Existen otras proteínas capaces de fosforilar a CTCF. Por ejemplo, la mediada por la proteína poloquinasa 1 (PLK1) en la serina 224 (Ser224-P) de CTCF se produce en la fase G2/M del ciclo celular (Tyler et al., 2019). CTCF también es un sustrato para la quinasa LATS, y la fosforilación de CTCF inhibe su actividad de unión al ADN. Por lo tanto, las señales externas pueden modular la estructura del genoma 3D mediante la fosforilación de CTCF (Tyler et al., 2019).

## 7. INTERACCIÓN DE CTCF CON OTRAS PROTEÍNAS

CTCF es capaz de reconocer e interactuar con gran cantidad y variedad de proteínas, denominadas *partners* (Ziatanova & Caiafa, 2009). Algunas de ellas se unen a CTCF mediante secuencias específicas, pero otras no lo hacen en ningún dominio en particular. Dichas proteínas pueden dividirse en varios grupos según su función (Figura 4) (Arzate-Mejía et al., 2018).

- Proteínas de unión a ADN: incluyendo factores de transcripción y cofactores. Algunos de ellos son: YB-1, YY1, Kaiso, RFX5, CIITA, SMAD3, OCT1 y 4, TAF3 y UBF.
- Proteínas arquitectónicas: ADN topoisomerasa II y cohesina, principalmente.
- Modificadores epigenéticos y modeladores de la cromatina: SIN3A, NURF, CDH8, BRD2, Suz12, Wdr5, DNMT1 y H2A.Z, entre otros.
- Proteínas de transcripción basal: ARN polimerasa I, TFII-I y TAF 3.
- Otras que no pueden clasificarse en ninguno de los grupos anteriores: PARP1, CENP-E, LDB1 y NPM1, entre otros.



**Figura 4: CTCF interactúa con gran variedad de proteínas.** (A) Dominios estructurales de CTCF. (B) Algunas se unen a dominios específicos en CTCF. Multitud de ellas interactúan con el dominio de dedos de zinc. Solo ARNP II, cohesina, Kaiso y TFII-1 con el dominio C-terminal. Muy pocas lo hacen con el dominio N-terminal (Arzate-Mejía et al., 2018).

A continuación, se describen brevemente algunos de los *partners* que han demostrado ser esenciales para la función de CTCF o que están implicados en su papel en cáncer.

## 7.1. Proteínas de unión a ADN

### 7.1.1. YY1

YY1 desempeña un importante papel en la inactivación del cromosoma X (Ziatanova & Caiafa, 2009). Además, la sobreexpresión y activación de YY1 están relacionadas con la pérdida del control de la proliferación celular, produciendo diversos efectos sobre p53 y PARP1 (Sun et al., 2022).

### 7.1.2. UBF

UBF es un factor de transcripción esencial para la síntesis del ARNr por la ARN pol I y fue identificado por nuestro grupo como el primer *partner* descrito capaz de interactuar tanto con CTCF como con BORIS (Van De Nobelen et al., 2010).

### 7.1.3. SMAD3 y SMAD4

SMAD es una familia de proteínas transductoras de señales intracelulares. SMAD3 participa en la vía de señalización del factor de crecimiento transformante beta (TGF- $\beta$ ) y transmite señales desde la superficie celular al núcleo, regulando la actividad genética y la proliferación celular (Xu et al., 2016). Esta proteína forma un complejo con otras

proteínas SMAD, como SMAD4 y se unen al ADN, funcionando como factor de transcripción y supresor de tumores. La desregulación de la señalización de SMAD3 y SMAD4 se ha implicado en cáncer (Jeon et al., 2023).

#### 7.1.4. OCT-4 y OCT-1

Respecto a OCT4, es una proteína cuyo papel principal, junto a otros factores, es mantener la pluripotencialidad de las células madre embrionarias. La interacción de OCT4 y CTCF es conocida por su papel en la promoción de la autorrenovación y la inhibición de la diferenciación (Patel & Parchem, 2022).

#### 7.1.5. MAZ

MAZ es un factor de transcripción que puede actuar como activador o represor transcripcional. MAZ forma un complejo con c-MYC y se ha observado una expresión aberrante de MAZ en varios tipos de cáncer (Deen et al., 2021). Además, MAZ participa en la regulación de procesos celulares incluida la progresión del ciclo celular, la apoptosis y la diferenciación. MAZ comparte funciones con CTCF, como son la actividad de aislamiento y la interacción con la subunidad Rad21 de la cohesina (Sun et al., 2022). Además, MAZ estabiliza la unión de CTCF a sus sitios de unión (Xiao et al., 2021).

### 7.2. Proteínas arquitectónicas

#### 7.2.1. Cohesina

La cohesina es un complejo de proteínas que desempeña un papel fundamental en el ciclo celular, al mantener en contacto a las dos cromátidas hermanas desde la fase S hasta que se separan en la anafase de la mitosis. Además, también participa en la recombinación homóloga y la formación de bucles de ADN, e interviene en su reparación (Li et al., 2020).

El complejo está compuesto por 4 subunidades: SMC1, SMC3, Scc1 y Scc3 (SA1 y SA2 en humanos). SMC1 y SMC3 son miembros de la familia de proteínas de mantenimiento estructural de los cromosomas (SMC), mientras que se cree que Scc1 y Scc3 (subunidad del complejo de cohesina 1 y 3) participan en la formación de una estructura de anillo alrededor de las dos cromátidas (Ishiguro, 2019). En los mamíferos, se requieren otras dos proteínas que no pertenecen al grupo SMC, Scc2 y Scc4, para cargar el complejo de cohesina en el ADN (Hirano, 2005).

El componente que interactúa directamente con CTCF es el SA1. Los sitios de unión CTCF/cohesina están definidos y compuestos por 4 módulos (1-4). Cada dedo de zinc de CTCF reconoce un módulo concreto: ZF3, 4-7 y 9-11 reconocen los módulos 4, 3-2 y el 1, respectivamente, y el ZF8 se adapta a distancias variables entre los módulos 1 y 2 (Li et al., 2020).

Aproximadamente el 70% de los sitios de unión a CTCF y cohesina son compartidos por ambas proteínas. A pesar de esto, hay ciertos sitios ocupados exclusivamente por CTCF o exclusivamente por la cohesina (Parelho et al., 2008).

### 7.2.2. TOP2A

La topoisomerasa II (TOP2) es una enzima que afecta la estructura topológica del ADN al interactuar con el ADN de doble hélice, desempeñando así un papel importante en la replicación, transcripción, recombinación, condensación y segregación del ADN. TOP2A se encuentra sobreexpresada en células en proliferación y en numerosos tumores humanos, incluidos los cánceres de colon, gástricos y de mama, linfomas, etc. Además, su localización es prevalente en los sitios de unión de CTCF/cohesina que flanquean a los TAD. Las topoisomerasas tipo II son diana de algunos fármacos quimioterápicos como las antraciclinas (doxorubicina y epirubicina) (Delgado et al., 2019).

### 7.3. Modificadores epigenéticos

La epigenética es el campo de estudio que se encarga de los cambios en la estructura de la cromatina que no producen alteraciones en la secuencia del ADN subyacente. Los cambios epigenéticos pueden ayudar a determinar si los genes están activados o desactivados y, por tanto, influyen en la producción de proteínas de las células (Klemm et al., 2019). La estructura y empaquetamiento de la cromatina es un proceso altamente regulado, crucial para el control de la expresión génica. Los reguladores epigenéticos incluyen proteínas que modulan las modificaciones postraduccionales de las histonas, la metilación del ADN y la remodelación de la cromatina. Algunos de los mecanismos epigenéticos incluyen modificaciones de las histonas como: metilación de la histona H3 en la lisina 4 (H3K4me), marcador de regiones promotoras activas; trimetilación de la histona H3 en la lisina 27 (H3K27me3), una marca represiva de elementos potenciadores; o acetilación de la histona H3 en lisina, asociada con genes transcripcionalmente activos.

En lo que respecta a CTCF, varios de sus *partners* pertenecen a dicho grupo de proteínas (Sun et al., 2022). Entre ellos se encuentran:

#### 7.3.1. H2A.Z

Además de las clásicas histonas, se han encontrado otras variantes, como H2A.X, H2A.Z, y otras (Cutter & Hayes, 2015). H2A.Z es una variante de histona que influye en la accesibilidad y la estructura de la cromatina, afectando así a la unión de los factores de transcripción y la actividad de la maquinaria transcripcional (Giaino et al., 2019). La coexistencia de sitios de unión de CTCF y H2A.Z en ciertos loci genómicos puede indicar un papel conjunto en el mantenimiento de la arquitectura de la cromatina (Wen et al., 2020).

#### 7.3.2. SUZ12

SUZ12 es un componente del complejo represivo 2 de Polycomb (PRC2), que desempeña un papel en el silenciamiento de genes durante el desarrollo y la diferenciación. Las mutaciones o desregulación de SUZ12 se han implicado en diversos cánceres, especialmente en los tumores del estroma endometrial (Hrzenjak, 2016).



### 7.3.3. WDR5

WDR5 forma parte de los complejos COMPASS y MLL, complejos con actividad histona metiltransferasa, que producen modificaciones en las histonas asociadas a la activación transcripcional y que están alterados en algunas neoplasias hematológicas. Su función incluye la remodelación de la cromatina y la regulación de la expresión genética. El adecuado funcionamiento de WDR5 es fundamental para el desarrollo embrionario, la diferenciación celular y el mantenimiento de la pluripotencialidad de las células madre (Sun et al., 2022).

## 7.4. Proteínas de transcripción basal y otras

### 7.4.1. ARN polimerasa

Como representante de este grupo encontramos la ARN polimerasa II, que interactúa con CTCF regulando la transcripción. Pero no es el único, también interactúa con varios factores de transcripción para regular la estructura tridimensional del genoma, que es esencial para la autorrenovación y diferenciación de las células madre embrionarias (Sun et al., 2022).

### 7.4.2. PARP1

La enzima central para la producción de PAR en las células y el objetivo principal de la poli(ADP-ribosil)ación durante el daño del ADN es la poli(ADP-ribosa) polimerasa 1 (PARP1). Las alteraciones de PARP1 están asociadas con la resistencia al inhibidor de PARP1, que regula la función de las células Treg y la expresión de PDL1 en las células tumorales; y la sobreexpresión de PARP1 se asocia significativamente con un comportamiento agresivo y resistencia a la quimioterapia en varios tumores (Zhang et al., 2021).

### 7.4.3. NPM1

La nucleofosmina 1 (NPM1) es una proteína que desempeña un papel en varios procesos celulares, incluido el ensamblaje de ribosomas, la duplicación de centrosomas y la regulación de la proteína supresora de tumores p53. NPM1 se encuentra principalmente en el nucleolo, pero se desplaza entre el nucleolo y el nucleoplasma. Las mutaciones de NPM1 representan la lesión genética más común en la leucemia mieloide aguda (LMA) en adultos. Debido a sus características únicas, la AML con mutación en NPM1 se reconoce como una entidad distinta en la clasificación de neoplasias hematopoyéticas de la OMS (Falini et al., 2020).

Finalmente, hemos utilizado la base de datos CanProVar (ver Métodos) para analizar las proteínas que interaccionan con CTCF y que están relacionados con su papel en el cáncer. En la tabla 1 se recogen los resultados obtenidos.

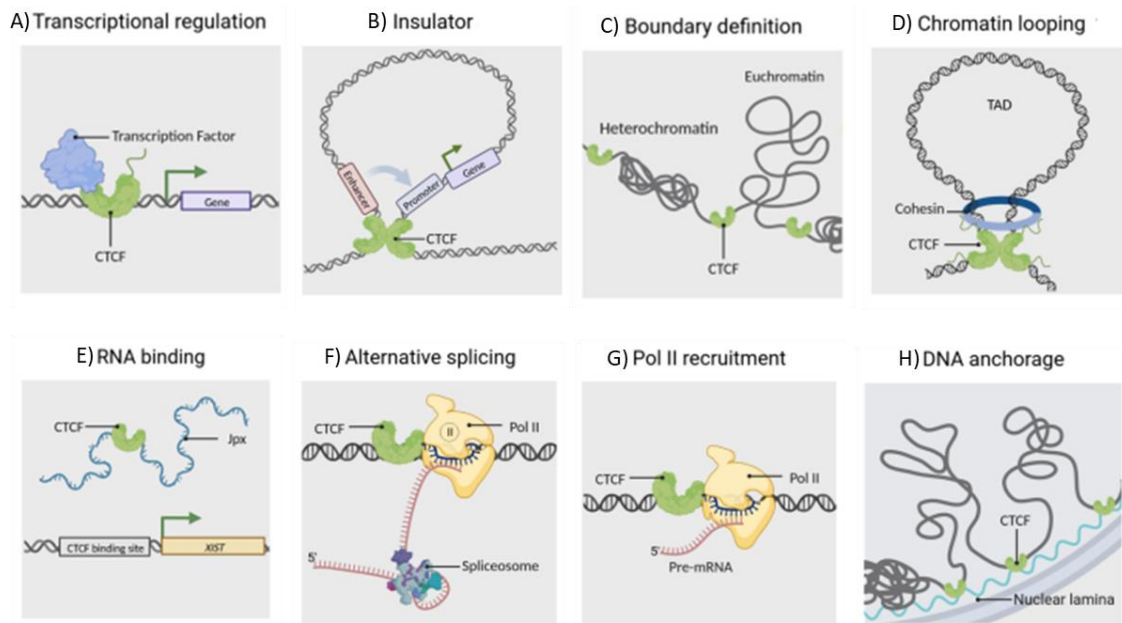
**Tabla 1: *Partners de CTCF en cáncer analizados con CanProVar.*** En la tabla extraída de CanProVar se muestran aquellas proteínas que interactúan con CTCF en algún tipo de cáncer. Las más representativas han sido descritas en el apartado *partners* y se mencionarán en los tumores correspondientes.

