



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

ESTUDIO DEL SILENCIAMIENTO EPIGENÉTICO  
Y SU BLOQUEO POR “*INSULATORS*”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

EDGAR FERNÁN GONZÁLEZ BUENDÍA

TUTOR PRINCIPAL

Dr. FÉLIX RECILLAS TARGA  
Instituto de Fisiología Celular

MIEMBROS DEL COMITÉ TUTOR

Dr. Carlos Rosales Ledezma  
Instituto de Investigaciones Biomédicas

Dr. Felipe Cruz García  
Facultad de Química

MÉXICO, D. F. Septiembre, 2015



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

## Reconocimientos académicos

El presente trabajo fue elaborado en el Departamento de Genética Molecular del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México en el laboratorio dirigido por el Dr. Félix Recillas Targa, con los apoyos de la Dirección General de Asuntos del Personal Académico UNAM (DGAPA-PAPIIT) IN209403, IN203811 y IN201114 y del Consejo Nacional de Ciencia y Tecnología (CONACyT) 42653-Q, 128464 y 220503.

Los estudios de doctorado de Edgar Fernán González Buendía, con número de CVU 207989, se realizaron con el apoyo de la Beca Nacional CONACYT para posgrados y con el Programa de Apoyo a los Estudios del Posgrado (PAEP).

Se extiende un sincero agradecimiento al comité tutor de este proyecto, constituido por:

Dr. Félix Recillas Targa	Instituto de Fisiología Celular
Dr. Carlos Rosales Ledezma	Instituto de Investigaciones Biomédicas
Dr. Felipe Cruz García	Facultad de Química

Se agradece el excelente apoyo académico y técnico de la Dra. Karin Meier, de la Biol. Georgina Guerrero Avendaño y el Dr. Fernando Suaste Olmos.

Asimismo, se reconoce el trabajo de la laboratorista Gianelli Cortés González, LCC. Gabriela Valdés Silva del departamento de Genética Molecular, el apoyo del Ing. Armando Canto Canto, Jefe del Departamento de Servicios Generales IFC-UNAM, de la Dra. Laura Ongay Larios, M. en C. Minerva Mora Cabrera, Biol. Ma. Guadalupe Códiz Huerta de la Unidad de Biología Molecular IFC-UNAM. El apoyo del Ing. Juan Manuel Barbosa Castillo e Ing. Ivett Rosas Arciniega de la Unidad de Cómputo IFC-UNAM.

Adicionalmente, se reconoce la asesoría académica del Dr. Roberto Coria Ortega, el apoyo de M. en C. Norma Trejo Medina y de la Sra. Leticia García Gutiérrez de la Coordinación del Programa de Posgrado Maestría y Doctorado en Ciencias Bioquímicas de la UNAM.

Finalmente, se agradece a los miembros del jurado, quienes enriquecieron este trabajo:

Dra. Victoria Chagoya de Sánchez	IFC	PRESIDENTE
Dr. Emilio Rojas del Castillo	IIB	SECRETARIO
Dra. Viviana del Carmen Valadez Graham	IBT	VOCAL
Dr. Diego González Halphen	IFC	VOCAL
Dra. Leonor Pérez Martínez	IBT	VOCAL

## ABREVIATURAS

ADN	Ácido Desoxirribonucleico
ARN	Ácido Ribonucleico
<i>FOLR1</i>	Gen Receptor de Folato 1
HSB-HSA'	Sitio de Hipersensibilidad a la DNasa I B - A'
CTCF	Factor de transcripción CCCTC
TFs	Factores Transcripcionales
PEV	Efecto de Posición Variegado
PcG	Complejo proteico represor Polycomb
TrxG	Complejo proteico activador Trithorax
LINE	Secuencia de ADN Repetida larga
SINE	Secuencia de ADN Repetida corta
LTR	Secuencia de ADN Repetida larga terminal
HP1	Proteína de la Heterocromatina 1
TPE	Efecto de Posición Telomérico
TAD	Dominios Topológicamente Asociados
shRNA	Ácido Ribonucleico pequeño tipo tallo-asa

# ÍNDICE

<b>RECONOCIMIENTOS ACADÉMICOS</b> .....	<b>2</b>
<b>ABREVIATURAS</b> .....	<b>3</b>
<b>RESUMEN</b> .....	<b>7</b>
<b>ABSTRACT</b> .....	<b>9</b>
<b>INTRODUCCIÓN</b> .....	<b>10</b>
¿CÓMO SE REGULA LA TRANSCRIPCIÓN DE LOS GENES?.....	12
<i>El genoma de las células eucariontes</i> .....	12
<i>Factores transcripcionales</i> .....	12
<i>Elementos genéticos que regulan la transcripción</i> .....	13
<i>La estructura de la cromatina y la regulación de la transcripción</i> .....	15
<i>La regulación epigenética</i> .....	17
¿QUÉ ES EL SILENCIAMIENTO EPIGENÉTICO?.....	18
<i>La eucromatina y la heterocromatina</i> .....	18
<i>El silenciamiento epigenético: formación mantenimiento y propagación de la heterocromatina</i> .....	21
<i>Mecanismo de silenciamiento del complejo SIR</i> .....	26
<i>Mecanismos del silenciamiento telomérico (Efecto de Posición Telomérico)</i> .....	28
<i>Mecanismos del silenciamiento epigenético en Schizosaccharomyces pombe</i> .....	30
<i>Mecanismo del ARN de interferencia</i> .....	32
¿CÓMO SE ESTABLECE UN DOMINIO TRANSCRIPCIONALMENTE ACTIVO DE EUCROMATINA?.....	37
<i>Factores transcripcionales pioneros</i> .....	37
<i>El factor nuclear CTCF es una proteína “insulator”</i> .....	41
<i>CTCF y datos genómicos que apoyan el concepto de la organización del genoma en dominios</i> .....	44
<b>ANTECEDENTES PARTICULARES</b> .....	<b>48</b>
EL DOMINIO CROMATÍNICO B-GLOBINA DE POLLO.....	48
EL DOMINIO DE LAS 16 KB DE HETEROCROMATINA LOCALIZADO ENTRE EL GEN <i>FOLR1</i> Y LOS GENES B-GLOBINA.....	49
EL DOMINIO FORMADO POR EL GEN RECEPTOR DE FOLATO <i>FOLR1</i> .....	50
<b>PLANTEAMIENTO DEL PROBLEMA</b> .....	<b>51</b>
<b>HIPÓTESIS</b> .....	<b>51</b>
<b>OBJETIVO GENERAL</b> .....	<b>52</b>
<b>OBJETIVOS PARTICULARES</b> .....	<b>52</b>
<b>MATERIAL Y MÉTODOS</b> .....	<b>53</b>
PLÁSMIDOS Y CONSTRUCCIONES.....	53
LINEAS CELULARES Y CULTIVO CELULAR.....	53
TRANSFECCIONES TRANSITORIAS Y ESTABLES.....	55
PCR-CUANTITATIVA.....	55
PCR DUPLEX SEMICUANTITATIVA.....	56
INMUNOPRECIPITACIÓN DE LA CROMATINA-PCR DÚPLEX.....	56

CONVERSIÓN DE ADN CON BISULFITO DE SODIO .....	57
GENERACIÓN DE UN ARN DE INTERFERENCIA CHCTCF SHRNAI. ....	58
GENERACIÓN DE PARTÍCULAS LENTIVIRALES PARA LA REDUCCIÓN DE LOS NIVELES DEL FACTOR NUCLEAR CTCF.....	59
ENSAYO DE HIPERSENSIBILIDAD AL CORTE POR LA DNASA I.....	59
EMSA .....	59
EL ENSAYO DE PROTECCIÓN CONTRA EFECTOS DE POSICIÓN CROMOSOMAL (CPE) .....	60
ENSAYO DE DETECCIÓN DE ACTIVIDAD “INSULATOR” <i>IN VIVO</i> EN EL PEZ CEBRA <i>DANIO RERIO</i> ZEBRAFISH (ZED).....	61
GENERACIÓN DEL ANTICUERPO ANTI-CCTCF DE POLLO.....	61
<b>RESULTADOS.....</b>	<b>62</b>
1. CARACTERIZACIÓN DE UN ELEMENTO FRONTERA “INSULATOR” LOCALIZADO ENTRE EL DOMINIO DEL GEN <i>FOLR1</i> Y LAS 16 KB DE HETEROCROMATINA. ....	62
<i>Cambios en la expresión del gen FOLR1 correlacionan con cambios en el estado de la cromatina.</i> .....	62
<i>Identificación de un nuevo sitio de unión a CTCF in vivo en el límite entre el dominio del gen FOLR1 y el dominio de las 16 kb de heterocromatina.</i> .....	67
<i>El elemento HSB-A´ tiene la capacidad de bloquear señales de regulación ajenas al dominio que enmarca.</i> .....	71
<i>El elemento HSB-HSA´ protege a un transgen de los efectos de su entorno cromatínico.</i> .....	75
<i>El enriquecimiento in vivo de CTCF sobre el elemento HSB-HSA´ coincide con la expresión del gen FOLR1.</i> .....	77
2. CONFIRMAR SI LAS 16 KB DE HETEROCROMATINA SITUADAS RÍO ABAJO DEL DOMINIO DEL GEN <i>FOLR1</i> PUEDEN AFECTAR LA ESTRUCTURA DE LA CROMATINA DEL DOMINIO Y LA EXPRESIÓN DEL GEN <i>FOLR1</i> DE POLLO. ....	80
<i>El aumento y propagación en la metilación del ADN en la región regulatoria del gen FOLR1 coincide con su silenciamiento.</i> .....	80
<i>El “insulator” HSB-HSA´ puede proteger a un transgen de las señales de silenciamiento epigenético provenientes de las 16 kb de heterocromatina constitutiva.</i> .....	82
3. CARACTERIZAR Y ESTUDIAR LOS MECANISMOS POR LOS CUALES UN “INSULATOR” CONTRARRESTA LA PROPAGACIÓN DE LA HETEROCROMATINA Y COMO ÉSTE CONTRIBUYE A LA FORMACIÓN DE UN DOMINIO TANTO <i>IN VITRO</i> COMO <i>IN VIVO</i> . ....	84
<i>La disminución en los niveles de la proteína CTCF afecta la expresión del gen FOLR1.</i> .....	84
<b>DISCUSIÓN.....</b>	<b>87</b>
<b>PERSPECTIVAS.....</b>	<b>93</b>
<b>ANEXO A .....</b>	<b>94</b>
GENERACIÓN DE UN ANTICUERPO POLICLONAL ANTI-CTCF DE POLLO. ....	94
<b>ANEXO B .....</b>	<b>97</b>
ANÁLISIS BIOINFORMÁTICO DE LA REGIÓN 3´ DEL GEN <i>FR</i> EN BÚSQUEDA DE SITIOS DE UNIÓN AL FACTOR DE TRANSCRIPCIÓN CTCF. ....	97
.....	99
<b>ANEXO C .....</b>	<b>100</b>
ANÁLISIS <i>IN VITRO</i> E <i>IN VIVO</i> DE LA UNIÓN DEL FACTOR NUCLEAR CTCF AL ELEMENTO HSB-HSA´.....	100
<b>ANEXO D .....</b>	<b>103</b>

ANÁLISIS IN VIVO DEL EFECTO DE LA DISRUPCIÓN DE FACTOR NUCLEAR USF1 EN EL “INSULATOR” 5´CHS4 DEL DOMINIO B-GLOBINA EN FUNCIÓN DE LA POSIBLE PROPAGACIÓN DE LAS 16 KB DE HETEROCROMATINA. ....	103
1) <i>Presencia del sitio de hipersensibilidad 5´cHS4 sobre el “insulator” en la línea celular HD3.</i> .....	103
2) <i>Contribución de USF1 a la estructura de la cromatina en el “insulator” 5´cHS4.</i> .....	105
3) <i>Generación de una línea estable deficiente de USF1 mediante ARN interferencia y ensayo de hipersensibilidad a la DNasa I en la línea celular HD3-USF1Kd.</i> .....	106
<b>PUBLICACIONES.....</b>	<b>111</b>
<b>BIBLIOGRAFÍA.....</b>	<b>112</b>

## RESUMEN

El genoma de las células eucariontes se encuentra organizado en dominios de eucromatina y heterocromatina. Los “insulators” son secuencias de ADN que forman y mantienen a éstos dominios dependiendo de la unión de factores nucleares como CTCF para su función. Nuestro estudio se centró en el dominio que incluye al gen *FOLR1*, el cual es regulado a través de la diferenciación eritroide del pollo. Utilizando herramientas de biología molecular, biología celular y bioinformática, demostramos la presencia de un nuevo elemento “insulator”, el cual llamamos HSB-HSA’ y permite la expresión del gen *FOLR1* durante la etapa de pre-eritroblastos. Este nuevo elemento tiene un motivo de unión al factor nuclear CTCF que marca una frontera entre el dominio de eucromatina del gen *FOLR1* y el dominio contiguo de 16 kb de heterocromatina constitutiva. HSB-HSA’ tiene las propiedades “insulator” clásicas tanto *in vitro* como *in vivo*, actuando como un elemento barrera contra la propagación de heterocromatina y bloqueando las señales regulatorias de un enhancer o potenciador ajeno al dominio que enmarca. En el contexto endógeno, la ausencia de CTCF en HSB-HSA’ coincide con el silenciamiento del gen *FOLR1* y el aumento de los niveles de metilación del ADN en sus regiones regulatorias, resultado de la propagación de las 16 kb de heterocromatina contigua. Con base en estos resultados, proponemos que el silenciamiento del gen *FOLR1* permite junto con otros procesos celulares, avanzar en la diferenciación eritroide. El estudio de la formación y mantenimiento de dominios cromatínicos a una escala local, permite entender a una escala global, el

establecimiento de programas de expresión de los genes durante la diferenciación celular.

## ABSTRACT

The eukaryotic genome is organized into active and inactive chromatin domains. Insulators are DNA elements with the capacity to generate and maintain these chromatin domains. Our study focused on the chicken folate receptor gene (*FOLR1*) which is regulated through the erythroid differentiation. Using bioinformatics, molecular and cellular biology experimental strategies we demonstrate that a novel insulator element, called HSB-HSA', contribute to *FOLR1* regulation of gene expression. This chromatin insulator presents a CTCF binding site palyng a boundary role between *FOLR1* active chromatin domain and the adjacent 16 kb constitutive heterochromatin domain. The HSB-HSA' element possess both classical insulator properties *in vitro* and *in vivo*; protects a transgene against chromosomal position effects and enhancer blocking activity. In the endogenous context, the absence of CTCF in HSB-HSA' results in irreversible *FOLR1* gene silencing with an increased levels of DNA methylation in its gene-regulatory regions. We propose that the spreading of 16 kb constitutive heterochromatin results in an epigenetic silencing, which allows together with other cellular processes the erythroid differentiation advance. Understanding the formation and maintenance of chromatin domains at a local level, provides new insights about the establishing of gene expression programs during cell differentiation.

## INTRODUCCIÓN

El avance tecnológico en la investigación científica, ha permitido realizar observaciones y conclusiones cada vez más precisas, de una gran cantidad de procesos dentro de la biología.

A inicios del siglo XX, con el descubrimiento de la estructura del ADN, se dio un avance significativo en el entendimiento de la genética y también, se dio el surgimiento de la biología molecular, teniendo como base al dogma central de la biología, el cual establece un flujo de información genética, desde el ADN hasta la síntesis de proteínas. Complementando estas observaciones, a mediados del siglo XX, Conrad Waddington propuso el término epigenética, para “clasificar a todos los eventos del desarrollo que van desde el cigoto fertilizado hasta el organismo maduro, es decir, todos los procesos regulados que, a partir del material genético, dan forma al producto final” (1). Fue a finales del mismo siglo, que investigadores pioneros en el estudio de los mecanismos epigenéticos, describieron “cambios en la expresión de los genes, que no podían ser explicados por alteraciones en la secuencia del ADN” (2).

A principios del siglo XXI, se completó el primer borrador de la secuencia del ADN humano, dándose un avance significativo en el estudio de la función a nivel global, del genoma y la epigenética (3, 4). Esto ha permitido acumular evidencias, que han revolucionado el entendimiento de los procesos epigenéticos asociados a la regulación de la expresión de los genes (5). En esta nueva era de la Biología se han replanteado preguntas como: ¿Cuáles son los mecanismos por

los que un solo genoma, puede dar origen a todos los tipos celulares que constituyen a un organismo multicelular?, ¿Por qué individuos con el mismo genotipo, pueden exhibir diferencias en su fenotipo?, ¿Cómo influye el medio ambiente en la expresión de los genes?, ¿Cuáles son las bases genéticas y epigenéticas del envejecimiento?, ¿Cuál es la relación entre el medio ambiente, el genoma y el comportamiento social?. Es claro que hoy, no se puede entender a la genética sin la epigenética, por tanto, no pueden abordarse éstas preguntas desde una sola perspectiva.

Con el estudio de las bases moleculares de la epigenética, podremos seguir avanzando y resolviendo preguntas fundamentales, que involucran íntimamente al genoma y la expresión de los genes. El conocimiento adquirido podrá traducirse en el perfeccionamiento del diagnóstico y tratamiento de enfermedades relacionadas con los genes, incluyendo el uso de la terapia génica en humanos.

Nuestro grupo de investigación se ha dado a la tarea de entender la función de las regiones del ADN no-codificante y su contribución a la regulación de la expresión génica. En este trabajo de tesis, abordaremos la propagación del silenciamiento epigenético durante la diferenciación eritroide del pollo, y estudiaremos la región regulatoria del gen *FOLR1*, para comprender la influencia del contexto genómico y el ambiente cromatínico sobre la transcripción de los genes.

## **¿Cómo se regula la transcripción de los genes?**

### **El genoma de las células eucariontes**

El genoma está constituido por ADN y contiene toda la información genética de las células eucariontes, tiene un tamaño aproximado de 2 metros y es empaquetado dentro de un núcleo de apenas unos cuantos micrones de diámetro (6). En el genoma humano, existen aproximadamente 20,687 genes que codifican para proteínas (7), los cuales representan apenas el 2-3% de la información del ADN (8), el resto, fue considerado inicialmente como ADN-basura (9).

Actualmente sabemos que más del 90% del genoma es no-codificante y es transcrito en ARN (8). En humanos, se ha propuesto que la mayoría de dichos transcritos participan en al menos un evento bioquímico ó epigenético (5). En otras palabras, la vasta mayoría del genoma es potencialmente funcional. Nuestro estudio involucra al ADN estructurado en cromatina. Comenzaré describiendo a los factores transcripcionales, a los elementos genéticos codificados en el ADN, definiré la estructura de la cromatina y a los procesos epigenéticos asociados.

### **Factores transcripcionales**

Los factores transcripcionales (TF) son proteínas nucleares que tienen un dominio de unión al ADN con la capacidad de reconocer una combinación de nucleótidos específica (10). Son necesarios para llevar a cabo la transcripción, la replicación y la reparación del ADN.

Existen principalmente dos tipos de factores transcripcionales, los generales para la transcripción y los que son tejido específico. Los factores transcripcionales pueden, una vez que se han unido al ADN, reclutar elementos remodeladores de la cromatina o a la maquinaria necesaria para realizar la transcripción a través de su interacción con co-factores. Su regulación permite la expresión orquestada de los genes, en consecuencia, alteraciones en la función de éstos pueden dar origen a enfermedades de tipo genético e incluso epigenético.

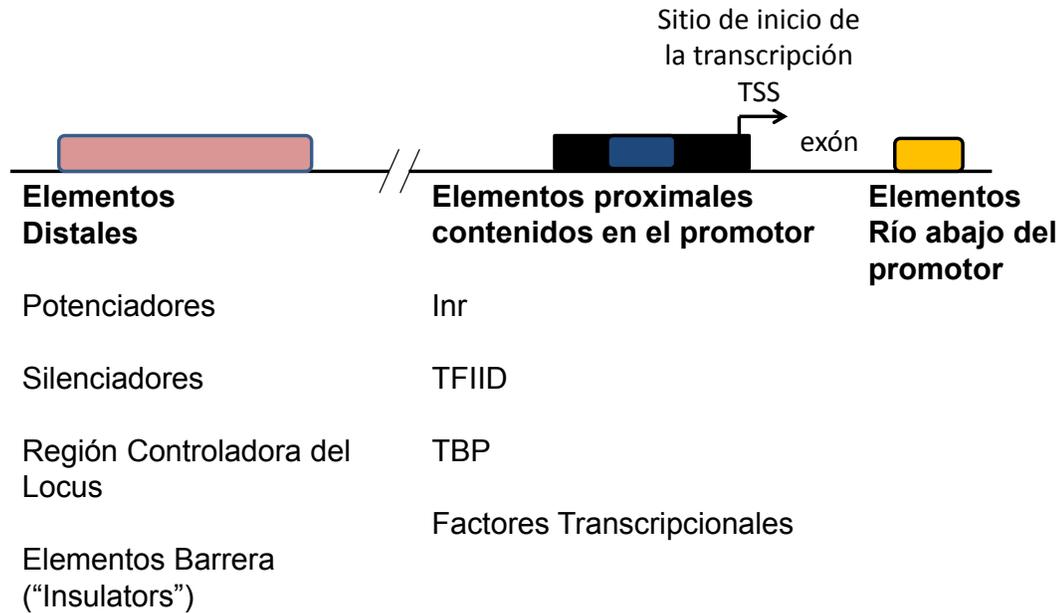
### **Elementos genéticos que regulan la transcripción**

Los elementos genéticos que regulan la transcripción son: a) promotores, b) potenciadores (enhancers), silenciadores (silencers), elementos controladores de locus (LCRs) y los elementos barrera (insulators) (Fig. 1). Pueden clasificarse en base a su posición en el ADN respecto al inicio de la transcripción en: a) elementos proximales del promotor, b) elementos distales (10, 11) (Fig. 1).

Un promotor se define como la secuencia continua de ADN que tiene la capacidad de unir y dirigir a la maquinaria de la transcripción mediada por la ARN polimerasa II (10, 11). La arquitectura de un promotor comprende a los elementos proximales y a los elementos "core" (Fig. 1). Entre los elementos proximales encontramos a los sitios de unión a factores de transcripción, el sitio de unión al factor de transcripción TBP ó TRF, a los elementos de reconocimiento del factor TFIIB y a los elementos iniciadores (inr) (11, 12) (Fig. 1).

Los elementos distales son los potenciadores (enhancers), los silenciadores (silencers), los elementos controladores de locus o LCR (Locus Control Region) y los elementos frontera (insulators) (13) (Fig. 1). Todos ellos actúan de manera orquestada en la regulación de la transcripción de uno o varios genes.

Los elementos barrera o “insulators” han despertado el interés de muchos grupos de investigación incluyendo al nuestro porque establecen dominios genómicos favoreciendo el entorno en el que se da la transcripción de los genes, en particular, nos interesa su capacidad para contener contra la propagación de la heterocromatina. Son tema central en este trabajo de tesis, por tanto, haré una descripción detallada de sus funciones más adelante. A continuación describiré de qué manera el genoma se encuentra estructurado y organizado en cromatina.

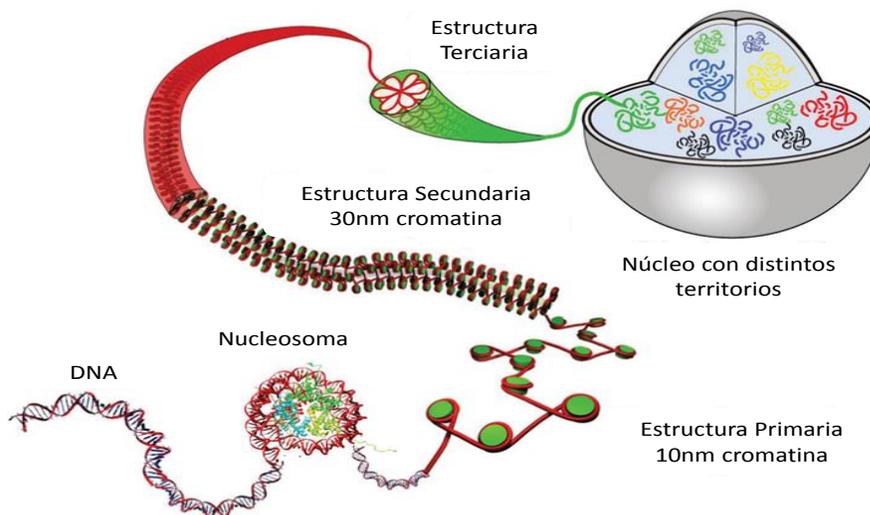


**Figura 1. Elementos reguladores genéticos codificados en *cis* que regulan la transcripción.** En levaduras, por ejemplo, la mayoría de los Factores Transcripcionales situados en el promotor, se unen ~115 pares de bases arriba del sitio de inicio de la transcripción TSS (14), en cambio, los elementos distales pueden encontrarse a decenas de kilo bases de distancia respecto al inicio de la transcripción.

## La estructura de la cromatina y la regulación de la transcripción.

A finales del siglo XIX existía un gran interés en el estudio de los seres vivos a escala microscópica, en esta época, se lograron avances importantes en la preparación de tinciones celulares y de especímenes para observar al microscopio. En este contexto, en 1878, Walther Flemming describió a los componentes estructurales observados al interior del núcleo y se refirió a ellos como: la cromatina (15, 16). En 1904 Boveri ya relacionaba a la cromatina con las fases del ciclo celular: "Cromatina es la sustancia que se transforma en cromosomas durante la mitosis" (17).

