



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
ZARAGOZA**

CTCF regula la expresión de *Notch1*

TESIS

**QUE PARA OBTENER EL TÍTULO DE
BIÓLOGO**

PRESENTA:

Ángel Josué Cerecedo Castillo



Director de Tesis: Dr. Félix Recillas Targa

Asesor Interno: Dr. Juan José Rodríguez Mercado

Diciembre 2016

Ciudad de México



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS ACADÉMICOS

Este trabajo se realizó bajo la dirección y tutoría del Dr. Félix Recillas Targa en el laboratorio 122 Norte del Departamento de Genética Molecular del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México.

Se contó con el apoyo y la asesoría del LCG. Rodrigo Arzate Mejía, la Dra. Karin Meier y la Biol. Georgina Guerrero Avendaño. Además se contó con el apoyo de la laboratorista Gianelly Cortés González.

Adicionalmente se reconoce la asesoría académica de todos y cada uno de los miembros del laboratorio del Dr. Félix Recillas Targa.

Se agradece también a Guadalupe Cortés de la Unidad de Biología Molecular del Instituto de Fisiología Celular por la secuenciación de las muestras.

El trabajo contó con el financiamiento de DGAPA-PAPIIT, UNAM (IN209403, IN203811 e IN201114), CONACyT (42653-Q, 128464 y 220503) y de Fronteras de la Ciencia 2015 CONACyT número 290.

Finalmente se agradece a los miembros del jurado que revisaron y enriquecieron con sus observaciones este trabajo:

Dr. Juan José Rodríguez Mercado

Dra. Elia Roldán Reyes

Dra. Lucila Álvarez Barrera

M. en C. Rodrigo Aníbal Mateos Nava

DEDICATORIAS

A mi familia, en especial a mi abuelita Socorro

A mi hermana Irly

A Sus, por los increíbles momentos

A Daniel, por ser mi socio y amigo

A Rodrigo, pieza fundamental en este trabajo

ÍNDICE

1. Índice de figuras	5
2. Lista de abreviaturas	7
3. Resumen	8
I. INTRODUCCIÓN	9
a. La vía de señalización de Notch y su impacto en la proliferación y diferenciación celular	9
b. El gen <i>Notch1</i> humano	12
c. Mecanismos de regulación transcripcional y epigenética	13
d. Organización tridimensional del genoma	16
e. El factor nuclear CTCF	19
II. ANTECEDENTES PARTICULARES	22
a. CTCF regula la expresión de proto-oncogenes a través de la formación de asas de cromatina	22
b. Las células K562 como un modelo de estudio para la regulación transcripcional del gen <i>Notch1</i>	26
III. PLANTEAMIENTO DEL PROBLEMA	27
IV. HIPÓTESIS	27
V. OBJETIVO GENERAL	27
VI. OBJETIVOS PARTICULARES	27
VII. MATERIALES Y MÉTODOS	28
a. Análisis bioinformáticos	28
b. Cultivo celular	29
c. Extracción de ácidos nucleicos	29
d. Diseño y estandarización de oligonucleótidos	30
e. Retrotranscripción seguida de PCR en tiempo real (RT-qPCR) ...	30
f. Análisis estadístico	31
g. Inmunoprecipitación de la cromatina (ChIP)	31
h. <i>Knockdown</i> de CTCF	32
i. Sistema de edición genómica CRISPR-Cas9	33

VIII.	RESULTADOS	40
a.	CTCF se une a la región intergénica 5' y 3' del gen <i>Notch1</i> en células K562	40
b.	El abatimiento de CTCF afecta la expresión del gen <i>Notch1</i> en células K562	43
c.	La eliminación de los sitios de unión a CTCF por CRISPR-Cas9 afecta la expresión del gen <i>Notch1</i> en células K562	45
IX.	DISCUSIÓN	49
X.	CONCLUSIONES	55
XI.	PERSPECTIVAS	55
XII.	REFERENCIAS	55
XIII.	INFORMACIÓN SUPLEMENTARIA	60
XIV.	ANEXOS	65

ÍNDICE DE FIGURAS

Fig. 1. Ligandos y receptor de la vía de Notch.

Fig. 2. Vía de señalización de Notch.

Fig. 3. Organización del genoma en compartimentos y en dominios asociados topológicamente.

Fig. 4. Dominios topológicamente asociados (TADs).

Fig. 5. Unión de CTCF a su motivo en el ADN.

Fig. 6. Orientación convergente de sitios de unión a CTCF.

Fig. 7. Deleciones recurrentes en pacientes con T-ALL.

Fig. 8. Activación de proto-oncogenes por pérdida de asas de cromatina.

Fig. 9. Sistema de knockdown mediante un shRNAi.

Fig. 10. Complejo formado por el ARN guía y la endonucleasa Cas9.

Fig. 11. Principales mecanismos de reparación del ADN.

Fig. 12. Diseño y clonación de ARNs guías en el vector lentiviral lentiCRISPR-V2.

Fig. 13. Transducción, genotipificación y dilución clonal.

Fig. 14. CTCF se asocia en orientación convergente al locus del gen *Notch1* en siete líneas celulares.

Fig. 15. El locus del gen *Notch1* constituye un asa de cromatina.

Fig. 16. CTCF se une a la región intergénica 5' y 3' del gen *Notch1* en células K562.

Fig. 17. El *knockdown* de CTCF afecta la expresión del gen *Notch1* en células K562.

Fig. 18. Sitios de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1*.

Fig. 19. La eliminación del sitio de unión a CTCF #1 por CRISPR-Cas9 afecta la expresión del gen *Notch1* en células K562.

Fig. 20. La eliminación del sitio de unión a CTCF #2 por CRISPR-Cas9 afecta la expresión del gen *Notch1* en células K562.

Fig. 21. Modelo propuesto del asa de cromatina que incorpora al gen *Notch1* y el efecto de la eliminación de un sitio de unión a CTCF.

Figuras suplementarias

Fig. S1. Integridad del ARN extraído de células K562.

Fig. S2. Estandarización de oligonucleótidos para el ARNm del gen *Notch1*.

Fig. S3. Tamaño de los fragmentos de la cromatina después de la sonicación.

Fig. S4. Estandarización de los oligonucleótidos que amplifican los sitios de unión a CTCF observados por ChIP-seq en células K562.

Fig. S5. Estandarización de los oligonucleótidos utilizados para observar las eliminaciones en los sitios de unión a CTCF río arriba del gen *Notch1*.

Fig. S6. Mapa de interacciones de Hi-C en células K562 a una resolución de 5 kb.

Tablas

Tabla S1. Datos de ChIP-seq utilizados obtenidos por ENCODE.

Tabla S2. Secuencias de oligonucleótidos utilizados en cada uno de los métodos.

Tabla S3. Secuencias de los ARNs guías utilizados para las eliminaciones con CRISPR-Cas9.

Tabla S4. Sitios de unión a CTCF observados por ChIP-seq con sus coordenadas y la secuencia y orientación del motivo de unión a CTCF.

LISTA DE ABREVIATURAS

ADN: Deoxyribonucleic Acid.

ALL: Acute Lymphoblastic Leukemia.

ARN: Ribonucleic Acid.

ChIA-PET: Chromatin Interaction Analysis by Paired-End Tag Sequencing.

ChIP: Chromatin Immunoprecipitation.

CLL: Chronic Lymphocytic Leukemia.

CML: Chronic Myeloid Leukemia.

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats.

CTCF: CCCTC-binding Factor.

EGF: Epidermal Growth Factor.

ENCODE: The Encyclopedia of DNA Elements.

HDR: Homology-directed Repair.

Hi-C: High-throughput sequencing Chromosome Conformation Capture.

ICN: Intracellular Notch1.

NHEJ: Nonhomologous End Joining.

PAM: Protospacer Adjacent Motif.

PCR: Polymerase Chain Reaction.

RT-qPCR: Reverse Transcription quantitative Polymerase Chain Reaction.

sgRNA: single guide RNA.

shRNAi: short hairpin RNA.

TAD: Topologically Associating Domain.

TCs: Chromosome Territories.

RESUMEN

Desde el descubrimiento del gen *Notch* en 1913 y su caracterización en 1983 (Artavanis-Tsakonas *et al.*, 1983), miles de trabajos han investigado el papel de esta vía de regulación en diversos organismos y tipos celulares, sin embargo poco se sabe sobre los mecanismos moleculares que controlan su transcripción durante el desarrollo y en patologías como el cáncer. Recientemente se ha reportado que la organización tridimensional del genoma al interior del núcleo celular es capaz de afectar la regulación de la expresión génica (Dixon *et al.*, 2016), aunque a la fecha es poco claro cómo se establece y se mantiene dicha organización. Por ejemplo, la pérdida de asas de cromatina, causada por eliminaciones de sitios de unión a la proteína CTCF, fue capaz de activar la expresión de proto-oncogenes (Hnisz *et al.*, 2016). En este trabajo analizamos el papel de la proteína arquitectural CTCF en la regulación de la expresión del gen *Notch1* en células K562. Por inmunoprecipitación de la cromatina validamos que CTCF se une a las regiones intergénicas río arriba y río abajo del locus del gen *Notch1* y por datos publicados de captura conformacional de cromosomas se identificó la presencia de un asa de cromatina en el locus con las secuencias de unión al factor CTCF en orientación convergente en sus límites. Evaluamos la participación de CTCF en la expresión del gen *Notch1* mediante un *knockdown* de CTCF con un shRNAi, lo cual ocasionó un incremento en la expresión del gen. Adicionalmente eliminamos de manera precisa dos sitios de unión a CTCF mediante la técnica de edición genómica CRISPR-Cas9 localizados en la región intergénica río arriba del inicio de la transcripción del gen *Notch1*. La eliminación de cada sitio resultó en un incremento significativo de los niveles de ARNm para el gen *Notch1*. Por lo tanto, los resultados de este trabajo sugieren que CTCF regula la expresión del gen *Notch1* en células K562.

INTRODUCCIÓN

La diversidad entre diferentes especies y tipos celulares parece haber evolucionado a través del uso combinatorio de un relativamente pequeño número de vías de señalización conservadas (Kopan e Ilagan, 2009). Entre esas, la vía de señalización de Notch en metazoarios permite la comunicación de corto alcance entre células, lo cual puede promover o suprimir la proliferación celular, muerte celular, diferenciación y activación de diferentes programas de expresión génica (Yatim *et al.*, 2012). Pese a la gran importancia de la vía de Notch poco se sabe sobre los mecanismos moleculares que regulan la expresión génica de los diversos elementos que conforman a la vía.

La vía de señalización de Notch y su impacto en la proliferación y diferenciación celular

La vía de señalización de Notch es una vía conservada entre los metazoarios y participa en una variedad de procesos fundamentales que permiten el desarrollo de un organismo (Bray, 2006). Una característica de la vía de Notch que la separa de otras vías de señalización conservadas es su mecanismo de transducción de señales. Tanto los receptores (*Notch*) como los ligandos (*Delta* y *Jagged*) son proteínas *trans*-membranales con grandes dominios extracelulares (Schweisguth, 2004).

En mamíferos los receptores Notch son cuatro, codificados por genes parálogos, los cuales son *Notch1-4*, y muestran tanto funciones únicas como redundantes (Bray, 2006). El dominio extracelular de todos los receptores Notch contiene regiones parecidas al factor de crecimiento epidermal (EGF), algunas de las cuales regulan las interacciones con el ligando. Seguido de dichas regiones se encuentra una región regulatoria negativa (NRR), la cual previene la activación del receptor en ausencia de ligandos. En el dominio intracelular existen un dominio RAM y siete repetidos de anquirina (ANK) y finalmente en el carboxilo terminal se encuentran

dominios PEST que regulan la degradación del receptor (**Fig. 1A**) (Le Borgne *et al.*, 2005).

Para el caso de los ligandos, en mamíferos son dos, denominados Delta y Jagged, los cuales son parecidos estructuralmente al dominio extracelular del receptor Notch, ya que contienen varias regiones repetidas del *EGF*, los cuales son reconocidos por las mismas regiones presentes en los receptores (**Fig. 1B**) (Wright *et al.*, 2004).

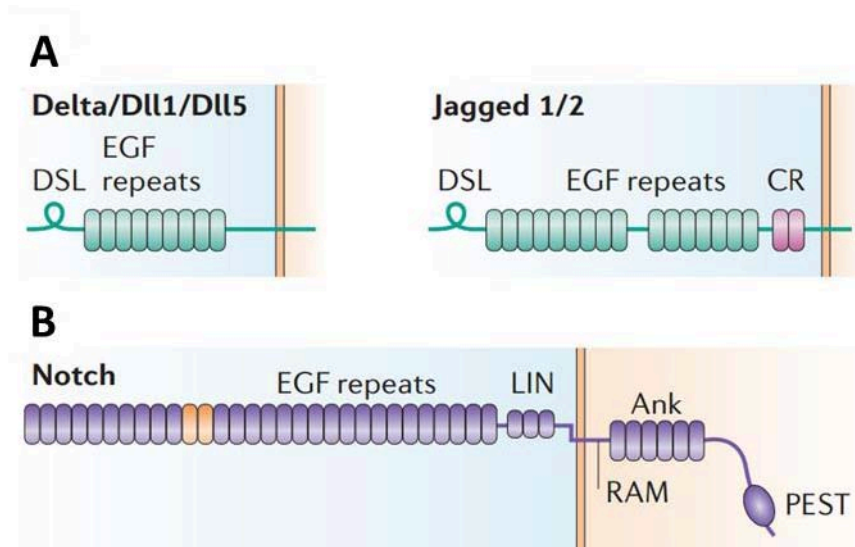


Fig. 1. Ligandos y receptor de la vía de Notch. (A). Esquema representativo de los ligandos Delta y Jagged, los cuales muestran secuencias repetidas del EGF. **(B).** Esquema representativo del receptor Notch con sus dominios extracelular e intracelular. Imagen tomada de Bray, 2006.

De forma general, durante la señalización por la vía de Notch, la interacción física entre el ligando y el receptor promueve dos cortes proteolíticos en el receptor Notch, llevados a cabo por una metaloproteínasa y la γ -secretasa; estos cortes liberan al dominio intracelular de Notch (Intracellular Notch Domain; *ICN*), el cual se transloca al núcleo celular y se une con la proteína CSL, la cual es una proteína de unión a ADN, y así promueve la transcripción de los genes blanco de la vía los cuales son principalmente *Hes1*, el cual es un factor transcripcional, y componentes de la vía de Myc y mTOR (**Fig. 2**) (Schweisguth, 2004; Bray, 2006, Kopan e Ilagan, 2009).

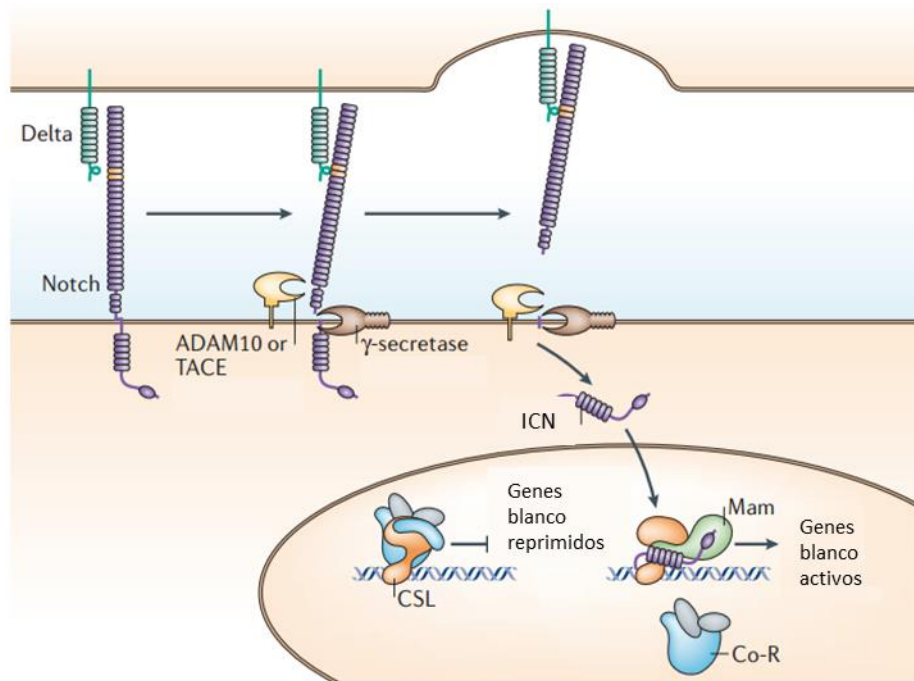


Fig. 2. Vía de señalización de Notch. Cuando el ligando y el receptor interactúan físicamente se activan cortes proteolíticos del receptor y el dominio intracelular se transloca al núcleo donde interactúa con la proteína CSL para activar la transcripción de sus genes blanco. Imagen tomada de Bray, 2006.

La vía de Notch funciona durante diversos procesos fisiológicos y del desarrollo, los cuales pueden ser subdivididos dentro de tres categorías: la inhibición lateral (Artavanis-Tsakonas *et al.*, 1999), las decisiones de linaje (Louvi y Artavanis-Tsakonas, 2006) y la inductiva o de fronteras (Micchelli y Perrimon, 2006). La comunicación celular a través de esta vía es dependiente de contacto, ya que receptores y ligandos están en las membranas celulares (Bray, 2006).

Dado que la vía de Notch juega un papel crítico en varios procesos fundamentales y en un amplio rango de tejidos celulares, no es sorprendente que la ganancia o la pérdida aberrante de los componentes de la vía ha sido directamente relacionado a múltiples patologías humanas, incluyendo síndromes del desarrollo (Garg *et al.*, 2005), enfermedades neurológicas (Louvi y Artavanis-Tsakonas, 2006) y cáncer (Kopan e Ilagan, 2009).

La mayoría de los procesos patológicos y fisiológicos dependientes de la vía de Notch recaen en la habilidad del dominio intracelular de su receptor (ICN) para convertir a la proteína de unión a ADN CSL (también conocida como RBPJ) de un represor a un activador transcripcional (Yatim *et al.*, 2012).

A pesar de que existen cuatro receptores de la vía de Notch en mamíferos, el gen *Notch1* humano ha sido el más estudiado, esto debido a que es el principal receptor de la vía, siendo responsable de aproximadamente más del 80% de la actividad de la vía de Notch (Bray, 2006).

El gen *Notch1* humano

El gen *Notch1* codifica para una proteína *trans*-membranal que actúa como receptor de la vía de Notch. *Notch1* es también determinante en los procesos de carcinogénesis, actuando como un supresor de tumores o un oncogen dependiendo del tipo celular y el contexto específico (Radtke y Raj, 2003; Roy *et al.*, 2007; South *et al.*, 2012).

Un caso interesante de la dualidad en la función del gen *Notch1* y la vía de señalización de Notch en los procesos de carcinogénesis se da en células hematopoyéticas. En el linaje linfoide el gen *Notch1* actúa como un oncogen, promoviendo a través de su vía de señalización la proliferación y la desdiferenciación celular lo cual afecta el crecimiento y maduración de linfocitos B y T (Schwanbeck y Just, 2011). Por otro lado, en el linaje mieloides el gen *Notch1* actúa como un gen supresor de tumores, suprimiendo la proliferación celular de células mieloides como granulocitos y eritrocitos (Suresh e Irvine, 2015). La diferencia de función de Notch1 en estos casos es aún desconocida.