| NO. | ProteinA | ProteinB   |
|-----|----------|------------|
| 1   | CTCF     | KPNA2      |
| 2   | RAD21    | CTCF       |
| 3   | CTCF     | KPNA1      |
| 4   | LMNA     | CTCF       |
| 5   | CTCF     | HIST2H2AA4 |
| 6   | PARP1    | CTCF       |
| 7   | CTCF     | CHD8       |
| 8   | YY1      | CTCF       |
| 9   | CTCF     | SMAD3      |
| 10  | CTCF     | SMAD4      |
| 11  | CTCF     | SUZ12      |
| 12  | CTCF     | HIST2H2AC  |
| 13  | HMGB1    | CTCF       |
| 14  | CTCF     | SMC1A      |
| 15  | CTCF     | SUMO2      |
| 16  | NPM1     | CTCF       |
| 17  | KPNA4    | CTCF       |
| 18  | RXRA     | CTCF       |
| 19  | POU5F1   | CTCF       |
| 20  | TGFB1    | CTCF       |
| 21  | UBC      | CTCF       |
| 22  | CTCF     | CHD7       |
| 23  | TOP2A    | CTCF       |
| 24  | TAF1     | CTCF       |
| 25  | CTCF     | SIN3A      |
| 26  | CTCF     | YBX1       |
| 27  | CTCF     | HSPB9      |
| 28  | POLR2A   | CTCF       |
| 29  | H2AFZ    | CTCF       |

## 8. FUNCIONES MOLECULARES DE CTCF

Originalmente CTCF se describió como un clásico factor de transcripción, inicialmente únicamente como represor, y después también como activador. Posteriormente se supo que CTCF no solo es un regulador transcripcional, sino que desempeña un papel muy importante en diversas funciones (Figura 5). Se le ha considerado una proteína arquitectónica por mediar el mantenimiento de los TAD (dominios topológicamente asociados), definiendo los límites entre la eucromatina y la heterocromatina (Kentepozidou et al., 2020). Por otro lado, regula el anclaje del ADN principalmente a la lámina nuclear, pero también a otras estructuras celulares (Del Moral-Morales et al., 2023), y funciona como una proteína de andamio para factores de transcripción y epigenéticos (Klemm et al., 2019). Además, actúa como un aislante controlando las interacciones entre potenciadores y promotores (Del Moral-Morales et al., 2023).

Recientemente se ha demostrado que está también involucrado en otros procesos como son la interacción con el ARN y la unión a la ARN polimerasa II (Del Moral-Morales et al., 2023).



**Figura 5: Descripción general de las funciones de CTCF:** A) Regulador transcripcional, B) Proteína de unión a *Insulators*, C) Definición de límites de la cromatina, D) Formación de bucles de cromatina, E) Unión al ARN, F) *Splicing* alternativo, G) Unión a ARN polimerasa II, H) Anclaje del ADN. Modificado de (Del Moral-Morales et al., 2023).

### 8.1. Regulador transcripcional

Originalmente CTCF se describió como un factor de transcripción (Figura 5A) que reprimía al gen *c-myc* en los fibroblastos embrionarios del pollo (Klenova et al., 1993); esta función también represora se describió para el gen de la lisozima del pollo (Baniahmad et al., 1990), y para el gen *C-MYC* en humanos y ratones (Filippova et al., 1996).

Hasta entonces, todos los estudios apuntaban a que era un represor transcripcional. Sin embargo, en las células HeLa se encontró por primera vez que CTCF se une *upstream* del sitio de iniciación de la transcripción del gen precursor de la proteína beta amiloide, esta vez actuando como activador transcripcional (Vostrov & Quitschke, 1997).

Además, CTCF regula varios genes implicados en la progresión del ciclo celular, como el gen *RB* (De La Rosa-Velázquez et al., 2007), *p14ARF*, *PIM1* humano y de ratón, y la poloquinasa 1, entre otros (Ohlsson et al., 2010).

### 8.2. Proteína de unión a *insulators*

Los *insulators* o aislantes de cromatina son secuencias de ADN específicas que desempeñan un importante papel en la regulación de la expresión génica, ya que actúan como barreras físicas y funcionales para bloquear o limitar la propagación de señales reguladoras a lo largo de la cromatina, previniendo así interacciones génicas inapropiadas (Figura 5B) (Klemm et al., 2019). Los aislantes pueden clasificarse en dos

grupos según su función principal. El primero es el de los bloqueadores de los potenciadores o *enhancers*, para ello se sitúan entre el potenciador y el promotor o silenciador, para bloquear así su interacción y con ello la transcripción. El segundo es ejerciendo como barrera, para evitar la propagación de la heterocromatina hacia los dominios activos (Klemm et al., 2019; Ohlsson et al., 2010).

El papel de CTCF como proteína de unión a secuencias *insulators* se describió por primera vez para el elemento HS4 del pollo en el límite 5' del locus de la globina beta (Bell et al., 1999). Posteriormente, también se identificaron sitios de unión a CTCF en H19 (ICR), donde actúa como un aislante que resulta en el silenciamiento de la expresión de IGF2 del alelo materno, lo cual vinculó también a CTCF con el mecanismo de la impronta genómica (Kanduri et al., 2000).

Respecto a su función como barrera de la cromatina, esta consiste en regular el equilibrio entre las marcas de histonas activas y represivas (Figura 5C). La cromatina menos condensada o abierta se denomina eucromatina y se caracteriza por contener genes transcripcionalmente activos. Por otro lado, la cromatina más condensada o compacta es inaccesible para la transcripción y se conoce como heterocromatina. Los límites entre la eucromatina y la heterocromatina están marcados por histonas. Como ejemplo, la histona H3 trimetilada en la lisina 27 (H3K27me3) marca el límite de la cromatina inactiva (heterocromatina), y la histona H2A acetilada en la lisina 5 (H2AK5ac) marca el límite en la cromatina activa (eucromatina). Una porción de los sitios de unión a CTCF coincide con estos límites (Cuddapah et al., 2009).

### 8.3. Regulador epigenético

La epigenética desempeña un papel muy importante en el desarrollo normal y en algunas enfermedades, incluido el cáncer. El papel de CTCF como regulador epigenético se caracterizó por primera vez por su capacidad para compensar el silenciamiento genético en el promotor del gen *RB* humano. También para proteger al gen promotor de *TP53* contra marcas represivas de histonas (De La Rosa-Velázquez et al., 2007).

La unión de CTCF puede prevenir la propagación de la metilación del dinucleótido CpG (citosina-fosfato-guanina) y proteger a los promotores del silenciamiento, manteniéndolos libres de la metilación del ADN en un tejido específico. Además, las secuencias consenso de unión del CTCF al ADN contienen CpG susceptibles a la metilación del ADN y la metilación de CpG previene la unión de CTCF, que se une preferentemente a secuencias no metiladas (De La Rosa-Velázquez et al., 2007).

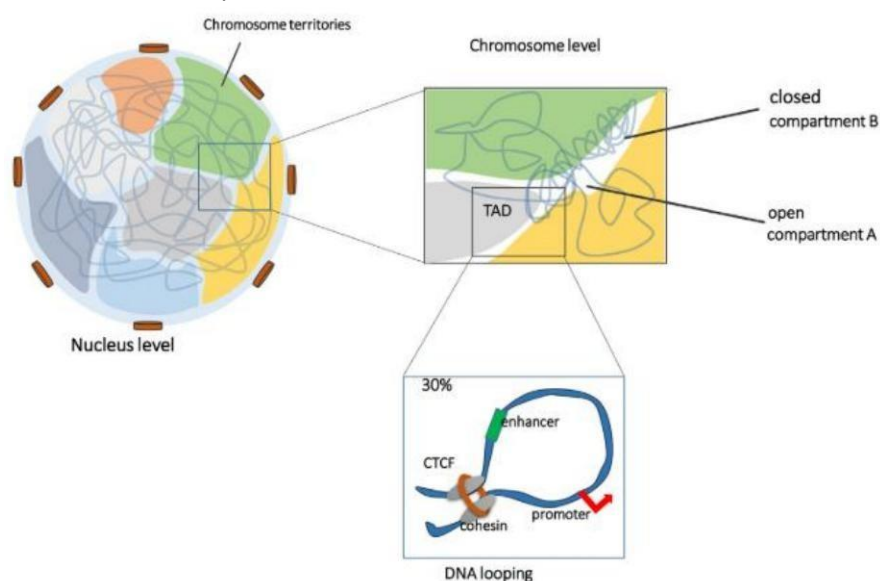
CTCF regula genes implicados en el cáncer como *MYC*, *TP53*, *p16* o *RB* mediante mecanismos epigenéticos. Por ejemplo, la unión de CTCF en el locus *MYC* humano es necesaria para la protección contra la metilación del ADN, lo que conduciría a una pérdida progresiva de la actividad transcripcional (Gombert & Krumm, 2009). Nuestro grupo ha demostrado que CTCF se une y regula epigenéticamente al gen *BCL6*, fuertemente implicado en linfomas agresivos de células B (Batlle-López et al., 2015) y CTCF puede jugar un papel en el tratamiento epigenético de estos linfomas (Cortiguera et al., 2019).

La conclusión que obtenemos es que la relación entre CTCF y cromatina está regulada de forma compleja a través de una variedad de mecanismos moleculares, como son la sensibilidad de los CTS a los cambios en la metilación del ADN (Monteagudo-Sánchez et al., 2024), el posicionamiento de nucleosomas, la transcripción de ARN no codificante, y la regulación del equilibrio entre marcas de cromatina activas y represivas.

#### 8.4. Organización tridimensional de la cromatina

El ADN es capaz de lograr una enorme condensación gracias a sus diferentes niveles estructurales de compactación. Durante la interfase, los cromosomas descondensados ocupan ciertas regiones específicas dentro del núcleo, llamados territorios cromosómicos (Figura 6), pequeñas porciones de cromatina que tienden a interactuar preferentemente entre sí. De este modo, la posición de los genes en el núcleo y su relación con otros genes y estructuras es capaz de modular la función genética. El genoma puede segregarse espacialmente en distintas regiones transcripcionalmente activas (tipo A) o inactivas (tipo B) denominados compartimentos (Gilbert & Allan, 2019).

A su vez, cada compartimento está formado por autoasociaciones de segmentos cromosómicos denominados TADs (dominios topológicamente asociados) (Figura 6), regiones genómicas que interactúan consigo mismas. Por lo tanto, las secuencias de ADN dentro de un TAD interactúan entre sí con más frecuencia que con secuencias de fuera del TAD (Dixon et al., 2012).

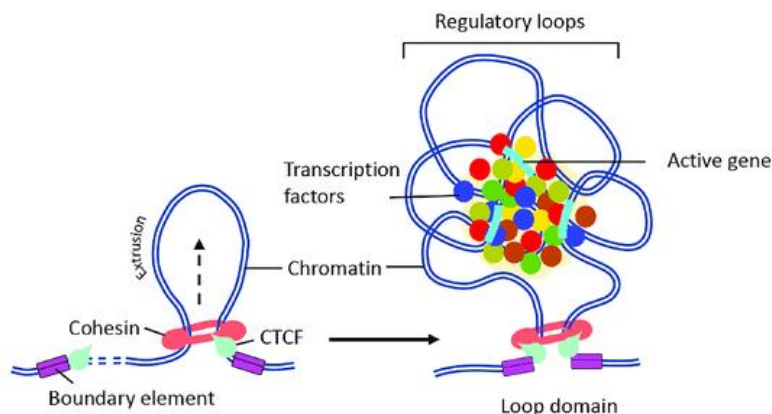


**Figura 6: Organización tridimensional del genoma.** Los cromosomas se organizan en compartimentos seguidos de dominios topológicamente asociados (TAD) y bucles. CTCF y cohesina participan en la formación de bucles de cromatina y el 30% de los bucles reúnen promotores y potenciadores (Wang et al., 2019).

Los sitios de unión de CTCF no sólo se identificaron dentro de los TAD, mediando las llamadas interacciones intraTAD de corto alcance, sino también en sus fronteras (Figura 5D). Los límites de estos dominios están altamente enriquecidos en CTCF y sitios de unión a cohesina (Chen & Lei, 2019). En la Figura 7 se esquematiza el mecanismo de formación del bucle de cromatina por los complejos CTCF/cohesina. Cuando se silencia

CTCF, se interrumpe el aislamiento entre TADs vecinos. Y cuando se están formando los bucles, el silenciamiento de CTCF afecta al bucle comprendido entre sus sitios de unión (Zuin et al., 2014).

Sin embargo, el bucle de cromatina requiere, no sólo la unión de CTCF a la cohesina, sino también a ARN no codificantes (ncRNA), que están implicados en la curvatura de los bordes del dominio y en la formación de grandes regiones de cromatina represiva o activa (Yamamoto & Saitoh, 2019).



**Figura 7: Mecanismo de formación del bucle de cromatina.** Los TADs contienen un número variable de bucles generados mediante extrusión de bucles por complejos CTCF/cohesina. La formación de bucles de cromatina facilita las interacciones entre los elementos promotores y potenciadores (Pongubala & Murre, 2021).