La cromatina tiene como unidad mínima de organización al nucleosoma, el cual está formado por la interacción de 147 pares de bases (pb) de ADN alrededor de un octámero de histonas (formado por un par de histonas H3-H4 y otro par H2A-H2B) (18) (Fig. 2). El siguiente nivel de organización de la cromatina está dado por la incorporación de una quinta histona, la histona H1, a la región internucleosomal de aproximadamente 40 a 60 pb de ADN (variable entre diferentes tipos celulares incluso de un mismo organismo) facilitando la formación de la fibra de 30 nm ó solenoide (19). Actualmente el modelo de solenoide enfrenta mucha controversia, por ejemplo, no queda claro que la fibra de 30 nm sea la conformación preferencialmente adoptada por la cromatina al interior del núcleo en interfase. Sin embargo, la mayor parte de los procesos de remodelación de la cromatina permiten transitar de manera regulada al nivel de la fibra de 10 nm durante el proceso de la transcripción (18, 20-22) (Fig. 2).



**Figura 2. El genoma se estructura en cromatina.** La unidad fundamental de la cromatina es el nucleosoma. Se muestran los diferentes niveles de estructuración de la cromatina en el núcleo. La incorporación de la histona H1 permite formar la estructura de 30 nm, la cual se le considera heterocromatina.

En este contexto, la transcripción de un gen es regulada en tres niveles: el primer nivel requiere de la secuencia nucleotídica o información genética codificada en la molécula del ADN. El segundo nivel involucra la estructura de la cromatina donde la regulación de la expresión de un gen no solo depende de la información codificada en la secuencia del ADN sino también de su organización en cromatina y que se engloba en lo que se ha llamado regulación epigenética (Fig. 2) y el tercer nivel, tiene un vínculo estrecho con la organización tridimensional del genoma al interior del núcleo, el cual retomaré más adelante en esta tesis (Fig. 2).

### **La regulación epigenética**

La regulación epigenética, se refiere a todos aquellos procesos que influyen en la expresión heredable de los genes sin que ocurran cambios en la secuencia del ADN (1, 2). De manera general consideramos que los procesos asociados a la regulación epigenética son básicamente seis: 1) modificaciones post-traduccionales de las histonas, 2) la metilación del ADN, 3) los complejos de remodelaje ATP-dependientes, 4) los complejos proteicos llamados Polycomb y Trithorax (PcG y TrxG), 5) ARNs no-codificantes y 6) la organización tridimensional de la cromatina junto con la dinámica nuclear. Cabe señalar la existencia de una importante interdependencia entre todos estos procesos (1).

A continuación explicaré cuales son los tipos de cromatina y sus características generales.

## ¿Qué es el silenciamiento epigenético?

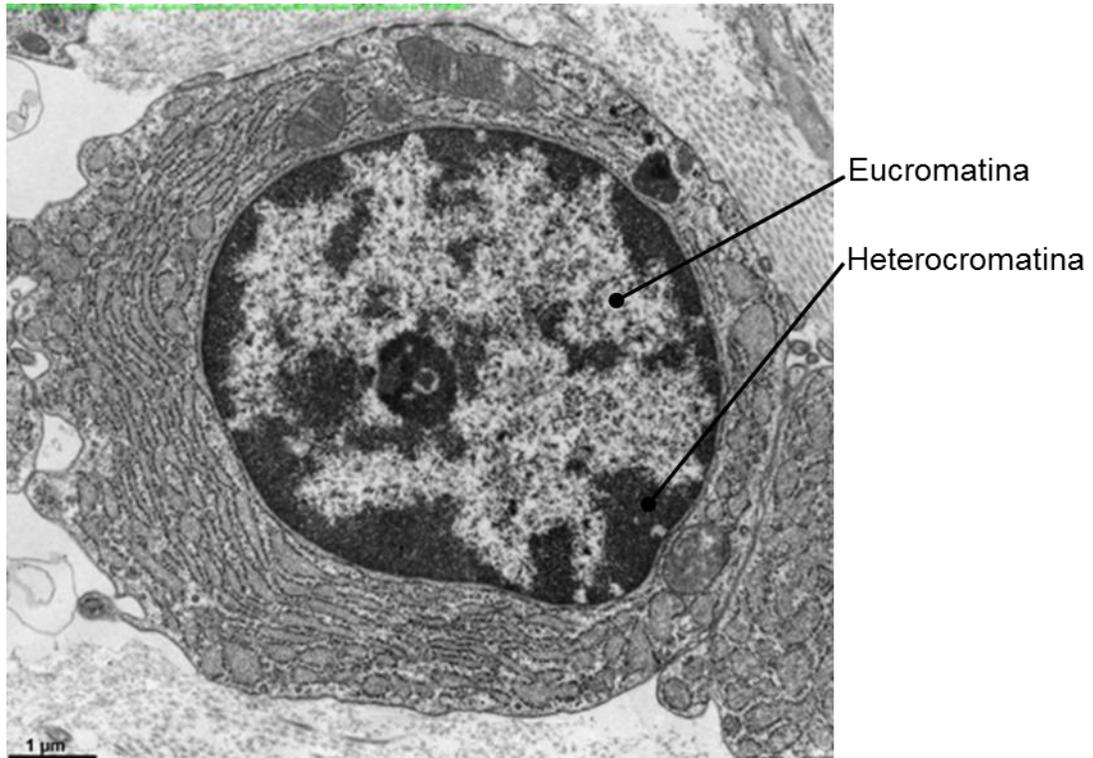
### La eucromatina y la heterocromatina

En 1928 Emil Heitz observó que existían dentro de la cromatina, zonas teñidas diferencialmente. Determinó entonces, que existían dos tipos de cromatina y utilizó el término heterocromatina y eucromatina para distinguir entre el estado condensado y el estado de-condensado de la cromatina (23). Un año después, en 1929 Emil Heitz hipotetizó que la heterocromatina contiene menos genes o es “génicamente pasiva” en contraste con la eucromatina, un argumento que perdura hasta la actualidad (Fig. 3) (23).

El genoma eucarionte está organizado en dos tipos distintos de cromatina. La eucromatina está constituida por regiones ricas en genes, pobre en secuencias transponibles y/o repetidas, accesible a la digestión por parte de la endonucleasa DNasa I, presenta una cromatina laxa con abundancia de la ARN Polimerasa II (ARN Pol-II) y se replica temprano durante la fase S del ciclo celular (Fig. 3).

En contraste, la heterocromatina se caracteriza por contener regiones con una baja o nula densidad de genes, abundante en diversos tipos de secuencias repetidas (*LINE*, *SINE*, *LTRs*), una cromatina altamente compacta e inaccesible a la digestión por parte de la DNasa I, una baja presencia de la ARN Pol-II, con una

característica particular: la presencia de la proteína de heterocromatina HP1, finalmente, la heterocromatina mantiene su estructura compacta a través de las fases del ciclo celular  $G_1$ ,  $G_2$  y M, replicándose tarde durante la fase S (Fig. 3).



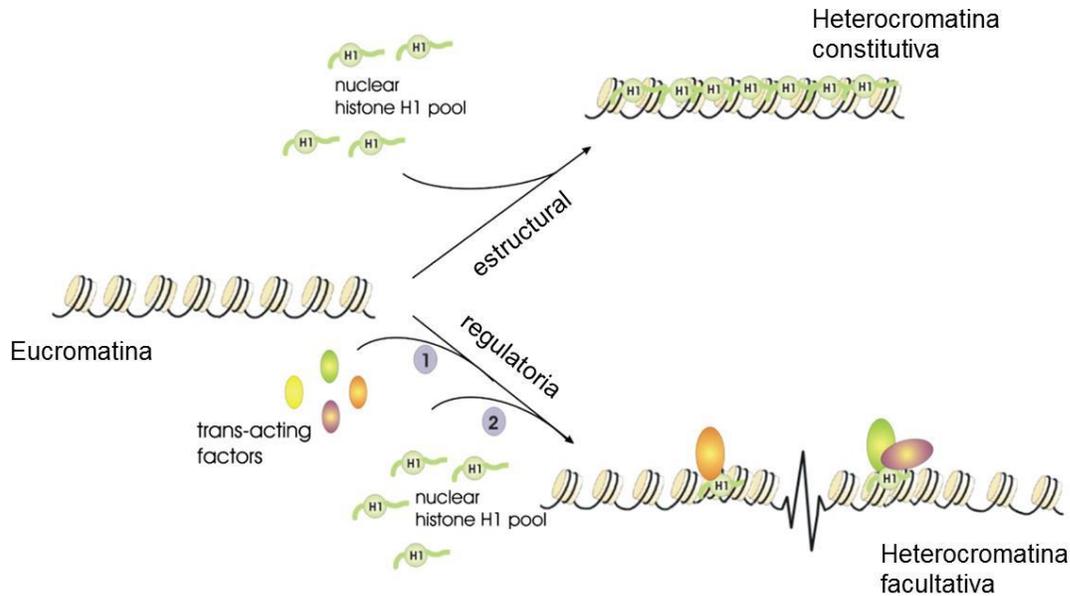
**Figura 3. Micrografía electrónica de una célula en interfase, se muestra a la eucromatina y a la heterocromatina.** Como se puede observar, la heterocromatina es más electrodensa y se distribuye hacia la periferia nuclear, en contraste, la eucromatina suele distribuirse en las regiones más centrales del núcleo. Micrografía tomada del sitio [http://kobiljak.msu.edu/cai/virtual\\_microscope/Cytology/Obj\\_04\\_02a.html](http://kobiljak.msu.edu/cai/virtual_microscope/Cytology/Obj_04_02a.html)

A su vez, existen dos tipos de heterocromatina: a) la heterocromatina constitutiva que tiene una función estructural y encontramos en centrómeros y telómeros, necesarios para la segregación y la integridad de los cromosomas respectivamente y, b) la heterocromatina facultativa que tiene una función regulada de silenciamiento génico y que encontramos durante el desarrollo de un organismo, por ejemplo, la heterocromatinización de un alelo, en uno de los

cromosomas homólogos, o como resultado de la diferenciación celular y el establecimiento de programas de expresión génica (Fig. 4) (24, 25).

El mecanismo de silenciamiento de la heterocromatina, es diferente al que involucra factores transcripcionales en las regiones promotoras de los genes. La heterocromatina puede silenciar genes a través de su propagación compactando las regiones regulatorias y codificantes, en otras palabras, impidiendo el acceso a la información del ADN tanto a los factores, como a la maquinaria de la transcripción, de aquí que ha sido englobada en los fenómenos epigenéticos. En consecuencia la heterocromatina permite: a) mantener la estabilidad del genoma, silenciando a las secuencias repetidas y en general, al ADN exógeno, que en algunos casos, puede generar mutaciones por transposición y, b) en el caso de las células diferenciadas, heredar de manera clonal un programa de expresión génico o epigenoma, pues, sus características estructurales impiden la transcripción activa del ADN, en este caso, de la información de los genes que no deben expresarse.

A continuación detallaré los mecanismos de la propagación de la heterocromatina.

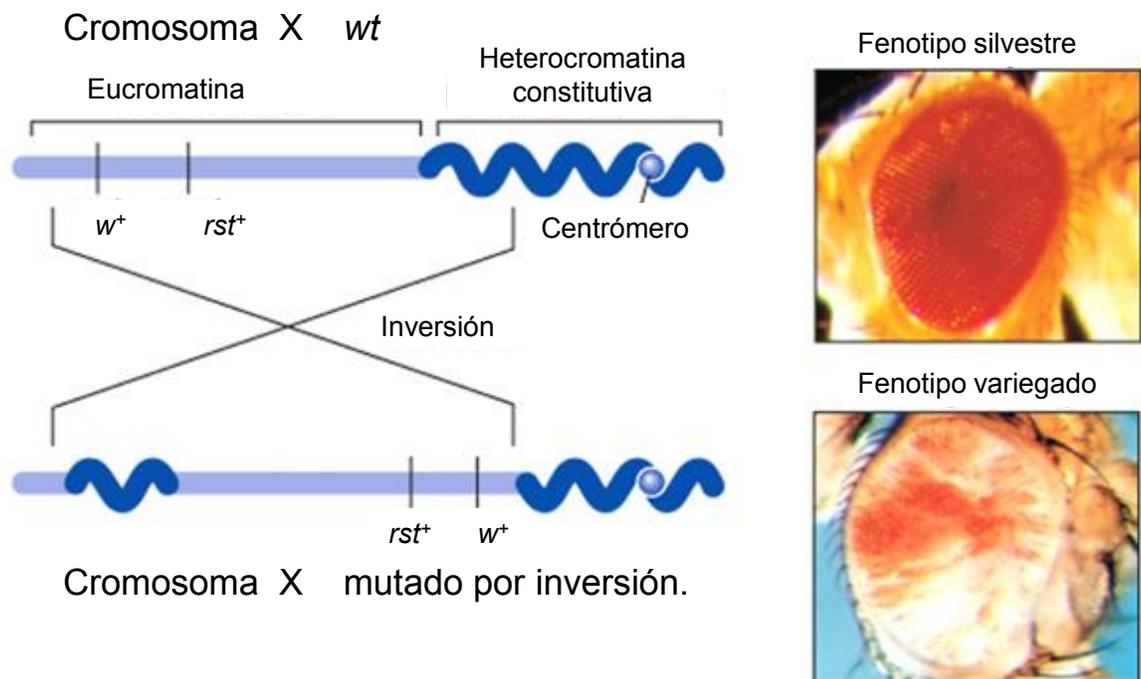


**Figura 4. Tipos de cromatina.** La heterocromatina constitutiva está relacionada con funciones estructurales del genoma como las que realizan los centrómeros y telómeros. La heterocromatina facultativa está relacionada con el silenciamiento de genes en células diferenciadas. Modificada de (25).

## El silenciamiento epigenético: formación mantenimiento y propagación de la heterocromatina

Hermann Muller en 1930 observó un fenómeno interesante mientras hacía estudios genéticos en embriones de la *Drosophila melanogaster* (26). Tras irradiar con luz ultravioleta embriones, generó un fenotipo variegado de expresión del gen “White”  $w^{m4}$ , debido a una inversión en el cromosoma X que acerca a éste gen hacia la heterocromatina pericentromérica. Es importante señalar que la heterocromatina se establece muy temprano en el desarrollo, por lo que aún cuando se hayan irradiado embriones, el fenotipo se observa en el adulto (Fig. 5). Normalmente el ojo de una mosca adulta es color rojo, pero en  $w^{m4}$  el fenotipo es un mosaico de color rojo-blanco (Fig.5). Muller propuso que el fenómeno era

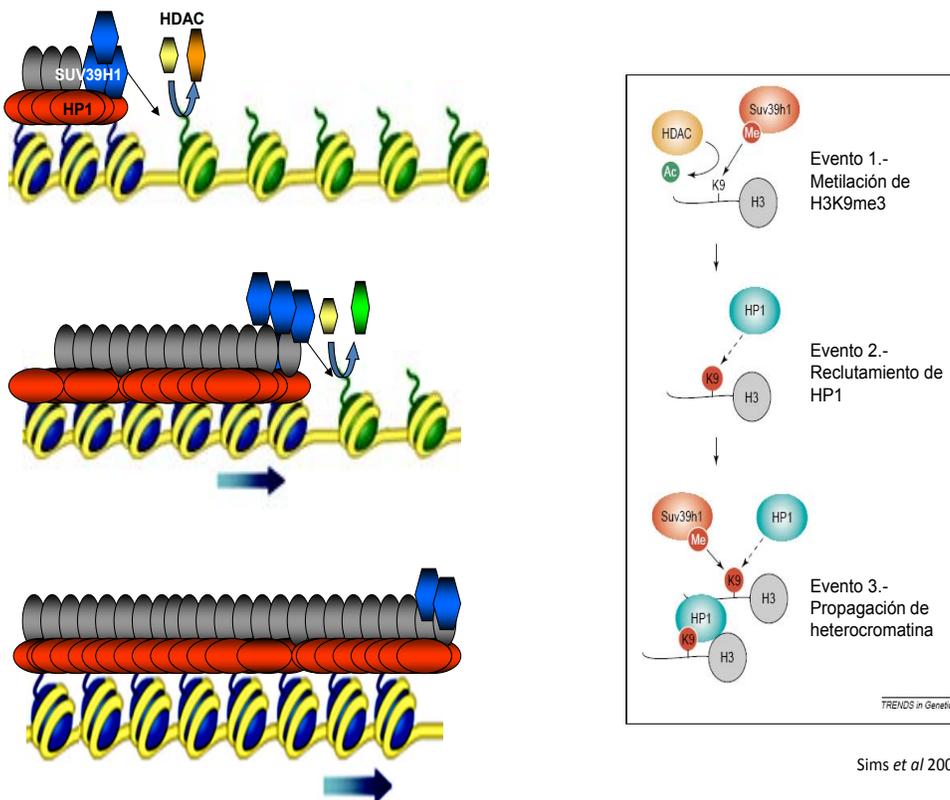
dependiente de la propagación de heterocromatina dando la primera evidencia experimental de que la heterocromatina silencia a los genes que se encuentran en su proximidad, un fenómeno también observado por Bárbara McClintock en 1965 en plantas (27). A este fenómeno se le conoce como “Position Effect Variegation” (PEV) (26) (Fig. 5).



**Figura 5. Efecto de posición cromosomal (PEV).** Este fenómeno demuestra que la heterocromatina constitutiva tiene la capacidad de propagarse silenciando permanentemente los genes que se encuentren cercanos sin afectar la secuencia del ADN, en otras palabras, es un fenómeno de silenciamiento epigenético, por tanto heredable. Tomado de <http://www.discoveryandinnovation.com/BIOL202/notes/lecture18.html>

Brevemente, el mecanismo PEV refleja la desacetilación de las histonas y el reconocimiento de la histona H3 modificada post-traduccionalmente por

trimetilación del residuo de lisina K9 (H3K9me3) por la proteína de la heterocromatina (HP1). Dicho reconocimiento se repite nucleosoma a nucleosoma traduciéndose en la propagación de la heterocromatina sobre el dominio marcado (Fig. 6). Es importante mencionar que algunas proteínas (como HP1) que tiene un papel central en la propagación de la heterocromatina, se encuentran conservadas en todos los organismos estudiados.



**Figura 6. Mecanismo de PEV.** Se muestran los tres eventos involucrados en la propagación de la heterocromatina. Modificado de (28).

Los mecanismos de silenciamiento por heterocromatina aún se encuentran bajo intenso estudio. La mayor parte de avances en este campo, proviene de

estudios en las levaduras debido a la versatilidad de este modelo para generar cepas mutantes de los componentes moleculares y generar individuos haploides.

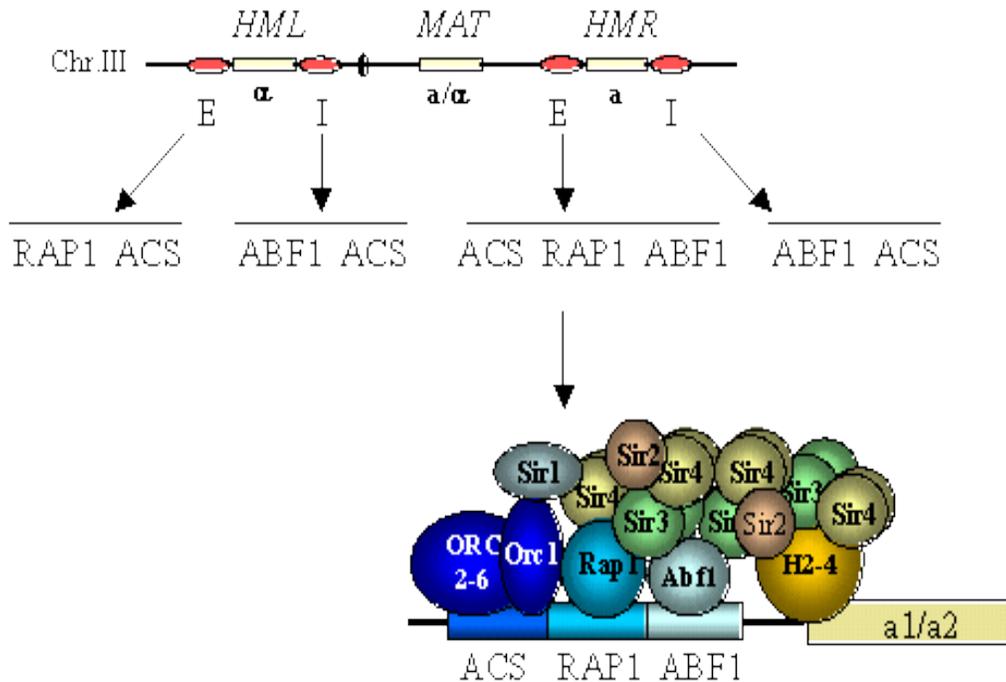
Se han disecado dos tipos de mecanismos moleculares para establecer y propagar el silenciamiento epigenético: el mecanismo que involucra elementos de regulación genéticos llamados silenciadores relativamente bien entendido en *Saccharomyces cerevisiae*, y el mecanismo que involucra la transcripción de secuencias repetidas relativamente bien entendido en *Schizosaccharomyces pombe*.

Comenzaré por describir el silenciamiento epigenético del “mating type” en *Saccharomyces cerevisiae* el cual representa un ejemplo de como la heterocromatina permite mantener una identidad, en este caso, del sexo de una levadura haploide. También detallaré el mecanismo por el cual la heterocromatina se ensambla y se propaga nucleosoma a nucleosoma, lo cual, me permitirá más adelante, exponer las funciones de los elementos con actividad barrera o “insulators” que son motivo de estudio en este trabajo de tesis.

### **El silenciamiento epigenético en *Saccharomyces cerevisiae***

El ciclo de vida de *Saccharomyces cerevisiae* puede generar individuos haploides y diploides. Existen dos tipos de individuos haploides, llamados  $a$  ó  $\alpha$  dependiendo del sexo de cada individuo. Dos individuos haploides de sexo diferente pueden conjugarse y generar un individuo diploide  $a/\alpha$  (29). La información que codifica

para la feromona  $\alpha$  y  $\alpha$ , se encuentra silenciada o dicho de otra manera, existe un proceso de represión de la transcripción en los genes del “*mating type*” *HML $\alpha$*  y *HML $\alpha$*  vía heterocromatinización específica en estos locus (Fig. 7). En el cromosoma III de las levaduras se encuentra el loci del “*mating type*”. En el brazo izquierdo a 12 kb del telómero se encuentra el dominio silenciado (heterocromatina) de *HML $\alpha$*  y contiene la información que codifica para la feromona  $\alpha$ . En el brazo derecho a 23 kb del telómero se encuentra el dominio silenciado (heterocromatina) de *HMR $\alpha$*  y contiene la información que codifica para la feromona  $\alpha$ . El locus *MAT* es transcripcionalmente activo (eucromatina) y puede contener información  $\alpha$  o  $\alpha$  (Fig. 7). Ambos locus contienen a los elementos genéticos regulatorios llamados silenciadores o “*silencer E (essential)*” y “*silencer I (important)*”, los cuales contienen sitios de unión a los factores transcripcionales Rap1, Abf1y el complejo multiprotéico ORC (29, 30) (Fig. 7). El locus *MAT* contiene la información que se expresa en el “*mating type*”. Todas las levaduras pueden intercambiar la información del locus *MAT* a través de la participación de la endonucleasa HO que promueve la recombinación de la información, tomando como templado *HML $\alpha$*  en el 85% de los casos ó de *HMR $\alpha$*  en el 90% de los casos (31).



**Figura 7. El silenciamiento del locus de los genes “mating type” en *Saccharomyces cerevisiae*.** Participan tres elementos principalmente a) las proteínas del complejo SIR, b) los extremos amino terminales de las histonas H3 y H4 c) Los elementos silenciadores E y I. Tomado de <http://edoc.hu-berlin.de/dissertationen/geissenhoener-antje-2004-07-13/HTML/chapter1.html>

### Mecanismo de silenciamiento del complejo SIR

En el silenciamiento epigenético de *S. cerevisiae* participan esencialmente tres factores:

- Las proteínas Sir (de sus siglas en inglés: Silencer Information Regulation).
- Los extremos amino terminales de las histonas.
- Los silenciadores *E* ó *I* del *HM locus*.