Notch1 es el gen que más frecuentemente se ve afectado en la leucemia linfocítica crónica (CLL) (Durinck *et al.*, 2014). Diversas mutaciones han sido encontradas tanto en regiones codificantes del gen *Notch1* como en regiones no codificantes presentes en el locus (Puente *et al.*, 2015). Además, el gen *Notch1* se ha encontrado

mutado en más del 50% de todos los casos de leucemia linfoblástica aguda (ALL) (Weng *et al.*, 2004).

Los estudios del gen *Notch1* se han centrado en el control de la vía de señalización a nivel post-traducciona, incluyendo el procesamiento y maduración del receptor, la activación por sus ligandos en la superficie celular y la modificación y degradación de la proteína (Bray, 2006), por esta razón existen pocos estudios enfocados en la regulación transcripcional del gen *Notch1*, a pesar de que datos sugieren que una importante forma de regulación del gen *Notch1* se da a nivel de la transcripción del gen y su entorno cromatínico, ya que su expresión es diferencial en contextos celulares específicos y puede estar alterada en el desarrollo del cáncer (Rizzo *et al.*, 2008).

Mecanismos de regulación transcripcional y epigenética

Todas las células de un organismo contienen la misma información genética, sin embargo, presentan funciones y fenotipos distintos resultado del establecimiento de patrones transcripcionales específicos para cada tipo celular. La complejidad regulatoria adquirida por el genoma eucariota a través de la evolución ha requerido el desarrollo de múltiples niveles de regulación molecular que permiten controlar la expresión genética durante el desarrollo y para mantener la homeostasis celular (Recillas-Targa y Escamilla del Arenal, 2004).

El genoma lineal en humanos tiene una longitud de aproximadamente dos metros y el núcleo celular tiene un diámetro de sólo 10 micrómetros, por lo que este ADN lineal en células de mamíferos debe ser compactado. Al interior del núcleo, el ADN se compacta a través de interacciones proteína-ADN formando a la cromatina, la cual permite contener el genoma en el interior de éste (Rincón-Arano y Recillas-Targa, 2008). Los principales componentes proteicos de la cromatina son proteínas llamadas histonas, las cuales compactan al ADN en distintos niveles, siendo el primero el nucleosoma, también conocido como la unidad fundamental de la cromatina, el cual está constituido por un octámero de histonas, dos de cada una

de las proteínas H2A, H2B, H3 y H4, alrededor del cual interaccionan 147 pares de bases de ADN (Felsenfeld y Groudine, 2003).

Existen múltiples niveles de compactación del ADN siendo el cromosoma en metafase el nivel más elevado, aunque este nivel sólo se alcanza solo durante la división celular; en interfase la cromatina se encuentra compacta en los niveles de nucleosoma, solenoide y asas de cromatina. Adicionalmente, la cromatina está distribuida a través del núcleo en regiones condensadas, denominadas heterocromatina, y regiones menos condensadas, llamadas eucromatina (Felsenfeld y Groudine, 2003).

Dado el elevado nivel de compactación del genoma, la estructura de la cromatina se vuelve un obstáculo para la actividad transcripcional, es decir, para la expresión de los genes. Por lo tanto, el entender cómo se expresa un gen depende en gran medida de la capacidad de remodelar la estructura de la cromatina tanto al nivel de sus regiones control como de sus secuencias codificantes y no codificantes (Recillas-Targa y Escamilla del Arenal, 2004; Allis y Jenuwein, 2016). Por lo tanto, la regulación transcripcional se ha dividido en tres niveles que *in vivo* son interdependientes: el primero lo encontramos a nivel de las secuencias nucleotídicas en la molécula de ADN representado por los motivos de unión a factores de transcripción; el segundo involucra la estructura de la cromatina, donde la regulación de la expresión de un gen no sólo depende de la información codificada en la secuencia del ADN, sino también de su organización en cromatina y la regulación epigenética asociada a ésta; y el tercer nivel involucra la organización espacial tridimensional del genoma al interior del núcleo (Recillas-Targa y Escamilla del Arenal, 2004; Allis *et al.*, 2015; Dekker y Misteli, 2015).

A nivel de la secuencia, la expresión génica es controlada por elementos de regulación proximales conocidos como promotores necesarios para iniciar la transcripción (Kadonaga, 2004), además de secuencias regulatorias distales como los potenciadores (enhancers), silenciadores y aisladores (insulators), los cuales en conjunto conducen la expresión génica controlada (van Driel *et al.*, 2003). Todos

estos elementos regulatorios son ricos en motivos de unión a factores transcripcionales (Allis *et al.*, 2015).

Por otro lado, la regulación epigenética se da en torno a la estructura de la cromatina y su efecto regulador sobre la expresión de los genes. El estudio de los mecanismos de regulación de la expresión génica por mecanismos epigenéticos permite comprender las diferencias en expresión de genes observadas durante el desarrollo, la diferenciación celular y patologías cuando no existe una diferencia en la secuencia de los genes estudiados (Recillas-Targa y Escamilla del Arenal, 2004). Los mecanismos epigenéticos básicos que se han caracterizado son la metilación del ADN, lo cual se asocia comúnmente a represión transcripcional (Allis *et al.*, 2015), y las modificaciones post-traduccionales de histonas, las cuales permiten identificar regiones de cromatina abierta o compacta, así como posibles elementos regulatorios tales como enhancers y promotores (Allis y Jenuwein, 2016).

A menudo los enhancers u otras regiones regulatorias no se encuentran linealmente cercanos a los promotores que regulan, por lo que necesitan del establecimiento de interacciones de largo alcance para ejercer su función en la regulación transcripcional (Li *et al.*, 2016). Además, se sabe que los genes se organizan en dominios de la cromatina que se encuentran abiertos en un tiempo y espacio coincidente con la expresión de genes. Dichos dominios están delimitados por elementos frontera o barrera, también conocidos como *insulators*, compuestos por secuencias a las que se unen proteínas nucleares y permiten la formación de dominios de expresión génica (Recillas-Targa y Razin, 2001; West *et al.*, 2002; Dixon *et al.*, 2016; Ali *et al.*, 2016).

Por lo anterior, la regulación de la expresión génica no puede ser entendida sin tomar en consideración a la estructura y organización tridimensional de la cromatina (Pombo y Dillon, 2015). Porque además, se ha demostrado que existe una correlación entre el establecimiento de contactos cromatínicos y la actividad transcripcional (Lieberman-Aiden *et al.*, 2009), la cual depende de interacciones físicas entre elementos regulatorios como enhancers y promotores, los cuales no se encuentran cercanos en el genoma lineal (Gorkin *et al.*, 2014).

Organización tridimensional del genoma

Los genomas eucariotas deben estar altamente organizados para ser contenidos en el núcleo celular y desde hace algunos años se ha evidenciado que este plegamiento está altamente organizado y que podría estar estrechamente ligado al control transcripcional (Sexton y Cavalli, 2015). A la fecha, varios estudios han demostrado una correlación entre la topología de la cromatina y la actividad génica, sin saber aún si la organización de los cromosomas es una causa o consecuencia de funciones genómicas (de Laat y Doboule, 2013; Cavalli y Misteli, 2013).

La topología y la actividad transcripcional parecen ligadas a diferentes escalas dentro del núcleo, ya que los cromosomas están caracterizados por una compartimentación estructural en varios niveles (Bickmore, 2013). A nivel del núcleo, los cromosomas individuales ocupan posiciones definidas en el espacio nuclear tridimensional, conocidos como territorios cromosómicos (TCs) (Cremer y Cremer, 2010). Éstos son estables a través de la interfase pero cambian su ubicación después de la mitosis, por lo que la localización específica de los cromosomas es aleatoria y varía entre tipos celulares y especies (Gibcus y Dekker, 2013). Sin embargo, la posición de algunas regiones dentro de los TCs no es aleatoria y está relacionada con la actividad transcripcional, ya que regiones ricas en genes tienden a estar en la periferia de los TCs, lo cual probablemente facilita el acceso de la maquinaria transcripcional (Boyle *et al.*, 2011).

A nivel de mega bases, dentro de los TCs, los cromosomas están estructurados en dominios sub-cromosomales caracterizados por interacciones entre compartimentos funcionalmente equivalentes, ya que a pesar de la variación célula a célula de los TCs, hay algunas características que permanecen constantes en una población celular. Por ejemplo, hay regiones genómicas que tienden a contactar con otras con similar actividad transcripcional, es decir, regiones activas interactúan en el espacio con otras regiones activas y regiones que carecen de actividad transcripcional interactúan con otras regiones inactivas (Lieberman-Aiden *et al.*, 2009). Dichas redes de interacción activas e inactivas son referidas como compartimentos A y B, respectivamente, los cuales presumiblemente reflejan una

tendencia global de la eucromatina y heterocromatina a asociarse en el espacio (Fig. 3A) (Gorkin *et al.*, 2014).

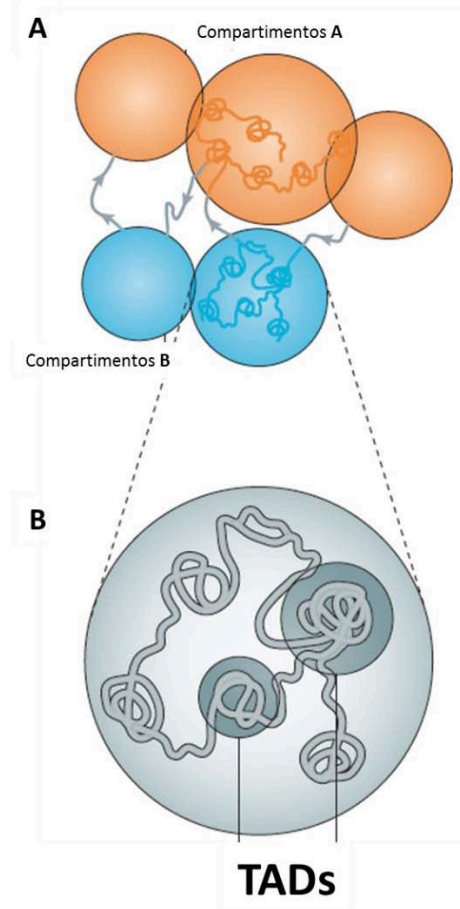


Fig. 3. Organización del genoma en compartimentos y en dominios asociados topológicamente (TADs). (A). Las regiones con similar actividad transcripcional tienden a asociarse en el espacio, formando a los compartimentos A y B, los cuales representan cromatina abierta y cerrada, respectivamente. (B). Cada compartimento está conformado por varios dominios topológicos, los cuales representan la unidad básica de la organización tridimensional del genoma. Imagen tomada y modificada de Dekker *et al.*, 2013.

Por ejemplo, se ha mostrado que los genes transcritos por la ARN polimerasa II se agrupan en zonas de actividad transcripcional conocidos como fábricas de transcripción (Papantonis y Cook, 2013), y las regiones transcripcionalmente inactivas se encuentran enriquecidas en la periferia nuclear (Padeken y Heun, 2014).

A una escala menor y mediante mapas de alta resolución de interacciones cromatínicas presentes en todo el genoma (Hi-C), se observó que el genoma se encuentra organizado en dominios topológicamente asociados (TADs) (**Fig. 3B**), los cuales son definidos por una alta frecuencia de interacciones dentro de un dominio y una baja frecuencia de interacciones entre dominios adyacentes (**Fig. 4**) (Dixon *et al.*, 2012; Nora *et al.*, 2012; Sexton *et al.*, 2012; Hou *et al.*, 2012).

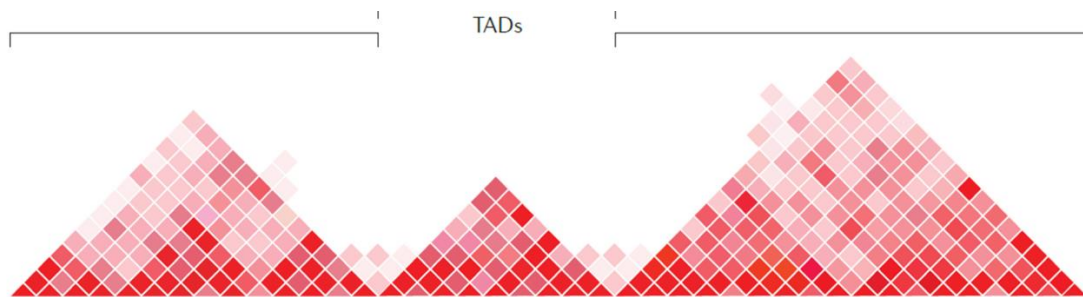


Fig. 4. Dominios topológicamente asociados (TADs). Mapas de interacciones obtenidos por Hi-C que muestran la organización del genoma en TADs, los cuales muestran interacciones frecuentes dentro de ellos y poco frecuentes con dominios adyacentes. Imagen tomada de Ong y Corces, 2014.

Todos los estudios anteriores apoyan un modelo jerárquico compuesto por múltiples niveles donde la expresión de un gen puede influenciar o ser influenciado por sus interacciones de cromatina locales, por sus asociaciones con otros genes y por el ambiente regulatorio provisto por su localización nuclear (Sexton y Cavalli, 2015).

Se ha postulado que las proteínas arquitectónicas juegan un papel importante en la organización tridimensional del genoma y en el establecimiento de contactos cromatínicos, regulando así la expresión de genes. Algunas de las proteínas arquitectónicas mejor estudiadas en vertebrados son CTCF y las cohesinas. Estas proteínas tienen dos funciones interrelacionadas: la compartimentación del genoma y facilitar las interacciones entre elementos reguladores, ya que están involucradas en la formación de los TADs, los cuales permiten contactos e interacciones de largo alcance entre secuencias de ADN (Dixon *et al.*, 2012; Cubeñas-Potts y Corces, 2015; Kim *et al.*, 2015, Ghirlando y Felsenfeld, 2016). CTCF es la proteína

arquitectural mejor caracterizada y que a diferencia de las cohesinas puede asociarse de forma específica a motivos de unión en el ADN.

El factor nuclear CTCF

El factor de unión CCCTC (CTCF) fue caracterizado inicialmente como un represor transcripcional del oncogen *c-myc* (Klenova *et al.*, 1993; Filippova *et al.*, 1996). Posteriormente, fue identificado como una proteína que al unirse a una secuencia *insulator* localizada entre un enhancer y un promotor impide la comunicación entre estos y la expresión del gen reportero es bloqueada (Bell *et al.*, 1999). En humano CTCF es una proteína de 82 kDa que está conformada por 727 aminoácidos organizados en tres dominios: el dominio N-terminal, el C-terminal y una región central de unión al ADN con 11 dedos de zinc (**Fig. 5A**). La unión de CTCF al ADN ocurre mediante diferentes combinaciones de sus 11 dedos de zinc y puede verse afectada por la metilación del ADN (Filippova *et al.*, 1996) así como también por modificaciones post-traduccionales como fosforilación, sumoilación y poly-ADP-ribosilación (Vostrov *et al.*, 2002).

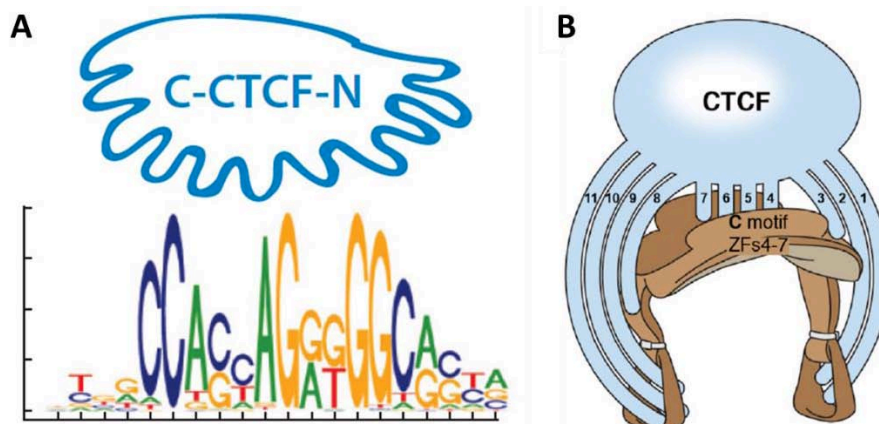


Fig. 5. Unión de CTCF a su motivo en el ADN. (A). Unión del factor CTCF a su motivo consenso en el ADN en orientación sentido. Imagen tomada de Merckenschlager y Nora, 2016. **(B).** Modelo de unión de CTCF con sus dedos de zinc centrales unidos al motivo y los dedos de zinc restantes le proporcionan estabilidad a la unión. Imagen tomada de Nakahashi *et al.*, 2013.

En particular, CTCF reconoce a su motivo de unión al ADN con los dedos de zinc centrales (4-7) y los dedos restantes tienen un papel estabilizador en la unión (**Fig. 5B**) (Nakahashi *et al.*, 2013). Datos de inmunoprecipitación de la cromatina seguida de secuenciación masiva (ChIP-seq) sugieren que CTCF tiene aproximadamente 50,000 sitios de unión en el genoma humano (Chen *et al.*, 2008), de los cuales el 50% son intergénicos, el 35% intragénicos y el resto se encuentran cerca de promotores (Chen *et al.*, 2012). CTCF se expresa de forma ubicua y está altamente conservado desde *D. melanogaster* hasta humano (Moon *et al.*, 2005) lo cual sugiere que CTCF es importante para la fisiología celular y el desarrollo de los organismos. Por ejemplo, ratones *knockout* para CTCF mueren antes de la implantación (Splinter *et al.*, 2006) y la eliminación materna de CTCF en ovocitos interrumpe la progresión normal al estadio de blastocisto (Fedoriw *et al.*, 2004). Además se ha mostrado que el *knockdown* de CTCF por ARNs pequeños de interferencia (siRNAs) en líneas celulares de mamíferos afecta su crecimiento, proliferación, diferenciación e induce a apoptosis (Torrano *et al.*, 2005).

Se ha demostrado que CTCF se une al promotor del gen retinoblastoma (*Rb*), tanto *in vitro* como *in vivo*. La ausencia de CTCF provoca represión transcripcional ya que protege al promotor de *Rb* contra el silenciamiento epigenético progresivo (De La Rosa-Velázquez *et al.*, 2007). Por otra parte, CTCF está involucrado en la regulación epigenética del gen *p53*, también al proteger contra el silenciamiento epigenético (Soto-Reyes y Recillas-Targa, 2010). Adicionalmente, CTCF también puede regular la expresión de genes que codifican a microRNAs como lo es el caso del miR-181c al proteger su promotor contra el silenciamiento epigenético en células de glioblastoma (Ayala-Ortega *et al.*, 2016).

Todos estos datos indican que CTCF es fundamental para la regulación de la expresión génica (Phillips y Corces, 2009).

A la fecha se han descrito para CTCF una gran variedad de funciones ligadas a diversos procesos celulares que incluyen su función como factor de transcripción, como proteína asociada a *insulators*, su papel en el procesamiento alternativo de ARNs mensajeros, en la impronta genómica y en la inactivación del cromosoma X,

aunque es posible que todas estas funciones estén basadas en su capacidad para organizar del genoma al interior del núcleo mediante la formación de asas cromatínicas, demostrando su papel como proteína arquitectónica (Ong y Corces, 2014).