### 8.5. Unión al ARN

Además de poder unirse al ADN, CTCF también tiene la capacidad de hacerlo con el ARN (Figura 5E), teniendo incluso una mayor afinidad con este último. Los principales dominios de unión al ADN son los dedos del 3 al 7, mientras que los de unión al ARN son el 1, el 10 y el 11.

Algunos ARN no codificantes de cadena larga (lncRNAs) interactúan con CTCF y modulan su unión a la cromatina, lo cual podría afectar a la formación de bucles de cromatina y por consiguiente a los patrones de expresión genética (Gabriele et al., 2022). Estos lncRNAs son capaces de actuar como guías a través de un nuevo motivo de unión a ARN (AGAUNGGA), para dirigir a CTCF hacia loci específicos de unión a lo largo del genoma (Kuang & Wang, 2020).

Además, CTCF participa en el procesamiento del ARN, concretamente en el proceso de *splicing* alternativo (Figura 5F) (Sun et al., 2022). Otras de sus funciones son el reclutamiento de la ARN polimerasa II para regular la transcripción (Figura 5G), y regular el anclaje del ADN a ciertas estructuras celulares como la lámina nuclear (Figura 5H).

## 9. DESREGULACIÓN DE CTCF EN CÁNCER

Numerosos estudios señalan la importante función de CTCF en la regulación de procesos tales como la proliferación celular, diferenciación y apoptosis. La desregulación de estos procesos celulares constituye algunos de los más cruciales *Hallmarks* del cáncer, descritos por Hanahan y Weinberg (Hanahan, 2022; Hanahan & Weinberg, 2011). Los niveles de CTCF tienen un gran impacto en las funciones celulares de forma contexto-dependiente. Por ejemplo, su sobreexpresión en algunos modelos resulta en disminución de la proliferación e inducción de la diferenciación celular (Torrano et al., 2005), mientras que en otros casos puede causar apoptosis. Además, tal y como se ha descrito en los apartados anteriores, CTCF desempeña un papel clave en la regulación genética mediante diversos mecanismos moleculares. Por ello, una disfunción en su actividad puede desencadenar una expresión incorrecta de oncogenes y genes supresores de tumores.

La primera vez que se estudiaron las mutaciones de *CTCF* fue en los tumores de mama y próstata (Filippova et al, 1998). El gen *CTCF* mapea en el locus 16q22.1, región que está delecionada en algunos tumores humanos como mama, próstata, ovario y otros (Figura 8), lo que sugirió originalmente su papel como gen supresor de tumores (Klenova et al., 2002).

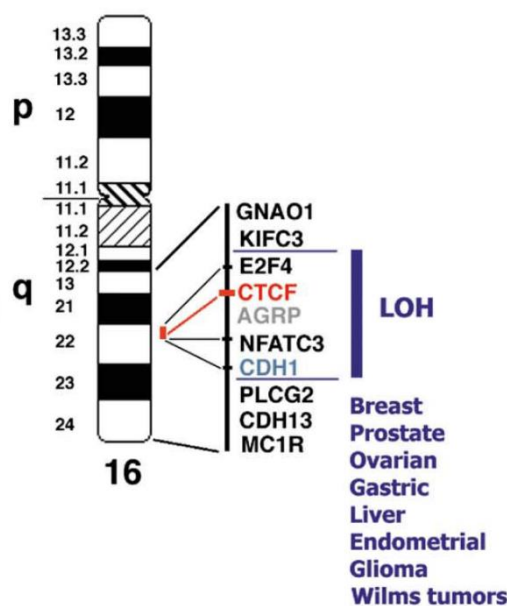


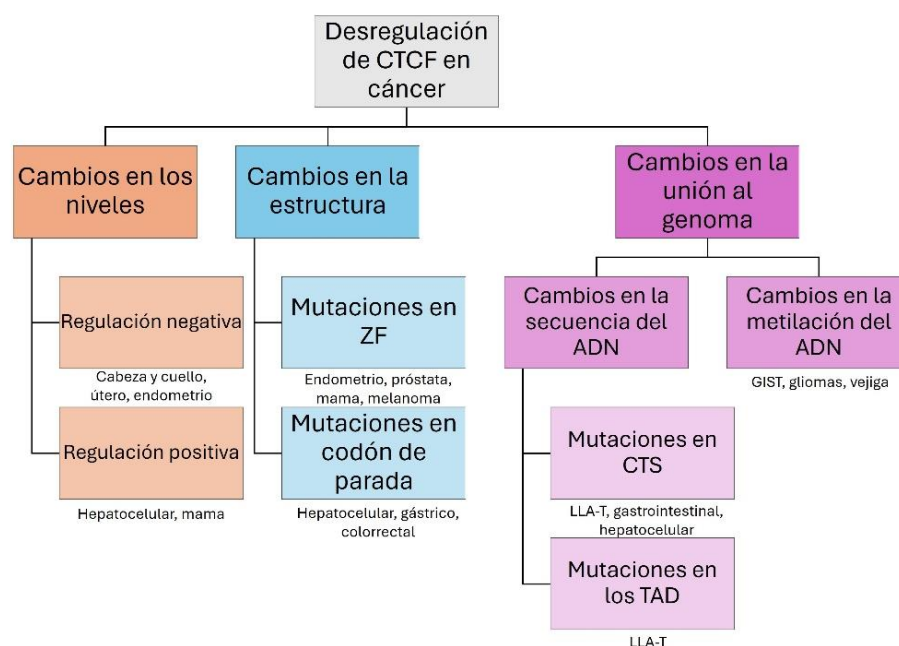
Figura 8: *CTCF* mapea en el cromosoma 16q22, en una región asociada al cáncer donde se produce frecuentemente pérdida de heterocigosidad (LOH). Modificada de (Klenova et al., 2002).

Posteriormente, la desregulación de CTCF en el cáncer se ha relacionado con mutaciones de todo tipo, con frecuencia variable en varios tipos de tumores (Dehingia et al., 2022).

Nuestro estudio de las alteraciones genómicas de CTCF utilizando la plataforma cBioportal donde se analizan 10967 muestras agrupadas en 32 estudios extraídos de TCGA (The Cancer Genome Atlas), indica que las alteraciones más frecuentes fueron variaciones estructurales, deleciones profundas, amplificaciones génicas, mutaciones truncantes y mutaciones *missense* (Anexo 2 Figura A2).



Para describir las alteraciones de CTCF en cáncer, las clasificaremos en tres categorías (Figura 9): i) variaciones en los niveles de CTCF, ii) alteraciones en su estructura, y iii) modificaciones en la unión de CTCF al genoma (Segueni & Noordermeer, 2022).



**Figura 9: Esquema de la desregulación de CTCF en cáncer.** Se muestran los distintos tipos de alteraciones de CTCF y los principales tumores donde se han descrito dichas alteraciones (esquema elaborado a partir de las citas bibliográficas indicadas más adelante).

### 9.1. Variaciones en los niveles de CTCF

Los niveles adecuados de CTCF son esenciales para la función celular normal, y su alteración está relacionada con muchas enfermedades, desde trastornos del desarrollo hasta diversos tipos de cáncer. CTCF se encuentra suprimido en varias neoplasias, y tal como se ha indicado, se considera un gen supresor de tumores (Filippova et al., 1998) (Fiorentino & Giordano, 2012).

Dependiendo del tipo de cáncer y de las transformaciones genómicas asociadas, tanto la disminución como el aumento de la dosis de CTCF pueden estar vinculados con la proliferación celular descontrolada característica de los tumores (Bailey et al., 2021). Sin embargo, en las células normales especialmente durante el desarrollo, CTCF es esencial para la proliferación y supervivencia. Por ello, la proliferación celular anómala en el cáncer requiere además otras transformaciones (Segueni & Noordermeer, 2022).

En la mayoría de los casos, lo que ocurre es la eliminación de un alelo de CTCF, como ocurre en el cáncer de cabeza y cuello, útero, endometrio, etc. Por ejemplo, en el caso de cáncer de endometrio, la haploinsuficiencia de CTCF da como resultado la regulación negativa de los genes supresores de tumores y la regulación positiva de los genes sensibles a los estrógenos (Marshall et al., 2017). Por tanto, una reducción de los niveles de CTCF puede favorecer aún más la capacidad de proliferación y la transformación celular neoplásica. Además, la haploinsuficiencia de CTCF en los tumores de endometrio



y mama exhiben un patrón diferente de metilación del ADN, lo que puede estar asociado con diferentes resultados clínicos y distintas oportunidades terapéuticas (Kemp et al., 2014).

No obstante, también está descrita una regulación positiva de CTCF en algunos cánceres, como ocurre en el carcinoma hepatocelular. En este se observa un aumento en la dosis de CTCF, correlacionada además con un peor pronóstico (Segueni & Noordermeer, 2022). Por su parte, en el cáncer de mama se sugiere que la sobreexpresión de CTCF confiere una ventaja proliferativa que protege a las células cancerosas de la apoptosis. Además, también en este mismo tumor, cambios en la cantidad de CTCF se han relacionado con la transición epitelio-mesénquima reversible, donde la regulación positiva de CTCF favorece el fenotipo mesenquimal, mientras que su regulación negativa favorece el epitelial (Johnson et al., 2022).

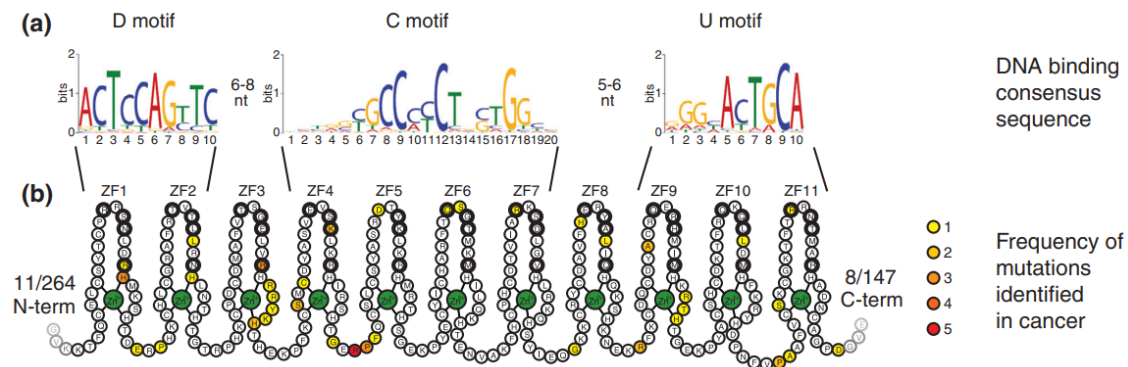
Los resultados utilizando la base de datos GEPIA, que analiza la expresión diferencial de CTCF en muestras tumorales vs normales, indican que en algunos casos no hay cambios significativos en la expresión, mientras que en otros aparece una mucha mayor expresión de CTCF en las muestras tumorales que en las normales (ver Anexo 3, Figura A7).

## 9.2. Alteraciones en la estructura de CTCF

Además de las mutaciones que eliminan por completo la función de CTCF (mutaciones nulas), existen otras mutaciones que pueden provocar cambios en su funcionalidad, las mutaciones *missense*. Se han identificado mutaciones *missense* recurrentes en varios cánceres, y que afectan sobre todo a los dedos de zinc de CTCF (Figura 10). Estas mutaciones pueden clasificarse en dos categorías: aquellas que afectan la estructura de un ZF, y aquellas que afectan a la unión a motivos específicos de ADN (Segueni & Noordermeer, 2022).

En primer lugar, se han descrito mutaciones de los ZF en varios cánceres, como el cáncer de endometrio (ZF 4-5 región espaciadora R377C), cáncer de próstata (mutación ZF3 H345R), tumor de Wilms (ZF3 R339W y mutaciones ZF7 R448Q) y cáncer de mama (mutación ZF3 K344E) (Oh et al., 2017). El residuo R377 es un punto crítico de mutaciones cancerosas que pueden afectar a la unión de CTCF al ADN, especialmente en cáncer de útero, melanoma y gastrointestinales (Dehingia et al., 2022). Además, en la leucemia relacionada con el síndrome de Down se han visto deleciones o mutaciones heterocigotas de CTCF (Yoshida et al., 2013). En el caso del carcinoma de endometrio, ciertas mutaciones puntuales en los ZF que conducen a una ganancia de función pueden mejorar la supervivencia celular al bloquear la apoptosis (Dehingia et al., 2022).

La caracterización extensa de dichas mutaciones indica que la unión al ADN se ve directamente afectada, particularmente en los ZF responsables de la unión de consenso central, ZF 3–7 (Marshall et al., 2014).



**Figura 10: Los dedos de zinc de CTCF están frecuentemente mutados en cáncer.** a) Motivos de consenso de unión al ADN de CTCF: D (*downstream*), C (central) y U (*upstream*). El motivo D está unido por los ZF 1-2, el motivo C está unido por los ZF 4-7 y el motivo U está unido por los ZF 9-11. b) Localización y frecuencia de mutaciones *missense* conocidas en CTCF en diversos cánceres. Las frecuencias de mutación sin sentido dentro de los extremos N y C se muestran en la figura como “número de mutaciones/número de aminoácidos”. Los aminoácidos en círculo en negrita se unen directamente al ADN. Modificado de (Marshall et al., 2014).

Por otro lado, pueden producirse mutaciones que generan codones de parada, distribuidas en las tres regiones de CTCF. Las mutaciones en residuos que coordinan la unión del zinc en los ZF que entran en contacto con el ADN, llevan a la pérdida de unión a CTCF (Debaugny & Skok, 2020).