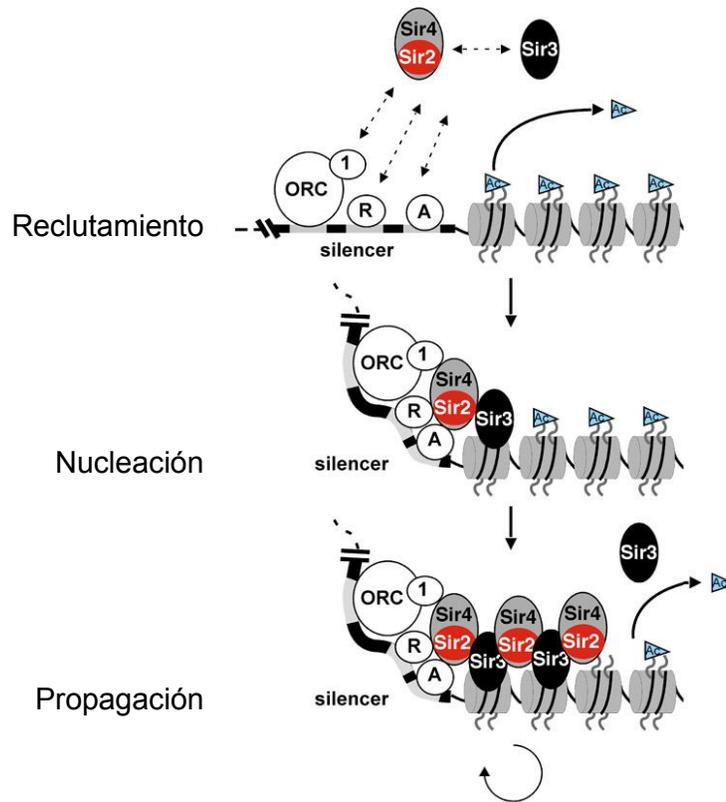
Podemos dividir en dos grandes eventos la formación de la heterocromatina, por un lado la parte genética que involucra a los factores

transcripcionales esenciales Rap1, Orc1 que reclutan a un complejo multiproteico en sus sitios de unión en el ADN conocido como ORC (*Origin Recognition Complex*).

Por otro lado la parte epigenética, que incluye a las proteínas del silenciamiento involucradas en la formación de heterocromatina que no son factores transcripcionales y por tanto no son capaces de unirse a un motivo en el ADN, por tanto, tienen que ser reclutados y son conocidos como el complejo SIR (Sir1-2-3-4) (32). Aunque Sir1 no es esencial para la formación de heterocromatina, se sabe que interactúa directamente con Orc1 y dispara la heterocromatinización, aumentando la posibilidad de que se reclute el complejo SIR (Sir2-3-4) a los complejos ORC (32). Una vez que, Sir 1 es reclutado, se une a Orc1 y puede reclutar a Sir4 y Rap1. Sir4 recluta a Sir2 y se establece el complejo Sir2-4. Sir4 también recluta a Sir3 al “silencer”. Sir3 une a Rap1, Sir4 y Abf1. Sir2 y Sir3 pueden entonces asociarse y finalmente Sir2 desacetila a las histonas de los nucleosomas adyacentes. Sir2 es una proteína desacetilasa NAD<sup>+</sup> dependiente, conservada desde bacterias hasta humano. La unión de Sir2 y Sir4 potencia la actividad desacetilasa de Sir2. Sir2 desacetila la H4 lisina K16 convirtiendo NAD<sup>+</sup> en O-acetil-ADP-ribosa lo que involucra una actividad ADPribosiltransferasa. Además de desacetilar a H4 K16 tanto *in vitro* como *in vivo*, se ha mostrado que desacetila a H3K9 y H3K14 (33).

La propagación del silenciamiento epigenético, involucra el mecanismo anteriormente descrito mediante la modificación de las marcas pos-traduccionales

de las histonas principalmente la desacetilación de H3K9 y H4K16, nucleosoma a nucleosoma.



**Figura 8. Mecanismo del silenciamiento epigenético en *S. cerevisiae*.** Existen tres elementos necesarios para la formación de heterocromatina: a) El complejo SIR, b) Los extremos amino terminales de las histonas, c) Los elementos silenciadores “silencers”. Tomado de <http://www.hhmi.org/research/epigenetic-control-gene-expression-and-genome-stability>.

### Mecanismos del silenciamiento telomérico (Efecto de Posición Telomérico)

Los telómeros de los cromosomas de *S. cerevisiae* son utilizados también como modelo de estudio y tienen un mecanismo similar al del silenciamiento epigenético que los genes del “*mating type*” (Fig. 8).

Los telómeros de *Saccharomyces cerevisiae* han sido estudiados insertando genes reporteros dentro o adyacentes a los telómeros para poder evidenciar el silenciamiento telomérico o “Telomeric Position Effect” (TPE), que involucra la propagación de heterocromatina en las regiones subtelméricas (34).

Brevemente, el silenciamiento telomérico tiene el siguiente mecanismo:

- a) La telomerasa Ku70/80 y las secuencias repetidas teloméricas *TG1-3* que contiene secuencias de unión a Rap1, reclutan al complejo SIR (30).
- b) El complejo SIR que contiene a Sir2, desacetila a las histonas.
- c) Se propaga el complejo SIR mientras los extremos aminoterminales de las histonas en los nucleosomas adyacentes, se encuentren desacetilados.
- d) Actualmente se ha sugerido que los telómeros de *S. cerevisiae* pueden, una vez heterocromatinizados, exhibir una estructura más compleja al doblarse sobre si mismos en un fenómeno conocido como “looping” telomérico (34) (Fig. 9).

En el mantenimiento de los telómeros, existe una ribonucleoproteína conocida como telomerasa Ku70/Ku80, que contiene en cadena sencilla, alrededor de 16-20 secuencias consenso para la unión del factor transcripcional Rap1. Este arreglo de sitios de unión para Rap1, permite mantener una caperuza sobre los extremos de cada cromosoma, y juega un papel crucial en el mantenimiento del tamaño de los telómeros, el cual, también está involucrado en el envejecimiento celular y senescencia. Por otra parte, Sir4 puede asociarse con componentes de la periferia del núcleo y estabilizar un locus silenciado (34) (Fig. 9).

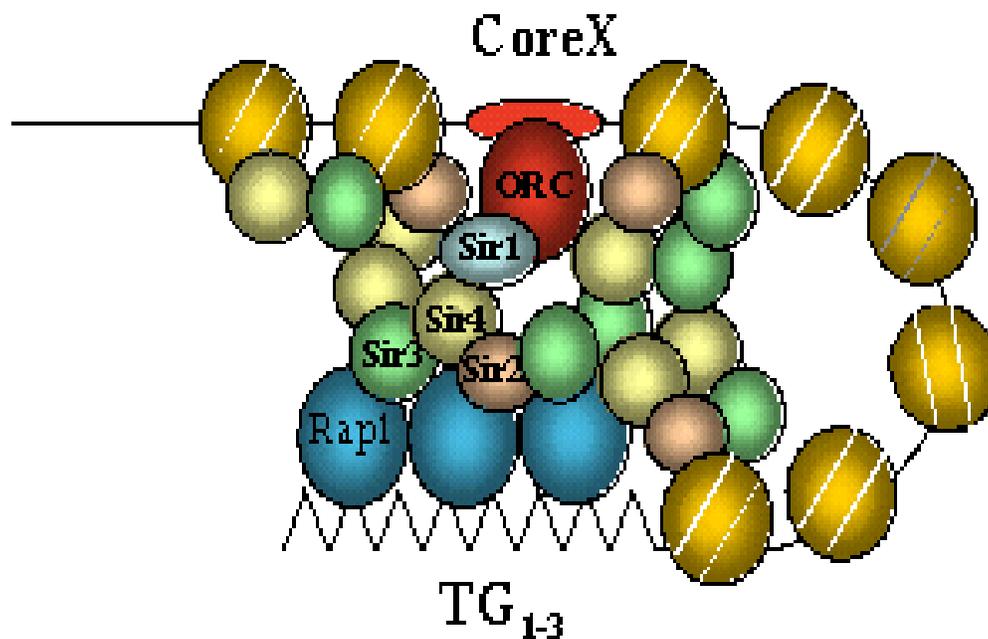


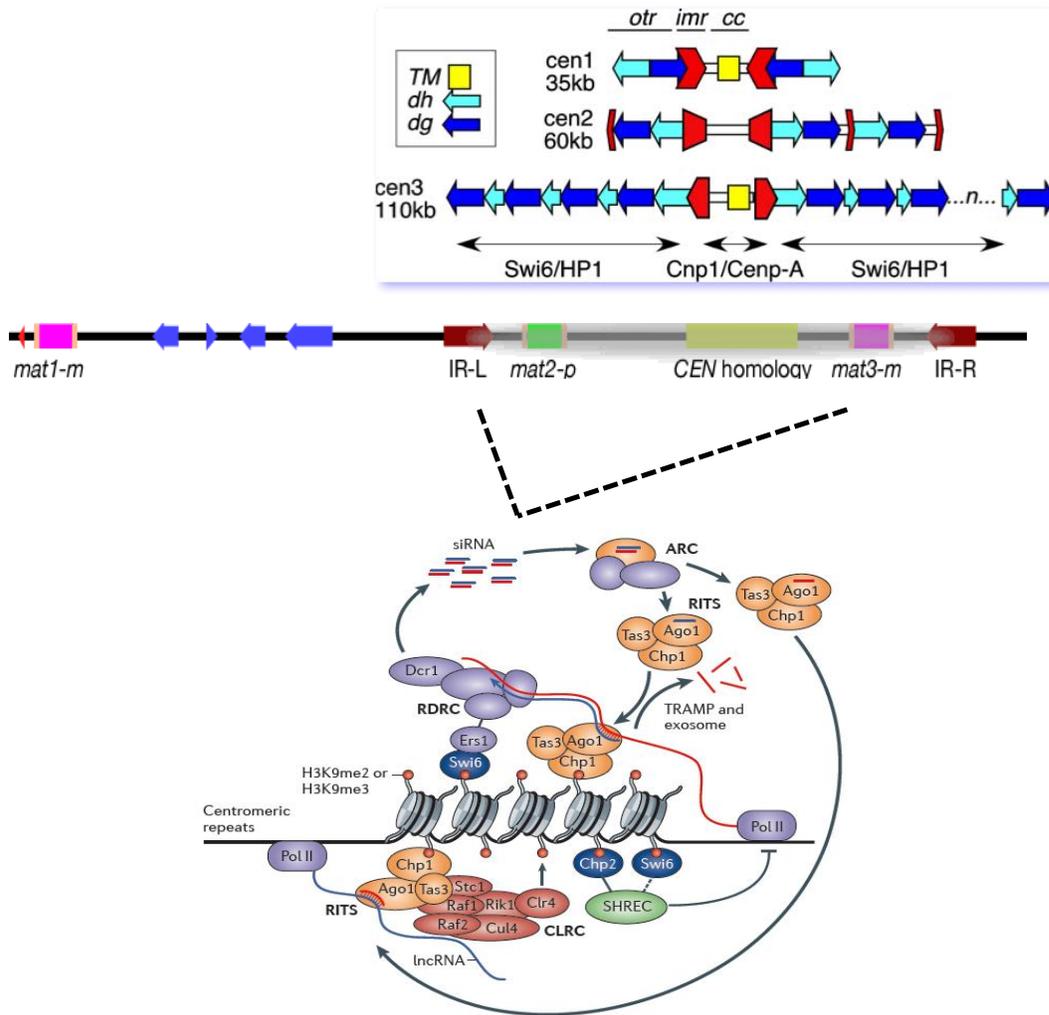
Figura 9. Modelo del mecanismo del silenciamiento telomérico de *Saccharomyces cerevisiae*. El silenciamiento de los telómeros está involucrado en envejecimiento y senescencia. Tomada de (30)

### Mecanismos del silenciamiento epigenético en *Schizosaccharomyces pombe*

La levadura de fisión es otro modelo utilizado para el estudio del ensamblaje, propagación y mantenimiento de heterocromatina. Al igual que *S. cerevisiae*, *S. pombe* puede intercambiar el sexo a través de regular la expresión del “mating type” en los subtipos: minus(-) o plus(+). La información para ambos sexos, yace en el genoma, silenciada por heterocromatinización de los genes *mat2-P(+)*, *mat3-M(-)*. Al igual que el locus *MAT* de *S. cerevisiae*, el locus *mat1* es activo (eucromatina) y puede contener información (+) ó (-) (Fig. 10). *S. pombe* ofrece 4 regiones en donde estudiar los procesos de propagación del silenciamiento

epigenético: a) “*mating type*” locus *mat2-mat3* en donde existe la región cenH, b) la región telomérica y subtelomérica, c) ADN-ribosomal, d) los centrómeros (Fig.10).

Los centrómeros es la estructura de los cromosomas implicada en la segregación cromosomal durante la división celular y, la formación de heterocromatina es necesaria para su función. A nivel de la secuencia del ADN los centrómeros en *S. pombe* están formados por: a) una región de repetidas llamadas “*outter repeats*” (subdivididas en elementos conocidos como *dg* y *dh* ó K y L respectivamente), las cuales enmarcan una región central del centrómero que incluye el conjunto de repetidas “*inner repeats*” (*imr* ó B) y, b) un dominio central (*cnt* ó CC) (35, 36) (Fig. 10). *S. pombe* tiene 3 cromosomas y por lo tanto los 3 centrómeros *cen1*, *cen2* y *cen3* que abarcan alrededor de 35, 60 y 110 pb en los cromosomas I, II, III, respectivamente. Dado que la presencia de secuencias repetidas involucra la presencia de heterocromatina, los centrómeros tienen la capacidad de reprimir la expresión génica. En *S. pombe* el silenciamiento de la expresión génica puede estudiarse a través de genes reporteros como *Ura4+* y *Ade6+*.



**Figura 10. Mecanismo del ARN de interferencia en las regiones centroméricas y pericentroméricas de *S. pombe*.** Tomado de (35) y modificado de <http://joelhuberman.net/HubermanLabArchives/BIR572/Fig4aYeastGenetics.html>

## Mecanismo del ARN de interferencia

Este mecanismo primero descrito en *Caenorhabditis elegans* como un mecanismo post-transcripcional de degradación de ARN, luego relacionado a la metilación del ADN y silenciamiento en *Arabidopsis thaliana*, ha sido descrito en *Schizosaccharomyces pombe* como un mecanismo transcripcional de defensa

contra la expresión de transposones y en general ADN exógeno (36). Este silenciamiento está asociado con la generación de ARN pequeños que tienen como función llevar a cabo la heterocromatinización de estas regiones y su mecanismo involucra principalmente la expresión de tres genes que tienen copia sencilla en éste organismo: *ago1*, *dcr1*, *rdp1*. La mutación por delección de cualquiera de éstos, genera inestabilidad genómica, pérdida de heterocromatina y una disminución en los niveles globales de la metilación de la histona H3K9, la cual, es una marca de histonas funcionalmente asociada a heterocromatina y conservada desde levaduras hasta plantas y humanos. Estos resultados muestran que los mecanismos del silenciamiento epigenético como la desacetilación y la trimetilación en el residuo K9 de la histona H3 juegan un papel central, al igual que HP1, en el proceso de formación, mantenimiento y propagación de heterocromatina en casi todos los organismos eucariontes (35).

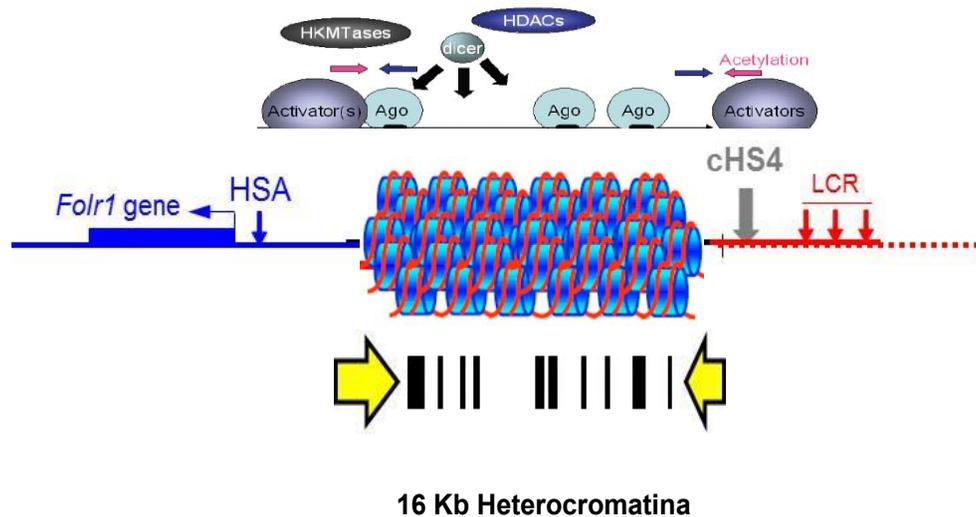
En las regiones pericentroméricas existen secuencias repetidas que se creen son remanentes de transposones y se mantienen en heterocromatina contribuyendo al ensamblaje y función del centrómero. La transcripción de éstas secuencias repetidas generan ARNs no-codificantes largos llamados cenARNs, que son procesados por la maquinaria del ARN de interferencia que incluye a Dcr1 y Ago1, en ARN pequeños interferentes (ARNi), los cuales reclutan al complejo RITS (“RNA Induced Transcriptional Silencing” formado por Tas3, Ago1, Chp1), el cual puede dirigirse hacia las secuencias repetidas vía la hibridación entre los ARNi y cenARNs-nacientes, esto recluta a los elementos que forman el complejo CLRC (Clr4-Rik1-Cul4). Clr4 es la H3K9 metil-transferasa que tri-metila a la K9 de

la histona H3 (H3K9me3) y puede ser reconocida por la proteína HP1 (Swi6), también es reconocida por los complejos desacetiladores de histonas.

HP1 entonces puede estabilizar la unión de RITS a la cromatina y asegurar el reclutamiento de RDRC que involucra a Rdrp1 (RNA-Dependent RNA polymerase Complex) y mantener la generación de ARN de doble cadena necesaria para que dcr1 y ago1 induzcan un nuevo ciclo de ARNi, generando un fenómeno de amplificación de la señal de heterocromatinización en la región marcada (35, 36). Cabe mencionar que aún se desconoce las diferencias entre la regulación de la transcripción de secuencias repetidas exógenas y la transcripción normal del ADN, la cual permite que no se generen eventos de heterocromatinización y silenciamiento aberrante en otras regiones diferentes a los centrómeros (Fig. 10).

Este mecanismo de formación, propagación y mantenimiento de heterocromatina, tiene una gran importancia en nuestras investigaciones, debido a que en el genoma del pollo existe un dominio de heterocromatina constitutiva de 16 kb con secuencias repetidas CR1, que se transcriben e involucran a la maquinaria del ARNi para la formación, mantenimiento y propagación de heterocromatina, es decir, el mecanismo de silenciamiento génico por ARN de interferencia anteriormente descrito, esta conservado en pollo y es el mecanismo por el cual, el dominio de heterocromatina constitutiva de 16 kb se ensambla y mantiene (37) (Fig. 11). En base a las características anteriormente descritas, nosotros nos preguntamos ¿Cómo es que el dominio de eucromatina

transcripcionalmente activo del gen *FOLR1*, se protege de la propagación del dominio de las 16 kb de heterocromatina contigua?



**Figura 11. Modelo de la formación y mantenimiento de las 16 Kb de heterocromatina.** La transcripción de las secuencias repetidas CR1 recluta a la maquinaria del ARNi. Cada caja negra es una secuencia repetida CR1 y las flechas amarillas son secuencias repetidas LTR. Esta heterocromatina constitutiva tiene la capacidad de propagarse y afectar al dominio del gen *FOLR1*. Modificado de (37).

Debido a que la heterocromatina puede propagarse, puede interferir con el establecimiento de los programas de expresión de genes, por tanto, dicha propagación debe estar regulada o contrarrestada. A continuación describiré los mecanismos de los elementos codificados en el ADN conocidos como “insulators” (38). Los “insulators” son secuencias de ADN capaces de formar y mantener dominios funcionales de eucromatina y contrarrestar la propagación de la heterocromatina. Iniciaré describiendo el papel de los factores transcripcionales pioneros, en el establecimiento de dominios de eucromatina en el genoma. Luego,

hablaré de las funciones de los “insulators” en la formación y mantenimiento de dominios de eucromatina y su papel en la organización de la cromatina al interior del núcleo. Finalmente, discutiré la contribución de los “insulators” en la expresión regulada de los genes.

## **¿Cómo se establece un dominio transcripcionalmente activo de eucromatina?**

### **Factores transcripcionales pioneros.**

Al interior del núcleo, el genoma se organiza y distribuye en dominios alternos de eucromatina (transcripcionalmente activos) y de heterocromatina (facultativa y constitutiva, que son transcripcionalmente inactivos), de manera tal que se facilita la correcta expresión coordinada de los genes que contienen para cada uno de los linajes celulares que conforman un organismo (39, 40).

Debido a que la cromatina tiende a obstruir parcial o totalmente la secuencia del ADN, los sitios de unión para los factores transcripcionales se encuentran ocultos, por ejemplo, por yuxtaposición entre los nucleosomas. Esto supone que debe haber un continuo remodelamiento en eucromatina. ¿Cuál es el factor primario que tiene la capacidad de reconocer su sitio blanco en el genoma y entonces reclutar al resto? Los factores transcripcionales pioneros, son capaces de reconocer su sitio de unión en el ADN, en un contexto de heterocromatina y permitir el establecimiento del resto de los factores transcripcionales ya sea reclutándolos o facilitando su unión. Analizando la presencia de algunos factores transcripcionales conocidos en sus sitios blanco en todo el genoma, se demostró que una baja proporción de ellos se encontraban ocupando sus sitios disponibles en el genoma, y que dicha unión era dependiente del estado de compactación de la cromatina (6). Complementando estas observaciones, ensayos de

reconstitución de cromatina *in vitro*, mostraron que la mayoría de los factores transcripcionales, no son capaces de unirse con una alta afinidad a su sitio blanco cuando existe un solo tipo de éstos, en cambio, cuando se encuentran en una mezcla que incorpora varios tipos diferentes de factores de transcripción, la afinidad por su sitio de unión en el ADN aumenta dramáticamente. Es decir, en cuestión de afinidad por su sitio blanco, los factores transcripcionales funcionan mejor cuando existe un modelo de “unión cooperativa”, en comparación con un modelo de “unión individual” (6) (Fig. 12).

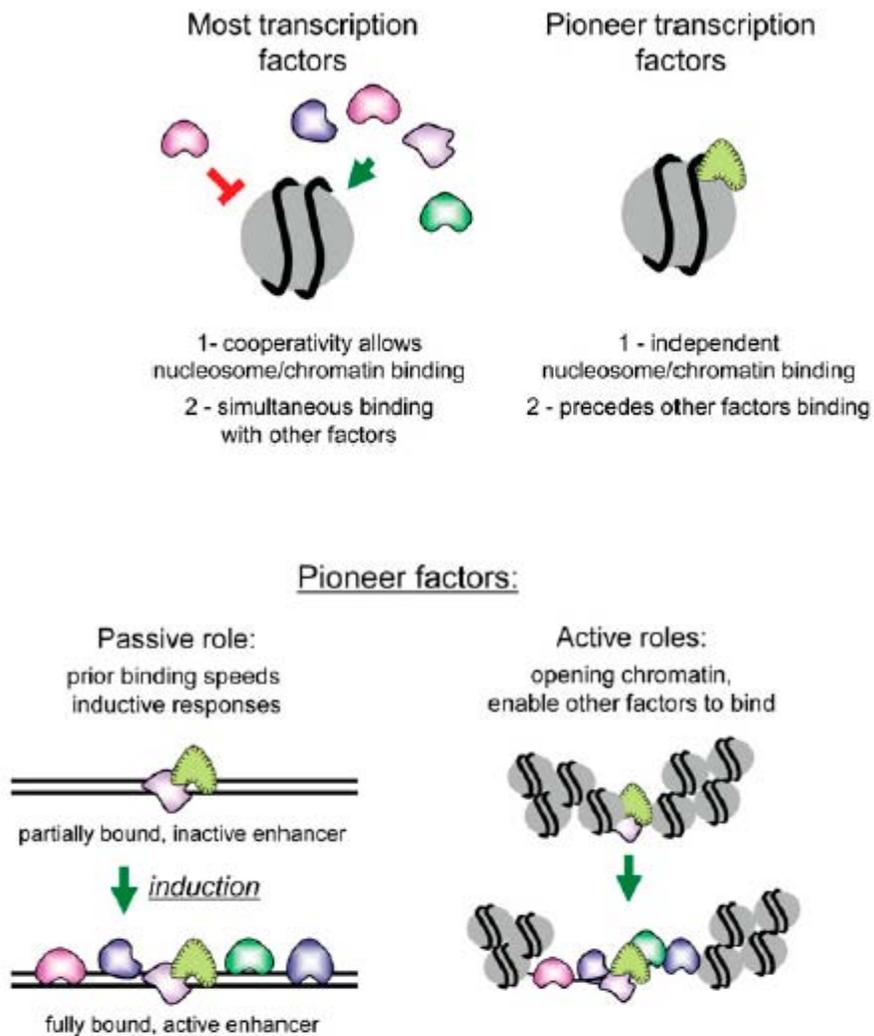
Por otro lado, es sabido que los ensayos de “footprinting” *in vivo*, permiten analizar secuencias de ADN que se encuentran protegidas por factores transcripcionales durante la digestión con endonucleasas como DNasa I. Mediante la secuenciación de los fragmentos de ADN no digeridos, se pueden intuir cuales son los factores transcripcionalmente unidos. Este fundamento, permitió observar que en ciertos elementos regulatorios del genoma, como los promotores y los enhancers, los factores transcripcionales FoxA y GATA, pueden unirse primero al ADN y facilitar el acceso al resto de los factores transcripcionales que tienen sitios de unión en la misma región regulatoria (6, 41, 42).

Con base en las observaciones anteriores, surgió el concepto de los factores transcripcionales primarios o pioneros, con la habilidad de, en el contexto de la cromatina cerrada (heterocromatina), reconocer su sitio blanco y así, dar paso al establecimiento de un dominio de eucromatina, mediante el reclutamiento de complejos de factores transcripcionales.