Adicionalmente, datos no publicados del laboratorio sugieren que CTCF a escala global se asocia a elementos regulatorios cercanos a genes que previamente han sido implicados en procesos de señalización celular.

A pesar de la gran cantidad de estudios enfocados en entender la relación entre la regulación de la expresión génica y la organización tridimensional del genoma, actualmente se sabe poco sobre qué factores están involucrados en la formación y/o mantenimiento de esta organización, y si es causa o consecuencia de la actividad transcripcional. Por lo anterior, es necesario demostrar cómo influye la organización del genoma en la expresión de genes, lo cual es el objetivo del presente trabajo, el cual además está enfocado en conocer la regulación de un gen de gran importancia tanto biológica como clínicamente como lo es el gen *Notch1*.

ANTECEDENTES PARTICULARES

- **CTCF regula la expresión de proto-oncogenes a través de la formación de asas de cromatina**

Dada la posible relación entre la organización tridimensional del genoma y la regulación en la expresión génica, varios grupos de investigación han tratado de demostrar dicha relación mediante distintas estrategias, la mayoría de ellas enfocadas en identificar los elementos que participan en la formación y/o mantenimiento de las asas de cromatina y en las consecuencias de la pérdida de dichas asas en la regulación transcripcional. Por ejemplo, recientemente se descubrió que la orientación de las secuencias que corresponden a los motivos de unión a CTCF es un factor determinante para la formación de asas de cromatina, ya que en más del 90% de las asas de cromatina presentes en todo el genoma humano los sitios de unión a CTCF se encuentran en orientación convergente (sentido y antisentido) (**Fig. 6**) (Rao *et al.*, 2014).

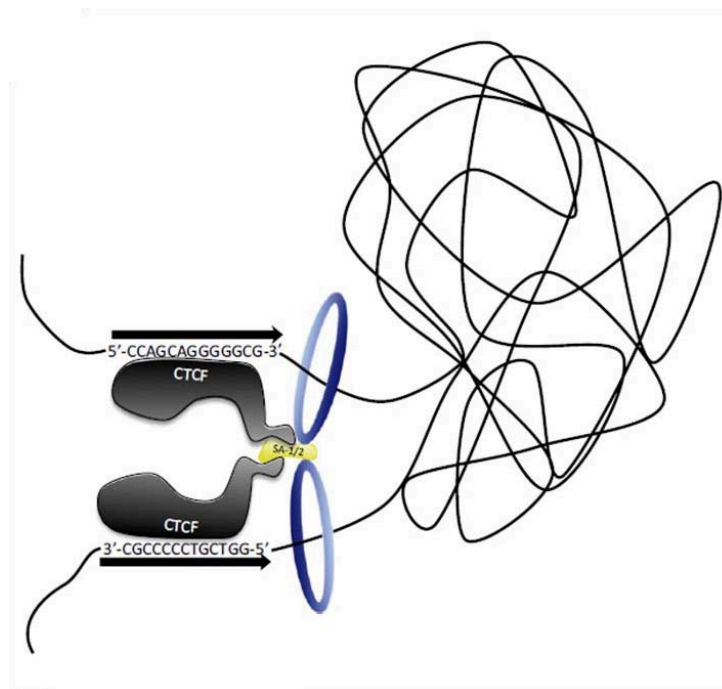


Fig. 6. Orientación convergente de sitios de unión a CTCF. Esta orientación con un sitio de unión a CTCF sentido y otro antisentido es la forma más frecuente en la que se observan los motivos de unión a CTCF en los límites de asas de cromatina, por lo que puede funcionar para predecir la presencia de un asa. Imagen tomada de Vietri Rudan y Hadjur, 2015.

En la leucemia linfoblástica aguda CTCF ha sido implicado en la regulación transcripcional de proto-oncogenes a través de la formación de asas de cromatina.

Richard Young y colaboradores demostraron que la mayoría de los proto-oncogenes en la leucemia linfoblástica aguda de células T (T-ALL), se encuentran dentro de asas de cromatina de aproximadamente 270 kb y con CTCF en sus límites (Hnisz *et al.*, 2016). De forma interesante, eliminaciones recurrentes en T-ALL afectan al menos uno de los límites o bases de asas de cromatina teniendo como consecuencia la ausencia del asa. La pérdida de las asas de cromatina a causa de las eliminaciones en T-ALL correlacionó con un incremento en la expresión de los genes al interior del asa (**Fig. 7**).

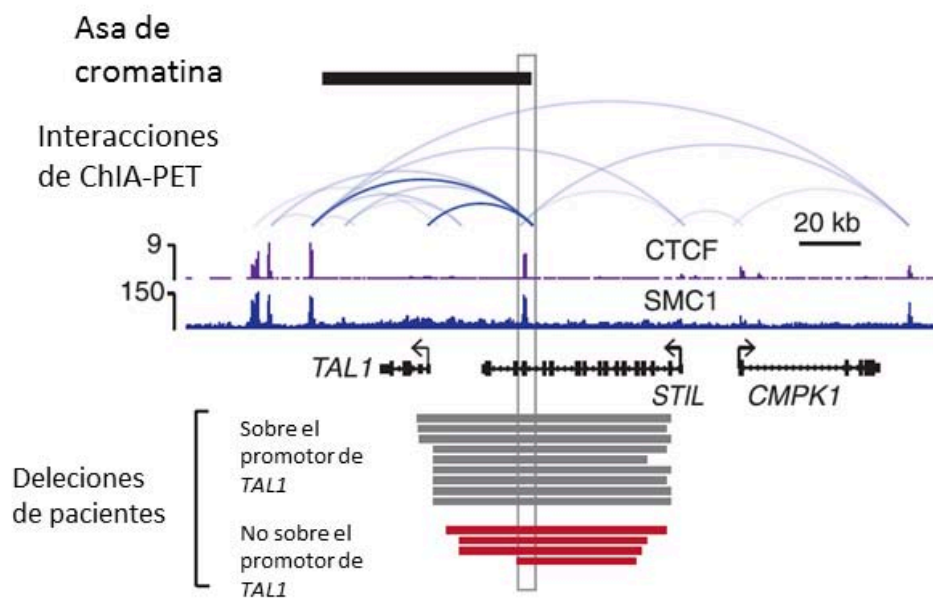


Fig. 7. Deleciones recurrentes en pacientes con T-ALL. Las deleciones se encuentran abarcando un sitio de unión a CTCF mostrado por ChIP-seq y que interactúa con otro sitio de CTCF para formar un asa de cromatina conteniendo al oncogen *TAL1*. Imagen tomada de Hnisz *et al.*, 2016.

Para demostrar que CTCF era importante para el mantenimiento de las asas de cromatina que contiene a proto-oncogenes, eliminaron sitios de unión a CTCF mediante la técnica de edición genómica CRISPR-Cas9 en los límites del asa en

células que tenían silenciados esos genes (**Fig. 8A**), y observaron que con dicha eliminación se perdieron las asas de cromatina y esto condujo a la activación de los proto-oncogenes (**Fig. 8B**).

Dado que elementos regulatorios como enhancers pueden potenciar la expresión de un gen, en ese estudio sugieren que estos elementos regulatorios podrían ser los causantes de la activación o sobreexpresión de los genes dentro de las asas, los cuales en condiciones normales no interactúan con el gen pero al perder la integridad del asa los enhancers los contactaron y potenciaron su expresión, dichos enhancers fuera de las asas fueron observados por marcas de cromatina asociada a éstos, como la acetilación de la lisina 27 de la histona H3 (**Fig. 8C**).

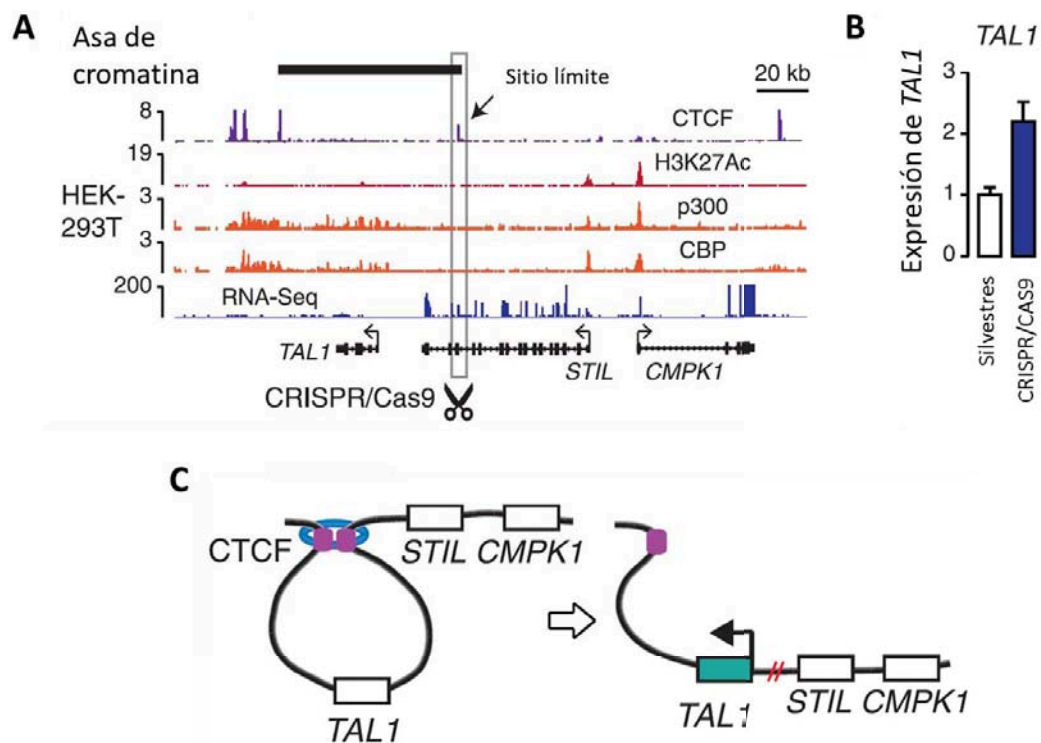


Fig. 8. Activación de proto-oncogenes por pérdida de asas de cromatina. (A). Eliminación mediante CRISPR-Cas9 de un sitio de unión a CTCF río arriba del inicio de la transcripción de *TAL1* en células HEK293, datos de ChIP-seq muestran el enriquecimiento de CTCF y de marcas de cromatina asociadas a enhancers. **(B).** Nivel de expresión relativa de *TAL1* obtenido por RT-qPCR en células HEK293 con y sin la eliminación del sitio de unión a CTCF. **(C).** Modelo representativo del asa de cromatina presente en el gen *TAL1* y el resultado de eliminar un sitio de CTCF en uno de sus límites. Imagen tomada de Hnisz *et al.*, 2016.

Además, se demostró recientemente que los sitios de unión a CTCF frecuentemente se encuentran mutados en cáncer, lo cual conlleva a una pérdida de la unión o del enriquecimiento de CTCF, afectando la arquitectura del genoma y llevando a una expresión génica anormal (Katainen *et al.*, 2015).

La pérdida del límite de asas de cromatina se ha visto también involucrada en diversas patologías como la disgenesia acropectorovertebral en donde una eliminación localizada justo en el límite de un asa ocasiona interacciones enhancer-promotor ectópicas y una expresión desregulada, lo cual causa malformaciones (Lupiáñez *et al.*, 2015).

En el caso de la leucemia mieloide aguda y el meduloblastoma también se ha observado ganancia de interacciones entre enhancers y promotores anormales lo cual se ha propuesto como uno de los principales mecanismos moleculares responsable de la desregulación de la expresión génica en dichas enfermedades (Groschel *et al.*, 2014; Northcott *et al.*, 2014).

Por otro lado, las mutaciones más frecuentes de ganancia de función del gen *Notch1* han sido caracterizadas en sus regiones intergénicas río arriba y río abajo, en secuencias no codificantes (Weng *et al.*, 2004; Puente *et al.*, 2015), lo cual sugiere que se podrían estar afectando elementos regulatorios del gen o sitios de unión a factores transcripcionales que regulan la expresión del gen *Notch1*.

Los antecedentes mencionados sugieren que el gen *Notch1* al ser considerado un oncogen en varios tipos celulares podría estar siendo regulado por asas de cromatina como en el caso de otros oncogenes en la T-ALL (Hnisz *et al.*, 2016). Dado que CTCF se encuentra frecuentemente mediando la formación de asas de cromatina, este factor podría regular la expresión del gen *Notch1* a través de la formación y mantenimiento de asas de cromatina.

- **Las células K562 como un modelo de estudio para la regulación transcripcional del gen *Notch1***

La línea celular K562, la cual corresponde a una leucemia mieloide crónica (CML), resulta un buen modelo para estudiar la regulación transcripcional de un gen como *Notch1* dado el papel relevante que muestra la vía de señalización de Notch en células mieloides, en donde participa como un gen supresor de tumores, lo cual ha resultado interesante y controversial para distintos grupos de investigación (South *et al.*, 2012; Suresh e Irvine, 2015). Además, desde un punto de vista práctico, el consorcio ENCODE ha generado una gran cantidad de datos de diferentes experimentos para las células K562, ya que es considerada prioridad para cualquier experimento que realicen (Dunham *et al.*, 2012), con lo cual se tiene una vasta información epigenómica sobre esas células. Por último, la línea K562 es fácil de manejar en cultivo y se puede utilizar para distintos experimentos dado que se puede hacer proliferar en grandes cantidades.

PLANTEAMIENTO DEL PROBLEMA

El factor CTCF participa en la regulación transcripcional, sin embargo, se desconoce su papel en la regulación de genes involucrados en el desarrollo como el gen *Notch1*, el cual además es el gen que más frecuentemente se encuentra afectado en la leucemia linfocítica crónica.

HIPÓTESIS

CTCF regula la expresión de diversos genes. Datos de ChIP-seq en células de humano sugieren que CTCF se asocia río arriba del inicio de la transcripción del gen *Notch1*, por lo que CTCF regula su expresión transcripcional.

OBJETIVO GENERAL

- Demostrar la regulación de la expresión del gen *Notch1* por el factor multifuncional CTCF en células humanas de eritroleucemia K562.

OBJETIVOS PARTICULARES

- Validar la presencia de CTCF en el locus del gen *Notch1* mediante ChIP.
- Evaluar la expresión del gen *Notch1* por RT-qPCR.
- Realizar un *knockdown* de CTCF y evaluar su efecto en la expresión del gen *Notch1*.
- Determinar el efecto de la eliminación precisa del sitio de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1* por CRISPR-Cas9, en la expresión del gen *Notch1*.

MATERIALES Y MÉTODOS

Análisis bioinformáticos

- Análisis de datos de Inmunoprecipitación de la Cromatina seguida por secuenciación masiva (ChIP-seq)

Los datos de ChIP-seq para CTCF y marcas de histonas usados en este trabajo provienen de los resultados obtenidos por el proyecto ENCODE en diferentes líneas celulares humanas (**Tabla S1**). Estos datos fueron visualizados en el Buscador de Genomas de la Universidad de California (UCSC Genome Browser: <https://genome.ucsc.edu/>) y en el visualizador de genomas IGV (Integrative Genomics Viewer: <https://www.broadinstitute.org/igv/>).

- Identificación de motivos de unión a CTCF

Para corroborar la presencia del motivo de unión a CTCF en los “picos” de enriquecimiento provenientes de los datos de ChIP-seq, así como para conocer la orientación de los motivos de unión a CTCF, se utilizó la base de datos JASPAR (http://jaspar.genereg.net/cgi-bin/jaspar_db.pl) y se usó el identificador de matriz: MA0139.1.

- Visualización de datos de Captura Conformacional de Cromosomas de Genoma Completo (Hi-C)

Los datos de Hi-C se visualizaron en el software Juicebox 1.5 (<http://www.aidenlab.org/juicebox/>), se utilizaron los datos de Rao y colaboradores de 2014 obtenidos por un experimento de Hi-C *in situ* en la línea celular GM12878 a una resolución de 1 kb, los cuales se cargaron en el software junto a los datos de la anotación de los genes y los datos de ChIP-seq de ENCODE para visualizar el enriquecimiento de CTCF.

- Visualización de datos de Análisis de Interacciones de Cromatina por secuenciación de extremos pareados (ChIA-PET)

Por último, en el buscador de la Universidad de Washington (WashU Epigenome Browser: <http://epigenomegateway.wustl.edu/browser/>) se visualizaron datos de

ChIA-PET de CTCF en las células K562 obtenidos por el consorcio ENCODE (<http://egg.wustl.edu/d/hg19/wgEncodeGisChiaPetK562>) para observar las interacciones mediadas por esta proteína.

Cultivo celular

Los experimentos mostrados en esta tesis se realizaron empleando la línea celular humana de leucemia mieloide crónica K562 (<http://www.atcc.org/products/all/CCL-243.aspx>). Las células se mantuvieron en cajas Petri de 100 mm con medio ISCOVE suplementado al 10% con Suero Fetal Bovino (SFB) y se incubaron a 37°C en una atmósfera húmeda saturante con 5% de CO₂. Para cada experimento se evaluó la viabilidad celular por el método de exclusión con azul de tripano y se realizó un conteo celular con la cámara de Neubauer para utilizar el número de células adecuado para cada experimento.

Extracción de ácidos nucleicos

1) ADN

Se extrajo ADN mediante el método de fenol-cloroformo. Brevemente, los *pellets* celulares se incubaron con Mix I adicionado con RNasa A (50 µg/ml) a 37°C por 30 minutos y posteriormente con Mix II adicionado con Proteinasa K (50 µg/ml) a 50°C durante toda la noche. Posteriormente se agregó un volumen de fenol-cloroformo-ácido isoamílico, se centrifugó y se lavó con un volumen de cloroformo. Posteriormente se precipitó el ADN con un volumen de isopropanol y por último se lavó el pellet de ADN con un volumen de etanol al 70%, se dejó secar y se resuspendió en agua inyectable. El ADN extraído se cuantificó en un NanoDrop-2000 y se ajustó la muestra a una concentración de 100 ng/µL.

2) ARN

El ARN se extrajo por el método de Trizol, primero se resuspendieron 3×10^6 células en 1 mL del reactivo Trizol, después se agregaron 200 µL de cloroformo para separar por centrifugación al ARN, se agregaron 500 µL de isopropanol y se dejó precipitando 10 minutos, se centrifugó la muestra y finalmente el pellet se lavó con

etanol al 70%, se dejó secar 10 minutos a temperatura ambiente y se resuspendió en 50 µL de agua inyectable. Para mejorar la solubilidad del ARN se incubó la muestra a 50°C durante 10 minutos, posteriormente se colocó en hielo y por último, el ARN extraído se cuantificó en un NanoDrop-2000.

Para determinar su integridad se corrieron 2 µg de ARN en un gel de agarosa desnaturante al 1.2% con 5% de formaldehído al 37% para evitar su degradación, el gel se corrió a 80 V durante 30 minutos y para confirmar su integridad se observaron las subunidades ribosomales 28S y 18S (**Fig. S1**).

Diseño y estandarización de oligonucleótidos

Para el caso de la PCR en tiempo real, la inmunoprecipitación de la cromatina y la edición con CRISPR-Cas9 se diseñaron pares de oligonucleótidos que permitieron amplificar los sitios elegidos, los cuales se obtuvieron en el programa Primer3 0.4 (<http://bioinfo.ut.ee/primer3-0.4.0/>). Las secuencias de las zonas de interés se obtuvieron en el Genome Browser y NCBI. También se utilizó el programa Primer-Blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) para determinar la especificidad de los oligonucleótidos obtenidos para cada sitio blanco.