En este caso destacamos como ejemplos los cánceres hepatocelular, gástrico y colorrectal. En el caso del cáncer gástrico y colorrectal esto ocurre con mucha frecuencia (25% y 19% respectivamente) y sobre todo implican cambios de base asociados con la unión de ZF 6, 8–10 (Debaugny & Skok, 2020).

En el anexo 2 se muestran los resultados de nuestros análisis utilizando la plataforma cBioportal. En la figura A4 se muestran la localización y el tipo de mutación en la estructura de CTCF. Dichas mutaciones no solo aparecen en el dominio central de dedos de zinc sino también a lo largo de toda la proteína.

Aparte de las mutaciones de CTCF descritas en la literatura, nuestro análisis de cBioportal muestra otras variantes estructurales en las que la alteración se produce por la fusión de CTCF con otra proteína (Figura A5).

### 9.3. Cambios en la unión de CTCF al genoma

Existe además otro mecanismo mediante el cual se modifica la función de CTCF, y es mediante la alteración de la unión de CTCF al ADN. Esta reorganización se puede producir por cambios en la secuencia del ADN, que abarcan desde sustituciones de un par de bases hasta reordenamientos más complejos, que pueden modificar la estructura de un TAD. Y también puede producirse por la disregulación de la metilación del ADN (Segueni & Noordermeer, 2022).

Si nos centramos en tumores concretos, por ejemplo, la unión alterada de CTCF en el locus 8q24 conduce a un aumento de la expresión del oncogén *MYC*, produciendo el

crecimiento y las metástasis del cáncer de mama. En el caso del cáncer de próstata también está implicado en la progresión de la enfermedad. Por su parte, en el cáncer colorrectal la unión alterada de CTCF puede alterar las interacciones potenciadores-promotores de largo alcance, lo que lleva a una desregulación de genes relacionados con la proliferación. Situaciones similares ocurren en el cáncer de páncreas, glioblastoma, y cáncer de pulmón. Este último se caracteriza por ser la primera causa de muerte por cáncer, por lo que el estudio de unión aberrante de CTCF cobra especial importancia en este caso (Bose et al., 2023).

#### 9.3.1. Cambios en la secuencia del ADN

La unión desregulada de CTCF puede alterar la función cromosómica y dar lugar a patrones anormales de expresión genética. El desarrollo de multitud de enfermedades, concretamente el cáncer, se ve afectado por cambios en los sitios de unión a CTCF, que promueven la inestabilidad cromosómica y la oncogénesis (Bose et al., 2023).

Se ha visto en distintos tipos de cáncer que los sitios de unión de CTCF a la cromatina (CTS) son puntos cruciales de mutación en tumores que exhiben inestabilidad cromosómica. La mayoría de las veces estas mutaciones coexisten en las células cancerígenas con otras anomalías cromosómicas (Bose et al., 2023). Las mutaciones que afectan a la unión de CTCF a los CTS provocan sobre todo sustituciones AT>GC en el motivo de unión al ADN o en bases adyacentes, como ocurre en la leucemia linfoblástica aguda de células T (LLA-T), cáncer de hígado y cánceres gastrointestinales (Guo et al., 2018).

Por su parte, las mutaciones en los límites de los TAD también están presentes en algunos tumores. En la LLA-T, se han visto microdeleciones cercanas a oncogenes, que eliminan los CTS de los límites de los TAD (Segueni & Noordermeer, 2022). Si no hay unión de CTCF en los límites de los TAD, esto facilita la fusión entre TAD contiguos que, en condiciones normales, deberían estar separados, formándose contactos inapropiados entre promotores y reguladores, llevando a la regulación positiva de ciertos oncogenes (Segueni & Noordermeer, 2022). Esto ocurre en la LLA-T, en la que la eliminación de aislantes unidos a CTCF conduce a una regulación positiva de los protooncogenes *TAL1* y *LMO2* (Fang et al., 2020).

Por último y en cuanto a las metástasis, está muy implicada la citocina *TGFB1*, y se ha descrito una mutación en el CTS próximo al locus de *TGFB1*, especialmente en las metástasis de melanoma (Dehingia et al., 2022).

#### 9.3.2. Cambios en la metilación del ADN

La unión de CTCF puede verse afectada también por cambios en la metilación del ADN, que con frecuencia está alterada en el cáncer, sobre todo por hipermetilación de los promotores en genes supresores de tumores, lo cual disminuye la unión de CTCF al ADN. Sin embargo, no siempre los CTS son sensibles a cambios en la metilación del ADN (Bose

et al., 2023). Este mecanismo afecta a varios tumores, como el cáncer colorrectal, de vejiga, de ovario y a los gliomas (Oh et al., 2017).

Este mecanismo ha sido estudiado sobre todo en tumores del estroma gastrointestinal (GIST) y en gliomas. Esta hipermetilación reduce la unión de CTCF a sus CTS. De la misma forma que las alteraciones en los CTS, los cambios en la metilación del ADN también pueden producir fusiones inapropiadas de TAD, activando diversos oncogenes. En el caso de los gliomas, las mutaciones en el gen de la IDH (isocitrato deshidrogenasa) inducen fenotipos metiladores de islas CpG, y esto lleva a la regulación positiva del oncogén *PDGFRA* (Flavahan et al., 2016; Monteagudo-Sánchez et al., 2024). Se ha comprobado en los tumores GIST que la hipermetilación de más de 600 sitios de unión de CTCF resultó en una pérdida de unión, lo que resultó en una regulación positiva del oncogén *FGF4* (Dehingia et al., 2022; Monteagudo-Sánchez et al., 2024). Además, en el *locus INK4B-ARF-INK4A (INK/ARF)*, compuesto por tres genes supresores de tumores, la regulación de la actividad génica está regulada por un sitio de unión de CTCF que es sensible a 5-meC, y que está hipermetilado en algunos tumores (Monteagudo-Sánchez et al., 2024).

En el caso del cáncer de vejiga, las alteraciones de la metilación del sexto sitio de unión a CTCF desempeñan un papel importante en la regulación de la expresión de *IGF2* y *H19*. En (Takai et al., 2001) se describe que la hipometilación en el alelo paterno en el cáncer de vejiga, que es bastante rara en tejido normal, podría ser más predominante que la metilación en el alelo materno en este locus, pudiendo participar en la sobreexpresión de *H19* en el cáncer de vejiga, especialmente en aquellos en estadio avanzado.

En conclusión, en este apartado se muestran los numerosos mecanismos moleculares por los que CTCF se altera en cáncer. Es de destacar que estas alteraciones modifican rutas de señalización intracelular fundamentales para la proliferación, destacando la desregulación del ciclo celular a través de la vía de TP53, como se demuestra en el análisis del cBioportal (anexo 2 Figura A6).

## 10. ALTERACIONES DE CTCF EN ALGUNOS TUMORES

Como se ha descrito en los apartados anteriores, la estructura o la función de CTCF pueden alterarse por diversos mecanismos, dando lugar a distintos procesos tumorales. A continuación, nos hemos centrado en describir tres ejemplos de tumores donde CTCF parece jugar un papel muy prevalente, tanto por los resultados descritos en la literatura como por nuestros análisis de cBioportal (anexo 2 Figura A3).

### 10.1. Cáncer de mama

El cáncer de mama constituye casi un 25% del total de cánceres diagnosticados en mujeres, y es el resultado de la interacción entre la genética y los factores ambientales. Con bastante frecuencia CTCF está eliminado o mutado en este tumor (Akhtar et al.,

2020). La disfunción de CTCF en cáncer de mama se produce tanto por mutaciones, como por la poli(ADP-ribosil)ación (PARilación) anormal (Akhtar et al., 2020). Algunas de las mutaciones detectadas son, por ejemplo, una mutación sin sentido, K344E, en el ZF3, afectando potencialmente la capacidad de CTCF para unirse a sitios promotores/aisladores de genes asociados con la proliferación celular, como *MYC*, *PLK*, *PIM-1*, *p19ARF* e *Igf2/H19* (Oh et al., 2017). Otros cambios incluyen la transición 240G→A en la UTR 5' y la mutación 1445C→T en el exón 4 del gen *CTCF* (Oh et al., 2017).

La PARilación de CTCF también contribuye a la regulación de la apoptosis y la inhibición del crecimiento celular mediante la translocación del nucleoplasma a los nucléolos en las células de cáncer de mama (Venkatraman & Klenova, 2015). La presencia de dos isoformas de CTCF, 130 kDa y 180 kDa, en el cáncer de mama, en comparación con un CTCF de 180 kDa formado mediante PARilación en tejidos normales, sugiere que el CTCF de 130 kDa podría servir como biomarcador de cáncer de mama. Se ha visto que el CTCF de 130 kDa se correlaciona negativamente con el tamaño y el estadio del tumor, y su aumento se asoció con una mayor proliferación celular (Oh et al., 2017).

En cuanto a los niveles de CTCF, se ha visto que su alteración también podría contribuir al desarrollo de este cáncer. Los niveles de CTCF están elevados en muchos tumores de mama (Docquier et al., 2005). La regulación del CTCF disminuyó significativamente en los cánceres de mama invasivos. Los niveles de expresión de CTCF se correlacionaron negativamente con la proliferación celular, especialmente en el carcinoma lobulillar invasivo (ILC) en comparación con otros tipos de cáncer de mama (Oh et al., 2017).

En cuanto a las alteraciones en la metilación del ADN, *BRCA1* es un gen que codifica una proteína supresora de tumores llamada *breast cancer type 1 susceptibility protein*. La expresión alterada o la ausencia de CTCF se ha relacionado con la inactivación epigenética de *BRCA1* a través de la metilación del promotor (Akhtar et al., 2020).

Además de la metilación, la expresión del gen CTCF podría estar sujeta a regulación diferencial por hormonas sexuales femeninas, ya que se han visto niveles reducidos de expresión de CTCF en la línea celular de cáncer de mama MCF-7 tratada con 17β-estradiol(E2). Estos resultados sugieren que la regulación hormonal puede afectar la expresión de CTCF en células de cáncer de mama, y que la regulación anormal de CTCF podría contribuir al crecimiento tumoral (Oh et al., 2017).

#### 10.2. Cáncer de endometrio

En el caso del cáncer de endometrio, las mutaciones de CTCF son más prevalentes en adenocarcinomas endometrioides, y, por su parte, la haploinsuficiencia de CTCF se encuentra más a menudo en carcinomas serosos. En el primero de los casos, la mayoría de las mutaciones son inactivadoras y resultan de cambios de marco de lectura, mutaciones *missense* y mutaciones en el sitio de empalme. Las localizaciones más frecuentes de estas mutaciones en CTCF (S345, K352, K365, G375, R377, P378, A387 y R567) se encuentran dentro del dominio central y están sobre todo concentradas en ZF 4. Otras mutaciones afectan al ZF 11 (Marshall et al., 2017).

En cambio, la mayoría de las pérdidas en el número de copias de CTCF se observan en tumores del subtipo seroso, siendo menos frecuentes en los de subtipo endometriode. Esta pérdida de copias se asocia con un pronóstico desfavorable, resultando en una menor supervivencia de las pacientes en comparación con los tumores diploides de CTCF. Además, la eliminación de CTCF es más probable en tumores de endometrio que experimentan recaída o metástasis en comparación con neoplasias malignas primarias (Marshall et al., 2017).

Estos hallazgos indican que la pérdida genética de CTCF, posiblemente en conjunto con la eliminación de otros genes en el cromosoma 16q, tiene consecuencias adversas para los pacientes (Marshall et al., 2017).

En el anexo 2 figura A3 se observa que el cáncer de endometrio es el que presenta un mayor porcentaje de alteraciones de CTCF, siendo la mayoría de ellas mutaciones, y en menor medida deleciones y otras.

### 10.3. Tumores hematológicos

En los diferentes tipos de leucemia son muy frecuentes las alteraciones estructurales del genoma, desde deleciones, translocaciones o variaciones en el número de copias. La hematopoyesis implica la diferenciación de las células madre hematopoyéticas pluripotentes en diversos tipos de células sanguíneas. La organización y modificaciones en el genoma desempeñan un rol crucial en este proceso, determinando los patrones de expresión génica en linajes celulares específicos (Qiu & Huang, 2020).

En la transición de la hematopoyesis fetal a la adulta, se fortalecen los límites de los TAD y la compartimentación de la cromatina, formándose bucles potenciadores/promotores dentro de los TAD en las células madre hematopoyéticas adultas. STAG1 y STAG2 son proteínas pertenecientes a la familia de las subunidades de cohesina que funcionan asegurando la estabilidad de los límites TAD definidos por CTCF, y facilitando las interacciones entre elementos reguladores dentro de estos límites (Viny et al., 2019).

CTCF presenta mutaciones en el 2% de los pacientes con leucemia mieloide aguda y en el 12% de los pacientes con leucemia megacarioblástica aguda relacionada con el síndrome de Down. Estas mutaciones son potenciales impulsores de la leucemia linfoblástica aguda (LLA) de células T, con una incidencia de aproximadamente el 10% en los pacientes (Fang et al., 2020).

Por su parte en la leucemia linfocítica crónica (LLC), en el subtipo hiperdiploide B la baja expresión de CTCF se asocia con debilidad en los límites de los TAD, con una reducción en la expresión de los genes cercanos a estos límites, y una expresión mayor de HOXA9, esencial en el desarrollo normal del sistema hematopoyético (Fang et al., 2020).