Entre los mecanismos por los que se puede dar la apertura de heterocromatina en el caso de FoxA1 se sabe, que además de su dominio de unión al ADN, tiene también en su extremo carboxilo terminal un dominio de unión a histonas que es imprescindible para su función en experimentos *in vitro* (42). Además, existen en el genoma, sitios de unión fuertes, medianos y débiles para FoxA1, lo cual es determinante para la función de abrir la cromatina (43) (Fig. 12).

Algunas características del genoma pueden contribuir con la función de los factores transcripcionales pioneros: a) existen regiones en el genoma que son desestabilizadoras de la cromatina a nivel local, por ejemplo, las secuencias abundantes en nucleótidos dA-dT (44), b) algunas secuencias promotoras pueden de manera intrínseca, desestabilizar a los nucleosomas (45), c) existen regiones del genoma que favorecen la incorporación variantes de histonas, por ejemplo, H2AZ y H3.3 que pueden desestabilizar a los nucleosomas a nivel local (46).

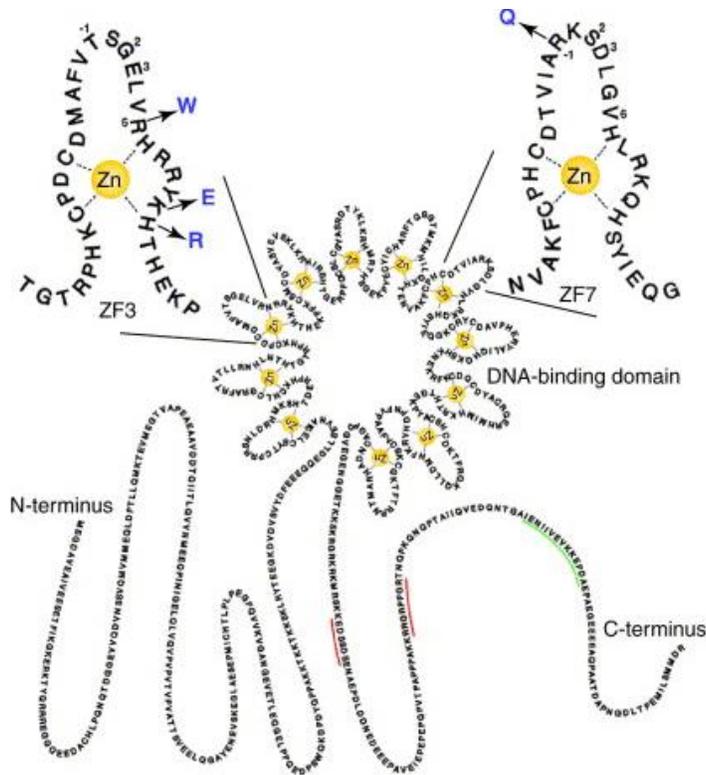
La lista de factores transcripcionales pioneros ha aumentado y se ha jerarquizado, es el caso de CTCF que es un factor nuclear con múltiples funciones en el genoma, una de ellas es la de proteína “insulator”, y se sabe que es un factor transcripcional pionero con una función en la cascada de eventos de apertura de la cromatina primaria a la función de FoxA1(47, 48).



**Figura 12. Diferencias entre los factores de transcripción “clásicos” y los factores transcripcionales pioneros.** La heterocromatina es el escenario al que los factores transcripcionales pioneros se enfrentan. CTCF es un factor nuclear multifuncional con actividad de factor transcripcional pionero. Tomado de (2).

## **El factor nuclear CTCF es una proteína “insulator”**

El factor nuclear CTCF, fue descrito por ser un factor transcripcional involucrado en la regulación del gen *c-myc* en el pollo (49-51). CTCF es ubicuamente expresado y altamente conservado en eucariontes (52, 53). La proteína CTCF tiene un extremo amino terminal, un extremo carboxilo terminal y un dominio central de 11 dedos de zinc como dominio de unión al ADN, estos tres dominios son sujetos de modificaciones post-traduccionales como fosforilación (54), poli-(ADP)-ribosilación (55) y SUMOilación (53). El “knockout” de CTCF es letal en ratón en etapas tempranas del desarrollo (E4.5-E5.5) (56). Nuestro grupo de investigación ha contribuido en el estudio de CTCF, dando evidencia de una región de unión a ARN (RBR), que es diferente a su dominio de unión al ADN y que permite que CTCF pueda multimerizarse, dando bases para las funciones de establecer dominios de cromatina *in vivo* (57) (Fig. 13). CTCF tiene múltiples funciones en el genoma y una de ellas es la de proteína “insulator”. En la actualidad se consideran a los “insulators” como parte de los elementos de regulación vinculados a la formación de dominios y la organización del genoma eucarionte (40).



**Figura 13. El factor nuclear CTCF.** Entre las funciones del factor nuclear CTCF están: Factor transcripcional, impronta génica, “insulator” y factor transcripcional pionero. Tomado de (58).

La cromatina no se organiza al azar, ha sido demostrado que tiene un patrón de organización tejido-específico, forma territorios cromosómicos y su organización nuclear se puede correlacionar con la actividad transcripcional. (59, 60). El genoma contiene genes que se transcriben de manera constitutiva y genes que lo hacen de manera regulada en un espacio y tiempo determinado. Los genes que se regulan en tiempo y espacio suelen organizarse en dominios cromatínicos (61-63). En algunos de estos dominios cromatínicos existen elementos que los delimitan llamados “insulators” o elementos barrera y son secuencias de ADN que promueven la formación y mantenimiento de un dominio “abierto” y evitan la acción inespecífica en los genes contenidos en él, por parte de otros elementos de

regulación ajenos a dicho dominio. Funcionalmente los “insulators” se han definido en base a dos propiedades: 1) tener la capacidad de bloquear la acción positiva de un enhancer sobre su promotor y 2) tener la capacidad de proteger a un transgen de los efectos de posición cromosomal, esto es, protege contra la propagación de la heterocromatina (64, 65).

En vertebrados el primer “insulator” descrito es cHS4 que define extremo 5´ del dominio  $\beta$ -globina de pollo y su función involucra la participación de CTCF (66). Al igual que cHS4 en pollo, también se demostró que CTCF se encontraba en el “insulator” del loci *Igf2/H19* en humano, controlando su impronta mediante la unión diferencial de CTCF al ICR (del inglés: Imprinting Control Region) controlado por su estado de metilación de ADN (67).

Funcionalmente, CTCF permite establecer contactos a larga distancia, por ejemplo, entre genes y promotores. Es el caso del enhancer *H19* y el promotor *Igf2* del cromosoma 7 en humanos, que pueden interactuar con otros genes activos de diferentes cromosoma como el locus *Wsb/Nf1* del cromosoma 11. Es decir, CTCF permite que colocalicen estos elementos génicos en el mismo espacio tridimensional del núcleo (68). La delección del sitio de CTCF en el ICR de *Igf2/H19*, tiene como consecuencia la pérdida de ésta interacción inter-cromosomal con cambios en la expresión de los genes de ambos locus (68). Por lo tanto, los “insulators” no solo contribuyen a la forma que tiene la cromatina en el núcleo sino también al establecimiento de programas de expresión génica de las células.

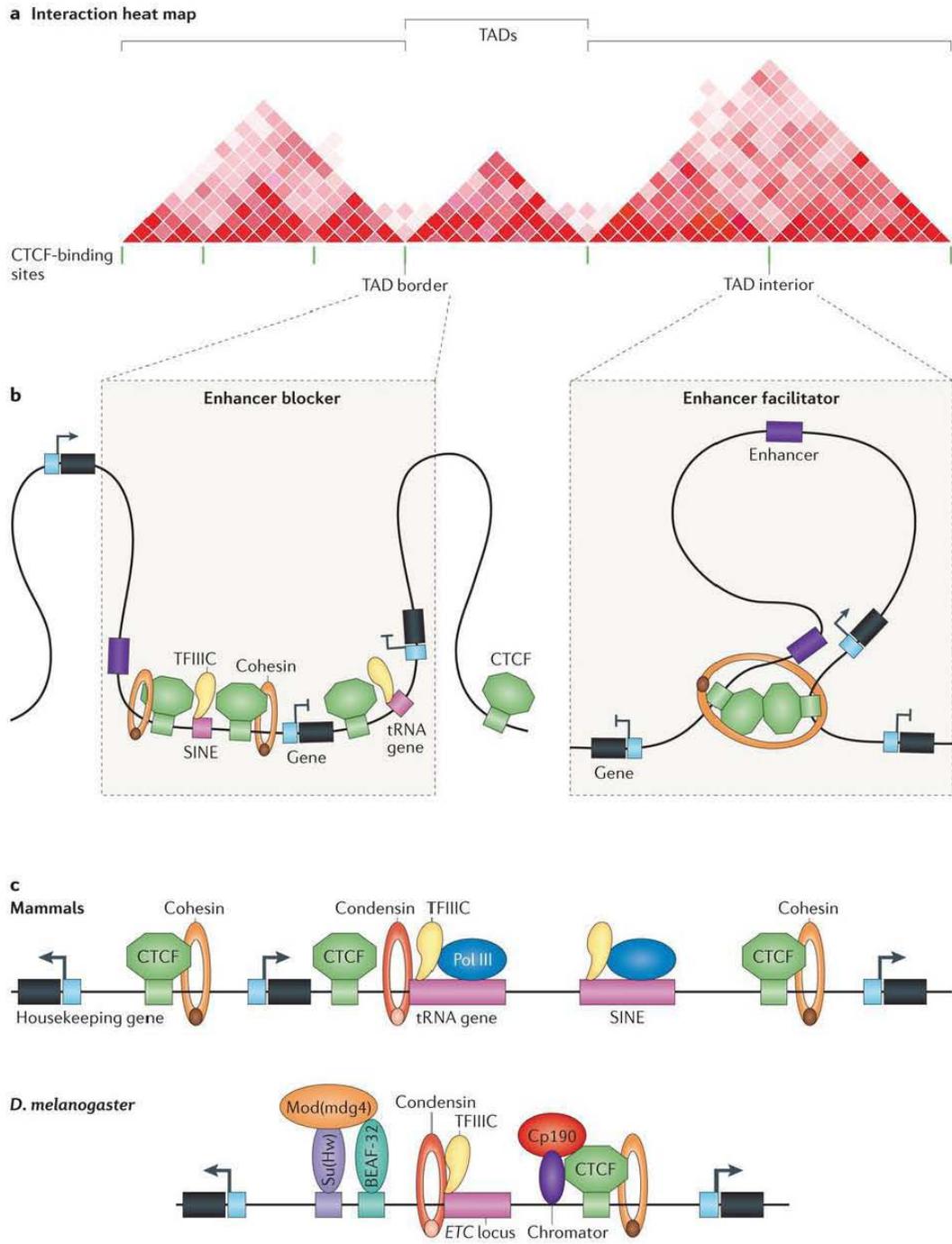
## **CTCF y datos genómicos que apoyan el concepto de la organización del genoma en dominios.**

El ensayo de inmunoprecipitación de la cromatina a escala global, permitió observar que CTCF puede unirse a múltiples sitios en el genoma. En mamíferos CTCF se une entre 55,000 a 65,000 sitios en el genoma (69), de estos sitios 50% son intergénicos, 35% son intragénicos y 15% son cercanos a los promotores (70). Algunos sitios de unión de CTCF están conservados en vertebrados (71, 72), y en humanos el 41% de los sitios de unión en el genoma están en regiones condicionalmente metiladas en su ADN (71).

CTCF puede unirse a la matriz nuclear, lo cual le permite establecer la arquitectura de la cromatina en el núcleo (63). Los “insulators” forman los límites de dominios de cromatina que pueden ser transcripcionalmente activos (eucromatina) o inactivos (heterocromatina), por ejemplo, algunos sitios CTCF que se encuentran en la lámina nuclear, tienen una transcripción muy baja (63).

Utilizando la información proveniente del ensayo conocido como Captura Conformacional de Cromosomas o 3C, se ha propuesto que existen Dominios Topológicamente Asociados (TADs) que no solo contribuyen con la arquitectura del genoma, sino que son importantes para su función (73, 74). Los TADs son funcionalmente y geoméricamente definidos por asas de cromatina (75, 76). Desde la perspectiva de la arquitectura nuclear, la estructura mínima del genoma son los dominios de cromatina “TADs”, su tamaño puede ir desde 3 Mb hasta 40 kb con una mediana de 185 kb (77), pueden contener uno o varios genes y

además contienen todos los elementos regulatorios necesarios para la correcta expresión de sus genes (78), por tanto, la mayoría de los dominios tienen subdominios generados por contactos entre los diferentes elementos contenidos (77, 79), finalmente, los dominios de cromatina están definidos por “insulators” como CTCF, tRNAs, SINEs y genes constitutivos (80). Los “insulators” en consecuencia, contribuyen funcionalmente a la expresión regulada de los genes. La pérdida de un “insulator” genera desregulación de los genes que contienen y genes que le rodean (74).



**Figura 14. Los elementos barrera o “insulators” contribuyen a la topología nuclear. Las asas de cromatina son la base funcional de los TADs**

En conclusión, el establecimiento de programas de expresión de genes es determinado por la propagación del silenciamiento epigenético de manera regulada, lo cual involucra a los elementos barrera o “insulators” que tienen la capacidad de formar y mantener dominios de eucromatina. El estudio de las bases moleculares de estos dos grandes procesos epigenéticos, nos puede permitir entender un poco más sobre cómo es que un solo genoma puede generar a todos los distintos epigenomas que constituyen a un organismo.

En este trabajo de tesis, nos enfocaremos en la función de las regiones del ADN no-codificante. Abordaremos la propagación del silenciamiento epigenético que se origina a partir del dominio de 16 kb de heterocromatina, y estudiaremos la región regulatoria del gen *FOLR1*, para comprender a escala local la formación y mantenimiento de un dominio de eucromatina durante la diferenciación eritroide del pollo.

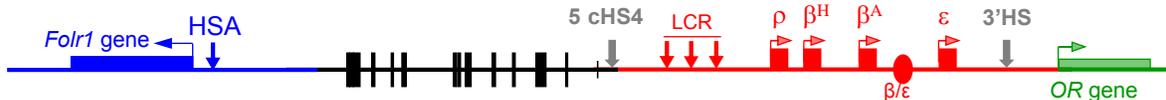
## Antecedentes particulares

¿Cómo se forman, mantienen y regulan los dominios cromatínicos en vertebrados?.

En nuestro grupo de investigación hemos adoptado como modelo de estudio el locus de los genes  $\beta$ -globina de pollo, el cual representa el paradigma de un grupo de genes con una expresión diferencial y que se organizan en un dominio cromatínico enmarcado por “insulators” (64, 81, 82).

### El dominio cromatínico $\beta$ -globina de pollo

El dominio  $\beta$ -globina abarca una región genómica de 33 kb (Fig. 15) (64). Cuenta con cuatro genes, un gen fetal ( $\rho$ ), un gen embrionario ( $\epsilon$ ) y dos genes que se expresan en etapas adultas ( $\beta^H$  y  $\beta^A$ ). Además, incluye un elemento de regulación tipo enhancer  $\beta^A/\epsilon$ , y una región de control del locus o LCR. El dominio está definido y enmarcado por dos “insulators” llamados 5'cHS4 y 3'HS (Fig. 15). Río arriba del dominio  $\beta$ -globina, se encuentra el dominio de 16 kb de heterocromatina y el dominio del gen receptor de folato *FOLR1* (81, 83).



**Figura 15. Región genómica que contiene al dominio  $\beta$ -globina de pollo y dominios adyacentes.** Se muestra en detalle los sitios hipersensibles a la DNasa I sobre la región regulatoria de *FOLR1* (HSA y HSA'); el dominio de las 16 kb de heterocromatina constitutiva, cada caja negra representa a una secuencia repetida tipo CR1, se muestra el sitio *HpaII* en rojo el cual corresponde con el límite 5' de éste dominio; el dominio de los genes  $\beta$ -globina enmarcado por los "insulators" 5'cHS4 y 3'HS los cuales tienen sitios de unión a CTCF, río abajo se encuentra el locus del gen receptor olfatorio con un programa de expresión cerebro-específico.

### **El dominio de las 16 kb de heterocromatina localizado entre el gen *FOLR1* y los genes $\beta$ -globina**

El dominio de las 16 kb de heterocromatina localizado entre el gen *FOLR1* y los genes  $\beta$ -globina presenta hiper-metilación del ADN evidente cuando se realiza una digestión genómica *in situ* con la enzima de restricción *HpaII*, hipoacetilación en las histonas H3 y H4, metilación de la lisina 9 de la histona H3 (H3K9me3) y es abundante en secuencias repetidas tipo CR1 enmarcadas por LTRs convergentes (37, 84) (Fig.11,15). Felsenfeld y colaboradores demostraron que las secuencias repetidas del dominio que abarca 16 kb de heterocromatina se transcriben generando ARNs no-codificantes induciendo el reclutamiento de la maquinaria de la vía del ARN de interferencia como mecanismo de formación, mantenimiento y propagación de heterocromatina (37, 84). Un aspecto complementario es que las 16 kb de heterocromatina representan un tipo particular y aparentemente único para su estudio dado que comparte todas las características de una

heterocromatina constitutiva, como la heterocromatina pericentromérica o telomérica (81, 83-86).

### **El dominio formado por el gen receptor de folato *FOLR1***

El locus del gen receptor de folatos (*FOLR1*) se encuentra río arriba de las 16 kb de heterocromatina, tiene un programa de expresión eritroide específico, el cual antecede a la activación de los genes  $\beta$ -globina, esto es, durante la etapa de pre-eritroblastos ó Unidad Formadora de Colonias Eritroides (CFU-E) en la diferenciación eritroide y el desarrollo del pollo (81). Contiene una secuencia tipo enhancer denominado HSA definido por un sitio fuerte de hipersensibilidad a la DNasa I, que no se observa en etapas avanzadas de la diferenciación eritroide (81) (Fig.15).

## **Planteamiento del problema**

Con base en todos estos antecedentes decidimos estudiar la regulación fina, a nivel epigenético, de la posible formación diferencial de dominios cromatínicos. En particular nos interesamos en tratar de entender los mecanismos involucrados en la propagación del silenciamiento epigenético a partir de un tipo particular de heterocromatina, en este caso en las 16 kb de heterocromatina, y la manera en que el dominio del gen *FOLR1* se protege de ésta para poder ser transcrito de manera regulada en una etapa específica de la diferenciación eritroide (Fig. 15).

## **Hipótesis**

Con base en la organización del genoma eucarionte en dominios cromatínicos, existe la presencia de un elemento frontera o “insulator” en el extremo 3’ del dominio del gen *Receptor de Folatos* el cual regula la propagación de las señales epigenéticas represivas provenientes de las 16 kb de heterocromatina contigua.

## Objetivo general

Determinar la presencia de un elemento tipo frontera (“insulator”) entre el dominio que incluye al gen *FOLR1* y el dominio conformado por 16 kb de heterocromatina con funciones regulatorias durante la diferenciación eritroide.

## Objetivos particulares

1. Demostrar la existencia de elementos frontera o “insulators” mediante la caracterización de la región intergénica localizada entre el dominio del gen *FOLR1* y las 16 kb de heterocromatina.
2. Confirmar si las 16 kb de heterocromatina situadas río abajo del dominio del gen *FOLR1* tienen la capacidad de propagarse y regular la estructura de la cromatina del dominio y la expresión del gen *FOLR1* de pollo.
3. Caracterizar y estudiar los mecanismos por los cuales un “insulator” contrarresta la propagación de la heterocromatina y como éste contribuye a la formación de un dominio tanto *in vitro* como *in vivo*.

## **Material y métodos**

### **Plásmidos y construcciones.**

El fragmento de ADN de 870 pb que contiene al “insulator” HSB-HSA', fue amplificado por PCR utilizando ADN genómico proveniente de cultivos primarios de eritrocitos de embriones de pollo de 10 días de desarrollo. Se utilizaron los oligonucleótidos Sentido: 5'-CCAGACACACACTGCTCCCAC y Antisentido: 5'-GGCATCCATGGGAAAAGGCTGC. El “insulator” HSB-HSA' fue clonado en el vector pGαD3, enmarcando al promotor  $\alpha^D$  y al cDNA de EGFP utilizando el sitio de restricción 5' EcoRI y el sitio 3' NheI-MluI respectivamente. Un sub-fragmento de 6.7 kb de ADN proveniente de 16 kb heterochromatina fue subclonado en pGαD3, utilizando a pSKHetero8.7 (amablemente donado por la Dra. M.N. Prioleau y el Dr. G. Felsenfeld) utilizando los sitios de restricción BglII-EcoRI.

### **Líneas celulares y cultivo celular.**

HD3 es una línea celular proveniente de eritroblastos de pollo (*Gallus gallus*) que fué transformada con el virus AEV (virus de la eritroblastosis aviar). Fue cultivada a 37°C, 5% CO<sub>2</sub>, con Dulbecco's Modified Eagle's Medium (DMEM, GIBCO) suplementado con 10% Fetal Bovine Serum, 2% de suero de pollo y antibióticos

(100 U Penicillin /0.1 mg Streptomycin/mL). 6C2 es una línea celular proveniente de pre-eritroblastos (CFU-E) de pollo (*Gallus gallus*) transformada con el virus AEV (virus de la eritroblastosis aviar). Fue cultivada a 37°C, 5% CO<sub>2</sub>, con  $\alpha$ -Minimum Essential Medium ( $\alpha$ -MEM, GIBCO) suplementado con 10% Fetal Bovine Serum, 2% de suero de pollo, 1mM HEPES (pH 7.2, SIGMA), 50 mM  $\beta$ -mercaptoethanol, y antibióticos (100 U Penicillin/0.1 mg Streptomycin/mL). HD24 es una línea celular multipotente eritroide-mieloide de pollo *Gallus gallus* transformada con el virus E26 y fue cultivada con Blastoderm medium (87). DT-40 es una línea celular de linfoma B de pollo *Gallus gallus* transformada con el virus ALV (virus de la leucosis aviar) y fue cultivada a 37°C, 5% CO<sub>2</sub>, Dulbecco's Modified Eagle's Medium (DMEM, GIBCO), con 2 mM L-glutamine y 50 mM  $\beta$ -mercaptoethanol, 10% tryptose phosphate broth, 10% Fetal Bovine Serum, 5% suero de pollo y antibióticos (100 U Penicillin/0.1 mg Streptomycin/mL). Los cultivos primarios de eritrocitos de embriones de pollo de 5 y 10 días (RBC) de desarrollo fueron obtenidos de vasos sanguíneos periféricos de embriones de huevos fertilizados (granjas ALPES, Tehuacán, Pue.). K562 es una línea celular humana eritroleucémica y fue cultivada a 37°C, 5% CO<sub>2</sub>, en RPMI medium con 2 mM Glutamine, 10% Fetal Bovine Serum, y antibióticos (100 U Penicillin/0.1 mg Streptomycin/mL).

## **Transfecciones transitorias y estables.**

1X10<sup>6</sup> células fueron resuspendidas en 800 µL de medio sin antibióticos y sin suero. Fueron sembrados en cajas 6-pozos y cada plásmido fue transfectado con lipofectamine 2000. Después de 48 horas de recuperación, las células fueron colectadas: a) En transfecciones transitorias se cuantificó la expresión de la Luciferasa midiendo su luminiscencia en un luminómetro y se utilizó a la Renilla como gen reportero para normalizar las transfecciones y la actividad de la luciferasa y b) En transfecciones estables fueron seleccionadas por antibióticos.

## **PCR-cuantitativa.**

El ARN de cada cultivo celular utilizado fue purificado con TRIzol Reagent (Invitrogen) y tratado con DNasa I (RQ1, Promega). Se utilizó Random Primer y oligo-dT para generar ADNc. La cuantificación del ARNm *FOLR1* se hizo mediante Real Time qPCR con SYBR Green (Sigma) y oligonucleótidos específicos (Sentido: 5'-GGATTGACCAGGCTGACAGC y Antisentido: 5'-GCGATTGGTTCCTGTTGCCC). El ARNn de *GAPDH* se utilizó como control normalizador (Sentido: 5'-TATCTTCCAGGAGCGTGA and Anti-sentido: 5'-TCAGATGAGCCCCAGCCTTC). Se utilizó el Sistema StepOne detection system

(Applied Biosystems) y se realizaron triplicados para cada muestra. Se utilizó el método comparativo  $\Delta\Delta C_t$  para cuantificar los niveles relativos de ARNm.

### **PCR duplex semicuantitativa.**

El ARN de cada cultivo celular fue aislado utilizando TRIzol Reagent (Invitrogen), y tratado con DNasa I (RQ1, Promega). El ADNc fue generado a partir de 5 $\mu$ g de ARN tratado con DNasa I (Reverse Transcription System, Promega) utilizando oligo-dT. Se realizó una PCR-radioactiva con [ $\alpha$ -P<sup>32</sup>] d-CTP radioactivo utilizando los siguientes oligonucleótidos específicos (FOLR1 Sentido: 5'-CCAGTTCTCTTTGCATGTGACG y Anti-sentido: 5'-GCAAGCGTCACTTCATCCAGG) se utilizó al gen constitutivo  $\beta$ -actina como un control endógeno de normalización de los resultados (B-ACT Sentido: 5'-CCAGACATCAGGGTGTGATG y Antisentido: 5'-GAACACGGTATTGTCACCAACTGG). Para cuantificar los niveles de transcritos, se utilizó un gel de acrilamida para separar los productos, el cual fue secado y expuesto en una pantalla sensible a [ $\alpha$ -P<sup>32</sup>] (Amersham-Biosciences). Se utilizó el programa ImageQuant para cuantificar los resultados.