Los oligonucleótidos fueron Desalt y se mandaron a diseñar a SIGMA. Una vez obtenidos los oligonucleótidos se estandarizaron realizando PCRs de punto final bajo distintas condiciones, variando la temperatura de alineamiento y la concentración de magnesio en la reacción, esto para conocer bajo qué condiciones se obtenía el producto esperado de amplificación de cada par de oligonucleótidos. Los oligonucleótidos utilizados en este trabajo se muestran en la **Tabla S2**, en donde se indican además las condiciones óptimas de amplificación de cada par de oligonucleótidos.

Retrotranscripción seguida de PCR en tiempo real (RT-qPCR)

Para determinar el nivel de ARNm del gen *Notch1* en las células humanas K562 se realizó primero un tratamiento con DNasa I al ARN purificado para eliminar el ADN residual y se realizó una retrotranscripción (RT) con 2 µg de ARN con la enzima

AMV (Promega) y con 2 μ L de Random Primers. La muestra se dejó 30 min a 37°C y 15 min a 65°C para obtener el ADN complementario (cDNA).

El producto obtenido se diluyó 1:5 con agua inyectable y se utilizó para realizar PCR cuantitativa (qPCR) con ayuda del equipo StepOne, y utilizando oligonucleótidos específicos, previamente estandarizados para el ARN mensajero del gen *Notch1* (**Fig. S2**). SYBR Green fue empleado para cuantificar el cDNA, así como la enzima KAPA Taq para amplificar el cDNA. Como control para normalizar se usaron oligonucleótidos para amplificar el gen constitutivo *GAPDH*. Para determinar la expresión relativa de los genes se utilizó el método comparativo $\Delta\Delta C_t$.

Análisis estadístico

Los datos se presentaron como la media \pm S.D. de al menos tres experimentos independientes con tres repeticiones y se analizaron estadísticamente usando una prueba t de Student con un nivel de significancia de $P < 0.05$. Los resultados obtenidos se analizaron con el software estadístico GraphPad Prisma 6.0.

Inmunoprecipitación de la cromatina

La inmunoprecipitación de la cromatina (ChIP) consiste en utilizar un anticuerpo específico contra la proteína de interés seguido de la extracción del ADN para conocer a qué secuencias en el genoma se encuentra unida la proteína.

Para esto, se utilizaron 20×10^6 células K562 y se agregaron directamente en el cultivo celular 2 mL de buffer de “cross-linking” en 20 mL de medio para obtener una concentración de 1% de formaldehído al 37% para fijar las interacciones ADN-proteínas que se estaban llevando a cabo en las células, se dejaron 10 min para la fijación. Después de la fijación las células se lavaron 3 veces con PBS frío y posteriormente se agregó a las células 1 mL de buffer de lisis con inhibidores de proteasas (PMSF, Leupeptina 1 mg/mL, Aprotinina 10 mg/mL y Pepstaína 1 mg/mL) y se realizó la sonicación de la muestra aplicando 12 pulsos de 30 segundos a 35% de amplitud para obtener fragmentos de aproximadamente 400-800 pb.

Para determinar el tamaño de los fragmentos obtenidos se tomó una alícuota de 25 μL , se extrajo el ADN con fenol-cloroformo y el ADN extraído se corrió en un gel de agarosa al 1% (**Fig. S3**).

Se utilizaron 50 μg de ADN para cada inmunoprecipitación, una muestra se incubó con 5 μL de un anticuerpo específico contra CTCF (Millipore), y como controles se utilizaron una muestra con anticuerpo IgG inespecífico y otra muestra sin ningún anticuerpo. Para precipitar las secuencias en las que estaba unida la proteína se utilizaron 30 μL de perlas Proteína G/A, ya que estas se unen a los anticuerpos utilizados. Después se realizaron 4 lavados con buffer de paro I y un lavado con buffer de paro II. Finalmente se agregó buffer de elución para separar el ADN de las perlas y se realizó extracción con fenol-cloroformo para obtener el ADN puro para realizar la PCR.

Las secuencias de ADN inmunoprecipitadas se amplificaron por PCR punto final utilizando oligonucleótidos específicos de las zonas de interés previamente estandarizados (**Fig. S4**), para así saber si esos sitios estaban enriquecidos con CTCF.

Knockdown de CTCF

El *knockdown* (KD) se realizó utilizando un ARN corto de interferencia con una estructura tallo-asa (shRNAi) previamente diseñado en el laboratorio (González-Buendía *et al*, 2014), el cual fue clonado en el vector lentiviral pLL3.7 para después ser transducido mediante partículas lentivirales a las células K562. El *knockdown* empleado no funciona hasta que se realiza el tratamiento con doxiciclina, la cual se agrega cada 48 hrs. Con este tratamiento el shRNAi se transcribe, reconoce al ARNm de CTCF e induce su degradación, con lo cual se evita que se traduzca a proteína y eso permite disminuir los niveles de CTCF en la célula (**Fig. 9**).

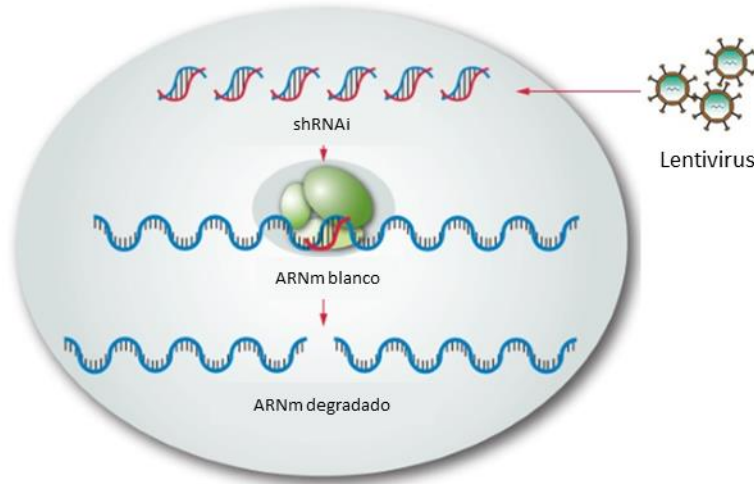


Fig. 9. Sistema de *knockdown* mediante un shRNAi. El shRNAi fue introducido a las células con partículas lentivirales. El shRNAi comienza a transcribirse cuando se trata a las células con doxiciclina, una vez transcrito, el shRNAi reconoce a su ARNm blanco, lo cual induce la degradación del ARNm.

Para validar la disminución de la proteína CTCF se realizó la técnica de Western Blot utilizando un anticuerpo específico contra CTCF y como control de carga se utilizó un anticuerpo contra GAPDH. Se realizó el Western Blot en células transducidas con el vector vacío con y sin el tratamiento con doxiciclina y en células con el vector con el shRNAi contra CTCF con y sin el tratamiento con doxiciclina. Una vez confirmado el KD se analizó el efecto que tuvo en la expresión del gen *Notch1* mediante RT-qPCR.

Sistema de edición genómica CRISPR-Cas9

El sistema CRISPR-Cas9 permite realizar edición genómica con gran precisión. El sistema consiste en utilizar a la endonucleasa Cas9 (en este caso la de *Streptococcus pyogenes*) y una secuencia que se conoce como guía única de ARN (sgRNA), la cual dirige a la Cas9 a su sitio blanco en el ADN para que pueda realizar cortes de doble cadena (Ran *et al.*, 2013). El sgRNA consta de 20 nucleótidos que le permiten reconocer a su sitio blanco en el genoma y también contiene una secuencia que reconoce a la Cas9 (tracker RNA) y permite que formen el complejo Cas9-sgRNA.

El complejo necesita reconocer además un motivo denominado motivo protoespaciador adyacente (PAM) para que la Cas9 realice el corte de doble cadena, este motivo debe estar inmediatamente río abajo del sitio blanco y en el caso del protocolo utilizado es de 3 nucleótidos que son 5'-NGG. El corte de doble cadena se realiza 3 pares de bases río arriba del PAM (**Fig. 10**).

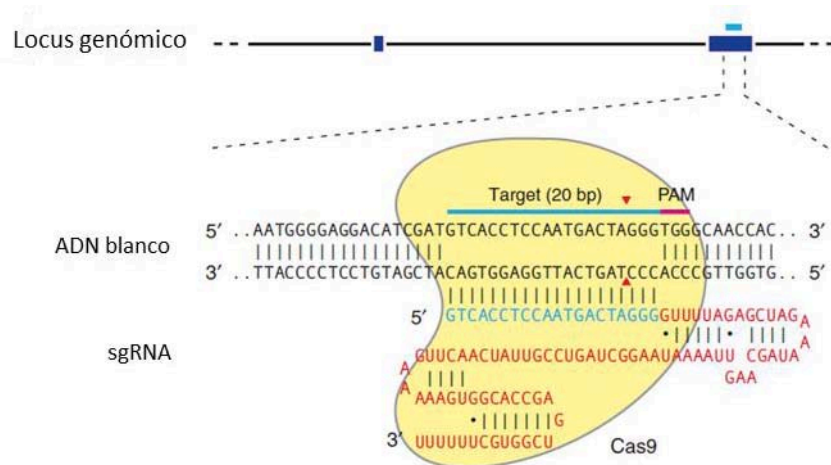


Fig. 10. Complejo formado por el sgRNA y la endonucleasa Cas9. El sgRNA le da la especificidad a la Cas9 para realizar un corte de doble cadena 3 pares de bases río arriba del PAM. Imagen tomada de Ran *et al.*, 2013.

Dicho corte de doble cadena activa uno de los dos principales mecanismos de reparación de ADN, ya sea el de unión de extremos no homólogos (NHEJ), el cual es el más frecuente, o el de reparación dirigida por homología (HDR), y mediante dichos mecanismos de reparación del ADN se pueden obtener eliminaciones, inserciones o inversiones (**Fig. 11**).

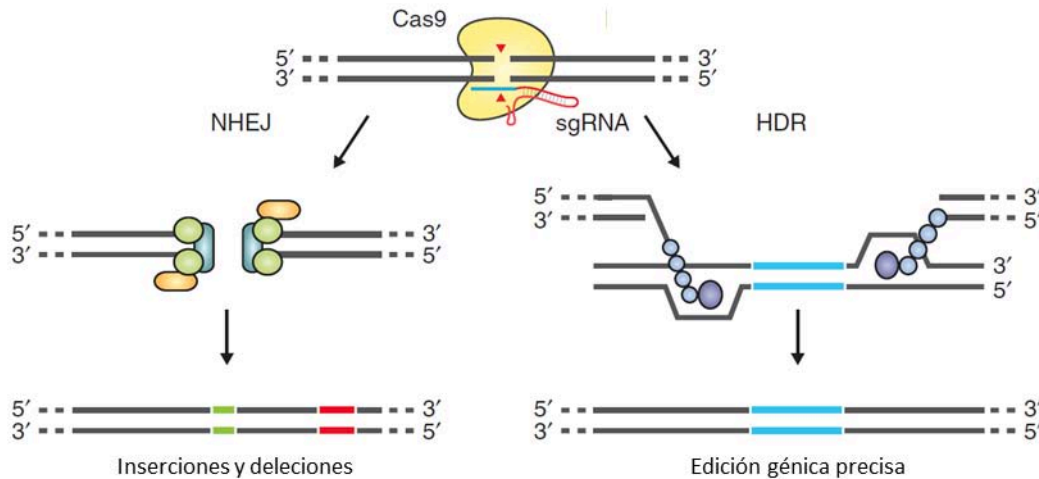


Fig. 11. Principales mecanismos de reparación del ADN. En el caso del de unión de extremos no homólogos (NHEJ) se reconoce el sitio de corte y al reparar sin homología se dan como resultado inserciones o deleciones y la reparación no es precisa. Por otro lado, el de reparación mediada por homología (HDR) utiliza un templado de ADN homólogo para reparar de forma precisa. Imagen tomada de Ran *et al.*, 2013.

En un experimento de edición genómica por CRISPR-Cas9 es necesario construir los plásmidos que contendrán los sgRNAs, introducirlos en las células de interés y evaluar con cuidado la lesión molecular. A continuación se describen cada uno de estos pasos.

1) Diseño y clonación de sgRNAs

Para generar una eliminación genómica empleamos una estrategia en la cual se usan dos sgRNAs que enmarcan la región a eliminar. En este trabajo, se diseñaron dos sgRNAs que enmarcaban los sitios de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1*, los sgRNAs se diseñaron con la herramienta CRISPR Design (<http://crispr.mit.edu/>), para lo cual se necesita una secuencia no mayor a 250 pb de la zona en donde se desea realizar el corte de doble cadena y elegir el organismo de interés. CRISPR Design Tool da una lista de sgRNAs cada uno con un puntaje que refleja la probabilidad de que ese sgRNA reconozca de forma inespecífica otros sitios. De forma general, sólo se seleccionaron sgRNAs que tuvieran un puntaje no menor a 80 y que enmarcaran el sitio de unión a CTCF que

se deseaba remover (**Fig. 12A**). Los sgRNAs seleccionados se muestran en la **Tabla S3**.

Los sgRNAs diseñados fueron clonados en el plásmido lentiviral lentiCRISPR-V2 el cual contiene la secuencia de la Cas9 (mejorada para el uso de codones de humano) y un gen de resistencia a la puromicina para poder seleccionar a las células infectadas (**Fig. 12B**).

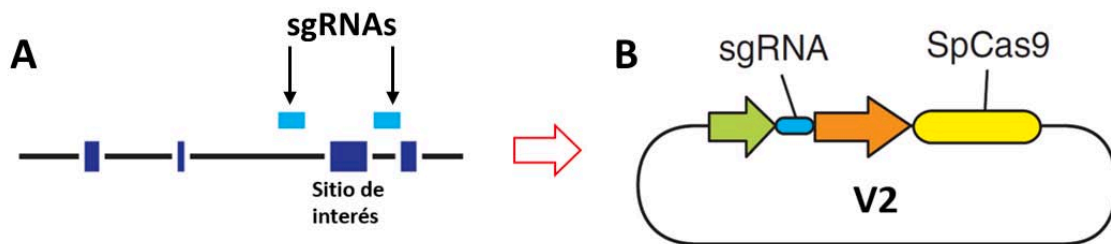


Fig. 12. Diseño y clonación de ARNs guías en el vector lentiviral lentiCRISPR-V2. (A). Representación de los sgRNAs enmarcando el sitio de interés para inducir una eliminación en dicha región. Para generar eliminaciones de más de 50 pb es recomendable usar esta alternativa de dos sgRNAs enmarcando la zona de interés. **(B).** Clonación y ligación del sgRNA en el plásmido lentiCRISPR-V2, junto a la secuencia que codifica a las Cas9. Con el tratamiento con puromicina se activa la expresión del plásmido lentiviral y se transcriben el sgRNA y la endonucleasa Cas9. Imagen tomada y modificada de Ran *et al.*, 2013.

Para confirmar la clonación y ligación de los sgRNAs en el plásmido, se realizó PCR en colonia y una vez verificados fueron expandidos por Midiprep y además se secuenciaron por Sanger para confirmar que las secuencias de los sgRNAs fueran las correctas.

2) Transducción con lentivirus

Brevemente, primero se clonaron los plásmidos lentivirales en el vector de expresión HPV17, los cuales se transfectaron en células HEK293 para la producción de los lentivirus. Después se tomó el sobrenadante de las células y se concentraron los virus centrifugándolos en una ultracentrífuga Beckman a 25,000 rpm por 90 min a 4 °C y el pellet se resuspendió en 50 µL de PBS frío.

Finalmente, para la transducción, se utilizaron 1×10^5 células K562 en 1.5 mL de medio y se les agregaron 10 μL de los lentivirus, se utilizaron 9 μL de polibreno (8 $\mu\text{g}/\text{mL}$) para mejorar la eficiencia de la transducción (**Fig. 13A**).

3) Selección y genotipificación

Pasadas 24 hrs de la transducción con lentivirus se cambió el medio de cultivo a las células y se agregó 1 $\mu\text{L}/\text{mL}$ de puromicina cada 48 hrs para seleccionar a las células que contenían el plásmido lentiviral.

Para confirmar la presencia de la eliminación se diseñaron oligonucleótidos al menos 100 pb alejados de los sitios de corte de la endonucleasa Cas9, los cuales se estandarizaron para obtener sólo el producto del tamaño predicho (**Fig. S5**). Después de 6 días en selección se extrajo ADN de las células infectadas con el cual se realizó PCR de punto final y se observó el producto de amplificación en un gel de agarosa.

Para confirmar que se obtuvo la eliminación esperada, primero, con los oligonucleótidos diseñados se determinó el tamaño esperado del amplicón eliminando la zona enmarcada por ambos sgRNAs, y después en el gel de agarosa es necesario observar productos de este tamaño para confirmar la funcionalidad del sistema (**Fig. 13B**).

4) Dilución clonal

Posteriormente, se realizó una dilución clonal para obtener poblaciones monoclonales que presentaran la eliminación predicha.

Para esto se realizaron diluciones seriales 1:10 partiendo de 5×10^5 células hasta obtener 50 células en 9.6 mL de medio con puromicina y se colocaron 100 μL en cada pozo en una placa de 96 pozos, resuspendiendo antes de colocar el volumen en los pozos. Después de una semana en cultivo, se revisaron las placas de 96 pozos para seleccionar las colonias que se formaron. Dichas monoclonas se expandieron, se extrajo ADN de cada una y se realizó PCR punto final para seleccionar las monoclonas homocigotas con la eliminación predicha (**Fig. 13C**).

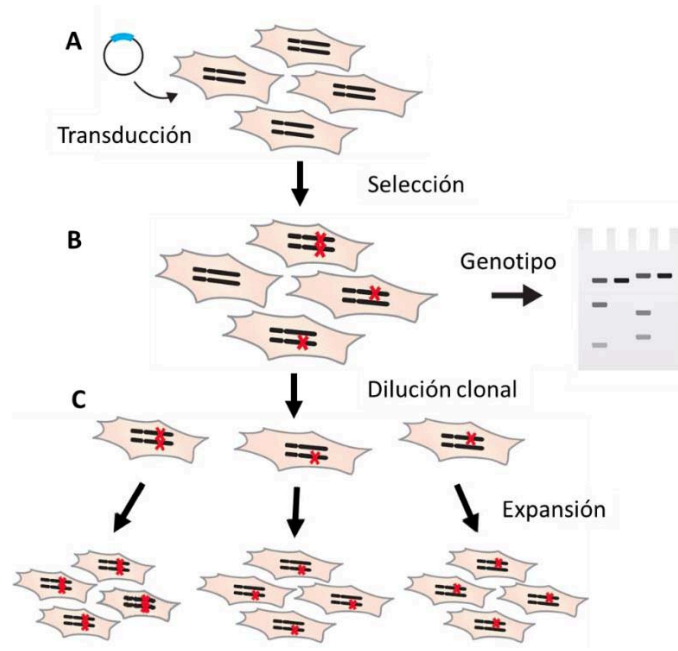


Fig. 13. Transducción, genotipificación y dilución clonal. (A). Transducción de partículas lentivirales con el sgRNA y la endonucleasa Cas9. **(B).** Las células se seleccionaron con puromicina y se observó el genotipo de la población para observar las eliminaciones. **(C).** La dilución clonal permitió aislar colonias monoclonales homocigotas y heterocigotas para posteriormente expandirlas. Imagen tomada y modificada de Ran *et al.*, 2013.