TAL1, al igual de HOXA9 es muy importante en la hematopoyesis, pero en algunas LLA, se ha observado una sobreexpresión anormal de TAL1, lo que puede contribuir al desarrollo y progresión de la enfermedad, concretamente hasta en el 60% de los pacientes con LLA-T. El locus *TAL1* tiene cuatro sitios de unión de CTCF que sirven para

separar TAL1 de sus genes vecinos, por lo que CTCF podría influir en su sobreexpresión (Qiu & Huang, 2020).

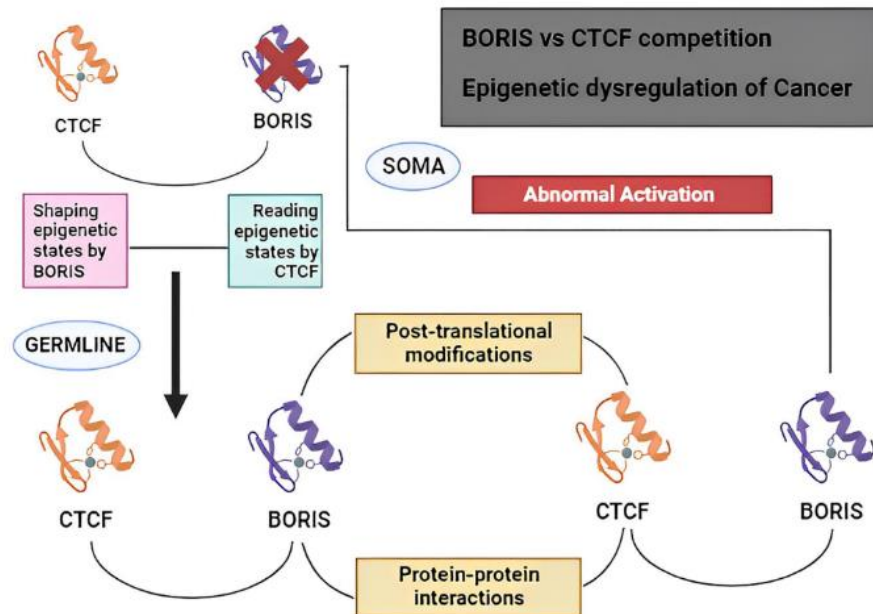
Por su parte TOP2B es una topoisomerasa de tipo II que interactúa tanto con CTCF como con cohesina en los límites de TAD (Uusküla-Reimand & Wilson, 2022). La inhibición de TOP2B puede producir daño y rotura en el ADN, contribuyendo también al desarrollo de la LMA (Fang et al., 2020).

## **11. CTCF Y BORIS COMO BIOMARCADORES**

En todas las enfermedades, pero especialmente en el cáncer, y sobre todo en el de pulmón, los biomarcadores tienen mucha importancia dado el valor de la detección temprana, las terapias personalizadas y la predicción del pronóstico de la enfermedad. El principal problema de estos es que carecen de sensibilidad, especificidad y reproducibilidad (Bose et al., 2023). Un biomarcador es una molécula que se puede encontrar en sangre, orina, saliva u otros fluidos o tejidos corporales, y cuya presencia es un signo de proceso anormal, afección o enfermedad. Puede usarse tanto para evaluar la presencia de una enfermedad, su progresión o su respuesta a un determinado tratamiento. El extenso conocimiento del papel de CTCF en la regulación epigenética ha hecho que se esté investigando su posible utilización como biomarcador para diagnosticar el cáncer y su implicación prediciendo la progresión de la enfermedad (Bose et al., 2023).

Con bastante frecuencia la expresión de BORIS (CTCF-L) está inducida en el cáncer, lo que podría llevar a pensar que su función es opuesta a la función de CTCF como supresor de tumores. Sin embargo, también se ha visto que inhibe la proliferación celular, por lo que no está del todo definida su actuación en las células cancerígenas (Marshall et al., 2014). Lo que sí que sabemos es que en BORIS se pueden observar múltiples mutaciones inactivadoras similares a las observadas en CTCF.

Se ha descubierto la expresión de BORIS en un número importante de mujeres con cáncer de mama. En este caso, cuando CTCF está aumentado, protege a las células de la apoptosis. Sin embargo, el silenciamiento de BORIS se ha relacionado con un aumento de la apoptosis (Akhtar et al., 2020). Todo esto nos sugiere que también este podría utilizarse como biomarcador de este tumor, e incluso en terapias como objetivo para modular la apoptosis celular en los cánceres de mama (Figura 11) (Bose et al., 2023).



**Figura 11: Representación de las funciones de CTCF-BORIS en las células normales y en las células cancerígenas** (Bose et al., 2023). BORIS es un factor específico de la línea germinal, pero cuando se activa erróneamente en células somáticas causa alteraciones y cambios epigenéticos.

Cuando BORIS se expresa en células sanas, provoca la activación de genes como *MAGE-A1* y *NY-ESO-1* del cáncer de testículo. Se sabe que desempeña una importante función en el control de los genes del cáncer de testículo, incrementando la expresión de los genes *MYC* y *BCRA1* (Bose et al., 2023). Por lo tanto, BORIS se considera un antígeno de testículos cancerosos, por lo que restringir su expresión podría representar un posible tratamiento dirigido de inmunoterapia en este tumor. Independientemente de su función protumoral o supresora de tumores, que en algunos cánceres aún no está del todo definida, BORIS representa un objetivo viable para la inmunoterapia contra el cáncer (Marshall et al., 2014).

A la hora de considerar CTCF como biomarcador y posible diana terapéutica, nos enfrentamos a varios desafíos, uno de ellos es su multifuncionalidad, ya que participa en tantos procesos que es muy difícil determinar cuál es su papel exacto en los tumores. Además, la unión de CTCF puede ser específica del tipo de célula, haciendo muy difícil la generalización de los hallazgos (Bose et al., 2023).

Por ejemplo, en el cáncer gástrico sí que se ha investigado una posible vía de señalización con posibilidad de tratamientos dirigidos. El aumento de la expresión de PD-L1 o de CTCF en las células del cáncer gástrico fomentan la quimiorresistencia, por lo que la vía de señalización CTCF-PD-L1 podría ser un medio para combatirla (Bose et al., 2023). Otro ejemplo es el carcinoma hepatocelular, uno de los más mortales, y en el cual se ha detectado CTCF, pero no en el tejido hepático normal. Su principal mecanismo en este tumor es que aumenta el crecimiento y las metástasis, al regular la expresión de la telomerasa transcriptasa inversa (TERT) y la FOXM1 (Oreskovic et al., 2021).



## **12. CONCLUSIONES**

En este trabajo se ha demostrado que CTCF es una proteína indispensable para las células, ya que desempeña funciones muy variadas, entre ellas la regulación de la expresión génica y de la estructura tridimensional del genoma. Cada vez vamos conociendo más acerca de las alteraciones de CTCF involucradas en la carcinogénesis y sobre los mecanismos moleculares subyacentes a su implicación en cáncer. Los hallazgos hasta el momento subrayan la importancia de investigar y comprender en profundidad dichos mecanismos, lo que podría llevar en el futuro al desarrollo de nuevas estrategias terapéuticas dirigidas específicamente a esta proteína, así como su uso como posible biomarcador diagnóstico, pronóstico y terapéutico.

## **13. AGRADECIMIENTOS**

A mis padres, mi hermano, y toda mi familia, por apoyarme siempre en todo y enseñarme a no rendirme pase lo que pase, no hubiera llegado hasta aquí sin vosotros. A Luis, por cuidarme, protegerme y confiar siempre en mí.

A mis amigas del colegio y del instituto, por aguantarnos tantos años y estar en las buenas y en las malas, por crecer juntas, y que siga siendo así siempre.

A todas las personas que me han acompañado durante estos 6 años maravillosos, en especial a David, Elena, Jose, Paula y Sarita, es un orgullo cumplir este sueño juntos, vais a ser los mejores médicos del mundo.

A mis tutoras, Dolores y Vanessa, desde que hablamos por primera vez supe que no podría haber elegido mejor, gracias por ayudarme y hacer que todo parezca más fácil, y por formar este equipazo.

## 14. BIBLIOGRAFÍA

- Akhtar, M. S., Akhter, N., Najm, M. Z., Deo, S. V. S., Shukla, N. K., Almalki, S. S. R., Alharbi, R. A., Sindi, A. A. A., Alruwetei, A., Ahmad, A., & Husain, S. A. (2020). Association of mutation and low expression of the CTCF gene with breast cancer progression. *Saudi Pharmaceutical Journal*, 28(5), 607–614. <https://doi.org/10.1016/j.jsps.2020.03.013>
- Arzate-Mejía, R. G., Recillas-Targa, F., & Corces, V. G. (2018). Developing in 3D: The role of CTCF in cell differentiation. *Development (Cambridge)*, 145(6), 1–16. <https://doi.org/doi:10.1242/dev.137729>
- Bailey, C. G., Gupta, S., Metierre, C., Amarasekera, P. M. S., O'Young, P., Kyaw, W., Laletin, T., Francis, H., Semaan, C., Sharifi Tabar, M., Singh, K. P., Mullighan, C. G., Wolkenhauer, O., Schmitz, U., & Rasko, J. E. J. (2021). Structure–function relationships explain CTCF zinc finger mutation phenotypes in cancer. *Cellular and Molecular Life Sciences*, 78(23), 7519–7536. <https://doi.org/10.1007/s00018-021-03946-z>
- Bastiaan Holwerda, S. J., & de Laat, W. (2013). CTCF: The protein, the binding partners, the binding sites and their chromatin loops. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1620), 1–8. <https://doi.org/10.1098/rstb.2012.0369>
- Batlle-López, A., Cortiguera, M. G., Rosa-Garrido, M., Blanco, R., Del Cerro, E., Torrano, V., Wagner, S. D., & Delgado, M. D. (2015). Novel CTCF binding at a site in exon1A of BCL6 is associated with active histone marks and a transcriptionally active locus. *Oncogene*, 34(2), 246–256. <https://doi.org/10.1038/onc.2013.535>
- 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), 387–396. [https://doi.org/10.1016/S0092-8674\(00\)81967-4](https://doi.org/10.1016/S0092-8674(00)81967-4)
- Bose, S., Saha, S., Goswami, H., Shanmugam, G., & Sarkar, K. (2023). Involvement of CCCTC-binding factor in epigenetic regulation of cancer. *Molecular Biology Reports*, 50(12), 10383–10398. <https://doi.org/10.1007/s11033-023-08879-3>
- Chen, D., & Lei, E. P. (2019). Function and regulation of chromatin insulators in dynamic genome organization. *Current Opinion in Cell Biology*, 58, 61–68. <https://doi.org/10.1016/J.CEB.2019.02.001>
- Cortiguera, M. G., García-Gaipo, L., Wagner, S. D., León, J., Batlle-López, A., & Delgado, M. D. (2019). Suppression of BCL6 function by HDAC inhibitor mediated acetylation and chromatin modification enhances BET inhibitor effects in B-cell lymphoma cells. *Scientific Reports*, 9(1), 1–15. <https://doi.org/10.1038/s41598-019-52714-4>
- Cuddapah, S., Jothi, R., Schones, D. E., Roh, T. Y., Cui, K., & Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Research*, 19(1), 24–32. <https://doi.org/10.1101/GR.082800.108>
- Cutter, A. R., & Hayes, J. J. (2015). A brief review of nucleosome structure. *FEBS Letters*, 589(20), 2914–2922. <https://doi.org/10.1016/j.febslet.2015.05.016>

- De La Rosa-Velázquez, I. A., Rincón-Arano, H., Benítez-Bribiesca, L., & Recillas-Targa, F. (2007). Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Research*, 67(6), 2577–2585. <https://doi.org/10.1158/0008-5472.CAN-06-2024>
- Debaugny, R. E., & Skok, J. A. (2020). CTCF and CTCFL in cancer. *Current Opinion in Genetics and Development*, 61, 44–52. <https://doi.org/10.1016/j.gde.2020.02.021>
- Deen, D., Butter, F., Daniels, D. E., Ferrer-Vicens, I., Ferguson, D. C. J., Holland, M. L., Samara, V., Sloane-Stanley, J. A., Ayyub, H., Mann, M., Frayne, J., Garrick, D., & Vernimmen, D. (2021). Identification of the transcription factor MAZ as a regulator of erythropoiesis. *Blood Advances*, 5(15), 3002–3015. <https://doi.org/10.1182/bloodadvances.2021004609>
- Dehingia, B., Milewska, M., Janowski, M., & Pękowska, A. (2022). CTCF shapes chromatin structure and gene expression in health and disease. *EMBO Reports*, 23(9), 1–22. <https://doi.org/10.15252/embr.202255146>
- Del Moral-Morales, A., Salgado-Albarrán, M., Sánchez-Pérez, Y., Wenke, N. K., Baumbach, J., & Soto-Reyes, E. (2023). CTCF and Its Multi-Partner Network for Chromatin Regulation. *Cells*, 12(10), 1–19. <https://doi.org/10.3390/cells12101357>
- Delgado, J. L., Hsieh, C.-M., Chan, N.-L., & Hiasa, H. (2018). Topoisomerases as Anticancer Targets. *Biochem J*, 475(2), 373–398. <https://doi.org/10.1042/BCJ20160583>
- Delgado, M. D., Chernukhin, I. V., Bigas, A., Klenova, E. M., & León, J. (1999). Differential expression and phosphorylation of CTCF, a c-myc transcriptional regulator, during differentiation of human myeloid cells. *FEBS Letters*, 444(1), 5–10. [https://doi.org/doi:10.1016/s0014-5793\(99\)00013-7](https://doi.org/doi:10.1016/s0014-5793(99)00013-7)
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature*, 485(7398), 376–380. <https://doi.org/10.1038/NATURE11082>
- Docquier, F., Farrar, D., D'Arcy, V., Chernukhin, I., Robinson, A. F., Loukinov, D., Vatolin, S., Pack, S., Mackay, A., Harris, R. A., Dorricott, H., O'Hare, M. J., Lobanenko, V., & Klenova, E. (2005). Heightened Expression of CTCF in Breast Cancer Cells Is Associated with Resistance to Apoptosis. *Cancer Research*, 65(12), 5112–5122. <https://doi.org/10.1158/0008-5472.CAN-03-3498>
- Falini, B., Brunetti, L., Sportoletti, P., & Martelli, M. P. (2020). NPM1-mutated acute myeloid leukemia: from bench to bedside. *Blood*, 136(15), 1707–1721.
- Fang, C., Rao, S., Crispino, J. D., & Ntziachristos, P. (2020). Determinants and role of chromatin organization in acute leukemia. *Leukemia*, 34(10), 2561–2575. <https://doi.org/10.1038/s41375-020-0981-z>
- Filippova, A., Lindblom, L. J., Meincke, E. M., Klenova, E. M., Neiman, P. E., Collins, S. J., Doggett, N. A., & Lobanenko, V. V. (1998). 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. *Genes*