### **Inmunoprecipitación de la cromatina-PCR dúplex.**

El ensayo de inmunoprecipitación de la cromatina ChIP, fue realizado como previamente se reportó en (88). Brevemente, 5X10<sup>7</sup> células fueron fijadas con

formaldehído al 1%, la cromatina fue fragmentada por sonicación y 4 µg de anticuerpo anti-CTCF fue utilizado (Anticuerpo generado en nuestro laboratorio) (89). El ADN inmunoprecipitado fue evaluado por PCR dúplex, utilizando un radioisótopo marcado con [ $\alpha$ -P<sup>32</sup>] d-CTP, utilizando oligonucleótidos específicos (HSA' Sentido: 5'-CCCACACTCACGGCCTTGTTG y Anti-sentido: 5'-GGCATCCATGGGAAAAGGCTGC; FII Sentido: 5'-GGTTGAAGAAAAGAAGCAGGC y Anti-sentido: 5'-CGTTCAGACGAAAGCGATCC y H1 Sentido: 5'-GGAAGTAGTTGGCAAGGTCCTCT y Anti-sentido: 5'-TCTTCTGCCCTGCCCGTAT). Para la cuantificación se resolvieron los productos de PCR en un gel de acrilamida que fue secado y expuesto en una pantalla sensible a [ $\alpha$ -P<sup>32</sup>] (Amersham-Biosciences). Se utilizó el programa ImageQuant para cuantificar los resultados. Se utilizó la siguiente fórmula: Enriquecimiento = (exp/control)Ab/(exp/control)IgG.

### **Conversión de ADN con bisulfito de sodio**

El ADN genómico fue aislado de cada cultivo celular utilizando el protocolo de fenol-cloroformo, 1.5 µg de ADN fue digerido con la enzima de restricción EcoRI antes de su conversión con bisulfito de sodio. La conversión con Bisulfito fue realizada como previamente se describió en (88). Los oligonucleótidos específicos para cada región amplificada fueron: promotor FOLR1 Sentido: 5'-GGGTTTTTGGTAGGTATTATTA-3' y Anti-sentido: 5'-

CTCCTAAAAATAAAATAACTCTC-3';	HSB	Sentido:	5'-
GGTTTTGGGTAGTTTGGGTTG-3'	y	Anti-sentido:	5'-
CCAAAAACAAACTATCACAACCTC-3';	HSC	Sentido:	5'-
GGGTATTGTTGTATAGGATAGT-3'	and	Anti-sentido:	5'-
ATAAAACAAACACAAATAATCCC-3';	Hetero1	Sentido:	5'-
GGGATTATTTGTGTTTTAT-3'	y	Anti-sentido:	5'-
CCACACTTAATTAACCTATTCTAC-3';	Hetero3	Sentido:	5'-
GTAGGTTTGGTATAATTTTAGTTA-3'	y	Anti-sentido:	5'-

AAAAATAACCTTAAAAAACATCC-3'). Los productos de PCR fueron clonados en el vector pGEMT-easy (Promega) y se secuenciaron al menos diez clonas utilizando oligonucleótidos específicos Sp6.

### **Generación de un ARN de interferencia chCTCF shRNAi.**

Se utilizó el programa <http://sirna.wi.mit.edu/> (90) para diseñar un shARNi chCTCF 5'-GCCAGCAGGGATACTTACA-3' utilizando el ARN-CTCF publicado en NCBI: M\_205332.4. El shARNi chCTCF fue clonado en el vector LL3.7 en los sitios de restricción HpaI-XhoI.

## **Generación de partículas Lentivirales para la reducción de los niveles del factor nuclear CTCF.**

El ARN de interferencia shARNi chCTCF fue utilizado para generar una reducción en los niveles globales del factor nuclear CTCF siguiendo el protocolo publicado en (91, 92). Brevemente células HEK293FT fueron utilizadas para producir partículas pseudolentivirales pLL3.7 control y shARNi chCTCF 5'-GCCAGCAGGGATACTTACA. Las células 6C2 fueron transducidas por 8 horas utilizando polybrene (8 µg/mL;Sigma). Los cultivos fueron seleccionados por 3 a 4 días con puromycin (5 µg/mL;Sigma) y utilizados en los experimentos que son descritos.

## **Ensayo de Hipersensibilidad al corte por la DNasa I.**

El ensayo de Hipersensibilidad a la DNasa I fue realizada en las líneas celulares 6C2 y HD3 como se publicó en (93).

## **EMSA**

El ensayo de retardo en la movilidad electroforética fue realizado como se describe en (89). Los oligonucleótidos utilizados en este ensayo son: HSA´ sentido 5'-

GACACCATTCTCACCCCATGGGATCGGAACCCAGGCGGGGGGAATCCCTGG



## **Ensayo de detección de actividad “insulator” *in vivo* en el pez cebra *Danio rerio* Zebrafish (ZED).**

El vector de Zebrafish Enhancer Detection ZED fue utilizado para probar la capacidad “insulator” de HSA´ *in vivo* (95). Se micro-inyectaron embriones de pez cebra en la etapa de una célula de desarrollo como se describe en (95). Se analizaron y cuantificaron 10-30 individuos de pez cebra por cada condición.

## **Generación del anticuerpo anti-cCTCF de pollo.**

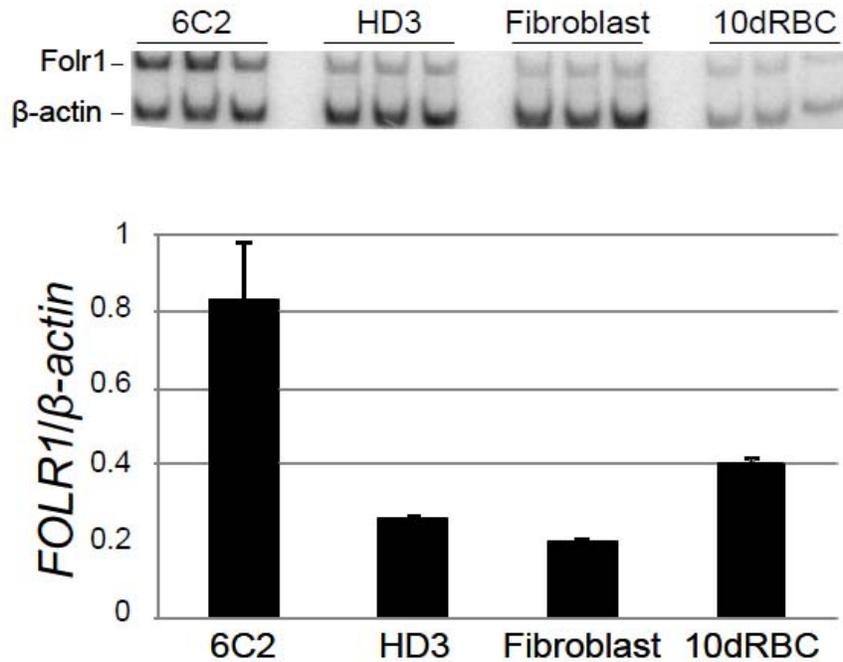
Se utilizó un fragmento del ADNc de CTCF de pollo que codifica para los residuos de aminoácidos 82 a 233 que fueron clonados en BamHI-XhoI en el plásmido pET28(b), luego fue expresado en *E. coli* cepa BL21(DE3) y se aisló y purificó la proteína recombinante, la cual se utilizó para inmunizar un conejo hembra White New Zeland con un esquema de tres refuerzos de 250 µg del fragmento recombinante chCTCF en los días 14, 28 y 42. El suero colectado fue analizado y validado en su inmunoreactividad anti-chCTCF y entonces purificado utilizando el kit AminoLink Plus Immobilization (Thermo Scientific). Ver Anexo A.

## Resultados

### 1. Caracterización de un elemento frontera “insulator” localizado entre el dominio del gen *FOLR1* y las 16 kb de heterocromatina.

#### ***Cambios en la expresión del gen FOLR1 correlacionan con cambios en el estado de la cromatina.***

Cuando un gen se expresa, sus elementos regulatorios se encuentran en un estado de la cromatina laxo (eucromatina). En el caso de *FOLR1*, nos preguntamos si: ¿existen cambios en la estructura de la cromatina que acompañan la expresión y el silenciamiento de éste gen? Para contestar esta pregunta, primero analizamos el patrón de expresión de *FOLR1* en la diferenciación eritroide representada por las líneas celulares 6C2 (CFU-E), HD3 (eritroblastos) y 10dRBC (eritrocitos), mediante RT-PCR semi-cuantitativa (Fig.16 y 17B). Consistente con estudios previos, confirmamos la expresión del gen *FOLR1* en la etapa de CFU-E (células 6C2), mientras que en la etapa de eritroblastos (células HD3) observamos una expresión basal del gen *FOLR1* (Fig. 16) (81).

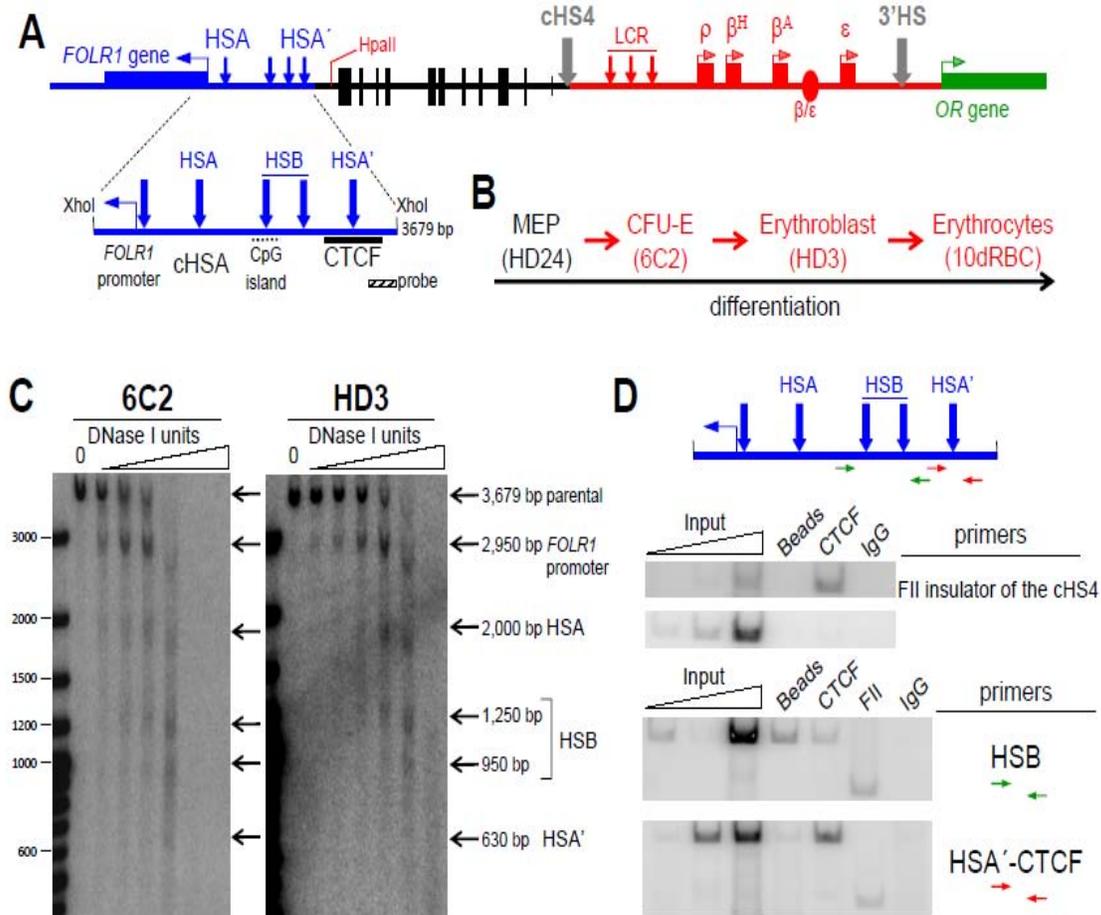


**Figura 16. Perfil de Expresión del gen *FOLR1* en diferentes tipos celulares.** Se muestra una PCR-dúplex representativa y se graficaron los resultados normalizados y su desviación estándar. Se realizaron tres replicas biológicas independientes, cada una por triplicado (n=3).

Con el objetivo de caracterizar el estado de compactación de la estructura de la cromatina (eucromatina o heterocromatina) realizamos un estudio comparativo, mediante un ensayo de hipersensibilidad al corte por la DNasa I abarcando la región intergénica entre el dominio del gen *FOLR1* y el dominio que comprende a las 16 kb de heterocromatina tanto en la etapa de expresión CFU-E (células 6C2) como en la etapa subsecuente de la diferenciación eritroide, en la cual, el gen *FOLR1* ya no se expresa (células HD3) (Fig. 17A y 17B). Aislamos núcleos de células 6C2 y HD3, los cuales fueron incubados *in situ* con concentraciones crecientes de la endonucleasa DNasa I (Fig. 17C). Utilizamos una

digestión genómica (*Xho*I) que nos permitió obtener el patrón de hipersensibilidades a la DNasa I en una resolución óptima y analizamos los resultados por Southern-blot utilizando una sonda que nos permitió identificar la región comprendida entre el promotor del gen *FOLR1* y el sitio HpaII localizado en el extremo 5' de las 16 kb de heterocromatina (Fig. 17A y 17C). En la línea celular 6C2 observamos cinco sitios de hipersensibilidad: tres correspondientes al promotor del gen *FOLR1*, y al enhancer HSA y al sitio HSA' (sin función) previamente descritas (81, 82). Por otra parte, dos nuevos sitios de hipersensibilidad que en conjunto describimos como HSB y fueron descubiertos en este trabajo de investigación (Fig. 17C). Resulta importante resaltar que en las células HD3 observamos que los mismos sitios de hipersensibilidad empiezan a aparecer cuando utilizamos concentraciones más altas de DNasa I, ver carril 0.3 vs carril 2.5, lo cual indica un estado de la cromatina más compacta en la etapa más avanzada de diferenciación eritroide representada por las células HD3 (Fig. 17C). En otras palabras, en las células HD3 la cromatina de la región regulatoria del gen *FOLR1* está más compacta (heterocromatina) con respecto a las células 6C2 (eucromatina). Estos datos sugieren que el gen *FOLR1* el cual se expresa únicamente durante la etapa temprana de la diferenciación eritroide representada por las células 6C2 (CFU-E), se encuentra con una configuración de la cromatina laxa y por lo tanto, permisiva para la transcripción (eucromatina) (Fig. 16 y 17C). En cambio, en una etapa posterior (en la línea HD3) éste gen y su región regulatoria adopta una configuración de la cromatina más compacta (heterocromatina) (Fig. 17B y 17C). De acuerdo a nuestros resultados el sitio de hipersensibilidad HSA' es el que mostró un arreglo más compacto de la estructura

de la cromatina seguida por los sitios de hipersensibilidad HSB en eritroblastos (células HD3), esto sugiere que el silenciamiento epigenético pudiera propagarse a partir del costado 5' de las 16 kb de heterocromatina. Una explicación alternativa es que el elemento "enhancer" HSA deja de propagar una estructura de la cromatina abierta, siendo primero afectados los elementos estructurales más cercanos a la heterocromatina, es decir HSA' seguido por HSB (Fig. 17A y 17C).

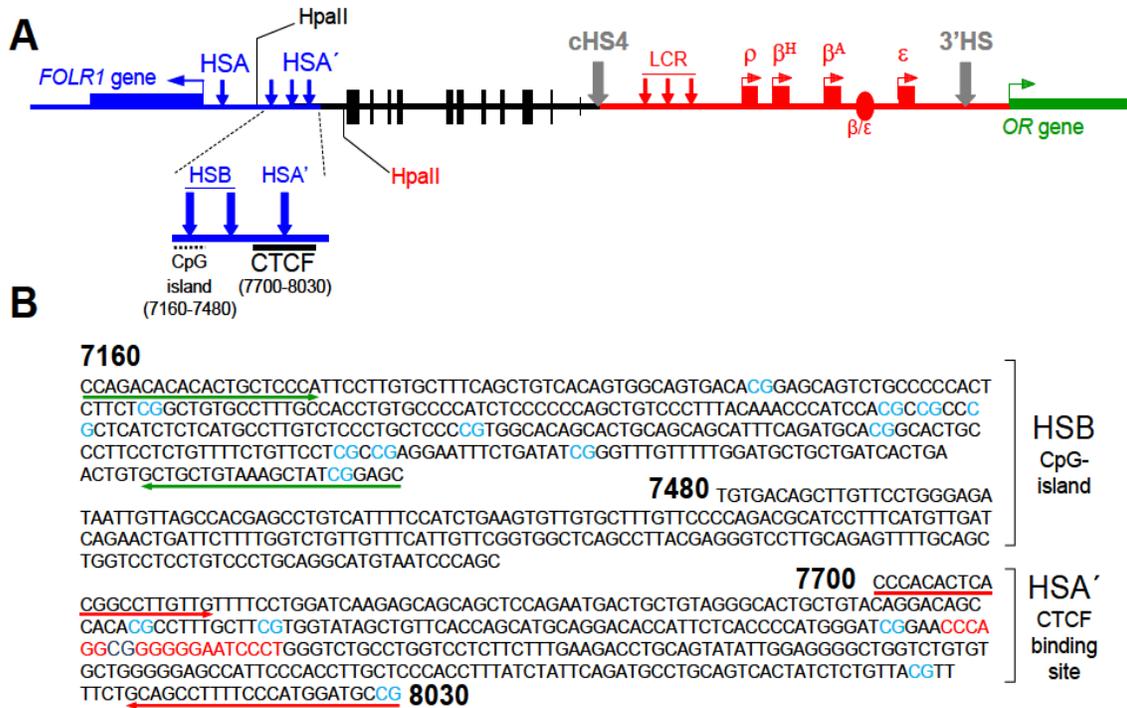


**Figura 17. La expresión del gen *FOLR1* coincide con un estado de la cromatina laxo, permisivo a la transcripción.** (A) Esquema de la región genómica del gen *FOLR1*. Se muestra en detalle los sitios hipersensibles a la DNasa I sobre la región regulatoria del gen *FOLR1* (HSA, HSB and HSA'); el dominio de las 16 kb de heterocromatina constitutiva, cada caja negra representa a una secuencia repetida tipo CR1, se muestra el sitio *Hpa*I en rojo el cual corresponde con el límite 5' de éste dominio; el dominio de los genes  $\beta$ -globina enmarcado por los "insulators" 5'cHS4 y 3'HS los cuales tienen sitios de unión a CTCF, río abajo se encuentra el locus del gen receptor olfatorio con un programa de expresión cerebro-específico. (B) Líneas celulares transformadas de pollo utilizadas en éste estudio y su correspondiente etapa en la diferenciación eritroide. En negro se muestran a las células HD24 que corresponden con un progenitor mieloide /eritroide y en rojo las líneas celulares arrestadas en diferentes etapas de la diferenciación eritroide. (C) Análisis comparativos de los sitios de hipersensibilidad a la DNasa I en núcleos aislados de células 6C2 y células HD3. Unidades de DNasa I utilizadas en cada condición son: 0.3; 0.7; 1.25; 2.5; 5 and 10. En (A) se muestra la sonda utilizada en éste análisis. (D) Inmunoprecipitación de la cromatina (ChIP) muestra el enriquecimiento *in vivo* del factor nuclear CTCF en el sitio HSA'. Como control positivo se analizó la región correspondiente al "insulator" 5'cHS4 del dominio  $\beta$ -globina (FII: sitio de unión a CTCF). El sitio HSB no se encuentra enriquecido de CTCF, consistente con datos previos. Se muestra un ChIP representativo de cuatro ensayos independientes.

**Identificación de un nuevo sitio de unión a CTCF in vivo en el límite entre el dominio del gen *FOLR1* y el dominio de las 16 kb de heterocromatina.**

El ensayo comparativo de hipersensibilidad al corte por la endonucleasa DNasa I, realizado con una digestión genómica (*XhoI*) nos permitió apreciar, en una alta resolución, que el elemento HSB-HSA' es el sitio regulatorio/estructural más cercano al límite del dominio de las 16 kb de heterocromatina (Fig. 17C). Los "insulators" se caracterizan por ser sitios estructurales e hipersensibles a la DNasa I, por lo tanto, HSA' puede ser el límite del dominio del gen *FOLR1*. CTCF es un factor nuclear multifuncional que en vertebrados se ha relacionado íntimamente con elementos tipo "insulators". Es importante recordar, que en el extremo 3' del dominio de las 16 kb de heterocromatina se encuentra el "insulator" 5'cHS4, un "insulator" muy bien caracterizado, al cual se une CTCF (sitio FII) y que marca el inicio del dominio de los genes  $\beta$ -globina (Fig. 15).

Analizamos la secuencia correspondiente a la región HSB-HSA' en búsqueda de sitios de unión a CTCF en el genoma del pollo. Utilizamos la base de datos de ChIP-seq realizado en células 10dRBC de pollo previamente publicada por nuestro laboratorio en la cual, se reportó el motivo de unión de CTCF en el genoma de pollo (72). Encontramos un motivo de unión de 20 pb en la secuencia de HSA' para CTCF (Fig.11) (72).



**Figura 18. Secuencia de la región HSB-HSA'.** (A) Esquema general de los tres dominios analizados. Se detallan las características de HSB and HSA'. (B) Secuencias correspondientes a las hipersensibilidades a la DNasa I HSB-HSA'. subrayados? En verde y rojo se representan los oligonucleótidos sintéticos utilizados en los ensayos de ChIP y se muestra el motivo de unión a CTCF en rojo.

En las secuencias asociadas a la región HSB no fue posible identificar un motivo de unión a CTCF. Para validar la unión *in vivo* de CTCF, realizamos ensayos de inmunoprecipitación de la cromatina (ChIP) utilizando un nuevo anticuerpo anti-CTCF de pollo generado en nuestro laboratorio recientemente (Fig. 19 y Anexo A) (89). El ensayo de ChIP fue realizado en células de pollo correspondiente a la etapa CFU-E para evaluar la presencia de CTCF en HSA' y como control interno analizamos la secuencia de HSB. Los resultados de este estudio muestran que CTCF se enriquece preferencialmente en la región asociada al sitio HSA' (Fig. 17D, 20, 23, 24). Adicionalmente se repitió el mismo ensayo de

manera cuantitativa en todas las etapas de la diferenciación eritroide HD24, 6C2, HD3, y cultivos primarios de eritrocitos de 5 días y 10 días de desarrollo del pollo (5dRBC y 17dRBC) (Fig. 20, 23 y 24). Utilizamos como control negativo una línea no-eritroide DT40 (línea celular linfoide). Los resultados muestran que solo en las células 6C2 existe un enriquecimiento *in vivo* significativo de CTCF sobre la hipersensibilidad HSA' (Fig. 20, 23 y 24).



En conclusión, toda esta serie de datos muestran que el sitio HSA' presenta un motivo de unión a CTCF y que su mayor enriquecimiento se observa en la etapa CFU-E, que corresponde a una etapa temprana de la diferenciación eritroide (células 6C2).

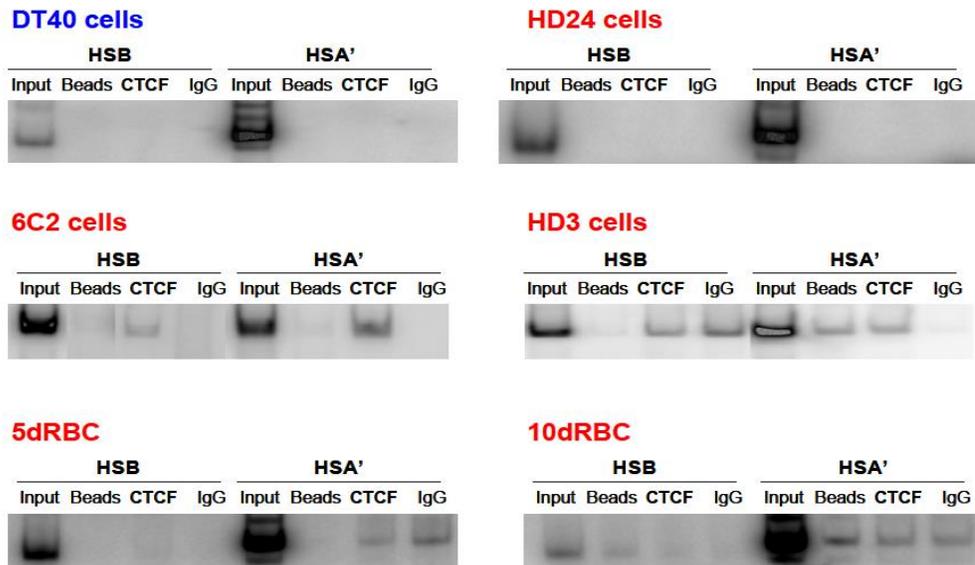


Figura 20. Análisis comparativo de inmunoprecipitaciones de la cromatina utilizando anticuerpo anti-CTCF en las regiones de HSB y HSA'. El enriquecimiento de CTCF se encuentra presente solo en la etapa de 6C2 y es eritroide específico. DT40 es una línea linfocítica utilizada como control no eritroide.