5) Secuenciación de clonas seleccionadas

Una vez seleccionadas las monoclonas con el tamaño de la eliminación predicha se realizó PCR de cada una, se separaron los productos en un gel de agarosa al 1% y se cortaron las bandas del tamaño esperado. Las muestras se purificaron para eliminar la agarosa y obtener sólo el ADN, el cual se clonó en el plásmido pGEM T-Easy. Se realizó una digestión de las muestras clonadas con la enzima de restricción EcoR1 para confirmar que el producto ligado en el plásmido fuera del tamaño esperado. Por último se purificó el ADN mediante columnas de Miniprep y se cuantificó el ADN obtenido para secuenciarlo.

Las muestras se secuenciaron por el método de Sanger y se usaron oligonucleótidos del promotor T7. Los electroferogramas obtenidos se analizaron con el software BioEdit 7.2 y se alinearon las secuencias en el UCSC Genome Browser.

RESULTADOS

- **CTCF se une a la región intergénica 5' y 3' del gen *Notch1* en células K562**

CTCF regula la expresión de diversos genes y se ha visto involucrado en el control de la transcripción de genes involucrados en cáncer como *p53*, *Rb* y de microRNAs como *miR-181c*, principalmente a través de la protección contra el silenciamiento por metilación de islas CpG (De La Rosa-Velázquez *et al.*, 2007; Soto-Reyes y Recillas-Targa, 2010; Ayala-Ortega *et al.*, 2016), sin embargo, poco se sabe sobre su papel en la regulación de genes importantes para el desarrollo e independiente de su papel como protector contra el silenciamiento por metilación.

Análisis de datos de ChIP-seq generados por el proyecto ENCODE muestran que CTCF se asocia a la región intergénica 5' y 3' del gen *Notch1* en al menos siete líneas celulares de humano. Interesantemente, algunos sitios de unión a CTCF son constitutivos en todas las líneas celulares y otros son específicos de algunas líneas (**Fig. 14**). La presencia de CTCF sugiere que podría regular la expresión del gen *Notch1*.

Recientemente, se ha demostrado que CTCF se localiza en la base de asas de cromatina en humano. Particularmente, la unión de CTCF en los extremos del asa de cromatina parece estar correlacionada con la orientación de sus motivos, siendo la conformación más frecuente la convergente, es decir, sentido y antisentido (**Fig. 6**) (Rao *et al.*, 2014). La presencia de motivos convergentes para CTCF enmarcando al gen *Notch1* podría sugerir la presencia de un asa de cromatina, dado que la mayoría de las asas tienen a los motivos de unión a CTCF en dicha orientación. Por lo anterior, se analizó la orientación de los sitios de unión a CTCF identificados por ChIP-seq usando JASPAR, se observó que los motivos se encuentran en una orientación convergente, es decir, la región 5' tiene un motivo de unión en dirección sentido, mientras la región 3' tiene un motivo en dirección antisentido (**Fig. 14**). En la **Tabla S4** se muestran las coordenadas, la secuencia y el puntaje de JASPAR de la orientación de cada sitio de unión a CTCF.

La unión de CTCF identificada por los datos de ChIP-seq y la orientación de los motivos de CTCF en el locus del gen *Notch1* sugieren que CTCF podría regular la expresión de *Notch1* a través de la formación de un asa de cromatina.

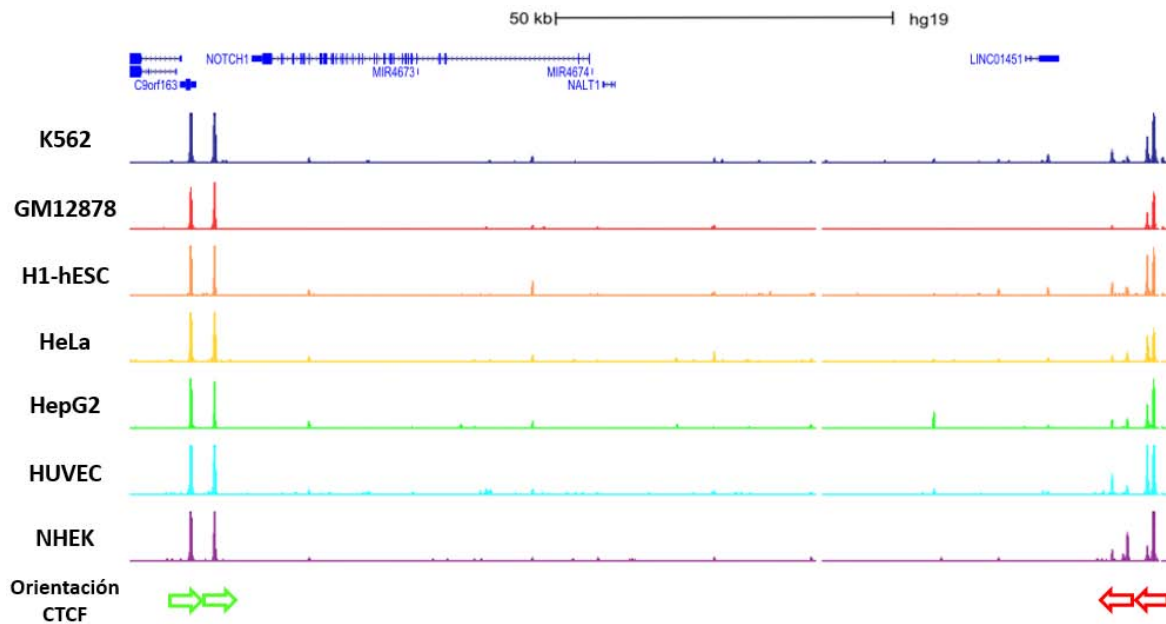


Fig. 14. CTCF se asocia en orientación convergente al locus del gen *Notch1* en siete líneas celulares. Datos de ChIP-seq de CTCF de 7 líneas celulares obtenidos por ENCODE, visualizados en el UCSC Genome Browser, con los cuales se pueden observar sitios de enriquecimiento para CTCF enmarcando al gen *Notch1*, ya que se observan sitios enriquecidos con CTCF en la región 5' y 3' del locus. La orientación de los sitios de CTCF obtenidos de JASPAR muestra una orientación convergente.

Con el objetivo de conocer todas las interacciones cromatínicas presentes en el genoma, Rao y colaboradores realizaron un experimento de Hi-C *in situ* en la línea celular GM12878 a una resolución de 1 kb (Rao *et al.*, 2014). Dado que es el mapa de interacciones con mayor resolución disponible a la fecha, se utilizaron esos datos para visualizar la presencia de un asa de cromatina en el locus del gen *Notch1*, y se pudieron observar interacciones frecuentes entre los sitios de unión a CTCF mostrados por ChIP-seq enmarcando al gen *Notch1*; además el algoritmo utilizado identificó en el locus del gen *Notch1* un asa de cromatina de aproximadamente 170 kb (**Fig. 15A**). En el mapa de interacciones generado en las células K562, aunque

es de una menor resolución (5 kb), también se observaron interacciones frecuentes entre los sitios de unión a CTCF en el locus del gen *Notch1* (Fig. S6).

Adicionalmente, mediante ChIA-PET es posible observar sólo aquellas interacciones cromatínicas que estén mediadas por una proteína de interés, en este caso CTCF (Goh *et al.*, 2012), y los datos observados de ChIA-PET de CTCF en las células K562 mostraron de igual forma una alta frecuencia de interacciones entre los sitios enriquecidos con CTCF en el locus que comprende al gen *Notch1* (Fig. 15B). Estos datos sugieren que CTCF puede ser el responsable del establecimiento de la asa cromatina que contiene al gen *Notch1* (Fig. 15C).

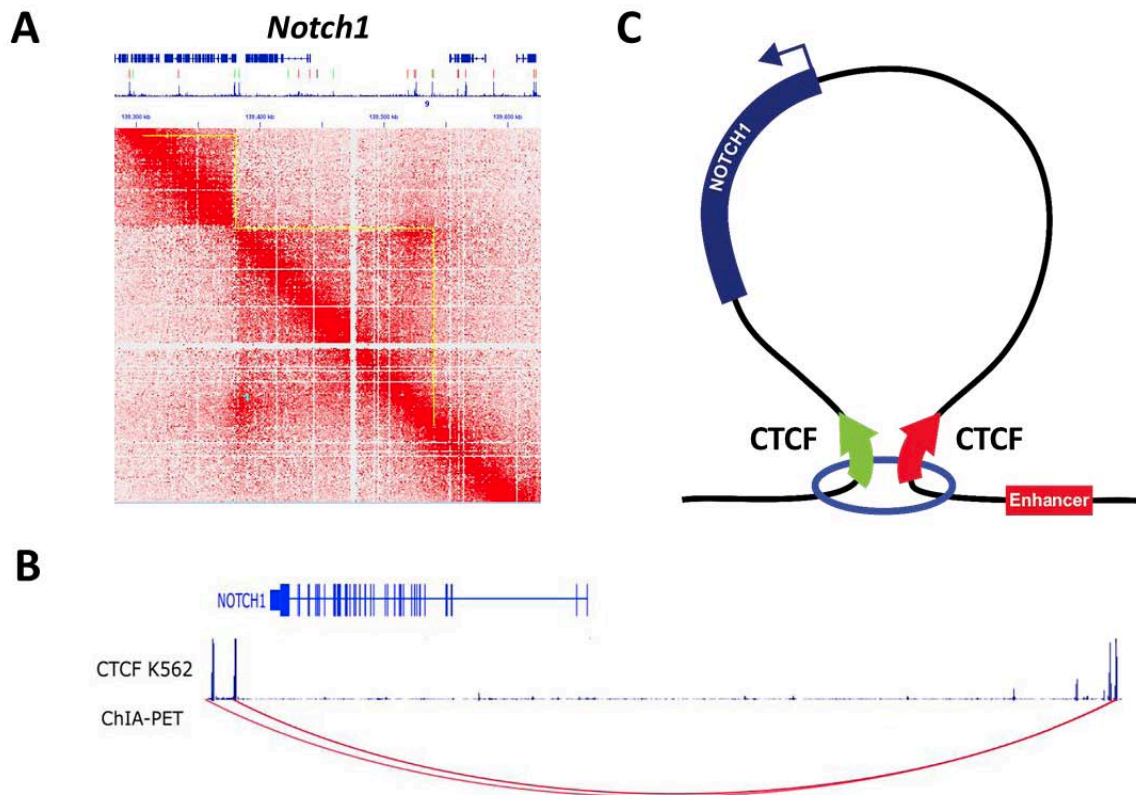


Fig. 15. El locus del gen *Notch1* constituye un asa de cromatina. (A). Mapa de frecuencia de interacciones de Hi-C *in situ* en las células GM12878 de 1 kb de resolución mostrando un asa de cromatina en el locus del gen *Notch1*, anclado por sitios enriquecidos con CTCF mostrados por ChIP-seq. **(B).** Interacciones mediadas por CTCF en las células K562 en el locus del gen *Notch1* obtenidas por ChIA-PET. **(C).** Esquema representativo del asa de cromatina observado en el locus del gen *Notch1*, teniendo CTCF en sus límites.

Los datos de ChIP-seq obtenidos a partir de las células de eritroleucemia K562 sugieren la presencia de CTCF en el locus del gen *Notch1*, sin embargo, se decidió validar esos datos mediante ChIP-PCR usando un anticuerpo altamente específico contra CTCF.

Se diseñaron y estandarizaron oligonucleótidos para amplificar los sitios de enriquecimiento observados, dos sitios en la región 5' del locus del gen *Notch1* y seis en la región 3' del locus, los cuales fueron sitios representados en todas las líneas celulares. En todos los sitios elegidos se pudo observar la presencia de CTCF, tanto en los sitios constitutivos como en los específicos de las células K562, incluso se pudo observar en sitios que mostraban poco enriquecimiento en los datos de ChIP-seq, como el sitio #8 (**Fig. 16**),

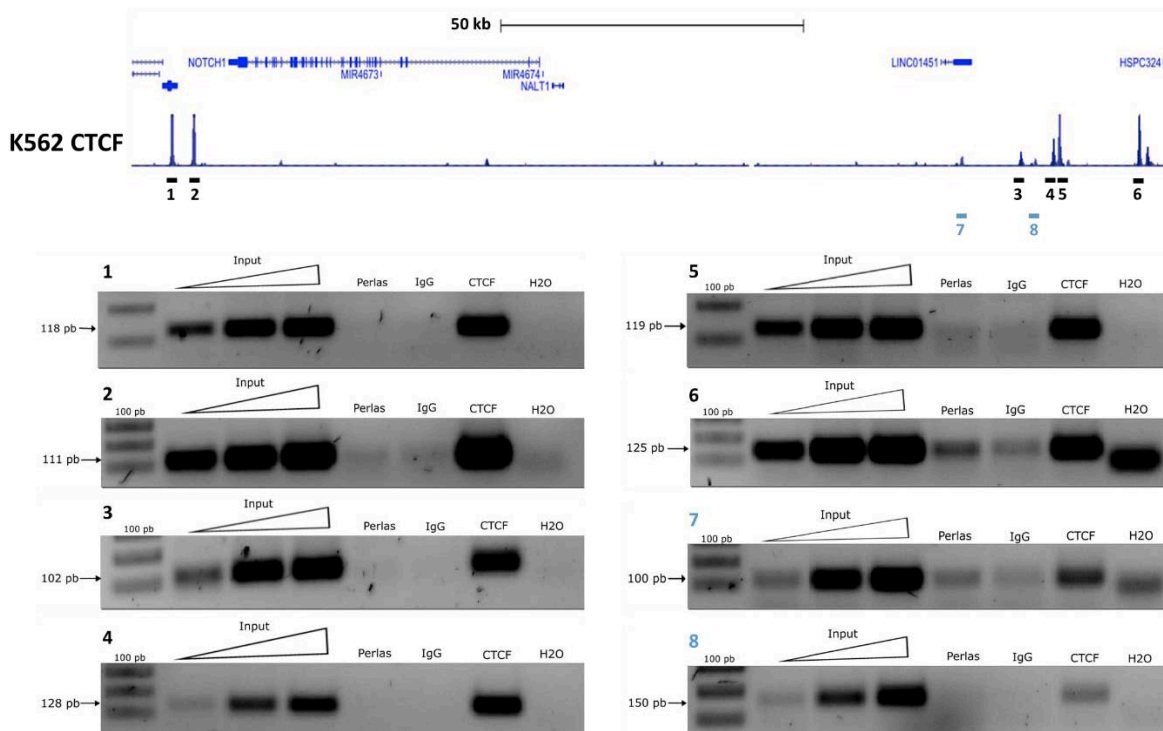


Fig. 16. CTCF se une a la región intergénica 5' y 3' del gen *Notch1* en células K562. ChIP-PCR de sitios de enriquecimiento de CTCF presentes tanto en la región 5' como la región 3' del locus del gen *Notch1* en células K562. En todos los sitios se observa claramente un producto más enriquecido con el anticuerpo específico contra CTCF comparado con el IgG inespecífico y con el control de perlas.

Por lo tanto, se validó la unión de CTCF en ocho diferentes regiones del locus del gen *Notch1* por ChIP-PCR. Además, la orientación de los motivos de unión para esos sitios, así como datos de Hi-C y ChIA-PET sugieren que el gen *Notch1* se localiza al interior de un asa de cromatina, la cual podría estar mediada por la unión de CTCF.

- **El abatimiento de CTCF afecta la expresión del gen *Notch1* en células K562**

La unión de CTCF al locus del gen *Notch1* así como datos de captura conformacional sugieren que podría regular su expresión a través de la formación de un asa de cromatina. Recientemente, diversos grupos han demostrado que CTCF puede regular la expresión de genes al inducir la formación y mantenimiento de asas de cromatina y en especial se ha visto que es importante para limitar la expresión de proto-oncogenes en leucemia de células T (Hnisz *et al.*, 2016). Para evaluar la contribución de CTCF en la expresión del gen *Notch1* se decidió abatir los niveles de CTCF empleando un ARN interferente mediante un vector lentiviral inducible bajo el tratamiento con doxiciclina (González-Buendía *et al.*, 2014).

Se evaluó la expresión del gen *Notch1* en células K562, para lo cual se diseñaron y estandarizaron dos pares de oligonucleótidos en diferentes secuencias del gen, un par diseñado en el exón tres del gen (Exón) y otro par en los exones ocho y nueve (Exón-Exón). Mediante RT-PCR se pudo confirmar la expresión del gen *Notch1* en células K562 (**Fig. 17A**). Se decidió utilizar el par de oligonucleótidos Exón-Exón para evaluar el cambio de expresión en los siguientes experimentos porque permitía diferenciar ARN de ADN genómico y mostró un producto de amplificación más enriquecido.

Para validar la disminución de la proteína en las células se realizó un Western Blot en células con el vector vacío y en células con el shRNAi, ambas con y sin el tratamiento con doxiciclina, la cual permite que se transcriba el shRNAi. En la **Fig. 17B** se muestra el Western Blot y se observó una disminución muy evidente de CTCF sólo en las células infectadas con el shRNAi y con el tratamiento con doxiciclina, lo cual confirmó el *knockdown*.

La expresión del gen *Notch1* fue evaluada en las células con el *knockdown* de CTCF mediante RT-qPCR usando los oligonucleótidos Exón-Exón, e inesperadamente se observó un incremento en la expresión del gen *Notch1*, sin embargo, este cambio no fue estadísticamente significativo (**Fig. 17C**). Este comportamiento se observó en dos experimentos de *knockdown* independientes.