*Chromosomes and Cancer*, 22, 26–36. [https://doi.org/https://doi.org/10.1002/\(SICI\)1098-2264\(199805\)22:1<26::AID-GCC4>3.0.CO;2-9](https://doi.org/https://doi.org/10.1002/(SICI)1098-2264(199805)22:1<26::AID-GCC4>3.0.CO;2-9)

- Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J., & Lobanenko, V. V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology*, 16(6), 2802–2813. <https://doi.org/10.1128/MCB.16.6.2802>
- Fiorentino, F. P., & Giordano, A. (2012). The tumor suppressor role of CTCF. *Journal of Cellular Physiology*, 227(2), 479–492. <https://doi.org/10.1002/JCP.22780>
- Flavahan, W. A., Drier, Y., Liao, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suvà, M. L., & Bernstein, B. E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*, 529(7584), 110–114. <https://doi.org/10.1038/NATURE16490>
- Gabriele, M., Brandão, H. B., Grosse-Holz, S., Jha, A., Dailey, G. M., Cattoglio, C., Hsieh, T.-H. S., Mirny, L., Zechner, C., & Hansen, A. S. (2022). Dynamics of CTCF-and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, 376, 496–501. <https://doi.org/10.1126/science.abn6583>
- Gaiotto, B. D., Ferrante, F., Herchenröther, A., Hake, S. B., & Borggrefe, T. (2019). The histone variant H2A.Z in gene regulation. *Epigenetics and Chromatin*, 12(1), 1–22. <https://doi.org/10.1186/s13072-019-0274-9>
- Gilbert, N., & Allan, J. (2019). The many length scales of DNA packaging. *Essays in Biochemistry*, 63(1), 1–4. <https://doi.org/10.1042/EBC20190040>
- Gombert, W. M., & Krumm, A. (2009). Targeted deletion of multiple CTCF-binding elements in the human C-MYC gene reveals a requirement for CTCF in C-MYC expression. *PLoS One*, 4(7), 1–8. <https://doi.org/10.1371/journal.pone.0006109>
- Guo, Y. A., Chang, M. M., Huang, W., Ooi, W. F., Xing, M., Tan, P., & Skanderup, A. J. (2018). Mutation hotspots at CTCF binding sites coupled to chromosomal instability in gastrointestinal cancers. *Nature Communications*, 9(1), 1–14. <https://doi.org/10.1038/s41467-018-03828-2>
- Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discovery*, 12(1), 31–46. <https://doi.org/10.1158/2159-8290.CD-21-1059>
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Hashimoto, H., Wang, D., Horton, J. R., Zhang, X., Corces, V. G., & Cheng, X. (2017). Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Molecular Cell*, 66(5), 711–720. <https://doi.org/10.1016/j.molcel.2017.05.004>
- Hirano, T. (2005). SMC proteins and chromosome mechanics: From bacteria to humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1455), 507–514. <https://doi.org/10.1098/rstb.2004.1606>

- Hrzenjak, A. (2016). JAZF1/SUZ12 gene fusion in endometrial stromal sarcomas. *Orphanet Journal of Rare Diseases*, 11(15), 1–8. <https://doi.org/10.1186/s13023-016-0400-8>
- Ishiguro, K. ichiro. (2019). The cohesin complex in mammalian meiosis. *Genes to Cells*, 24(1), 6–30. <https://doi.org/10.1111/gtc.12652>
- Jeon, H. Y., Pornour, M., Ryu, H., Khadka, S., Xu, R., Jang, J., Li, D., Chen, H., Hussain, A., Fazli, L., Gleave, M., Dong, X., Huang, F., Wang, Q., Barbieri, C., & Qi, J. (2023). SMAD3 promotes expression and activity of the androgen receptor in prostate cancer. *Nucleic Acids Research*, 51(6), 2655–2670. <https://doi.org/10.1093/NAR/GKAD043>
- Johnson, K. S., Hussein, S., Chakraborty, P., Muruganantham, A., Mikhail, S., Gonzalez, G., Song, S., Jolly, M. K., Toneff, M. J., Benton, M. L., Lin, Y. C., & Taube, J. H. (2022). CTCF Expression and Dynamic Motif Accessibility Modulates Epithelial-Mesenchymal Gene Expression. *Cancers*, 14(1), 1–26. <https://doi.org/10.3390/CANCERS14010209>
- Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R., & Lobanenko, V. V. (2000). Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Current Biology: CB*, 10(14), 853–856. [https://doi.org/10.1016/S0960-9822\(00\)00597-2](https://doi.org/10.1016/S0960-9822(00)00597-2)
- Kemp, C. J., Moore, J. M., Moser, R., Bernard, B., Teater, M., Smith, L. E., Rabaia, N. A., Gurley, K. E., Guinney, J., Busch, S. E., Shaknovich, R., Lobanenko, V. V., Liggitt, D., Shmulevich, I., Melnick, A., & Filippova, G. N. (2014). CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Reports*, 7(4), 1020–1029. <https://doi.org/10.1016/j.celrep.2014.04.004>
- Kentepozidou, E., Aitken, S. J., Feig, C., Stefflova, K., Ibarra-Soria, X., Odom, D. T., Roller, M., & Flicek, P. (2020). Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. *Genome Biology*, 21(1), 1–19. <https://doi.org/10.1186/S13059-019-1894-X>
- Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nature Reviews. Genetics*, 20(4), 207–220. <https://doi.org/10.1038/S41576-018-0089-8>
- Klenova, E. M., Chernukhin, I. V., El-Kady, A., Lee, R. E., Pugacheva, E. M., Loukinov, D. I., Goodwin, G. H., Delgado, D., Filippova, G. N., León, J., Morse, H. C., Neiman, P. E., & Lobanenko, V. V. (2001). Functional Phosphorylation Sites in the C-Terminal Region of the Multivalent Multifunctional Transcriptional Factor CTCF. *Molecular and Cellular Biology*, 21(6), 2221–2234. <https://doi.org/10.1128/mcb.21.6.2221-2234.2001>
- Klenova, E. M., Morse, H. C., Ohlsson, R., & Lobanenko, V. V. (2002). The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Seminars in Cancer Biology*, 12(5), 399–414. [https://doi.org/10.1016/S1044-579X\(02\)00060-3](https://doi.org/10.1016/S1044-579X(02)00060-3)
- Klenova, E. M., Nicolas, R. H., Paterson, H. F., Carne, A. F., Heath, C. M., Goodwin, G. H., Neiman, P. E., Lobanenko, V. V., Lobanenko, V. V., Nicolas, R. H., Adler, V. V., Paterson, H., Klenova, E. M., Polotskaja, A. V., & Goodwin, G. H. (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. *Molecular and Cellular Biology*, 13(12), 7612–7624. <https://doi.org/10.1128/mcb.13.12.7612-7624.1993>
- Kuang, S., & Wang, L. (2020). Identification and analysis of consensus RNA motifs binding to the genome regulator CTCF. *NAR Genomics and Bioinformatics*, 2(2), 1–13. <https://doi.org/10.1093/nargab/lqaa031>
- Lee, J. M., Hammarén, H. M., Savitski, M. M., & Baek, S. H. (2023). Control of protein stability by post-translational modifications. *Nature Communications*, 14(1), 1–16. <https://doi.org/10.1038/s41467-023-35795-8>
- Li, Y., Haarhuis, J. H. I., Sedeño Cacciatore, Á., Oldenkamp, R., van Ruiten, M. S., Willems, L., Teunissen, H., Muir, K. W., de Wit, E., Rowland, B. D., & Panne, D. (2020). The structural basis for cohesin–CTCF-anchored loops. *Nature*, 578(7795), 472–476. <https://doi.org/10.1038/s41586-019-1910-z>
- Lobanenko, V. V., Nicolas, R., Adler, V., Paterson, H., Klenova, E. M., Polotskaja, A. V., & Goodwin, G. H. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*, 5(12), 1743–1753.
- MacPherson, M. J., Beatty, L. G., Zhou, W., Du, M., & Sadowski, P. D. (2009). The CTCF Insulator Protein Is Posttranslationally Modified by SUMO. *Molecular and Cellular Biology*, 29(3), 714–725. <https://doi.org/10.1128/mcb.00825-08>
- Marshall, A. D., Bailey, C. G., Champ, K., Vellozzi, M., O'Young, P., Metierre, C., Feng, Y., Thoeng, A., Richards, A. M., Schmitz, U., Biro, M., Jayasinghe, R., Ding, L., Anderson, L., Mardis, E. R., & Rasko, J. E. J. (2017). CTCF genetic alterations in endometrial carcinoma are pro-tumorigenic. *Oncogene*, 36(29), 4100–4110. <https://doi.org/10.1038/onc.2017.25>
- Marshall, A. D., Bailey, C. G., & Rasko, J. E. J. (2014). CTCF and BORIS in genome regulation and cancer. In *Current Opinion in Genetics and Development*, 24(1), 8–15. <https://doi.org/10.1016/j.gde.2013.10.011>
- Monteagudo-Sánchez, A., Noordermeer, D., & Greenberg, M. V. C. (2024). The impact of DNA methylation on CTCF-mediated 3D genome organization. *Nature Structural & Molecular Biology*, 31(3), 404–412. <https://doi.org/10.1038/s41594-024-01241-6>
- Nishana, M., Ha, C., Rodriguez-Hernaez, J., Ranjbaran, A., Chio, E., Nora, E. P., Badri, S. B., Kloetgen, A., Bruneau, B. G., Tsirigos, A., & Skok, J. A. (2020). Defining the relative and combined contribution of CTCF and CTCFL to genomic regulation. *Genome Biology*, 21(1), 1–34. <https://doi.org/10.1186/s13059-020-02024-0>
- Oh, S., Oh, C., & Yoo, K. H. (2017). Functional roles of CTCF in breast cancer. *BMB Reports*, 50(9), 445–453. <https://doi.org/10.5483/BMBRep.2017.50.9.108>
- Ohlsson, R., Lobanenko, V., & Klenova, E. (2010). Does CTCF mediate between nuclear organization and gene expression? *BioEssays*, 32(1), 37–50. <https://doi.org/10.1002/bies.200900118>

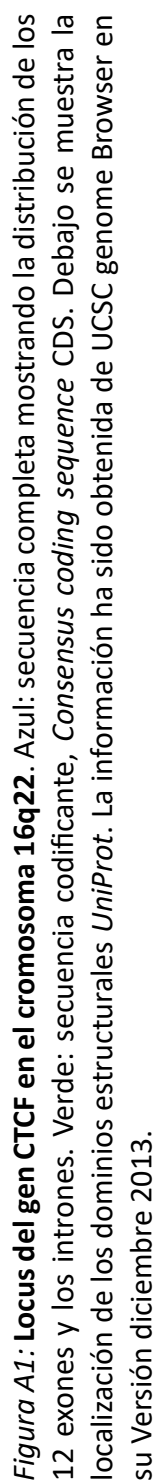
- Oreskovic, E., Wheeler, E. C., Mengwasser, K. E., Fujimura, E., Martin, T. D., Tothova, Z., Elledge, S. J., Levine, R., & Zou, L. (2022). Genetic analysis of cancer drivers reveals cohesin and CTCF as suppressors of PD-L1. *Proc Natl Acad Sci U S A*, 119(7), 1–10. <https://doi.org/10.1073/pnas.2120540119/-/DCSupplemental>
- Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B. S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A. G., & Merkenschlager, M. (2008). Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell*, 132(3), 422–433. <https://doi.org/10.1016/j.cell.2008.01.011>
- Patel, I., & Parchem, R. J. (2022). Regulation of Oct4 in stem cells and neural crest cells. *Birth Defects Research*, 114(16), 983–1002. <https://doi.org/10.1002/bdr2.2007>
- Pongubala, J. M. R., & Murre, C. (2021). Spatial Organization of Chromatin: Transcriptional Control of Adaptive Immune Cell Development. *Frontiers in Immunology*, 12, 1–11. <https://doi.org/10.3389/fimmu.2021.633825>
- Qiu, Y., & Huang, S. (2020). CTCF-mediated genome organization and leukemogenesis. *Leukemia*, 34(9), 2295–2304. <https://doi.org/10.1038/s41375-020-0906-x>
- Segueni, J., & Noordermeer, D. (2022). CTCF: A misguided jack-of-all-trades in cancer cells. *Computational and Structural Biotechnology Journal*, 20, 2685–2698. <https://doi.org/10.1016/j.csbj.2022.05.044>
- Sun, X., Zhang, J., & Cao, C. (2022). CTCF and Its Partners: Shaper of 3D Genome during Development. *Genes*, 13(8), 1–19. <https://doi.org/10.3390/genes13081383>
- Takai, D., Gonzales, F. A., Tsai, Y. C., Thayer, M. J., & Jones, P. A. (2001). Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. *Human Molecular Genetics*, 10(23), 2619–2626. <https://doi.org/10.1093/hmg/10.23.2619>
- Torrano, V., Chernukhin, I., Docquier, F., D'Arcy, V., León, J., Klenova, E., & Delgado, M. D. (2005). CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *The Journal of Biological Chemistry*, 280(30), 28152–28161. <https://doi.org/10.1074/JBC.M501481200>
- Torrano, V., Navascués, J., Docquier, F., Zhang, R., Burke, L. J., Chernukhin, I., Farrar, D., León, J., Berciano, M. T., Renkawitz, R., Klenova, E., Lafarga, M., & Delgado, M. D. (2006). Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosylation)-dependent mechanism. *Journal of Cell Science*, 119(9), 1746–1759. <https://doi.org/10.1242/jcs.02890>
- Tyler, J. K., Del Rosario, B. C., Kriz, A. J., Del Rosario, A. M., Anselmo, A., Fry, C. J., White, F. M., Sadreyev, R. I., & Lee, J. T. (2019). Exploration of CTCF post-translation modifications uncovers Serine-224 phosphorylation by PLK1 at pericentric regions during the G2/M transition. *ELIFE*, 8(42341), 1–29. <https://doi.org/10.7554/eLife.42341.001>