***El elemento HSB-A' tiene la capacidad de bloquear señales de regulación ajenas al dominio que enmarca.***

Los elementos tipo “insulator” tienen la capacidad de bloquear señales de regulación inespecíficas provenientes de otros dominios. Por esta razón decidimos explorar la capacidad de las secuencias HSB-HSA' para bloquear al enhancer  $\beta^A/\epsilon$  de la acción del promotor  $\beta^A$ , utilizando a la *Luciferasa* como gen reportero, en

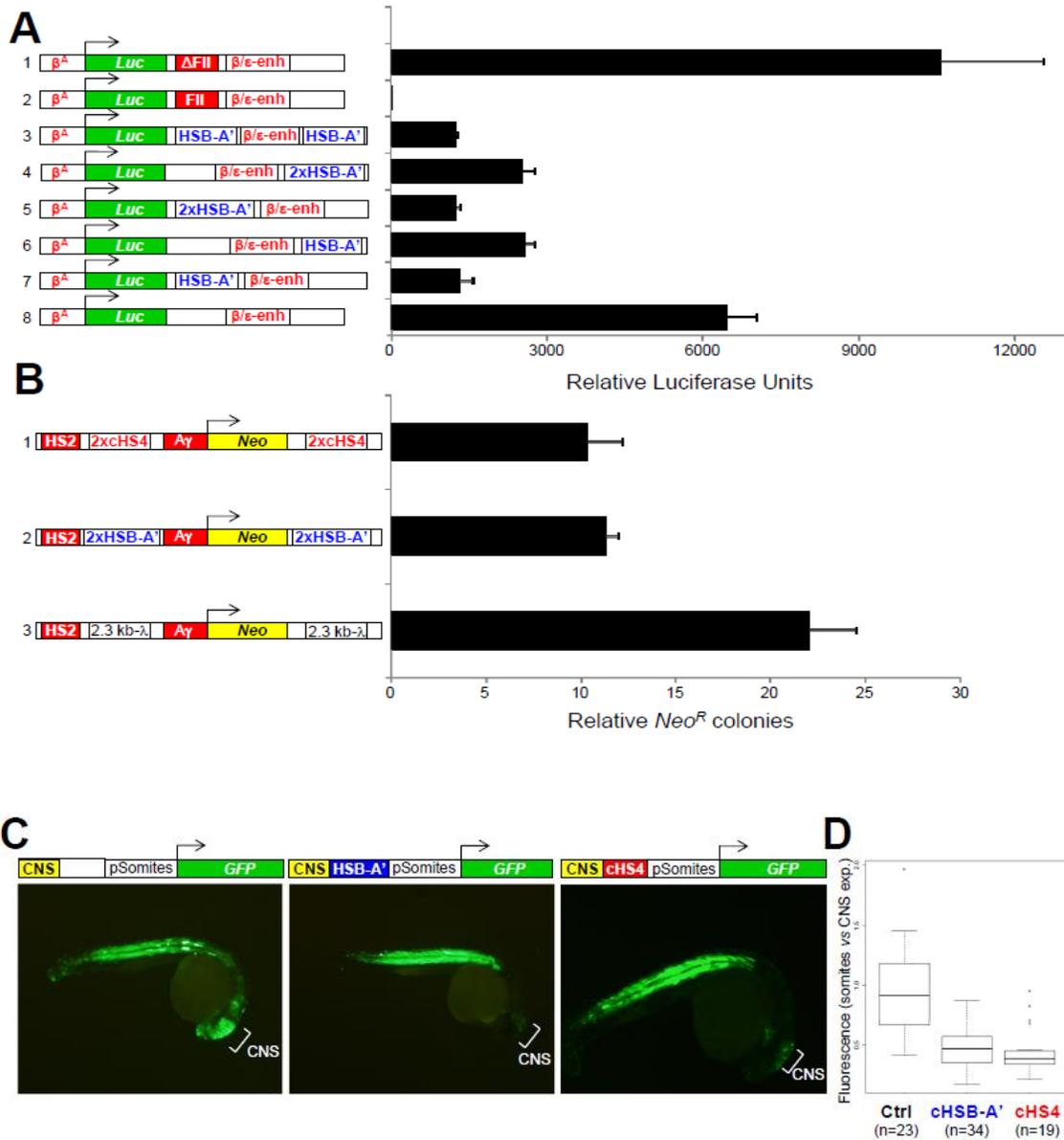
transfecciones transitorias realizadas en células 6C2. Por otra parte probamos el efecto de HSB-HSA' de manera sencilla (una copia), como 2xHSB-HSA' (dos copias), y lo colocamos interrumpiendo o enmarcando la comunicación entre el enhancer  $\beta^A/\epsilon$  y el promotor  $\beta^A$  (Fig. 21A). Como controles utilizamos al sitio FII (control positivo) que corresponde al sitio de unión a CTCF del "insulator" 5'cHS4 del locus  $\beta$ -globina y una versión de la secuencia mutada FII en el sitio de unión de CTCF (control negativo). El resultado de este experimento demostró que HSB-HSA' tiene la capacidad de bloqueo de enhancer (Fig. 21A).

Posteriormente decidimos analizar la capacidad de bloqueo de enhancer de HSB-HSA' utilizando el ensayo en colonia que nos permite analizar estas secuencias tomando en consideración el contexto cromatínico del sitio de inserción (39). Este ensayo fue realizado en la línea celular eritroleucémica humana K562 (66). Utilizamos al cassette de resistencia a la Neomicina como gen reportero bajo el control del promotor  $\beta^A$  y al enhancer HS2 del LCR  $\beta$ -globina humano (Fig. 21B). Como control negativo utilizamos una secuencia proveniente del genoma del fago lambda y como control positivo incorporamos al vector, el "insulator" 5'cHS4 ó el elemento HSB-HSA'. El resultado mostró que tanto nuestro control positivo 5'cHS4 como HSB-HSA' causan una reducción en el número de colonias resistentes al medio de cultivo con geneticina, lo que confirma una actividad de bloqueo de enhancer por parte del elemento HSB-HSA' (Fig 21B).

Para corroborar *in vivo* la función de bloqueo de enhancer del elemento HSB-HSA', realizamos este ensayo en peces cebra transgénicos (95). Para ello utilizamos el vector ZED (Zebrafish Enhancer Detection) que integra como reportero al gen *EGFP* controlado por el promotor de actina aislado de *Xenopus*

*laevis* y el enhancer irxZ48 del sistema nervioso central de *Danio rerio* (95). En este sistema el enhancer irxZ48 dirige una fuerte expresión de *EGFP* en el sistema nervioso central y en las somitas del pez cebra (Fig. 21C). En este ensayo utilizamos al elemento HSA'-HSB y al "insulator" 5'cHS4 como (control positivo). Se micro-inyectaron embriones de pez cebra en la etapa de una célula y 10-30 individuos fueron analizados y cuantificados los niveles de expresión de *EGFP* en cada condición tanto en sistema nervioso central como en las somitas (Fig. 21C y 21D). Cabe mencionar que la presencia de la EGFP en somitas permite normalizar nuestros datos de fluorescencia en relación con la señal asociada al sistema nervioso central (95). El resultado mostró que HSB-HSA' tiene la capacidad de bloquear a un enhancer *in vivo* (Fig. 21C y 21D).

En resumen, estos tres experimentos demostraron que el elemento HSB-HSA' tiene la capacidad de bloquear a un enhancer, es decir, tiene una de las dos propiedades funcionales que definen a los "insulators". A continuación decidimos explorar la segunda propiedad que tiene que ver con la capacidad de proteger a un transgen de los efectos de posición cromosomal (CPE).

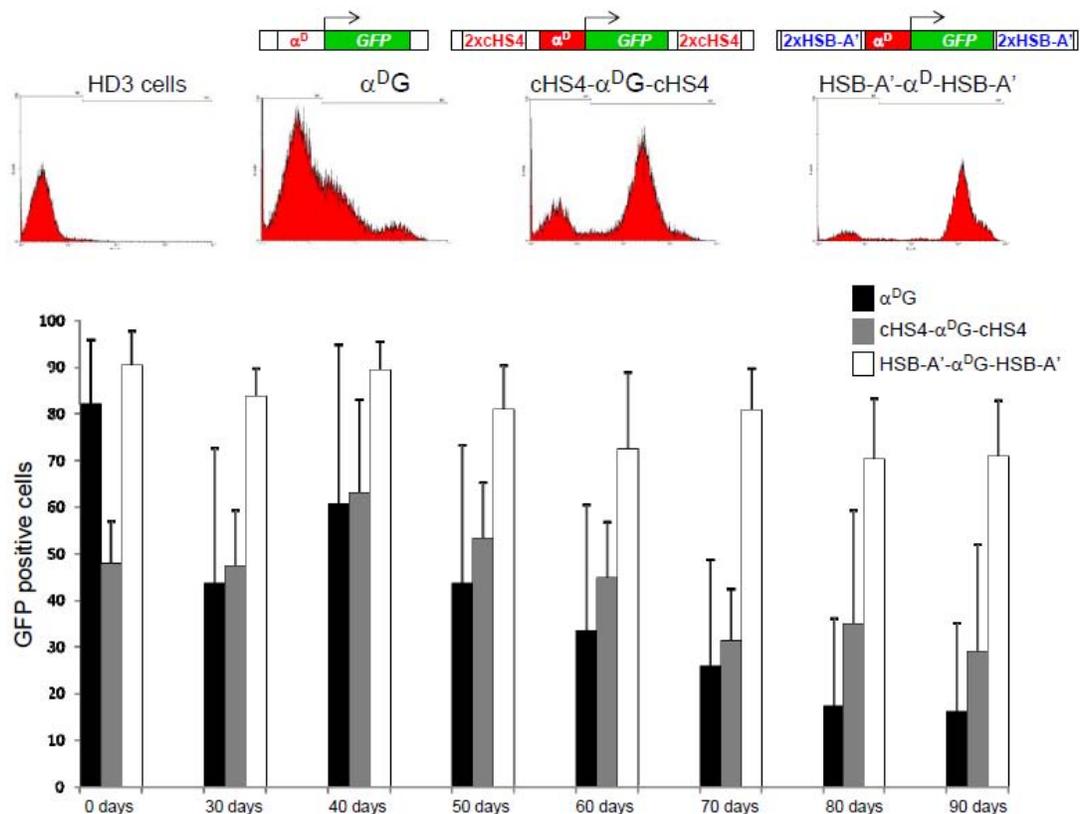


**Figura 21. HSB-HSA' tiene la capacidad de bloquear a un enhancer *In vitro* e *in vivo*.** (A) Ensayo de bloqueo de enhancer en células 6C2 transfectadas transitoriamente con plásmidos que contienen al gen reportero *Luciferasa* (*Luc*) dirigido por el promotor del gen β<sup>A</sup> y el enhancer β<sup>A</sup>/ε. Control positivo FII del "insulator" 5'cHS4 del dominio β-globina y como control negativo una mutante del sitio (ΔFII). (B) Ensayo en colonia en la línea celular eritroleucémica humana K562. Colonias resistentes a Neomicina fueron contadas y graficadas de tres transfecciones estables. Como control positivos se utilizaron dos copias del "insulator" 1.2 kb 5'cHS4 del dominio β-globina. Como control negativo se utilizaron 2.3 kb de λ-ADN. (C) Ensayo de bloqueo de enhancer *in vivo* utilizando peces cebra transgénicos. HSB-HSA' se colocó interfiriendo la comunicación del enhancer específico del sistema nervioso central CNS y el promotor de actina que dirige una expresión constante de *EGFP* en las somitas (pSomites). Como control positivo se utilizó al "insulator" 5'cHS4 del dominio β-globina. (D) Gráfica que muestra la intensidad de fluorescencia normalizada (somite-vs-CNS). Se indica el número de embriones micro-inyectados y analizados en este ensayo.

***El elemento HSB-HSA' protege a un transgen de los efectos de su entorno cromatínico.***

Otra de las propiedades que definen a los “insulators” es su capacidad para proteger a un transgen del entorno cromatínico ligado al sitio de inserción en el genoma y que se conoce como Efecto de Posición Cromosómico o CPE (por sus siglas en inglés (Chromosomal Position Effect) (96). Para estudiar la capacidad del elemento HSB-HSA' para proteger a un transgen, utilizamos al gen reportero *EGFP* bajo el control del promotor  $\alpha^D$ , enmarcado por 2x(HSB-HSA'), como control positivo 2x(5'cHS4) y como control negativo al transgen sin “insulators” (Fig. 22). Se aislaron al menos 10 clones independientes en las cuales se verificó el número de copias y su integridad por Southern-blot. Las líneas celulares caracterizadas se mantuvieron en cultivo continuo por más de 90 días y la expresión del gen reportero fue periódicamente analizada mediante la síntesis y abundancia relativa de la EGFP medida por citometría de fluorescencia (FACS). Los resultados obtenidos indican que las células que integraron al transgen sin “insulators” fueron progresiva y rápidamente disminuyendo la población de células GFP<sup>+</sup> positivas a lo largo del tiempo, esto debido al silenciamiento epigenético que actúa sobre el transgen. En cambio, las células que integraron el transgen enmarcado por HSB-HSA' ó 5'cHS4 mostraron una población constante de células GFP<sup>+</sup> positivas durante más de 90 días de cultivo continuo (Fig. 22).

En conclusión, HSB-HSA' tiene la capacidad de proteger a un transgen de la influencia de los efectos de posición cromosomales en distintos sitios de integración. Tomando en cuenta los resultados obtenidos a partir del ensayo de bloqueo de enhancer, podemos concluir además que el elemento HSB-HSA' cumple con las dos propiedades básicas que definen a un "insulator".

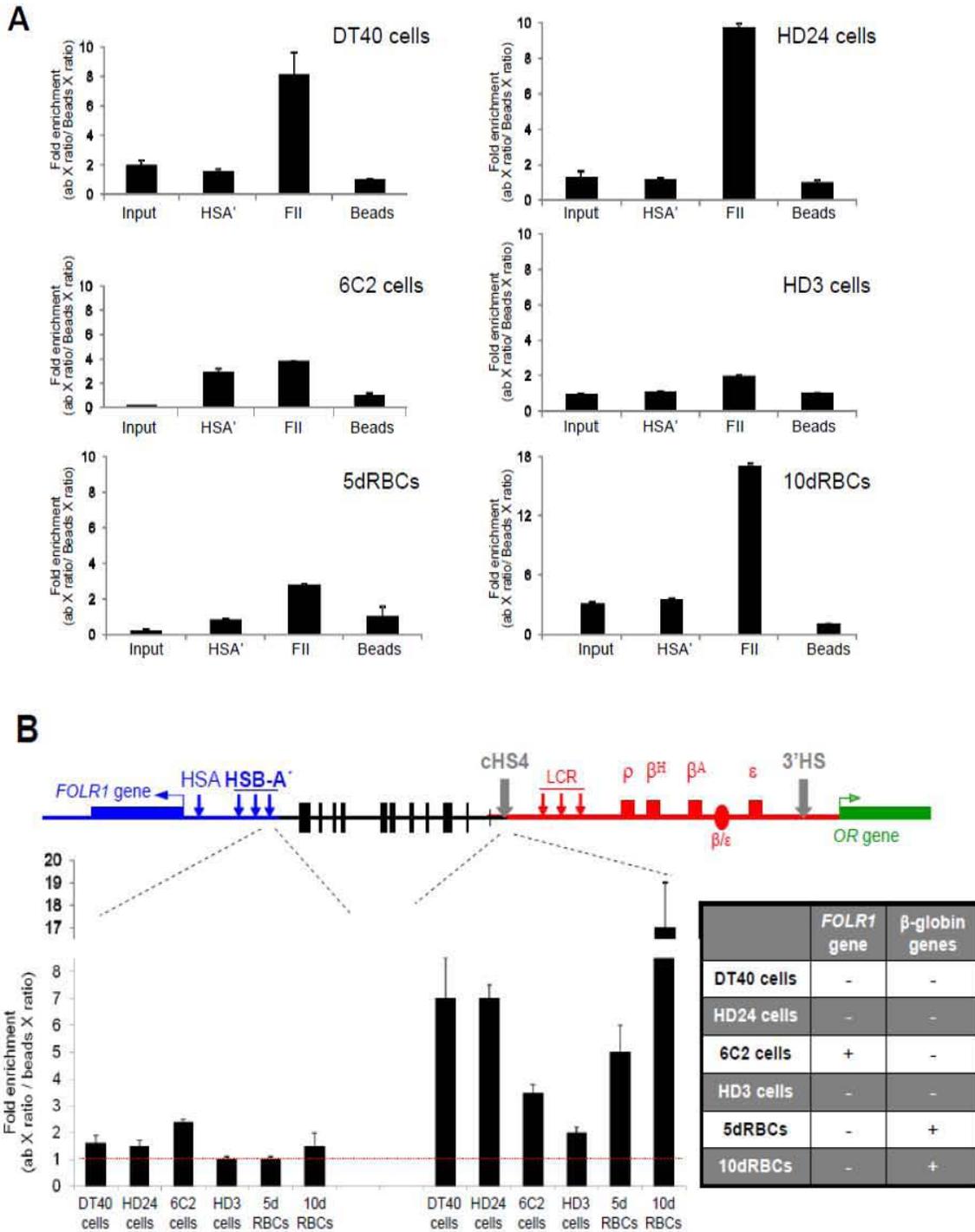


**Figura 22. Ensayo de protección de un transgen contra Efectos de Posición Cromosomales por parte del elemento HSB-HSA'.** Se aislaron al menos 10 líneas monoclonales independientes, se analizó el número de copias del transgen y se utilizaron únicamente aquellos transgenes que se integraron en copia sencilla. Cada línea celular se mantuvo en cultivo celular continuo por 90 días y se analizó el perfil de expresión de *EGFP* reflejado por la emisión de fluorescencia por parte de la proteína *EGFP* por citometría de flujo (FACS). Se muestra el histograma de una línea celular representativa al día 80. La gráfica muestra %GFP+ vs tiempo. Se muestran las desviaciones estándar de cada barra.

***El enriquecimiento in vivo de CTCF sobre el elemento HSB-HSA' coincide con la expresión del gen FOLR1.***

Con el objetivo de definir la dinámica de unión *in vivo* de CTCF al sitio HSA' y compararlo con el "insulator" previamente descrito 5'cHS4 del dominio  $\beta$ -globina, realizamos ensayos de ChIP con el anticuerpo anti-CTCF para determinar de manera cuantitativa los enriquecimientos de CTCF en ambas regiones genómicas. Los resultados mostraron que en el sitio HSA' existe un enriquecimiento de CTCF solo en la etapa de pro-eritroblastos (CFU-E) (células 6C2) en cambio en 5'cHS4 la unión de CTCF es constitutiva durante la diferenciación eritroide y el desarrollo (Fig. 23 y 24). Incluso dicho enriquecimiento está presente en una línea celular no-eritroide DT40 (Fig. 23B y 24B).

En conclusión, la unión de CTCF en el elemento HSA' tiene una dinámica que coincide con la expresión del gen *FOLR1*, con base a los datos acumulados hasta aquí, consideramos que el elemento HSA' tiene una aparente actividad barrera que podría contender contra la propagación del silenciamiento epigenético de las 16 kb de heterocromatina permitiendo la expresión regulada del gen *FOLR1*.



**Figura 23. Ensayo semi-cuantitativo de Inmunoprecipitación de la cromatina utilizando anticuerpo anti-CTCF a lo largo de la diferenciación eritroide y el desarrollo del pollo.** (A) Análisis comparativo de las diferentes etapas de la diferenciación y el desarrollo eritroide tanto de HSA' como del "insulator" 5'cHS4 del dominio  $\beta$ -globina. La línea celular DT40 es utilizada como control no-eritroide. (B) Gráfica que muestra los enriquecimientos de CTCF en la región genómica estudiada en diferentes líneas celulares. Se muestra un resumen del patrón de expresión tanto del gen *FOLR1* como del grupo de los genes  $\beta$ -globina.

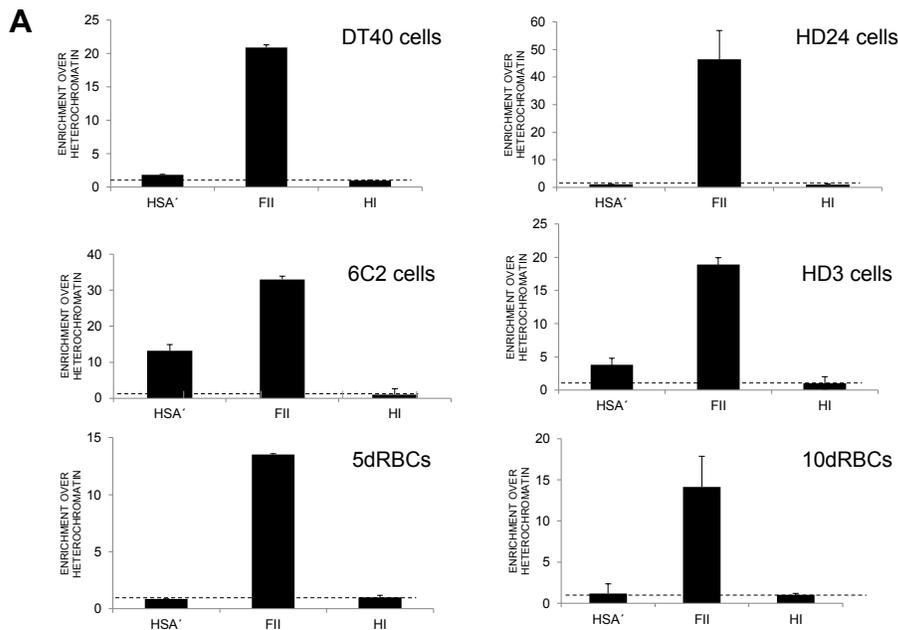


Figure 5

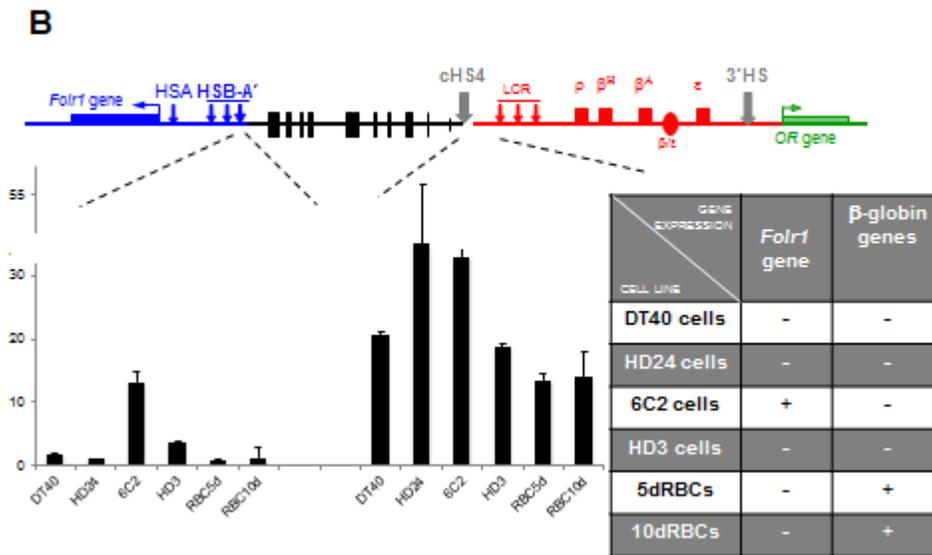


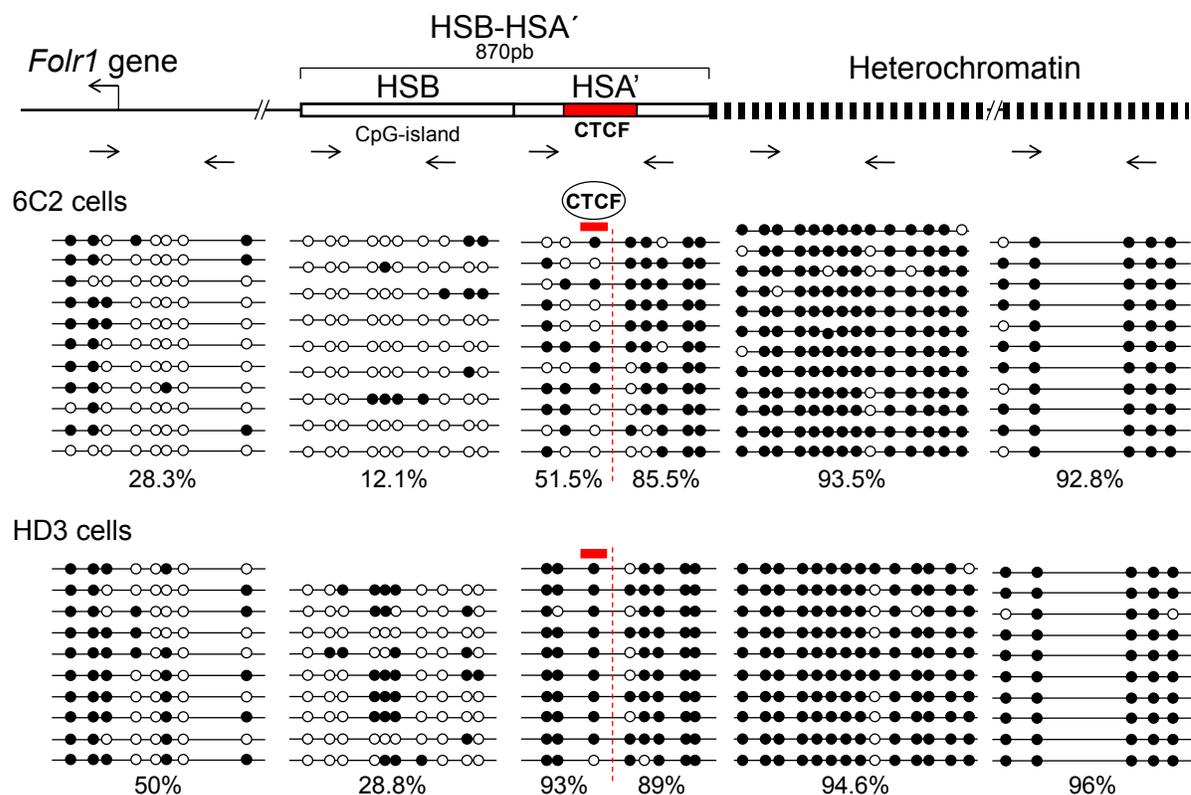
Figura 24. Ensayo cuantitativo con PCR en tiempo real de Inmunoprecipitación de la cromatina utilizando anticuerpo anti-CTCF a lo largo de la diferenciación eritroide y el desarrollo del pollo. (A) Análisis comparativo de las diferentes etapas de la diferenciación y el desarrollo eritroide tanto de HSA' como del "insulator" 5'cHS4 del dominio  $\beta$ -globina. La línea celular DT40 es utilizada como control no-eritroide. (B) Gráfica que muestra los enriquecimientos de CTCF en la región genómica estudiada en diferentes líneas celulares. Se muestra un resumen del patrón de expresión tanto del gen *FOLR1* como del grupo de los genes  $\beta$ -globina.