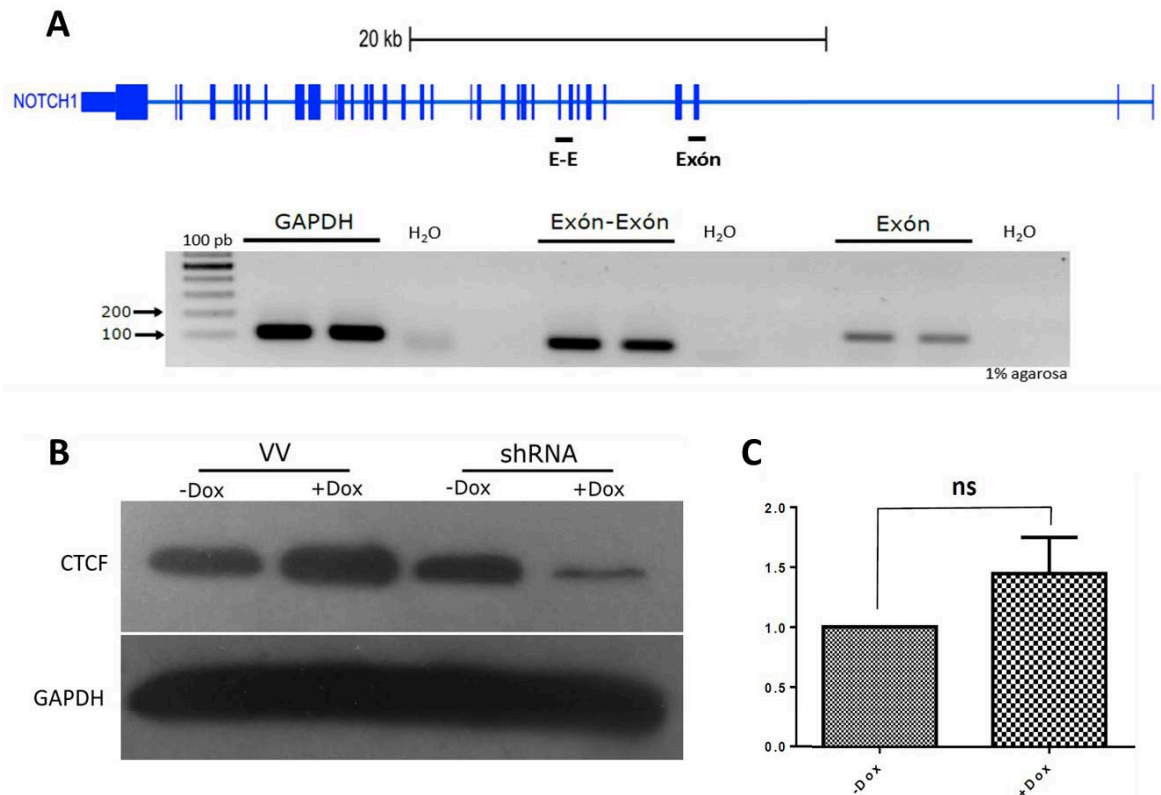


Fig. 17. El *knockdown* de CTCF afecta la expresión del gen *Notch1* en células K562. (A). Mapa del gen *Notch1* donde se muestran las regiones amplificadas por los oligonucleótidos utilizados y resultados de la expresión del gen *Notch1* por RT-PCR en células K562, *GAPDH* se utilizó como control endógeno. **(B).** Western Blot de células tratadas con vector vacío (VV) y con el shRNAi contra CTCF para inducir su *knockdown*, ambas con y sin tratamiento con doxiciclina. El control de carga utilizado fue *GAPDH*. **(C).** RT-qPCR que muestra los niveles de expresión relativa del gen *Notch1* normalizado con *GAPDH* en células K562 con el shRNAi contra CTCF con y sin doxiciclina. **ns:** no significativo.

Por lo tanto, la disminución de los niveles de CTCF mediante un shRNAi induce un incremento en la expresión del gen *Notch1* en células K562.

- **La eliminación de los sitios de unión a CTCF por CRISPR-Cas9 afecta la expresión del gen *Notch1* en células K562**

El *knockdown* contra CTCF en células K562 produce un incremento en la expresión del gen *Notch1*, sin embargo, el cambio no fue estadísticamente significativo.

Con base en los datos obtenidos y con el objetivo de evaluar de forma directa la contribución de esta proteína a la regulación transcripcional del gen *Notch1* se decidió realizar una eliminación precisa de los sitios de unión a CTCF en el locus del gen mediante la técnica de edición genómica CRISPR-Cas9. Se decidió eliminar los sitios de unión a CTCF que mostraron mayor enriquecimiento de señal en los ensayos de ChIP-seq. Estos sitios se localizan aproximadamente 80 kb río arriba del inicio de la transcripción del gen *Notch1* (**Fig. 18**).

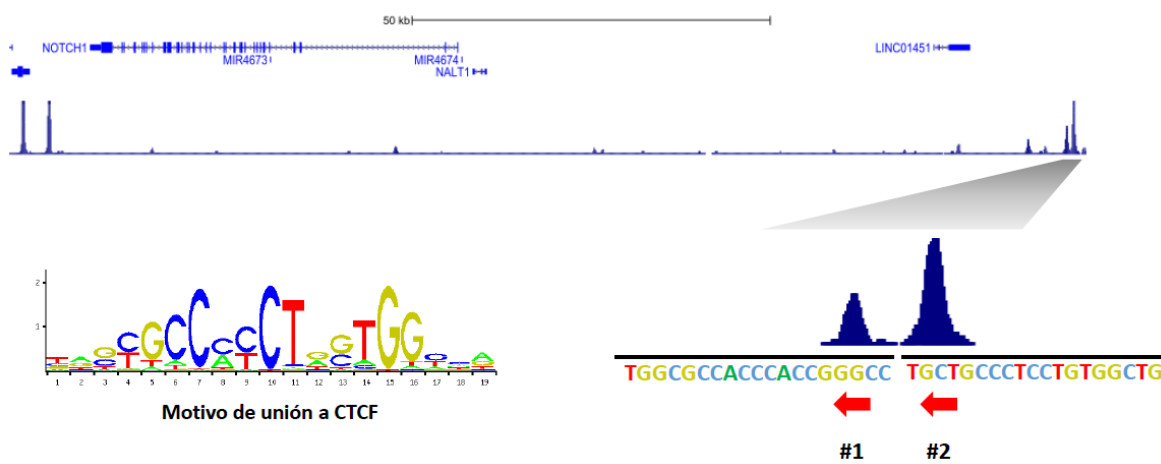


Fig. 18. Sitios de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1*. Los sitios con mayor enriquecimiento de señal observados en los datos de ChIP-seq muestran orientación antisentido. La secuencia de ambos sitios se obtuvo del genoma anotado Hg19 y también se muestra el motivo consenso de unión a CTCF en dicha orientación.

Para el sitio menos enriquecido río arriba del inicio de la transcripción de *Notch1*, el sitio #1, se diseñaron dos sgRNAs con los cuales se esperaba una eliminación de 119 pb (Ver métodos; **Fig. 19A**), y después de la selección y posterior dilución clonal se obtuvieron poblaciones monoclonales homocigotas y heterocigotas con el tamaño de producto que se esperaba luego de la eliminación (**Fig. 19B**).

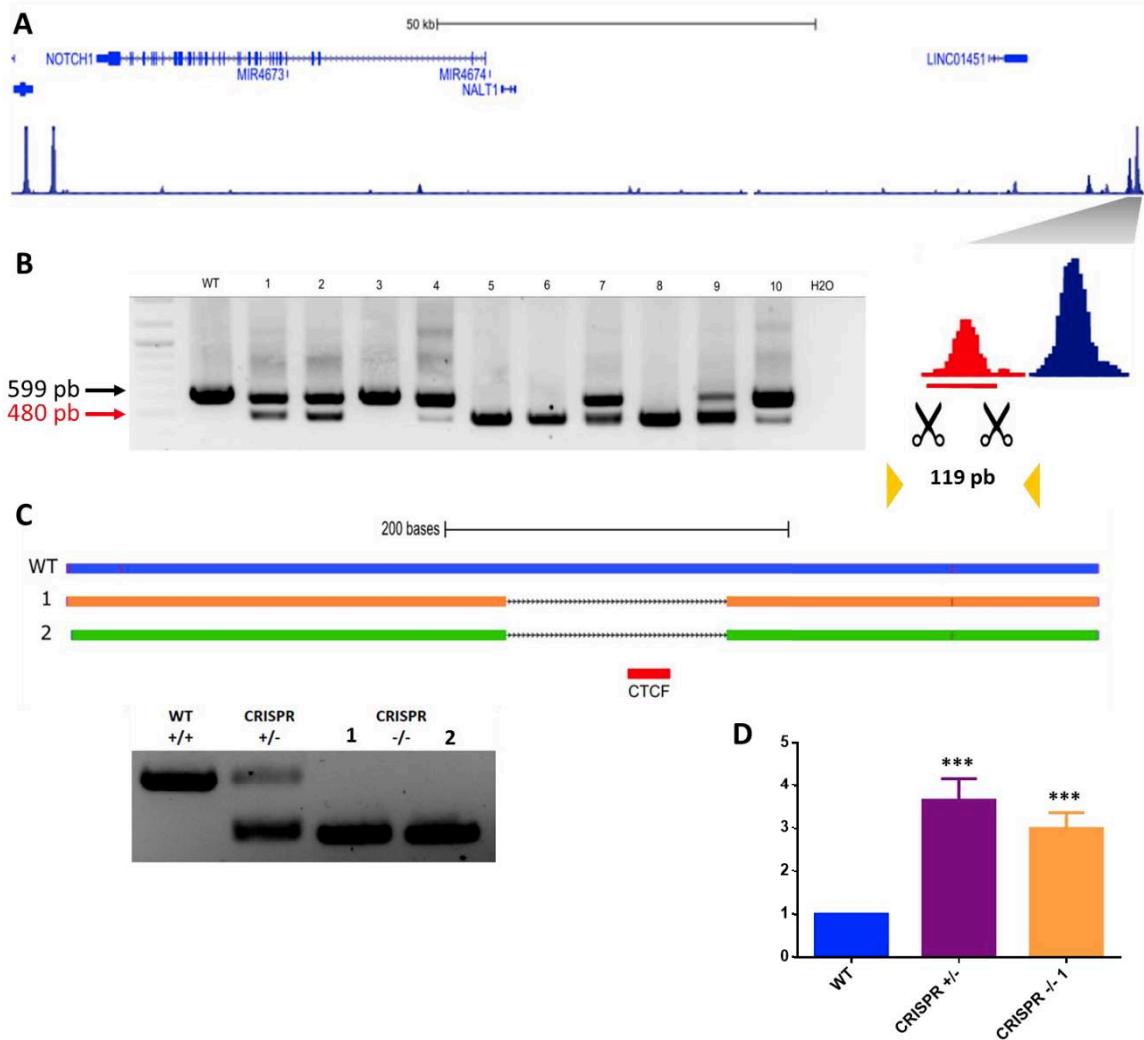


Fig. 19. La eliminación del sitio #1 por CRISPR-Cas9 afecta la expresión del gen *Notch1* en células K562. (A). Locus que comprende al gen *Notch1* con el sitio de unión a CTCF #1 río arriba del inicio de la transcripción del gen, la eliminación esperada mediante CRISPR-Cas9 fue de 119 pb. **(B).** Monoclonas obtenidas después de la selección y dilución clonal, el tamaño del producto de amplificación en las células silvestres (WT) fue de 599 pb y con la eliminación el tamaño fue de 480 pb, se observaron clonas heterocigotas como la 1, 2 y 7, y clonas homocigotas como la 5, 6 y 8. **(C).** Secuencias obtenidas de las células WT y de dos clonas homocigotas con la eliminación esperada, el sitio de unión a CTCF se eliminó exitosamente. **(D).** RT-qPCR que muestra los niveles de expresión relativa del gen *Notch1* en células K562 WT y en una clona heterocigota y dos homocigotas con la eliminación del sitio de CTCF. *** $P < 0.0007$.

Se eligieron para evaluar dos clonas homocigotas y una heterocigota, las cuales se secuenciaron para confirmar la naturaleza de la lesión molecular y se observó que se obtuvo la eliminación la cual remueve el motivo de unión a CTCF (**Fig. 19C**).

Se evaluó la expresión del gen *Notch1* en las células con la eliminación de CTCF y se observó que tanto en las células heterocigotas como en las homocigotas los niveles del gen *Notch1* se incrementan, presentando aproximadamente 2 veces más expresión del gen que en las células silvestres (**Fig. 19D**).

Adicionalmente, se diseñaron dos sgRNAs para eliminar el sitio más enriquecido río arriba del inicio de la transcripción del gen *Notch1*, el sitio **#2**, con los que se esperaba una eliminación de 160 pb (**Fig. 20A**). Sin embargo, después de la dilución clonal, para esta mutante sólo se obtuvieron poblaciones monoclonales heterocigotas que mostraron distintos tamaños de eliminaciones (**Fig. 20B**).

Por lo anterior, se eligió una clona heterocigota que mostró el tamaño de eliminación más cercano a lo esperado. Se secuenciaron los productos de amplificación para saber qué secuencias se habían eliminado; en las secuenciaciones se observó que en esa clona en un alelo no se eliminó el sitio de unión a CTCF, pero en el otro alelo sí se eliminó (**Fig. 20C**).

Al evaluar la expresión del gen *Notch1* en esas células, se observó una sobreexpresión, llegando a ser los niveles de expresión del gen *Notch1* casi 2 veces más que en las células sin la eliminación del sitio de unión a CTCF (**Fig. 20D**).

Estos datos sugieren que al eliminar cualquiera de los dos sitios enriquecidos de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1*, el gen tiene una sobreexpresión comparado con células sin la eliminación, lo cual sugiere que CTCF regula la expresión del gen *Notch1* en células K562.

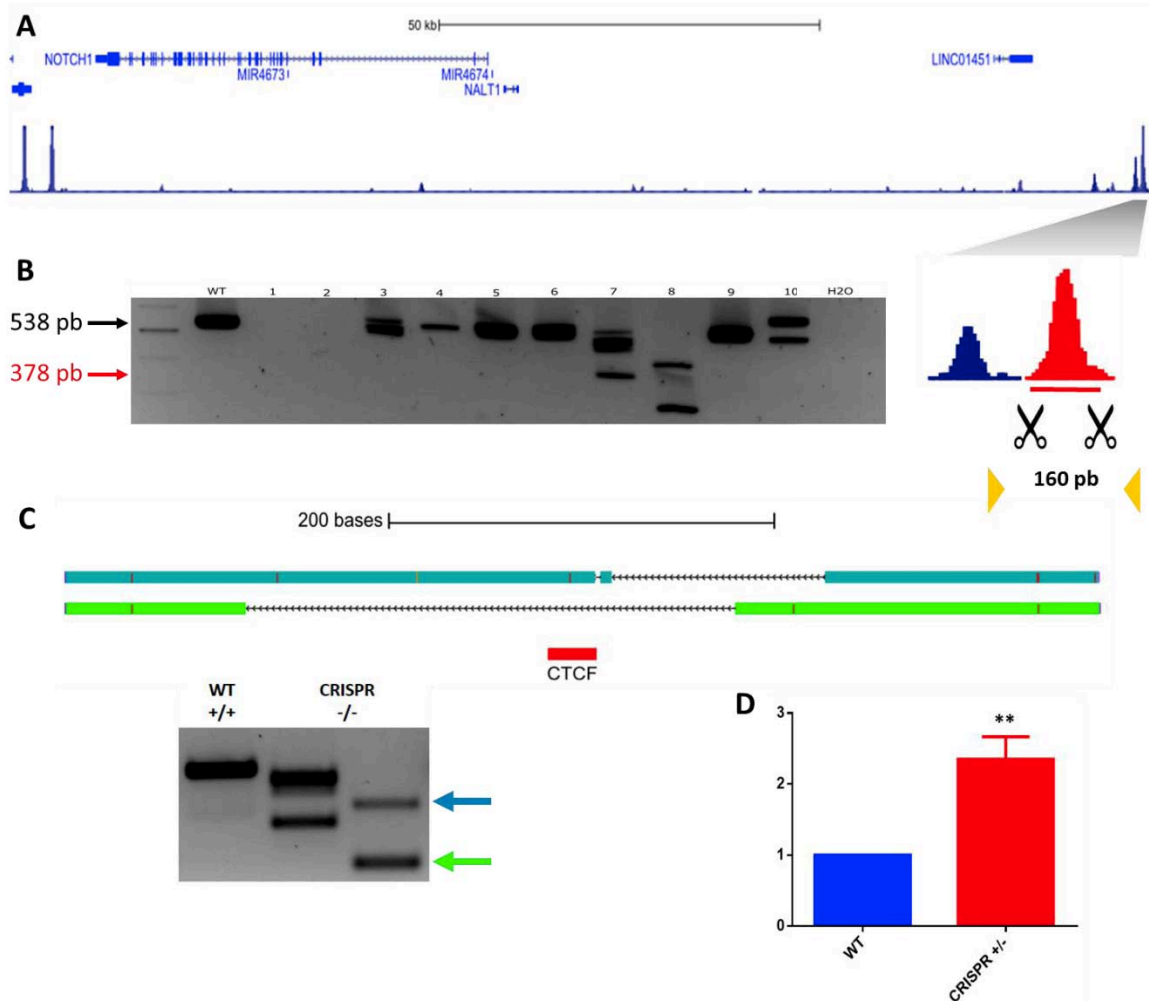


Fig. 20. La eliminación del sitio #2 por CRISPR-Cas9 afecta la expresión del gen *Notch1* en células K562. (A). Locus que comprende al gen *Notch1* con el sitio de unión a CTCF #2 río arriba del inicio de la transcripción del gen, la eliminación esperada mediante CRISPR-Cas9 fue de 160 pb. **(B).** Monoclonas obtenidas después de la selección y dilución clonal, el tamaño del producto de amplificación en las células silvestres (WT) fue de 538 pb y con la eliminación el tamaño fue de 378 pb, se observaron clonas con distintos tamaños de eliminaciones tanto homocigotas como heterocigotas. **(C).** Secuencias obtenidas de una clona heterocigota con un alelo sin la eliminación del sitio de unión a CTCF y otro alelo con la eliminación del sitio de CTCF. **(D).** RT-qPCR que muestra los niveles de expresión relativa del gen *Notch1* en células K562 WT y en la clona heterocigota secuenciada. **** P<0.0018.**

DISCUSIÓN

En la actualidad la organización tridimensional del genoma al interior del núcleo ha surgido como un regulador importante de la expresión de los genes, sin embargo, se desconoce cómo se forma y mantiene dicha organización, ya que se ha demostrado que existen múltiples niveles donde la expresión génica puede ser regulada por interacciones locales o por el ambiente provisto por el núcleo celular. Por lo cual, realizar este tipo de estudios resulta de gran relevancia para entender mejor el funcionamiento de los mecanismos de regulación transcripcional y epigenética, donde, las interacciones entre secuencias linealmente distantes y el aislamiento de regiones génicas mediante asas de cromatina resultan factores determinantes para la regulación de la expresión de genes.

En resumen, los datos experimentales y bioinformáticos obtenidos en este proyecto sugieren que CTCF está presente en el locus del gen *Notch1* participando en la formación de un asa de cromatina en esa región, tal como se determinó a partir de los datos de interacciones cromatínicas de Hi-C y ChIA-PET. La eliminación de un sitio de unión a CTCF en la región 3' del locus mediante CRISPR-Cas9 provocó una sobreexpresión del gen *Notch1*, sin embargo, la disminución global de la proteína CTCF, con un shRNAi, no afectó significativamente la expresión de *Notch1*.

Todos los sitios enriquecidos con CTCF en células K562 identificados por ensayos de ChIP-seq pudieron ser observados mediante inmunoprecipitación de la cromatina seguida por PCR de punto final realizada en el laboratorio confirmando la presencia de CTCF en los sitios esperados, y también se observó la expresión del gen *Notch1* en células K562, como lo indicaban datos de ENCODE de expresión de esta línea celular.

La línea celular K562 resultó ser un buen sistema para estudiar la regulación del gen *Notch1*, ya que este gen se expresa en dichas células y CTCF está unido en el locus, además de tener a la mano una gran cantidad de datos epigenómicos de esta línea celular, tales como datos de ChIP-seq y de interacciones de cromatina como Hi-C y ChIA-PET. En las células K562 el gen *Notch1* tiene una función interesante

en la supresión de tumores, con lo cual surge también un interés biomédico por estudiar este gen en esta línea celular.

El hecho de no observar un cambio significativo en la expresión del gen *Notch1* con el *knockdown* de CTCF fue algo interesante y no esperado, ya que se ha observado que la disminución global de CTCF ocasiona una ganancia de interacciones entre diferentes asas y una pérdida de interacciones dentro de asas, lo cual afecta la expresión de una gran cantidad de genes (Zuin *et al.*, 2014). Sin embargo, en otros estudios se ha observado que es difícil notar un cambio en un locus específico, ya que es necesario abatir completamente la unión de CTCF, lo cual no es posible lograrlo con un *knockdown*, el cual además puede presentar eficiencias variables (Hou *et al.*, 2010). Por lo cual, la mejor estrategia para determinar el papel estructural de CTCF en un locus específico es eliminar de forma precisa su sitio de unión.