- Uusküla-Reimand, L., & Wilson, M. D. (2022). Untangling the roles of TOP2A and TOP2B in transcription and cancer. *Science Advances*, 8(4920), 1–16. <https://doi.org/10.1126/sciadv.add4920>
- Van De Nobelen, S., Rosa-Garrido, M., Leers, J., Heath, H., Soochit, W., Joosen, L., Jonkers, I., Demmers, J., Van Der Reijden, M., Torrano, V., Grosveld, F., Delgado, M. D., Renkawitz, R., Galjart, N., & Sleutels, F. (2010). CTCF regulates the local epigenetic state of ribosomal DNA repeats. *Epigenetics and Chromatin*, 3(1), 1–21. <https://doi.org/10.1186/1756-8935-3-19>
- Venkatraman, B., & Klenova, E. (2015). Role of CTCF poly(ADP-Ribosyl)ation in the regulation of apoptosis in breast cancer cells. *Indian Journal of Medical and Paediatric Oncology*, 36(1), 49–54. <https://doi.org/10.4103/0971-5851.151784>
- Viny, A. D., Bowman, R. L., Liu, Y., Lavallée, V. P., Eisman, S. E., Xiao, W., Durham, B. H., Navitski, A., Park, J., Braunstein, S., Alija, B., Karzai, A., Csete, I. S., Witkin, M., Azizi, E., Baslan, T., Ott, C. J., Pe'er, D., Dekker, J., ... Levine, R. L. (2019). Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. *Cell Stem Cell*, 25(5), 682–696. <https://doi.org/10.1016/J.STEM.2019.08.003>
- Vostrov, A. A., & Quitschke, W. W. (1997). The zinc finger protein CTCF binds to the APB $\beta$  domain of the amyloid  $\beta$ -protein precursor promoter: Evidence for a role in transcriptional activation. *Journal of Biological Chemistry*, 272(52), 33353–33359. <https://doi.org/10.1074/jbc.272.52.33353>
- Wang, D. C., Wang, W., Zhang, L., & Wang, X. (2019). A tour of 3D genome with a focus on CTCF. *Seminars in Cell & Developmental Biology*, 90, 4–11. <https://doi.org/10.1016/J.SEMCDB.2018.07.020>
- Wen, Z., Zhang, L., Ruan, H., & Li, G. (2020). Histone variant H2A.Z regulates nucleosome unwrapping and CTCF binding in mouse ES cells. *Nucleic Acids Research*, 48(11), 5939–5952. <https://doi.org/10.1093/NAR/GKAA360>
- Xiao, T., Li, X., & Felsenfeld, G. (2021). The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7), 1–10. <https://doi.org/10.1073/pnas.2023127118>
- Xu, F., Liu, C., Zhou, D., & Zhang, L. (2016). TGF- $\beta$ /SMAD Pathway and Its Regulation in Hepatic Fibrosis. *Journal of Histochemistry and Cytochemistry*, 64(3), 157–167. <https://doi.org/10.1369/0022155415627681>
- Yamamoto, T., & Saitoh, N. (2019). Non-coding RNAs and chromatin domains. *Current Opinion in Cell Biology*, 58, 26–33. <https://doi.org/10.1016/j.ceb.2018.12.005>
- Yoshida, K., Toki, T., Okuno, Y., Kanazaki, R., Shiraishi, Y., Sato-Otsubo, A., Sanada, M., Park, M. J., Terui, K., Suzuki, H., Kon, A., Nagata, Y., Sato, Y., Wang, R. N., Shiba, N., Chiba, K., Tanaka, H., Hama, A., Muramatsu, H., ... Ogawa, S. (2013). The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nature Genetics*, 45(11), 1293–1301. <https://doi.org/10.1038/NG.2759>



- Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavoosidana, G., Mukhopadhyay, R., Kanduri, C., Oshimura, M., Feinberg, A. P., Lobanenkov, V., Klenova, E., & Ohlsson, R. (2004). Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature Genetics*, 36(10), 1105–1110. <https://doi.org/10.1038/ng1426>
- Zhang, X., Wang, Y., Qu, C., Chen, J., Slaney, C. Y., Yi, M., Yuan, Q., Zhao, Y., & Chen chenjie, J. (2021). Pan-Cancer Analysis of PARP1 Alterations as Biomarkers in the Prediction of Immunotherapeutic Effects and the Association of Its Expression Levels and Immunotherapy Signatures. *Frontiers in Immunology*, 12, 1–16. <https://doi.org/10.3389/fimmu.2021.721030>
- Ziatanova, J., & Caiafa, P. (2009). CTCF and its protein partners: Divide and rule? *Journal of Cell Science*, 122(9), 1275–1284. <https://doi.org/10.1242/jcs.039990>
- Zuin, J., Dixon, J. R., Van Der Reijden, M. I. J. A., Ye, Z., Kolovos, P., Brouwer, R. W. W., Van De Corput, M. P. C., Van De Werken, H. J. G., Knoch, T. A., Van Ijcken, W. F. J., Grosveld, F. G., Ren, B., & Wendt, K. S. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 111(3), 996–1001. <https://doi.org/10.1073/pnas.1317788111>

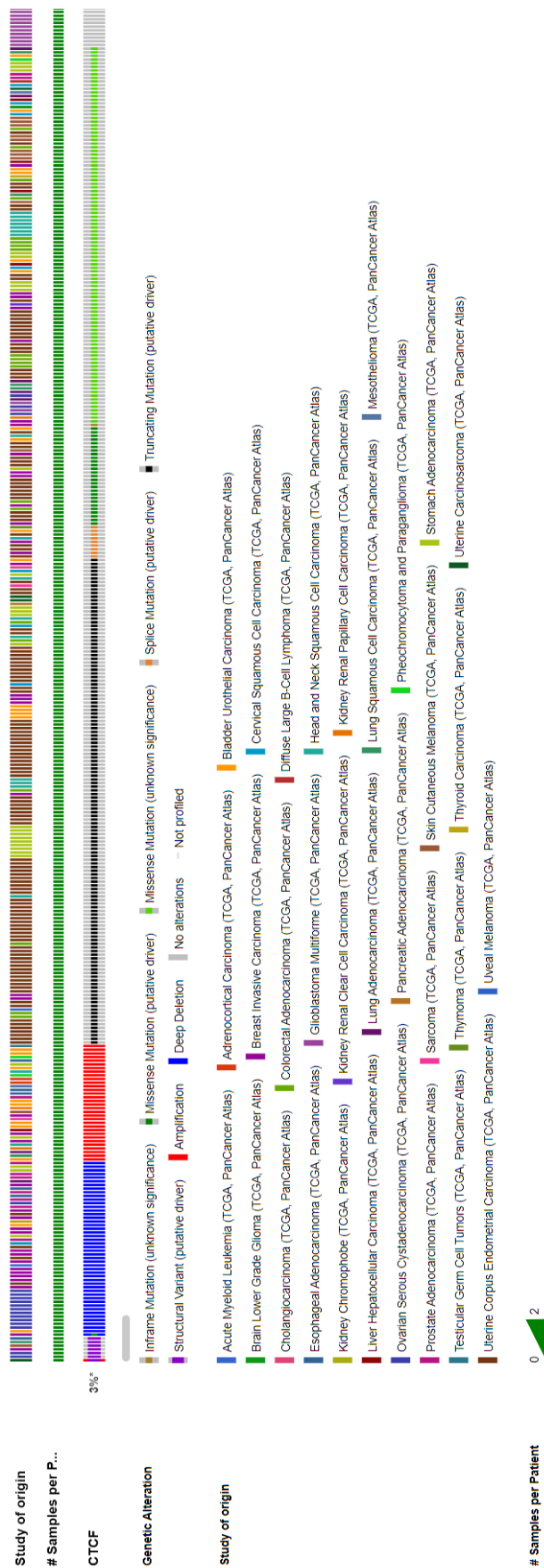
### Anexo 1: Representación del gen *CTCF* a partir del UCSC Genome Browser.



## Anexo 2: Análisis de la implicación de CTCF en cáncer usando cBioPortal

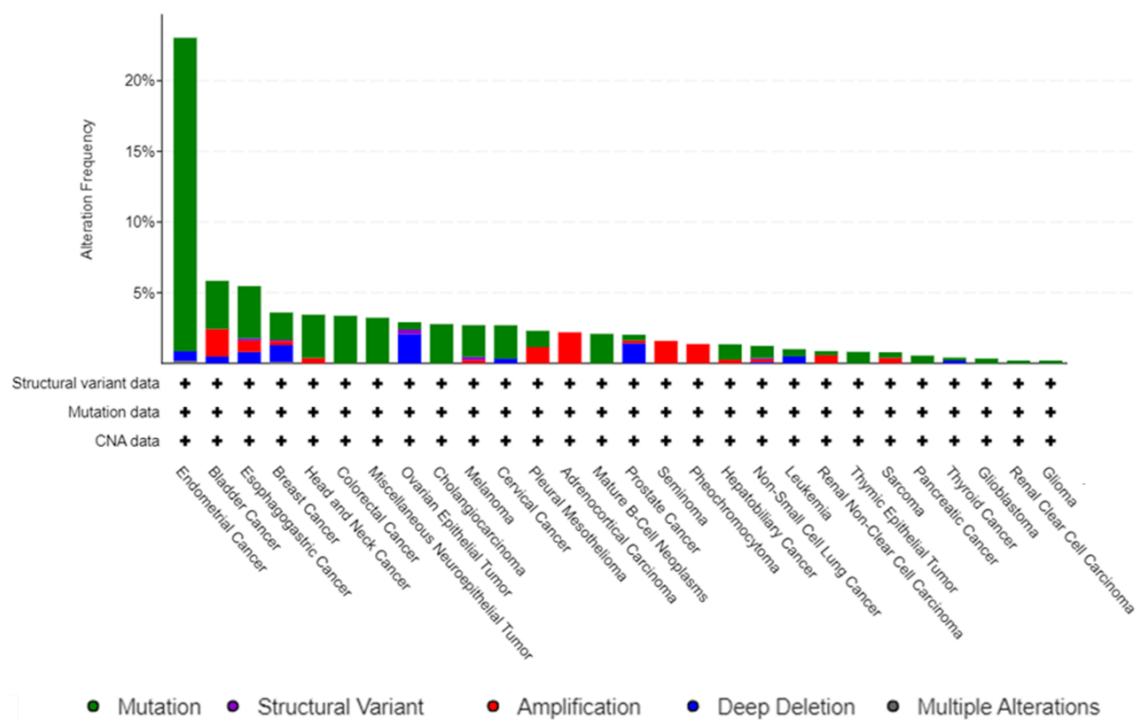
Para obtener más información sobre CTCF y su papel en el cáncer utilizamos la plataforma cBioPortal de Genómica del Cáncer (<https://www.cbioportal.org/>). cBioPortal proporciona visualización, análisis y descarga de conjuntos de datos genómicos del cáncer a gran escala. Elegimos datos del TCGA (The Cancer Genome Atlas), incluidos 32 estudios. Después de realizar la búsqueda "CTCF", seleccionamos las siguientes opciones:

- **OncoPrint:** Nos permite visualizar los distintos tipos de mutaciones (Figura A2).
- **Cancer Type summary:** datos sobre las alteraciones de CTCF y su frecuencia en cada tipo de cáncer (Figura A3).
- **Mutations:** análisis de las mutaciones de CTCF (Figura A4).
- **Structural variants:** función nueva que nos permite ver aquellos tumores en los que se alteran los genes por una fusión de CTCF con otra proteína (Figura A5).
- **Pathways:** nos enseña las rutas alteradas por mutaciones de CTCF (Figura A6).