**2. Confirmar si las 16 kb de heterocromatina situadas río abajo del dominio del gen *FOLR1* pueden afectar la estructura de la cromatina del dominio y la expresión del gen *FOLR1* de pollo.**

***El aumento y propagación en la metilación del ADN en la región regulatoria del gen *FOLR1* coincide con su silenciamiento.***

Uno de los procesos a nivel epigenético involucrados en la propagación de heterocromatina es la metilación del ADN la cual es sinónimo de silenciamiento transcripcional (97). Para contestar la pregunta de si las 16 kb de heterocromatina influyen negativamente sobre la expresión del gen *FOLR1* a través de su propagación, decidimos determinar el estado de metilación del ADN por medio de la conversión con bisulfito de sodio y secuenciación. Analizamos la región genómica que comprende desde el promotor del gen *FOLR1* hasta la región de heterocromatina contigua, incluyendo al sitio HSA' (sitio de unión a CTCF), tanto en la etapa de CFU-E (células 6C2), como en la etapa de eritroblastos (células HD3) (Fig. 25). Los resultados muestran que el sitio de unión a CTCF en el elemento HSA' marca una clara transición entre una cromatina laxa (menos metilada) y una cromatina compacta (más metilada) en la etapa CFU-E (Fig. 25). Recordemos que en la etapa de eritroblastos (células HD3), CTCF se disocia lo cual concuerda con un incremento de la metilación del ADN hacia el dominio que comprende al gen *FOLR1* (Fig. 25).

En resumen, la ganancia de metilación del ADN en el dominio *FOLR1* a manera de gradiente, desde el dominio de las 16 kb en la etapa de eritroblastos (células HD3) coincide con un aumento de la compactación de la estructura de la cromatina (HSA'), la ausencia de CTCF en el elemento HSA' y el silenciamiento transcripcional del gen *FOLR1* apoya nuestra hipótesis que sugiere la propagación del dominio de las 16 kb de heterocromatina.



**Figura 25. Análisis comparativo del estado de la metilación del ADN en la región intergénica localizada entre el dominio que comprende al gen *FOLR1* y el dominio de las 16 kb de heterocromatina constitutiva.** Los círculos blancos representan CpGs no-metiladas y los círculos negros representan CpGs-metiladas. Se muestran la posición de los cinco pares de oligonucleótidos utilizados en este análisis. Se muestra el sitio metilado HpaII (en rojo) que marca el límite 5' de las 16 kb de heterocromatina.

***El “insulator” HSB-HSA’ puede proteger a un transgen de las señales de silenciamiento epigenético provenientes de las 16 kb de heterocromatina constitutiva.***

Con la finalidad de determinar la capacidad del elemento HSB-HSA’ para proteger contra la heterocromatinización o el silenciamiento dependiente de las 16 kb de heterocromatina, se aisló un sub-fragmento de ADN de 6.7 kb proveniente de la región genómica incluida en las 16 kb heterocromatina y se incorporó a las construcciones utilizadas para el ensayo de protección contra efectos de posición cromosomal (Fig. 26). Se aislaron al menos 10 clonas independientes para cada transgen y se corroboró el número de copias y su integridad por Southern-blot (datos no mostrados). Se realizó un análisis periódico de la expresión del gen reportero *EGFP* por medio de FACS a lo largo de 90 días de cultivo continuo (Fig. 26). Los resultados demostraron que, incluso al día 0, que el fragmento de ADN asociado a la región de heterocromatina es capaz de silenciar rápidamente al transgen que no se encuentra protegido por los “insulators” (Fig. 26). En cambio, en presencia de los “insulators” HSB-HSA’ ó 5’cHS4 observamos una protección eficaz que se tradujo en una expresión sostenida del gen *EGFP* durante más de 90 días de cultivo celular continuo (Fig. 26). Estos resultados nos permiten concluir que el elemento HSB-HSA’ es un “insulator” con la capacidad de proteger eficientemente a un transgen contra el silenciamiento epigenético proveniente de un fragmento de ADN de 6.7 kb de heterocromatina (Fig.26).

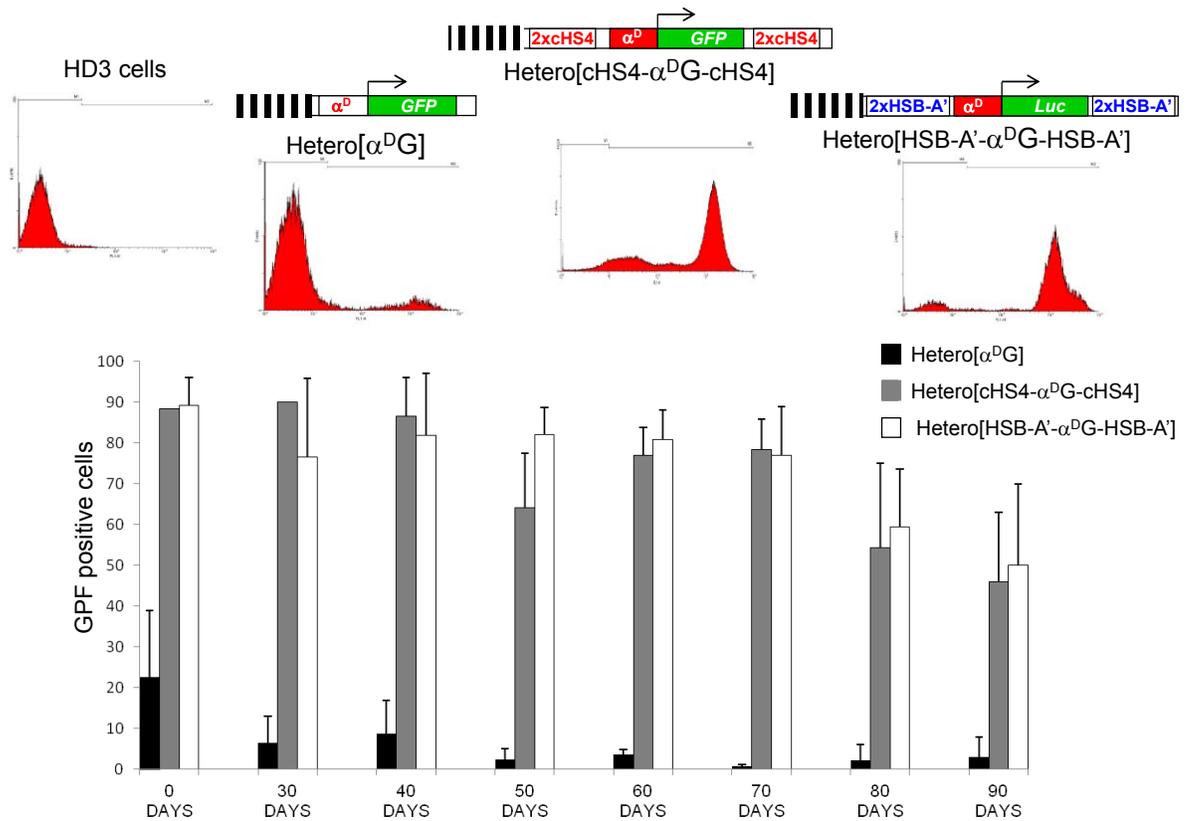


Figura 26. Los “insulators” cHS4 y HSB-HSA’ protegen a un transgen contra las señales de silenciamiento epigenético provenientes de un sub-fragmento (6.7 kb) de las 16 kb de heterocromatina constitutiva. Se utilizó la misma metodología de CPE. La inclusión de un sub-fragmento genómico de heterocromatina tiene una influencia negativa sobre la expresión de la *EGFP* desde el día 0.

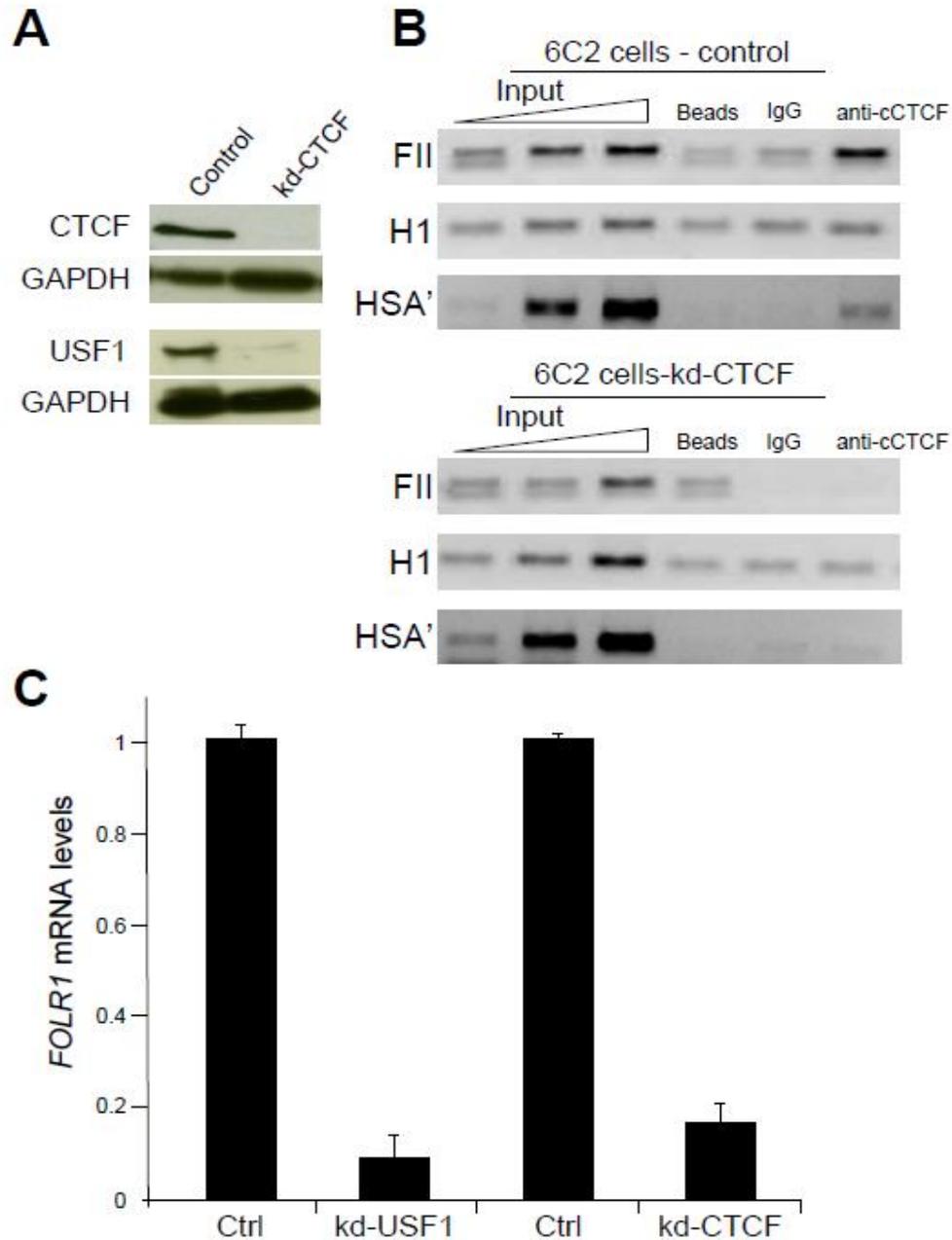
### **3. Caracterizar y estudiar los mecanismos por los cuales un “insulator” contrarresta la propagación de la heterocromatina y como éste contribuye a la formación de un dominio tanto *in vitro* como *in vivo*.**

#### ***La disminución en los niveles de la proteína CTCF afecta la expresión del gen FOLR1.***

Para contestar la pregunta de si CTCF contribuye con la regulación del gen *FOLR1*, realizamos un ensayo de pérdida de función utilizando partículas lentivirales que contienen un RNA interferente tipo tallo-asa (shRNAi) contra el mensajero de CTCF en células 6C2 (Fig. 27). Se evaluó la disminución en los niveles globales de la proteína CTCF por Western-blot, se corroboró la drástica disminución de los niveles de CTCF en el elemento HSA' por inmunoprecipitación de la cromatina (ChIP-CTCF) utilizando el anticuerpo anti-CTCF. A su vez se aisló ARNm para analizar la expresión del gen *FOLR1* por RT-PCR cuantitativa (Fig.27A). Los resultados muestran que la disminución de los niveles de CTCF causa un decremento notorio en la expresión del gen *FOLR1*. Por otra parte, los factores nucleares USF1/USF2 han sido descritos como responsables de reclutar complejos remodeladores de la cromatina en el “insulator” 5'cHS4. Además, recientemente se observaron diferentes niveles de enriquecimiento de éstos factores en las secuencias que corresponden en el sitio HSB (82). Con base en estos antecedentes, nos preguntamos si además de CTCF existen más factores

involucrados en el establecimiento y función del “insulator” HSB-HSA'. Para explorar la posible contribución de USF1, realizamos un ensayo de pérdida de función en la línea celular 6C2, utilizando 5 shRNAi distintos contra USF1 (Fig. 27A y 27C). Evaluamos la disminución de los niveles de USF1 por Western-blot y aislamos ARNm para evaluar la expresión de *FOLR1* por RT-PCR cuantitativa (Fig. 27C). Los resultados muestran que USF1 es importante para la expresión de *FOLR1* y que su disminución afecta drásticamente la expresión de este gen durante la etapa pre-eritroblastos (CFU-E) de la diferenciación eritroide.

En conclusión, la disminución de los niveles de los factores CTCF y USF1 sugieren que ambos son importantes para contrarrestar el silenciamiento epigenético originado a partir de las 16 kb de heterocromatina constitutiva.



**Figura 27. La disminución de los niveles proteicos de CTCF y USF1 afectan la expresión del gen *FOLR1* en células 6C2.** (A) Western-blot que muestra la disminución de la proteína CTCF y USF1. (B) ChIP que muestra la reducción en el sitio de unión de CTCF en HSA' en células 6C2-shRNAi-CTCF. Como control positivo se utilizó FII del "insulator" 5'cHS4 del dominio  $\beta$ -globina. Como control negativo se analizó una región central de las 16 kb de heterocromatina H1. Se muestra un ensayo representativo de tres experimentos independientes. (C) Gráfica que muestra el efecto de la disminución de los niveles proteicos de USF1 (kd-USF1) y CTCF (kd-CTCF) analizados por RT-PCR cuantitativa.

## Discusión

Una de las características de la diferenciación celular tiene que ver con el silenciamiento de genes y la ganancia de manera regulada de heterocromatina (65). Los resultados presentados en este trabajo apoyan el modelo en el cual los elementos llamados “insulators” contribuyen al mantenimiento de dos dominios transcripcionalmente activos: el de los genes  $\beta$ -globina y el del gen *FOLR1* protegiéndolos de la propagación del silenciamiento epigenético proveniente de las 16 kb de heterocromatina (Fig. 28). Ambos dominios génicos son eritroide específicos con diferentes programas de expresión que corresponden a distintas etapas de la diferenciación eritroide. El dominio que incluye al gen *FOLR1* tiene un programa de expresión temprano en la diferenciación eritroide, mientras que el dominio que comprende a los genes  $\beta$ -globina tiene un programa de expresión más tardío y que corresponde a eritrocitos diferenciados (Fig. 28). En este trabajo de investigación demostramos que el elemento HSB-HSA' actúa como un “insulator” capaz de establecer una transición entre regiones de heterocromatina y eucromatina. La pérdida de ésta frontera parece estar relacionada con la disociación de CTCF del elemento HSA' y el subsecuente silenciamiento epigenético a partir de la propagación de las 16 kb heterocromatina (Fig. 16, 17, 24, 25, 28). Este efecto contribuye al silenciamiento definitivo del gen *FOLR1* en etapas más avanzadas de la diferenciación eritroide (Fig. 25 y 28).

La propagación de la heterocromatina es un fenómeno relativamente bien comprendido en levaduras como se comentó en la introducción (98). La vía del ARN de interferencia en la formación y mantenimiento de cromatina puede ser

parte del mecanismo por el cual se establece y puede propagar el silenciamiento epigenético de las 16 kb de heterocromatina (37). Esta vía involucra la transcripción de las secuencias repetidas, distribuidas abundantemente en esta región genómica que comprende a las 16 kb de heterocromatina, para generar una doble cadena de ARN que dispara el reclutamiento de los demás elementos de la vía incluyendo cDicer (37). Los elementos HSB-HSA' y 5'cHS4 enmarcan la región de secuencias repetidas y guardan una sintenia similar incluyendo secuencias LTR cercanas. Dado que CTCF se encuentra en ambos "insulators", es posible que, como parte del mecanismo, CTCF acote la transcripción de estas secuencias repetidas bloqueando de esta manera la propagación de heterocromatina.

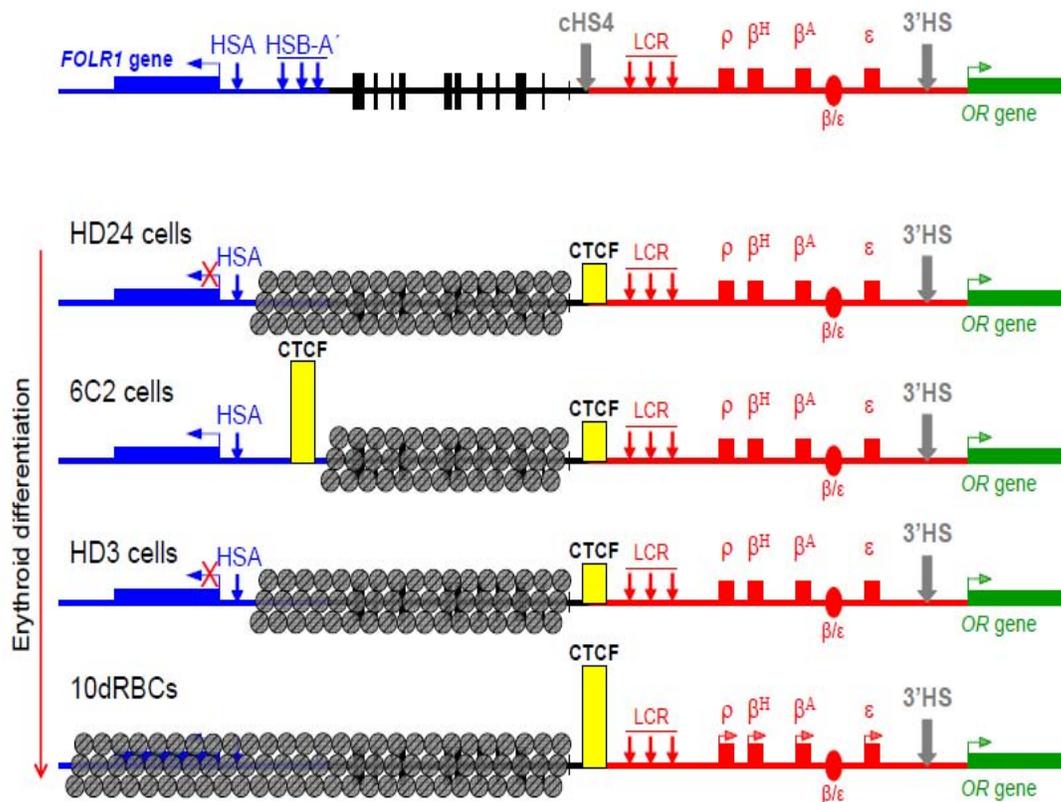
Como parte de éste modelo proponemos la formación de un arreglo tridimensional compatible como la generación de un asa cromatínica (40). Una vez disociado CTCF de HSA', el reclutamiento de proteínas involucradas en la formación de heterocromatina puede distribuirse, a manera de gradiente, desde las 16 kb de heterocromatina hacia el dominio que incluye al gen *FOLR1*, mediante un mecanismo de propagación nucleosoma a nucleosoma, heterocromatinizando unilateralmente en relación con el gen *FOLR1* (Fig. 28). Proponemos a este proceso como parte responsable de los eventos que permiten la transición a un estadio más diferenciado que contribuye a la expresión diferencial de genes contenidos en distintos dominios cromatínicos (Fig. 28). Como parte del mecanismo por los que actúa HSB-HSA' proponemos que: los factores CTCF en HSA' y USF1 en HSB pueden reclutar remodeladores de la

cromatina y que como demostramos, su ausencia lleva a un silenciamiento aberrante del gen *FOLR1* en la etapa CFU-E.

Por otra parte, de acuerdo a los datos del consorcio ENCODE, sabemos, que el 74.7% del genoma humano genera transcritos de ARN primarios no-codificantes (99), 2-3% generan transcritos que codifican para proteínas (100) y tomando en cuenta ARN de transferencia, microARN, lincARN y otros ARNs regulatorios se transcribe el 90% del genoma (100). Como parte colateral de este trabajo, contribuí en demostrar que CTCF tiene la propiedad de unir ARN (57). Se ha mostrado que en algunos casos el ARN puede regular la unión de CTCF en el genoma (101). Por ejemplo, recientemente se han descrito que existen ARNs pequeños no codificantes cercanos al sitio de inicio de la transcripción (tiRNA) y en las regiones promotoras (paRNAs) con funciones regulatorias que pueden silenciar o activar la expresión de genes codificantes y se han propuesto como una característica general en el genoma de humano, ratón, pollo y *Drosophila* (102-104). Estos transcritos son de una longitud de 20-90 nucleótidos con dos sitios de enriquecimiento localizados en la posición +50 pb y en la posición -250 pb respecto al sitio de inicio de la transcripción (TSS), y son transcritos por la ARN-polimerasa II (104). En contraste con la mayoría de los ARN pequeños no-codificantes, parecen no ser procesados por la maquinaria del ARN de interferencia (Dicer, Ago1, Ago2) pues se mantienen presentes en células que carecen de la proteína Dicer (102), los paRNAs tienen una baja tasa de transcripción por tanto un bajo número de copias, tienen un alto contenido de nucleótidos citosina y guanina (102) y su transcripción correlaciona con genes activos (103), el consorcio FANTOM4 describió además que son procesados es

decir son modificados en su extremo 5' con CAP, poliadenilados y degradados por el exosoma. Finalmente, se han encontrado a este tipo de transcritos asociados a sitios de unión de CTCF en el genoma. Particularmente se ha demostrado que estos transcritos forman parte de los mecanismos que regulan la unión de CTCF en su sitio de unión y se ha especulado que pueden ser un componente importante en la arquitectura nuclear (102, 105). También se ha asociado a la maquinaria del exosoma con la función “insulator” y CTCF (106). Por otro lado, existe evidencia de que los ARN no codificantes largos (>200 nt) pueden regular la unión de CTCF en el pollo (107). En conclusión, para el caso del elemento HSB-HSA' los ARNnc pueden ser parte de los mecanismos que determinan el enriquecimiento y la disociación de CTCF del sitio HSA'.

Actualmente se puede determinar si dos sitios lejanos en el genoma, pueden estar contactándose y se ha propuesto que CTCF puede formar dímeros y participar junto con algunas proteínas cohesinas en la organización del genoma formando asas cromatínicas (40, 76). Experimentos de Captura Conformacional de Cromosomas a escala de regiones amplias y/o del genoma completo (Hi-C) demuestran que CTCF se encuentra mayoritariamente en los sitios de contacto dentro de un cromosoma y además marca el límite entre una cromatina compacta y una cromatina abierta (80), por tanto, queremos descartar experimentalmente la posibilidad de que entre el “insulator” HSB-HSA' y el “insulator” 5'cHS4 exista un contacto directo, de manera que, la región de las 16 kb de heterocromatina se encuentre acotada en una asa de cromatina.