Los resultados obtenidos de la eliminación mediante el sistema CRISPR-Cas9 de los sitios de unión a CTCF en la región río arriba del inicio de la transcripción del gen *Notch1* concuerdan con los datos presentados recientemente por el grupo de Richard Young (Hnisz *et al.*, 2016), ya que eliminar un sitio de unión a CTCF causó una sobreexpresión del gen localizado dentro de un asa de cromatina. Estas observaciones sugieren que CTCF podría estar funcionando como el delimitador y como componente estructural del asa de cromatina.

Sin embargo, ellos basaron su estudio en proto-oncogenes en la leucemia linfoblástica aguda (ALL) y se ha demostrado que para el caso de la línea celular K562 el gen *Notch1* actúa como un gen supresor de tumores, con lo cual resulta interesante que se observe el mismo comportamiento de los oncogenes en líneas celulares de ALL al perturbar la integridad del asa de cromatina que los contiene. Por lo anterior, se cree que el efecto de la pérdida de la arquitectura del genoma en la regulación transcripcional depende del contexto genético y epigenético en el que se encuentre el gen *Notch1*.

Experimentos de edición genómica mediante CRISPR-Cas9 han sido realizados en los últimos años para caracterizar las alteraciones en las interacciones cromatínicas

y en la formación y mantenimiento de asas de cromatina cuando se eliminan sus límites. Por ejemplo, la eliminación de un sitio de unión a CTCF en el locus de *Xist* humano, un lncRNA clave para la inactivación del cromosoma X, resultó en la fusión parcial de dos asas adyacentes y en una activación génica ectópica (Nora *et al.*, 2012). De igual manera la eliminación de un sitio de unión a CTCF asociado con el límite del locus *HoxA* en *Drosophila* resultó en un cambio en la posición del límite y una propagación de cromatina activa (Narendra *et al.*, 2015).

También se han observado fenotipos malignos causados por la eliminación de límites de asas de cromatina, como la disgenesia acropectorovertebral, la leucodistrofia desmielinizante y la displasia mesomélica, entre otros (Spielmann *et al.*, 2013; Lupiáñez *et al.*, 2015; Giorgio *et al.*, 2015).

Por otra parte, en un estudio reciente se determinó que los sitios de unión a CTCF se encuentran frecuentemente mutados en cáncer (Katainen *et al.*, 2015). Dichas mutaciones se han visto enriquecidas en sitios de CTCF presentes en los límites de asas de cromatina constitutivos (Hnisz *et al.*, 2016). Estas mutaciones pueden crear cambios en el motivo consenso de unión a CTCF, lo cual podría llegar a afectar tanto el nivel de enriquecimiento como la unión de CTCF en esos sitios, conduciendo a la inestabilidad de las asas.

En todos los casos mencionados anteriormente, la pérdida de las asas de cromatina provocó la formación de interacciones anormales y eso condujo a una expresión génica desregulada. Dichas interacciones anormales se dieron principalmente entre promotores y enhancers, lo cual promovió una activación o sobreexpresión de los genes blanco. Este comportamiento podría explicar la sobreexpresión del gen *Notch1* descrita al eliminar alguno de los sitios de unión a CTCF en la región 3' del locus, ya que río abajo de esos sitios y fuera del asa de cromatina que incluye al gen *Notch1* se identificó una región marcada como enhancer, dado que presenta marcas cromatínicas asociadas a éstos, y dicho enhancer en condiciones normales podría no tener contacto con el promotor del gen *Notch1*, pero al eliminar un sitio de CTCF el enhancer podría ganar acceso al promotor e inducir así su sobreexpresión (**Fig. 21**).

Cabe resaltar que al estudiar a CTCF es necesario diferenciar los sitios de unión a CTCF entre regulatorios y estructurales, ya que aproximadamente el 15% de los sitios de unión a CTCF se sitúan en los límites de asas cromatínicas, y el resto de los sitios se distribuyen dentro de las asas (Dixon *et al.*, 2012). Otro porcentaje se localiza cercano a promotores, en donde CTCF podría estar no sólo involucrado en la formación de contactos entre elementos regulatorios sino también en el reclutamiento de diversas proteínas que regulan la expresión génica (Ong y Corces, 2014).

El estudio presentado aquí resalta la importancia de las regiones no codificantes en la regulación transcripcional, ya que en estas regiones no sólo existen elementos regulatorios como promotores y enhancers, sino también existen sitios de unión a una variedad de factores de transcripción, los cuales junto a los elementos regulatorios componen un complejo sistema que regula la expresión génica (Spielmann y Mundlos, 2016).

Adicionalmente, los resultados presentados sugieren una importante relación entre la arquitectura del genoma y la expresión de los genes, dándole un papel clave a los límites de las asas de cromatina y a los elementos que participan en su formación, los cuales son proteínas arquitectónicas tales como CTCF.

Por último, dada la importancia funcional que se ha demostrado recientemente de la organización tridimensional del genoma en la regulación transcripcional y que se ha observado que varias patologías se dan por la pérdida de la arquitectura del genoma, se ha sugerido que se debería considerar a la organización tridimensional del genoma con fines de diagnóstico y a largo plazo se podría pensar en su participación en aspectos terapéuticos (Hnisz *et al.*, 2016; Lupiáñez *et al.*, 2016).

Además, es importante tener en cuenta que patrones de expresión diferencial de los ligandos y receptores, como Notch1, no han sido suficientes para explicar las diferencias observadas en la actividad de la vía de señalización de Notch (Kopan e Ilagan, 2009; Bray, 2006), aunque existen pocos estudios al respecto, por lo que sería interesante demostrar una correlación entre la sobreexpresión del gen *Notch1* y la actividad de la vía de Notch.

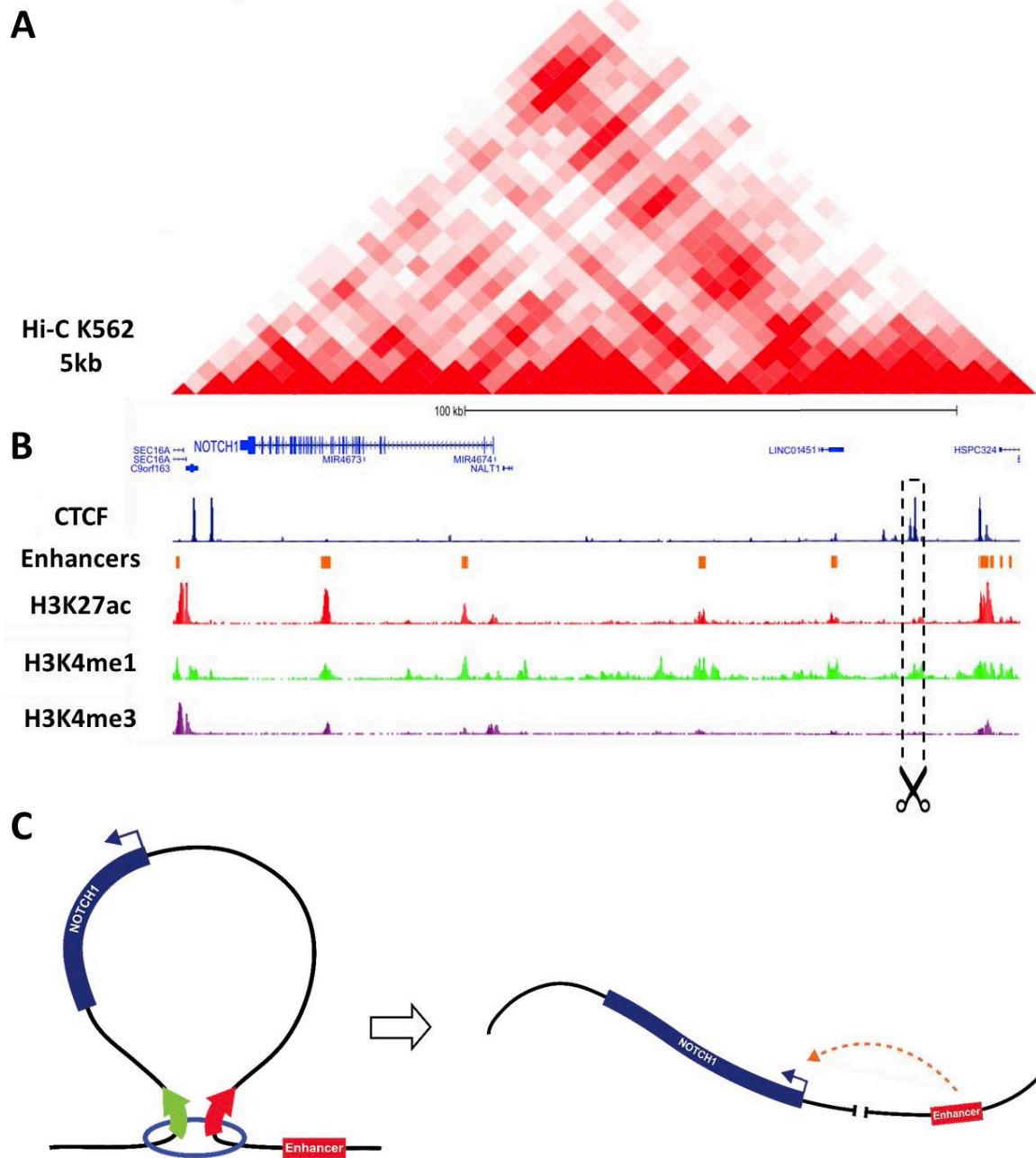


Fig. 21. Modelo propuesto del asa de cromatina que incorpora al gen *Notch1* y el efecto de la eliminación de un sitio de unión a CTCF. (A). Mapa de interacciones de Hi-C en células K562 a 5 kb de resolución obtenidos por Rao *et al.*, 2014. **(B).** Locus del gen *Notch1* con los sitios de enriquecimiento de CTCF, H3K27ac, H3K4me1 y H3K4me3 obtenidos por ENCODE mediante ensayos de ChIP-seq y regiones marcadas como enhancers por sus marcas de cromatina en células K562. **(C).** Modelo del asa de cromatina presente en el locus del gen *Notch1* con CTCF en sus límites y al eliminar un sitio de unión a CTCF el asa de cromatina se pierde y el enhancer en la región 3' del locus puede interactuar con el gen *Notch1*.

CONCLUSIONES

- CTCF se une en la región 5' y 3' del locus del gen *Notch1*.
- El gen *Notch1* se expresa en la línea celular de leucemia mieloide crónica K562.
- El *knockdown* de CTCF no afecta significativamente la expresión del gen *Notch1*, sin embargo se observa una tendencia hacia la sobreexpresión.
- La eliminación precisa de los sitios de unión a CTCF río arriba del inicio de la transcripción del gen *Notch1* mediante CRISPR-Cas9 induce una sobreexpresión significativa del gen.

PERSPECTIVAS

- ❖ Eliminar los sitios de unión a CTCF en la región 5' del locus de *Notch1* y determinar su efecto en la expresión del gen.
- ❖ Confirmar la pérdida de la unión de CTCF mediante ChIP en los sitios eliminados con CRISPR-Cas9.
- ❖ Realizar 3C- para determinar las interacciones presentes en el locus de *Notch1* en condiciones normales y después de la eliminación de los sitios de unión a CTCF.
- ❖ Determinar la actividad de la vía de Notch analizando la expresión de genes blanco, como *Hes1*, después de la eliminación de los sitios de unión a CTCF.
- ❖ Evaluar cambios fenotípicos, tales como proliferación y muerte celular en las células con la edición genómica.

REFERENCIAS

- Ali T, Rainer Renkawitz, Marek Bartkuhn. 2016. Insulators and domains of gene expression. *Curr Opin Genet Dev.* (37): 17-26.
- Allis D. and Jenuwein T. 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics* (17): 487–500.
- Allis, C. D., Caparros, M., Jenuwein, T. & Reinberg, D. 2015. Epigenetics. *Cold Spring Harbor.*
- Artavanis-Tsakonas S, Muskavitch MA, Yedvobnick B. 1983. Molecular cloning of Notch, a locus affecting neurogenesis in *Drosophila melanogaster*. *PNAS* 80(7):1977-81.
- Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. 1999. Notch signaling: cell fate control and signal integration in development. *Science* 284, 770–776.
- Ayala-Ortega E, Arzate-Mejía R, Pérez-Molina R, González-Buendía E, Meier K, Guerrero G, Recillas-Targa F. 2016. Epigenetic silencing of miR-181c by DNA methylation in glioblastoma cell lines. *BMC Cancer.* (16): 16:226.
- Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98: 387–396.
- Bickmore WA. 2013. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet.* (14): 67-84.
- Boyle Matthew J., Heather A., Jeffrey A., and Wendy A. Bickmore. 2011. Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res.* 19(7): 901–909.
- Bray SJ. 2006. Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* 7: 678–689.
- Cavalli G, Misteli T. 2013. Functional implications of genome topology. *Nat Struct Mol Biol.* 20(3): 290-9.
- Chen H, Tian Y, Shu W, Bo X, Wang S. 2012. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS One* 7: e41374.
- Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB et al. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133: 1106–1117.
- Cremer T, Cremer M. 2010. Chromosome territories. *Cold Spring Harb Perspect Biol.* 2(3): a003889.
- Cubeñas-Potts, Corcés VG. 2015. Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Letter.* <http://dx.doi.org/10.1016/j.febslet.2015.05.025>
- De La Rosa-Velazquez, I.A., Rincon-Arano, H., Benitez-Bribiesca, L., Recillas-Targa, F. 2007. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res* 67: 2577–2585.
- de Laat W, Duboule D. 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature.* 502(7472): 499-506.

- Dekker J, Martí-Renom MA y Mirny LA. 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet.* 14(6): 390-403.
- Dekker J, Misteli T. 2015. Long-Range Chromatin Interactions. *Cold Spring Harb Perspect Biol.* 7(10): a019356.
- Dixon JR, Gorkin DU, Ren B. 2016. Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell.* 62(5): 668-80.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485: 376–380.
- Dunham I. et al. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 489(7414): 57-74.
- Durinck K, Wallaert A, Van de Walle, Van Loocke, Volders, et al. 2014. The Notch driven long non-coding RNA repertoire in T-cell acute lymphoblastic leukemia. *Haematologica* 99(12): 1808-1816.
- Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS. 2004. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* 303: 238–240.
- Felsenfeld G, Groudine M. 2003. Controlling the double helix. *Nature* 421: 448-453.
- Filippova G, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G et al. 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. *Mol Cell Biol* 16:2802–2813.
- Garg, V., Muth, A.N., Ransom, J.F., Schluterman, M.K., Barnes, R., King, I.N., Grossfeld, P.D., and Srivastava, D. 2005. Mutations in NOTCH1 cause aortic valve disease. *Nature* 21, 180–184.
- Ghirlando R, Felsenfeld G. 2015. CTCF: making the right connections. *Genes Dev.* 30(8): 881-91.
- Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol Cell.* 49(5): 773-82.
- Giorgio, E. et al. 2015. A large genomic deletion leads to enhancer adoption by the lamin B1 gene: a second path to autosomal dominant adult-onset demyelinating leukodystrophy (ADLD). *Hum. Mol. Genet.* (24): 3143–3154.
- Goh Y, Fullwood MJ, Poh HM, Peh SQ, Ong CT, Zhang J, Ruan X, Ruan Y. 2012. Chromatin Interaction Analysis with Paired-End Tag Sequencing (ChIA-PET) for mapping chromatin interactions and understanding transcription regulation. *J Vis Exp.* (62): pii: 3770.
- González-Buendía E, Pérez-Molina R, Ayala-Ortega E, Guerrero G, Recillas-Targa F. 2014. Experimental strategies to manipulate the cellular levels of the multifunctional factor CTCF. *Methods Mol Biol.* (2014): 1165:53-69.
- Gorkin, Danny Leung and Ren B. 2014. The 3D Genome in Transcriptional Regulation and Pluripotency. *Cell Stem Cell.* 14(6): 762–775.
- Groschel, S. et al. 2014. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* (157): 369–381.
- Hnisz D et al. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science.* 351(6280): 1454-8.

- Hou C, Dale R, Dean A. 2010. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *PNAS*. 107: 3651-3656.
- Hou, C., Li, L., Qin, Z.S., and Corces, V.G. 2012. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell* (48): 471–484.
- Kadonaga JT. 2004. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* 116: 247-257.
- Katainen R, Dave K, Pitkänen E. et al. 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat Genet*. 47(7): 818-21.
- Kim S, Nam-Kyung Y, Bong-Kiun K. 2015. CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental and Molecular Medicine*. 47(6): e166.
- Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH et al. 1993. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol* 13: 7612–7624.
- Kopan R, Ilagan M. 2009. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* 137: 216–233.
- Le Borgne, R., Bardin, A. & Schweisguth, F. 2005. The roles of receptor and ligand endocytosis in regulating Notch signaling. *Development* 132, 1751–1762.
- Li W, Notani D, Rosenfeld MG. 2016. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet*. 17(4): 207-23.
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragooczy T, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326: 289–293.
- Louvi, A. & Artavanis-Tsakonas, S. 2006. Notch signalling in vertebrate neural development. *Nature Rev. Neurosci*. 7, 93–102.
- Lupianez, D.G. et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene–enhancer interactions. *Cell* (161): 1012–1025.
- Lupiáñez, Spielmann and Stefan Mundlos. 2016. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet*. 32(4): 225-37.
- Merkenschlager M, Nora EP. 2016. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu Rev Genomics Hum Genet*. (17) 10.1146.
- Micchelli, C. A. & Perrimon, N. 2006. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* 439, 475–479.
- Moon H, Filippova G, Loukinov D, Pugacheva E, Chen Q, Smith ST et al. 2005. CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep* 6: 165–170.
- Nakahashi, H. et al. 2013. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep* 3: 1678–1689.
- Narendra, V. et al. 2015. Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*. (347): 1017–1021.

- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* (485): 381–385.
- Northcott, P.A. et al. 2014. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* (511): 428–434.
- Ong C y Corcés VG. 2014. CTCF: an architectural protein bridging genome topology and function. *Nature Reviews genetics* 15: 234-246.
- Padeken J, Heun P. 2014. Nucleolus and nuclear periphery: velcro for heterochromatin. *Curr Opin Cell Biol.* 28: 54-60.
- Papantonis A, Cook PR. 2013. Transcription factories: genome organization and gene regulation. *Chem Rev.* 113(11): 8683-705.
- Phillips JE y Corces VG. 2009. CTCF: master weaver of the genome. *Cell* 137: 1194-1211.
- Pombo A, Dillon N. 2015. Three-dimensional genome architecture: players and mechanisms. *Nature Mol Cell Bio* 16: 245-257.
- Puente XS et al. 2015. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature.* 526(7574):519-24.
- Radtke, F. & Raj, K. 2003. The role of Notch in tumorigenesis: oncogene or tumour suppressor? *Nature Rev Cancer* (3): 756–767.
- Ran F, Hsu, Wright, Agarwala, Scott and Feng Zhang. 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols.* 8(11): 2281-2308.
- Rao, S.S.P., Huntley, M.H., Durand, N.C, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* (159): 1665–1680.
- Recillas-Targa F, Escamilla del Arenal M. 2004. Participación de la estructura de la cromatina en la regulación de la expresión génica. *Mensaje Bioquímico* Vol. XXVIII
- Recillas-Targa F, Razin S. 2001. Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. *Crit Rev Eukaryot Gene Expr* 11:227-242.
- Rincón-Arano H y Recillas Targa F. 2008. La arquitectura nuclear y su dinámica. *Rev Edu Bioquim.* (4): 130-137.
- Rizzo P, Osipo C, Foreman K, Golde T, Osborne B, et al. 2008. Rational targeting of Notch signaling in cancer. *Oncogene* 27: 5124–5131.
- Roy M, Pear WS, Aster JC. 2007. The multifaceted role of Notch in cancer. *Curr Opin Genet Dev* 17: 52–59.
- Schwanbeck R. and Just U. 2011. The Notch signaling pathway in hematopoiesis and hematologic malignancies. *Haematologica.* 96(12): 1735–1737.
- Schweisguth, F. 2004. Notch signaling activity. *Curr. Biol.* 14, R129–R138.
- Sexton T, Cavalli G. 2015. The role of chromosome domains in shaping the functional genome. *Cell.* 160(6): 1049-59.

- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. 2012. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* (148): 458–472.
- Soto-Reyes E, Recillas-Targa F. 2010. Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene* 29: 2217-2227.
- South, A.P., Cho, R.J., and Aster, J.C. 2012. The double-edged sword of Notch signaling in cancer. *Semin. Cell Dev Biol.* 23, 458–464.
- Spielmann M, Mundlos S. 2016. Looking beyond the genes: the role of non-coding variants in human disease. *Hum Mol Genet.* pii: ddw205.
- Spielmann M. and Mundlos S. 2013. Structural variations, the regulatory landscape of the genome and their alteration in human disease. *Bioessays.* (35): 533–543.
- Splinter E, Heath H, Kooren J, Palstra R-J, Klous P, Grosveld F, et al. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20: 2349–2354.
- Suresh S, Irvine AE. 2015. The NOTCH signaling pathway in normal and malignant blood cell production. *J Cell Commun Signal.* 9(1):5-13.
- Torrano V, Chernukhin I, Docquier F, D'Arcy V, Leon J, Klenova E y Delgado MD. 2005. CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J Biol Chem* 280: 28152–28161.
- van Driel R, Fransz PF, Verschurre PJ. 2003. The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci* 116: 4067-4075.
- Vietri Rudan y Hadjur S. 2015. Genetic Tailors: CTCF and Cohesin Shape the Genome During Evolution. *Trends Genet.* 31(11): 651-60.
- Vostrov AA, Taheny MJ, Quitschke WW. 2002. A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter. *J Biol Chem* 277: 1619–1627.
- Weng, A. P. et al. 2004. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306, 269–271.
- West AG, Gaszner M, Felsenfeld G. 2002. Insulators: many functions, many mechanisms. *Genes Dev* 16: 271–288.
- Wright, G. J., Leslie, J. D., Ariza-McNaughton, L. & Lewis, J. 2004. Delta proteins and MAGI proteins: an interaction of Notch ligands with intracellular scaffolding molecules and its significance for zebrafish development. *Development* 131, 5659–5669.
- Yatim A, Benne C, Sobhian B, Laurent-Chabalier S, Deas O, Judde JG, Lelievre JD, Levy Y, Benkirane M. 2012. NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. *Mol Cell.* 48(3):445-58.
- Zuin, J. et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.* (111): 996–1001.

INFORMACIÓN SUPLEMENTARIA

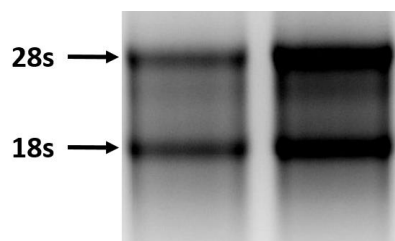


Fig. S1. Gel de agarosa para observar la integridad del ARN extraído, se muestran las subunidades ribosomales 28S y 18S.

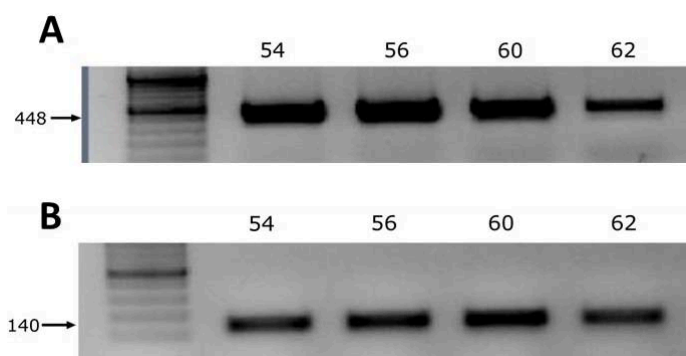


Fig. S2. Estandarización de oligonucleótidos para el ARNm del gen *Notch1*. **(A).** Oligonucleótidos del Exón 3. **(B).** Oligonucleótidos del Exón-Exón 8 y 9.

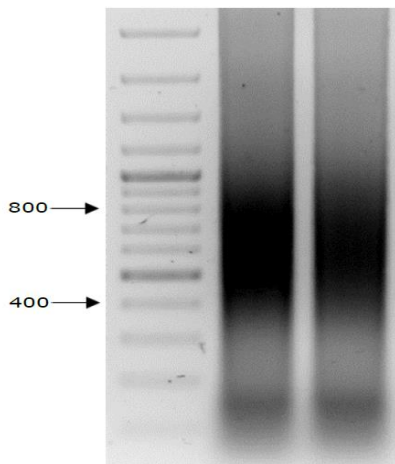


Fig. S3. Tamaño de los fragmentos de cromatina después de 12 pulsos de sonicación, se encuentran entre 400 y 800 pb.

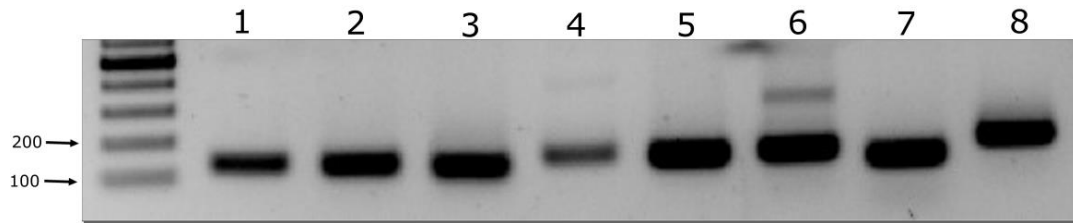


Fig. S4. Estandarización de los oligonucleótidos que amplifican los sitios de unión a CTCF observados por ChIP-seq en células K562.

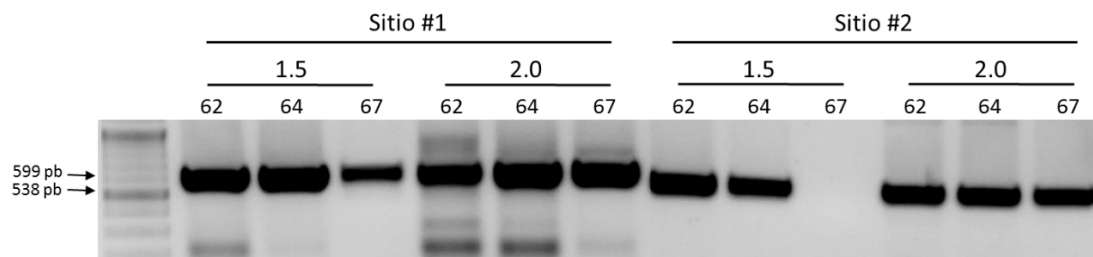


Fig. S5. Estandarización de los oligonucleótidos utilizados para observar las eliminaciones en los sitios de unión a CTCF río arriba del gen *Notch1*.

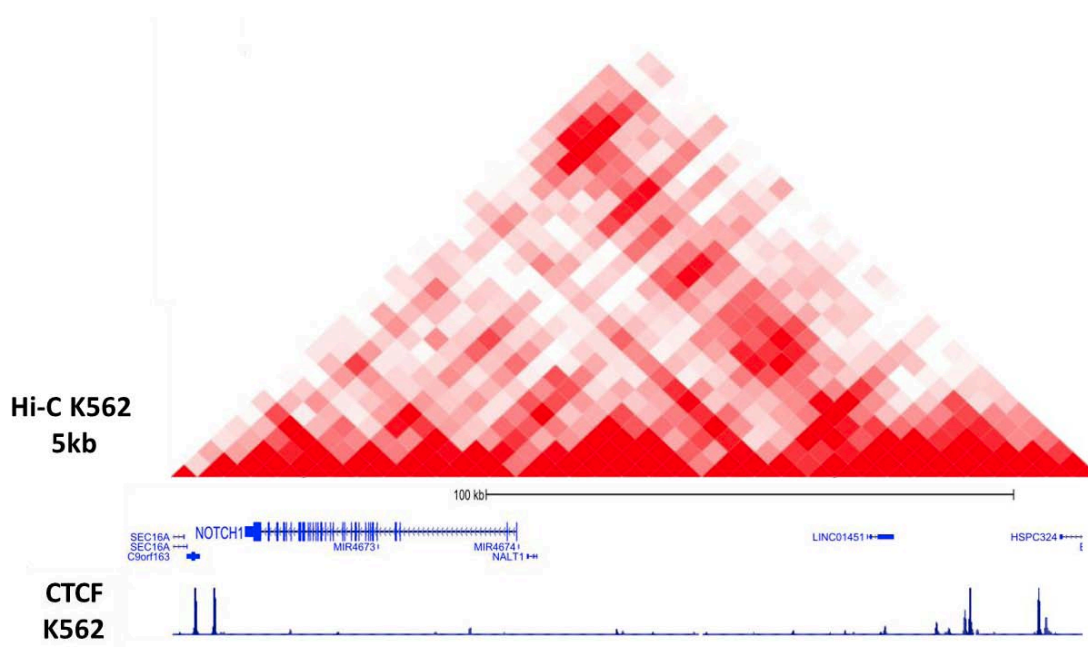


Fig. S6. Mapa de interacciones de Hi-C en células K562 a una resolución de 5 kb y los datos de ChIP-seq de CTCF muestran las mismas interacciones observadas en células GM12878 (Rao *et al.*, 2014).

Línea Celular	Factor de transcripción o Marca de histonas	Referencia
K562	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
	H3K27ac	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&c=chr9&q=wgEncodeRegMarkH 3k27ac
	H3K4me1	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&c=chr9&q=wgEncodeRegMarkH 3k4me1
	H3K4me3	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&c=chr9&q=wgEncodeRegMarkH 3k4me3
GM12878	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
H1-hESC	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
HeLa	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
HUVEC	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
HepG2	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html
NHEK	CTCF	https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=505581401_4etJVDrXdmiwKF D5zd21gH2V1Bqq&g=hub_4607_uniformTfbs#TRAC HTML http://genome.ucsc.edu/ENCODE/downloads.html

Tabla S1. Datos de ChIP-seq utilizados obtenidos por ENCODE y visualizados en el UCSC Genome Browser.

Método	Oligonucleótidos 5' -> 3'	Tamaño del producto	Condiciones de amplificación
qPCR Notch1 E-E	F: TGGACCAGATTGGGGAGTTC R: GCACACTCGTCTGTGTTGAC	82 pb	58 °C 1.5 mM MgCl ₂
qPCR Notch1 Exón	F: GAGGCGTGGCAGACTATGC R: CTTGTA CTCCGTCAGCGTGA	140 pb	58 °C 1.5 mM MgCl ₂
qPCR GAPDH	F: CCACTCCTCCACCTTTGAC R: ACCCTGTTGCTGTAGCCA	102 pb	58 °C 1.5 mM MgCl ₂
ChIP #1	F: CAGGGCAGGGATGTTAGACA R: CGTGCACTCAGGATTAGGGA	118 pb	56 °C 1.5 mM MgCl ₂
ChIP #2	F: CATTCTGCCCTGTGGTCAC R: AGCGTGACCTGTTACTTTG	111 pb	62 °C 1.5 mM MgCl ₂
ChIP #3	F: AACGGGGACACTGGGCTC R: GACCTGCCTCTCACCTG	102 pb	60 °C 1.5 mM MgCl ₂
ChIP #4	F: CCAGCGTCTTCCAGCAAAG R: CCAGGGCATTGAGGGCTA	128 pb	65 °C 2.0 mM MgCl ₂
ChIP #5	F: ACACACACCACACAAGCATG R: CATGGAGTTGTCAGCAGCC	119 pb	56 °C 1.5 mM MgCl ₂
ChIP #6	F: AGAGCCTTTGGAGAGAAGCC R: CGGACCAGTGGCCTCTAC	125 pb	62 °C 2.0 mM MgCl ₂
ChIP #7	F: AGAGATGGATGCAAAGCTCA R: GGATGACCGTGAGAAGCAC	100 pb	62 °C 2.0 mM MgCl ₂
ChIP #8	F: TCTCCTGGAAGCCCGAGAT R: CAATATGACTGGCGTCCTTGT	150 pb	60 °C 2.5 mM MgCl ₂
CRISPR #1	F: TAGGACCACTCGCTGCCT R: GAGCGTGGATGGGTGTGT	599 pb	64 °C 1.5 mM MgCl ₂
CRISPR #2	F: ACATACGCACGCAGGCAT R: CTTGATGGGGTCTGGAGCTC	538 pb	64 °C 1.5 mM MgCl ₂

Tabla S2. Secuencias de oligonucleótidos utilizados en cada uno de los métodos.

Sitio de interés	sgRNAs 5' -> 3'
CRISPR #1	sgRNA 1: AGGGACTGTTGGGGTGAAT sgRNA 2: TCAATGCCCTGGGGTGCTCA
CRISPR #2	sgRNA 1: CACAAGCATGCTCCACACA sgRNA 2: GAGGGCCACCATGGATGTCT

Tabla S3. Secuencias de los sgRNAs utilizados para las eliminaciones con CRISPR-Cas9.

Sitios de unión a CTCF	Coordenadas	Secuencia del motivo de unión a CTCF 5' -> 3'	Valor JASPAR	Orientación del motivo
#1	chr9:139379581-139379730	TCACCTGCAGGGGGCGCGG	0.899	Forward
#2	chr9:139383,201-139383350	GTGCCACCAGGGGGAGGCA	0.916	Forward
#3	chr9:139519601-139519750	ATTCCTGCAGATGGAGCTC	0.839	Reverse
#4	chr9:139524961-139525110	GGCCCGGTGGGTGGCGCCA	0.823	Reverse
#5	chr9:139,525,941-139,526,090	CAGCCACAGGAGGGCAGCA	0.879	Reverse
#6	chr9:139,539,041-139,539,190	CAGCCGGCAGATGGGGCGG	0.814	Reverse
#7	chr9:139,509,841-139,509,990	TGGCCATGTGAGGGCGCCT	0.807	Reverse
#8	chr9:139,521,921-139,522,070	CCTCCACCAGGCGGGGATG	0.797	Reverse

Tabla S4. Sitios de unión a CTCF observados por ChIP-seq con sus coordenadas y la secuencia del motivo de unión a CTCF junto a su score y orientación obtenidos por JASPAR.

ANEXOS

I. Soluciones y Reactivos

H ₂ O MiliQ	Tritón x-100 1X	Leupeptina 1 mg/mL
Formaldehído 37%	Glicina 2.5 M	Aprotinina 10 mg/mL
NaCl 5 M	NaHCO ₃ 1M	Pepstaina 1 mg/mL
EGTA 0.5 M	PBS 1X	Fenol-cloroformo
Hepes 0.5 M pH: 8	Proteína G/A	Cloroformo
SDS 10%	RNasa A	Etanol 100%
EDTA 0.5 M pH: 8	Proteinasa K	Etanol 75%
Tris-HCl 1 M pH: 8	PMSF	Agua inyectable

II. Buffers

- Mix I (10 mL)

Solución o Reactivo	Volumen	Concentración final
Tris-HCl 1 M	2 mL	200 mM
NaCl 5 M	800 µL	400 mM
H ₂ O	7.2 mL	-

- Mix II (10 mL)

Solución o Reactivo	Volumen	Concentración final
SDS 10%	400 µL	0.4 %
EDTA 0.5 M	200 µL	10 mM
H ₂ O	9.4 mL	-

- Buffer de “cross-linking” (10 mL)

Solución o Reactivo	Volumen	Concentración final
Formaldehído 37%	2.97 mL	11%
NaCl 5M	200 µL	100 mM
EGTA 0.5 M	10 µL	0.5 mM
Hepes 0.5 M	1 mL	50 mM
H ₂ O	5.82 mL	-

***Preparar justo antes de utilizar**

- Buffer de lisis (125 mL)

Solución o Reactivo	Volumen	Concentración final
SDS 10%	12.5 mL	1%
EDTA 0.5 M	2.5 mL	10 mM
Tris-HCl 1 M	6.25 mL	50 mM
H ₂ O	103.75 mL	-
Inhibidores de proteasas (por cada 5 mL de buffer de lisis)	-1 µL Leupeptina -0.1 µL Aprotinina -5 µL Pepstaina -50 µL PMSF	-

***Tomar una alícuota y agregar inhibidores justo antes de utilizar.**

- Buffer de dilución (125 mL)

Solución o Reactivo	Volumen	Concentración final
Tritón x-100 1X	1.25 mL	1%
EDTA 0.5 M	500 µL	2 mM
Tris-HCl 1 M	2.5 mL	20 mM
NaCl 5 M	3.75 mL	150 mM
H ₂ O	117 mL	-
PMSF	1% del volumen a utilizar	-

***Tomar una alícuota y agregar PMSF justo antes de utilizar.**

- Buffer de paro I (125 mL)

Solución o Reactivo	Volumen	Concentración final
SDS 10%	1.25 mL	0.1%
Tritón x-100 1X	1.25 mL	1%
EDTA 0.5 M	500 µL	2 mM
Tris-HCl 1 M	2.5 mL	20 mM
NaCl 5 M	3.75 mL	150 mM
H ₂ O	115.75 mL	-
PMSF	1% del volumen a utilizar	-

***Tomar una alícuota y agregar PMSF justo antes de utilizar.**

- Buffer de paro II (125 mL)

Solución o Reactivo	Volumen	Concentración final
SDS 10%	1.25 mL	0.1%
Tritón x-100 1X	1.25 mL	1%
EDTA 0.5 M	500 µL	2 mM
Tris-HCl 1 M	2.5 mL	20 mM
NaCl 5 M	12.5 mL	500 mM
H ₂ O	107 mL	-
PMSF	1% del volumen a utilizar	-

***Tomar una alícuota y agregar PMSF justo antes de utilizar.**

- Buffer de “descrosslinking” (120 µL p/10 muestras)

Solución o Reactivo	Volumen
NaCl 5 M	45 µL
Tris-HCl 1 M	45 µL
EDTA 0.5 M	22.5 µL
RNasa A	7.5 µL

***Preparar justo antes de utilizar**

- Buffer de elución (100 mL)

Solución o Reactivo	Volumen	Concentración final
SDS 10%	10 mL	1%
NaHCO ₃ 1 M	10 mL	100 mM
H ₂ O	80 mL	-