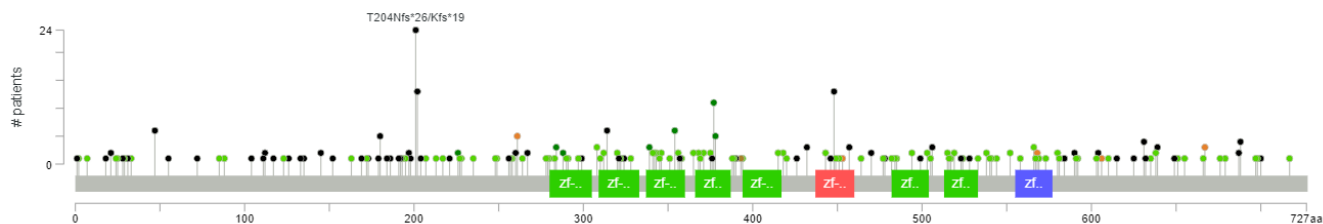


**Figura A2: Oncoprint.** Análisis de las alteraciones genómicas de CTCTF en todas las muestras. Se analizaron 10967 muestras agrupadas en 32 estudios extraídos de TCGA (The Cancer Genome Atlas). Cada una de las líneas verticales representa una muestra, y vemos que CTCTF se encuentra alterado en un 3% de ellas.

En cada muestra podemos ver el tipo de cáncer y el tipo de alteración genética como se expone en la leyenda. Las alteraciones más frecuentes fueron: variaciones estructurales (morado), deleciones profundas (azul), amplificaciones génicas (rojo), mutaciones truncantes (gris oscuro) y mutaciones *missense* (verde).



**Figura A3: Cancer Type summary.** Se muestran las alteraciones de CTCF en función del tipo de cáncer, ordenados de mayor a menor. Como podemos observar, el cáncer de endometrio es el que presenta alteraciones en CTCF con más frecuencia. Hemos seleccionado que solo aparezcan los cánceres con un mínimo de 10 casos, por lo que no se han tenido en cuenta aquellos con menos de 10 casos afectados. Además, la leyenda nos muestra los tipos de mutaciones por colores.



Mutation types and corresponding color codes are as follows:

- **Missense Mutations** (putative driver)
- **Missense Mutations** (unknown significance)
- **Truncating Mutations** (putative driver): Nonsense, Nonstop, Frameshift deletion, Frameshift insertion, Splice site
- **Truncating Mutations** (unknown significance): Nonsense, Nonstop, Frameshift deletion, Frameshift insertion, Splice site
- **Inframe Mutations** (putative driver): Inframe deletion, Inframe insertion
- **Inframe Mutations** (unknown significance): Inframe deletion, Inframe insertion
- **Splice Mutations** (putative driver)
- **Splice Mutations** (unknown significance)
- **Fusion Mutations**
- **Other Mutations** (putative driver): All other types of mutations
- **Other Mutations** (unknown significance): All other types of mutations

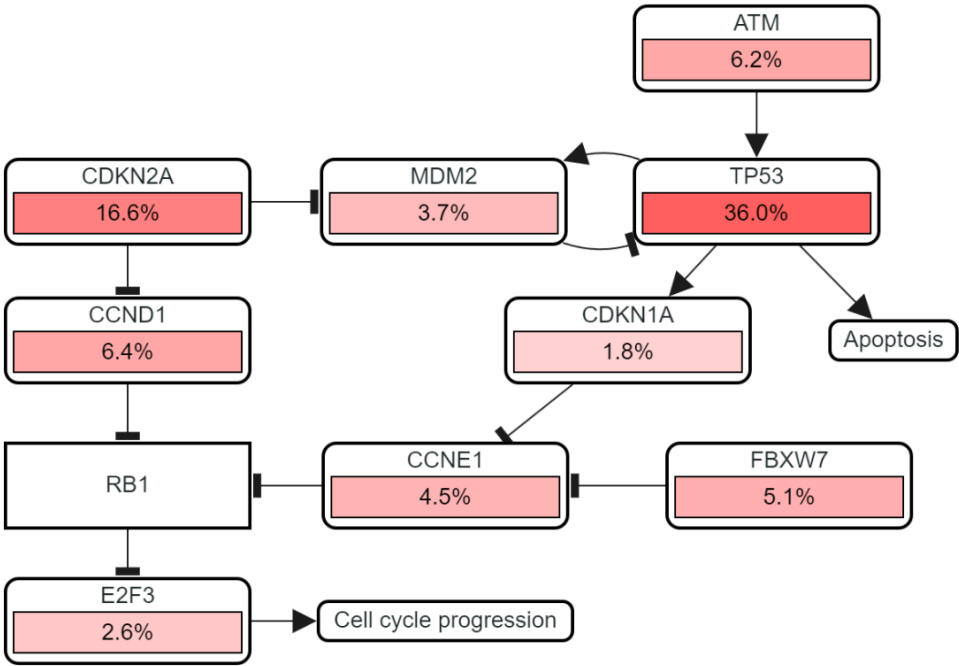
**Figura A4: Mutations.** Esta función nos permite ver la posición de las mutaciones de CTCF. Se nos muestra la estructura de CTCF del aminoácido 0 al 727. Cada bolita representa un sitio donde hay una alteración en CTCF, y la altura de las barras representa el número de casos, las barras más altas son aquellas regiones de CTCF que resultan más frecuentemente modificadas. En la leyenda vemos por colores cada tipo de mutación.

| Cancer Type Detailed                                           | Gene 1       | Gene 2         | Event Info                 |
|----------------------------------------------------------------|--------------|----------------|----------------------------|
| Esophageal Squamous Cell Carcinoma                             | <i>CTCF</i>  | <i>HNRNPDL</i> | <i>CTCF-HNRNPDL</i> fusion |
| Breast Invasive Ductal Carcinoma                               | <i>CTCF</i>  | <i>NUTF2</i>   | <i>CTCF-NUTF2</i> fusion   |
| Uterine Carcinosarcoma/Uterine Malignant Mixed Mullerian Tumor | <i>CTCF</i>  | <i>PSKH1</i>   | <i>CTCF-PSKH1</i> fusion   |
| Serous Ovarian Cancer                                          | <i>CTCF</i>  | <i>SPSB1</i>   | <i>CTCF-SPSB1</i> fusion   |
| Serous Ovarian Cancer                                          | <i>SF3B3</i> | <i>CTCF</i>    | <i>SF3B3-CTCF</i> fusion   |
| Lung Squamous Cell Carcinoma                                   | <i>CTCF</i>  | <i>E2F4</i>    | <i>CTCF-E2F4</i> fusion    |
| Cutaneous Melanoma                                             | <i>CTCF</i>  | <i>CES2</i>    | <i>CTCF-CES2</i> fusion    |

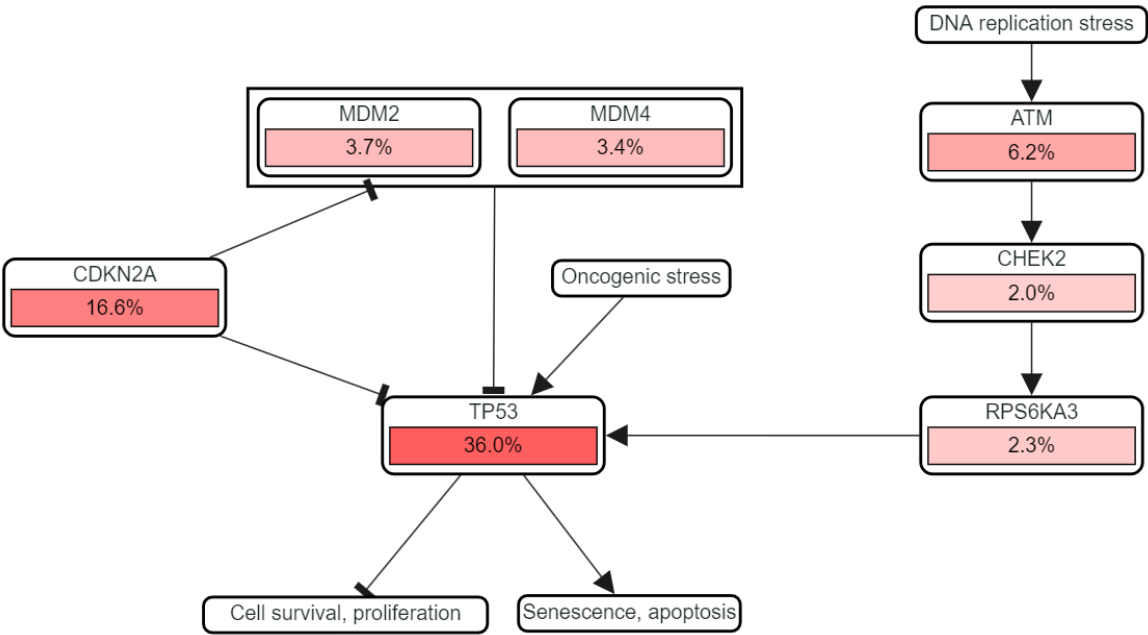
**Figura A5: Structural variants.** Se trata de una función nueva de cBioPortal que nos muestra en la tabla los tumores en los que la alteración se produce por la fusión de CTCF con otra proteína.

**Figura A6: Pathways.** Se muestran las rutas alteradas por mutaciones de CTCF, dentro de los tumores con CTCF mutado, el 36% tienen mutado TP53. De las 10 vías alteradas, se muestran el ciclo celular (A) y TP53 (B). Las otras vías son: WNT, MYC, TGF-Beta, RTK-RAS, PI3K, NRF2, NOTCH e HIPPO.

A) Ciclo celular

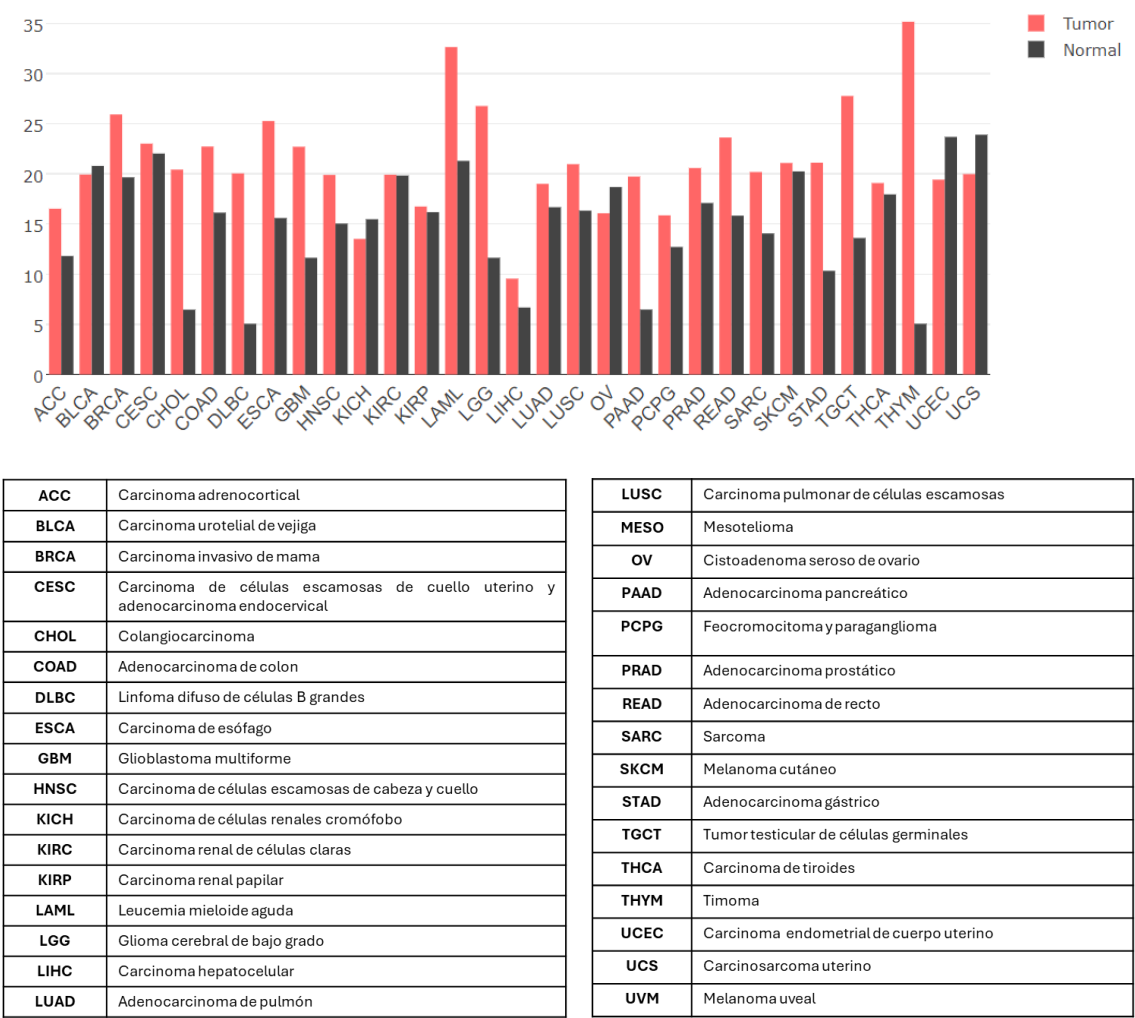


B) TP53





**Anexo 3: Análisis de la expresión diferencial de CTCF en células tumorales y normales con GEPIA.**



*Figura A7:* Expresión diferencial de CTCF en muestras normales (negro) y tumorales (rojo) en los diferentes tipos de tumores. La altura de las barras representa la media de la expresión en cada tipo de tejido o tumor.