**Figura 28. Modelo.** Se resumen nuestros resultados y se muestra como las 16 kb de heterocromatina participan en el proceso de la diferenciación eritroide silenciando al gen *FOLR1*. Las barras amarillas representan los enriquecimientos de CTCF en los “insulators” HSA y 5’cHS4.

En conclusión, nuestros resultados demuestran que existe un elemento barrera o “insulator” que contiene activamente contra la propagación de la heterocromatina, permitiendo formar y establecer el dominio de eucromatina que incluye al gen *FOLR1*, el cual es necesario para permitir su transcripción. Tomando en cuenta el “insulator” del dominio  $\beta$ -globina, queda demostrado que el dominio de las 16 kb de heterocromatina es acotado durante la etapa de expresión de *FOLR1*. De esta manera cuando la diferenciación avanza, el elemento barrera o “insulator” se pierde, con un consecuente silenciamiento del gen *FOLR1*.

Proponemos que la propagación de heterocromatina contribuye entonces, de manera regulada al silenciamiento definitivo del gen *FOLR1* y que, esto puede permitir establecer un nuevo programa de expresión a la célula eritroide acorde con su etapa de diferenciación.

## Perspectivas.

Entender el funcionamiento de los “insulators” y su contribución en la formación y mantenimiento de dominios cromatínicos, es una de los principales intereses de nuestro laboratorio. Los resultados obtenidos en este proyecto establecen un modelo de estudio que permite observar los cambios en la cromatina en los sitios regulatorios/estructurales de un dominio que debe ser silenciado como parte de los mecanismos de la diferenciación eritroide, y además permite estudiar los mecanismos asociados a la propagación de heterocromatina en vertebrados. Como perspectivas que surgen de éste proyecto proponemos estudiar los siguientes aspectos implicados en la función del “insulator” HSB-HSA’:

1. Estudiar la contribución de proteínas claves de la vía del ARN de interferencia como *cDicer* y *cAgo2* en los eventos de silenciamiento del dominio que comprende al gen *FOLR1*. Una primera aproximación será abatir los niveles de éstas proteínas utilizando vectores lentivirales inducibles que expresen ARNs interferentes contra *cDicer* y/o *cAgo2* monitoreando la expresión del gen *FOLR1* en células HD3 (37).
2. Analizar si en elemento HSB-HSA’ existe transcripción de ARNs no-codificantes que funjan como reguladores de la unión o disociación del factor CTCF en ésta frontera.
3. Estudiar el arreglo tridimensional que permite la función de HSB-HSA’ y 5’cHS4.

## ANEXO A

### Generación de un anticuerpo policlonal anti-CTCF de pollo.

Construcción plásmido pET28(b).

Un fragmento de ADN de 441 pb que abarca la secuencia codificante para la región amino-terminal no conservada entre pollo, ratón y humano de la proteína CTCF de pollo y que comprende los aminoácidos del residuo 86 al residuo 233 fue clonado en el marco de lectura abierto del vector pET28(b) utilizando los sitios de restricción 5´BamHI / 3´XhoI. Este vector nos permite adicionar una etiqueta de 6 histidinas en la región N-terminal del péptido de interés. El plásmido resultante se transformó en la cepa BL21(DE3) de *E. coli* y la expresión del péptido se indujo con IPTG.

Purificación del péptido recombinante.

Un cultivo de 10 mL de 12 horas de crecimiento de la cepa de *E. coli*/pET28(b)/CTCF<sub>86-233</sub>, se usó para re-inocular a un cultivo de 100 mL de LB utilizando el antibiótico kanamicina, se creció hasta fase exponencial, se agregó IPTG a una concentración final de 0.1 mM y se cultivó en agitación constante (200 rpm) a 37°C. Se colectaron a las bacterias por centrifugación a 6000 rpm, 10 minutos y se retiró el sobrenadante. El paquete celular se resuspendió en 5 mL de buffer PBS/10 mM Imidazol. Se lisaron las bacterias por sonicación 10 pulsos de 15 segundos cada uno a una amplitud de 35%. Se centrifugo la muestra a máxima velocidad (14000 rpm) durante 10 min. Se recuperó el sobrenadante y se purificó

al péptido recombinante utilizando una columna cargada con Níquel (Ni(+)). El péptido purificado His6x-CTCF<sub>86-233</sub>, se cuantificó por el método de Bradford y se utilizó para la inmunización de conejos hembra de la cepa New Zeland White Rabbit como se describe a continuación.

### Protocolo de Inmunización

Standard 70-Day Rabbit immunization protocol for rabbit polyclonal antibody production. Antigen preparation, such as peptide synthesis and/or immunogen conjugation, occurs before Day 0. Injection amounts are given for a conjugated peptide antigen (e.g., KLH) or protein immunogen; MAP-peptide antigens are injected at 0.5mg throughout. Protocol Days are approximate (± 2 days).		
Procedure	Protocol Day	Description
Control Serum Collection	Day 0	Pre-immune bleed (5mL per rabbit)
Primary Injection	Day 1	Immunize with 0.50mg of antigen in CFA, 10 SQ sites
1st Booster	Day 14	Boost with 0.25mg of antigen in IFA, 4 SQ sites
2nd Booster	Day 28	Boost with 0.25mg of antigen in IFA, 4 SQ sites
Serum Collection	Day 35	Bleed (~25mL per rabbit)
3rd Booster	Day 42	Boost with 0.25mg of antigen in IFA, 4 SQ sites
Serum Collection	Day 56, 58	Two bleeds (~50mL total per rabbit)
ELISA and Shipping	Day 60	ELISA titration (results available online); Verify disposition of rabbits; decide to continue or terminate
Instruction Due Date	Day 72	If no animal instructions are received, per diem charges will apply
Deliverables (Day 60): 12 vials crude antibody sera <ul style="list-style-type: none"> <li>• 2 vials: 1-2mL test aliquots pre-immune sera</li> <li>• 2 vials: 5mL per vial pre-immune sera</li> <li>• 4 vials: 1-2mL test aliquots post-inject sera (Day 35 and 58)</li> <li>• 4 vials: 20-50mL per vial bulk post-inject sera (Day 35 and 58)</li> </ul>		

His-6x-CTCF<sub>86-233</sub> aa

His-6x-CTCF<sub>86-250</sub> aa

250 µg	500 µg
250 µg	250 µg
250 µg	250 µg
250 µg	250 µg

El suero obtenido se utilizó para la detección de proteína CTCF de pollo en extractos totales de proteínas de las líneas celulares 6C2 y HD3 de pollo.

Una vez que se determinó la reactividad del suero anti- chCTCF (comparado con el suero preinmune), se realizó la purificación de las IgG contra la proteína CTCF

utilizando el kit de purificación AminoLink Plus Immobilization kit (Thermo Scientific)

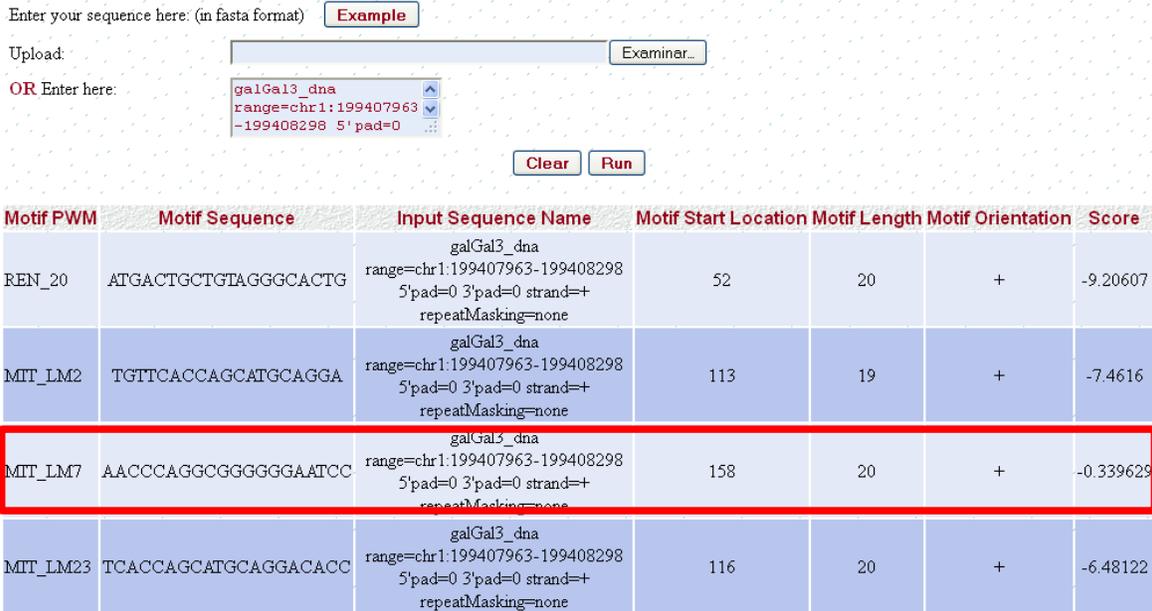
El anticuerpo purificado se cuantificó por Bradford y se utilizó para ensayos de WB, CHIP e Inmunofluorescencia (Fig. 19)

## **ANEXO B**

### **Análisis bioinformático de la región 3' del gen *FR* en búsqueda de sitios de unión al factor de transcripción CTCF.**

Debido a que los elementos “insulators” identificados a la fecha en vertebrados dependen de la presencia del factor nuclear CTCF para realizar su función, decidimos hacer una búsqueda y análisis de motivos de unión a este factor nuclear en la región regulatoria del dominio que incluye al gen *FOLR1* y que colinda con el dominio de las 16 kb de heterocromatina. El factor nuclear CTCF posee 11 dedos de zinc en su dominio de unión al ADN por lo que puede unirse a secuencias muy divergentes, por tanto se deben utilizar herramientas bioinformáticas basadas en matrices que pueden predecir los posibles sitios de unión de CTCF en el genoma. Utilizamos la base de datos: [insulatordb.utmem.edu](http://insulatordb.utmem.edu) y corroboramos nuestro análisis con las matrices de predicción de sitios de unión a CTCF generadas en nuestro laboratorio en la Fig. 29 se muestran los sitios de unión predichos.

[In silico CTCFBS prediction tool](#)



<http://insulatordb.uthsc.edu/>

**Figura 29. Sitios de unión al factor nuclear CTCF predichos en el elemento HSB-HSA´.** De los cuatro motivos relevantes que se muestran en nuestro análisis, el tercer motivo (resaltado en rojo) tiene el mejor score por lo que decidimos utilizar este motivo para corroborar la unión de CTCF *in vitro* e *in vivo*.

A continuación realizamos un análisis comparativo entre el sitio de unión a CTCF predicho en el elemento HSA´ con el motivo de unión de CTCF en el genoma de pollo utilizando la base de datos de ChIP-seq realizado en células 10dRBC de pollo previamente publicada por nuestro laboratorio en la cual, se reportó el motivo de unión de CTCF en el genoma de pollo (72) (Fig. 30).

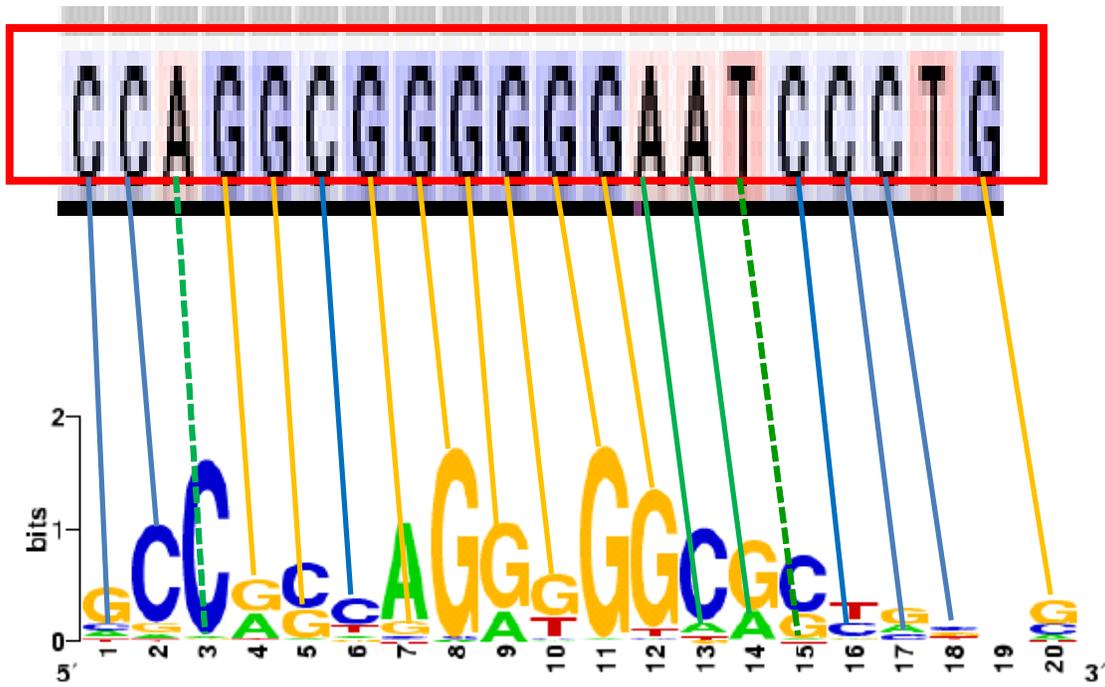


Figura 30. Alineamiento del motivo de unión de CTCF de HSA<sup>1</sup> predicho en la base de datos insulatordb con el motivo de unión de CHIP-seq de CTCF del genoma de pollo publicado por nuestro laboratorio.

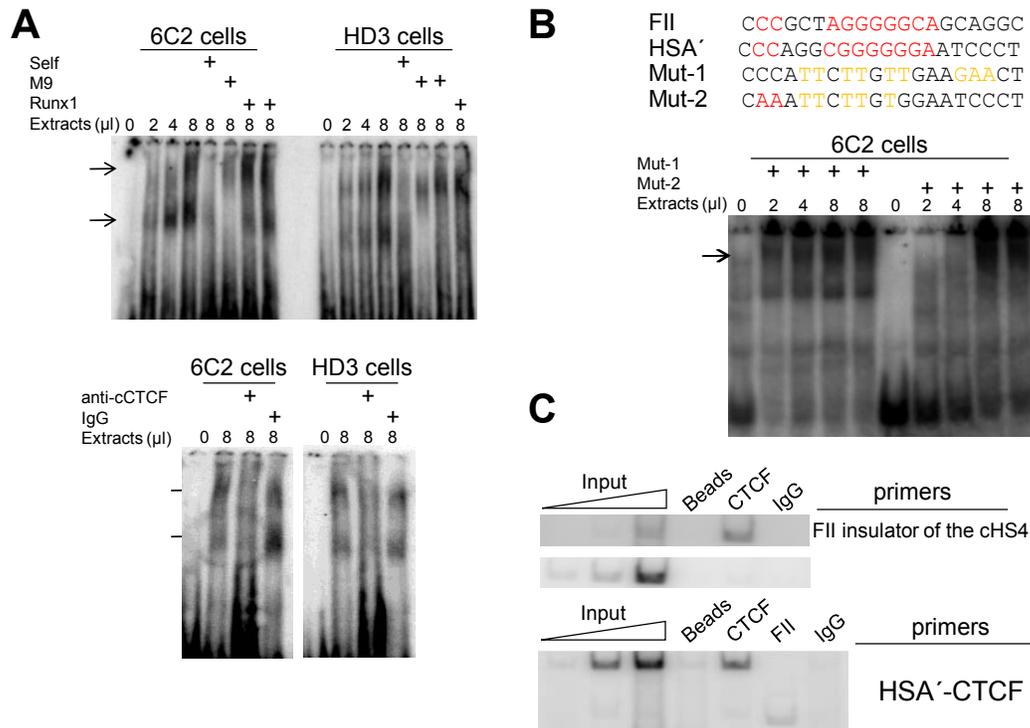
## ANEXO C

### **Análisis *in vitro* e *in vivo* de la unión del factor nuclear CTCF al elemento HSB-HSA´.**

La secuencia presentada en el anexo B es una predicción bioinformática que tiene un 77% de identidad con el motivo de unión de CTCF obtenido a partir de la secuenciación de todos los sitios que ocupa CTCF *in vivo* en el genoma de eritrocitos de pollo, sin embargo es necesario comprobar si CTCF se puede unir *in vitro* mediante un ensayo de retardo en gel (EMSA) utilizando como sonda a la secuencia de unión de CTCF predicha localizada en el elemento HSB-HSA´(Fig. 31). Realizamos el ensayo de EMSA utilizando extractos nucleares de células pre-eritroblastos 6C2 y eritroblastos HD3 (Fig. 31). Este ensayo nos indicó que la sonda HSB-HSA´ era capaz de unir a proteínas nucleares, las cuales son competidas con un oligonucleótido (M9) no marcado que contiene la secuencia correspondiente al del “insulator”  $\alpha$ EHS1.4 del dominio alfa globina de pollo y que tiene un sitio de unión a CTCF ya validado (89). Como control negativo, utilizamos una sonda no marcada con el sitio de unión al factor nuclear RUNX1 que no une al factor nuclear CTCF y observamos que no fue capaz de competir y evitar la formación del complejo. Para saber si el factor nuclear CTCF se une a la secuencia, a continuación realizamos el ensayo de retardo en gel utilizando un anticuerpo específico para CTCF de pollo, el cual hace desaparecer los complejos, esto significa que CTCF se encuentra presente en el complejo que retarda a la sonda HSB-HSA´ y que el anticuerpo titula específicamente a CTCF impidiendo la

formación de los complejos previamente observados, utilizamos como control negativo a un anticuerpo inespecífico (IgG), el cual no impide la formación de los complejos en la sonda HSB-HSA'. Finalmente generamos dos sondas HSB-HSA' que incorporan mutaciones en el sitio de unión de CTCF (Mut1 y Mut2) para saber que nucleótidos son importantes para la unión de CTCF en el elemento HSB-HSA'. En este caso observamos que Mut2 impide la generación de complejos proteicos que retardan la sonda del elemento HSB-HSA' (Fig. 31).

Para validar la unión *in vivo* de CTCF, realizamos ensayos de inmunoprecipitación de la cromatina (ChIP) utilizando un nuevo anticuerpo anti-CTCF de pollo generado en nuestro laboratorio recientemente (Fig. 19, Anexo A) (89). El ensayo de ChIP fue realizado en células 6C2 de pollo correspondiente a la etapa pre-eritroblastos (CFU-E) para evaluar la presencia de CTCF en HSA'. Los resultados de este estudio muestran que CTCF se enriquece preferencialmente en la región asociada al sitio HSA' (Fig. 17D, 20, 23, 24, 31 y Anexo C). Adicionalmente se repitió el mismo ensayo de manera cuantitativa en todas las etapas de la diferenciación eritroide HD24, 6C2, HD3, y cultivos primarios de eritrocitos de 5 días y 10 días de desarrollo del pollo (5dRBC y 17dRBC) (Fig. 20, 23 y 24).



**Figura 31. Ensayo de retardo en gel (EMSA) utilizando al sitio de unión de CTCF predicho en HSA' como sonda.** (A) EMSA utilizando extractos nucleares de la línea celular 6C2 y HD3. Se utilizaron los siguientes competidores: sonda HSA' sin marcar (self) como competidor, como control positivo utilizamos una secuencia de unión a CTCF previamente validado proveniente del dominio  $\alpha$ -globina de pollo (M9) y como control negativo una secuencia inespecífica para CTCF y que contiene al motivo de unión del factor transcripcional Runx1. Para el ensayo de super-retardo (super-shift) utilizamos al anticuerpo anti-cCTCF de pollo generado en nuestro laboratorio (Ver Materiales y Metodos, Fig. 19 y Anexo A), y como control negativo utilizamos un anticuerpo inespecífico (IgG). (B) Se muestra el diseño de dos sondas que contienen mutaciones para el sitio de unión de CTCF. Ensayo de retardo (EMSA) con las dos versiones mutadas del sitio de unión de CTCF en HSA'. Se muestran geles representativos de dos experimentos independientes. (C) Ensayo de Inmunoprecipitación de la cromatina (ChIP) que demuestra *in vivo* la unión de CTCF, como control positivo utilizamos oligonucleótidos específicos para el FII que contiene un sitio de unión para CTCF del "insulator" cHS4 del dominio  $\beta$ -globina de pollo. Como control negativo utilizamos oligonucleótidos específicos para una región de heterocromatina (16 kb) que no une a CTCF. Se muestran un experimento representativo de cuatro ensayos de ChIP independientes.

## ANEXO D

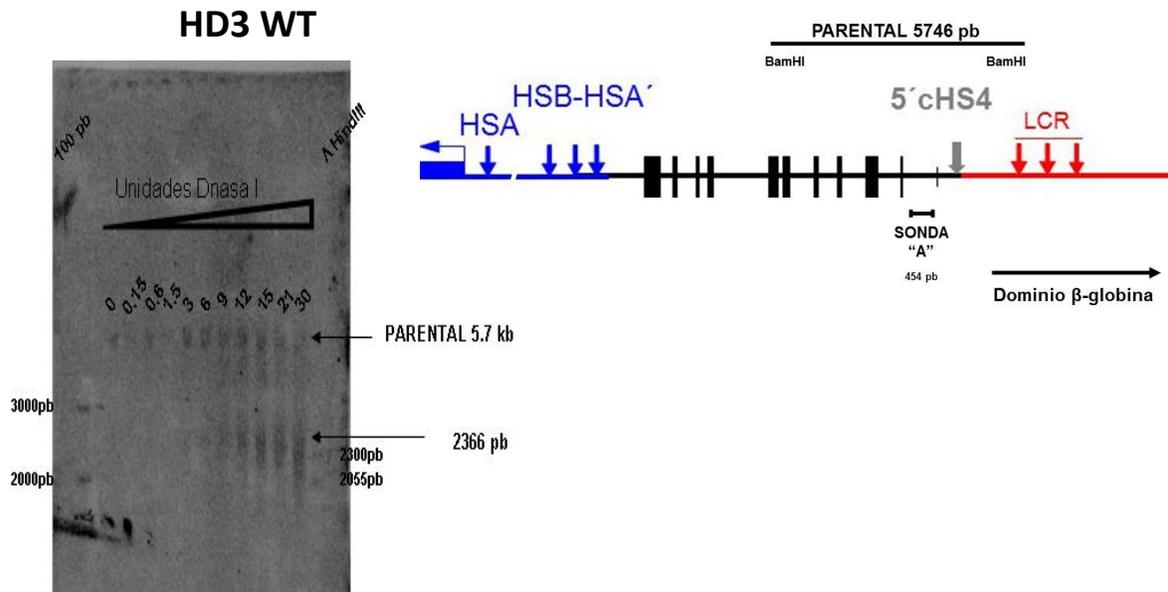
### **Análisis in vivo del efecto de la interrupción de factor nuclear USF1 en el “insulator” 5´cHS4 del dominio β-globina en función de la posible propagación de las 16 kb de heterocromatina.**

*1) Presencia del sitio de hipersensibilidad 5´cHS4 sobre el “insulator” en la línea celular HD3.*

Con la finalidad de establecer el protocolo para la identificación de sitios de corte a la DNasa I, planteamos corroborar la presencia del sitio hipersensible a la DNasa I número 4 correspondiente al “insulator” del dominio β-globina de pollo: 5´cHS4 (Fig. 32). Se diseñó primero en un mapa de restricción, una digestión con una sola enzima de restricción que generara un fragmento genómico que contuviera al “insulator”. Ya que esta digestión delimita la zona de estudio, la llamamos secuencia parental (BamHI-BamHI de 5746 pb; Fig. 32). Se escogió una sonda de 454 pb (sonda A454), que inicia justo en el sitio BamHI del extremo 5´ del fragmento parental y se extiende hasta el primer sitio SacI (respecto al sitio de restricción BamHI parental). El diseño experimental tuvo en consideración que la sonda no incluyera secuencias repetidas lo cual podría ocasionar hibridaciones inespecíficas. De un cultivo celular, se aislaron  $1 \times 10^9$  núcleos de células HD3, a partir de los cuales se utilizaron 11 muestras experimentales de  $1 \times 10^6$  núcleos y se hizo una serie de digestiones con cantidades crecientes de unidades de la

endonucleasa DNasa I (0, 0.15, 0.6, 1.5, 3, 6, 9, 12, 15, 21, 30). Se digirió con BamHI y se separó en un gel de agarosa al 1%, se realizó un Southern blot y se hibridó con la sonda marcada radioactivamente.

La autorradiografía reveló una banda que se enriquece conforme se incrementan las unidades de la enzima DNasa I (Fig. 32). Por su peso molecular dicha banda corresponde al sitio de hipersensibilidad a la DNasa I que coincide con el “insulator” 5'cHS4 en células HD3. En el ensayo se ve como la banda parental se digiere en 2 bandas por el corte de la DNasa I, generándose un fragmento de 2366 pb (Fig. 32).



**Figura 32. Ensayo de Hipersensibilidad al corte por la DNasa I en células HD3.** Observación de 5'cHS4. La hipersensibilidad a la DNasa I aparece a partir de 9 U/1x10<sup>6</sup> núcleos, carril 7. Se presenta el esquema de la región genómica analizada.

Este resultado tiene dos implicaciones importantes para en el proyecto; por un lado demostramos la presencia del sitio hipersensible a la DNasa I

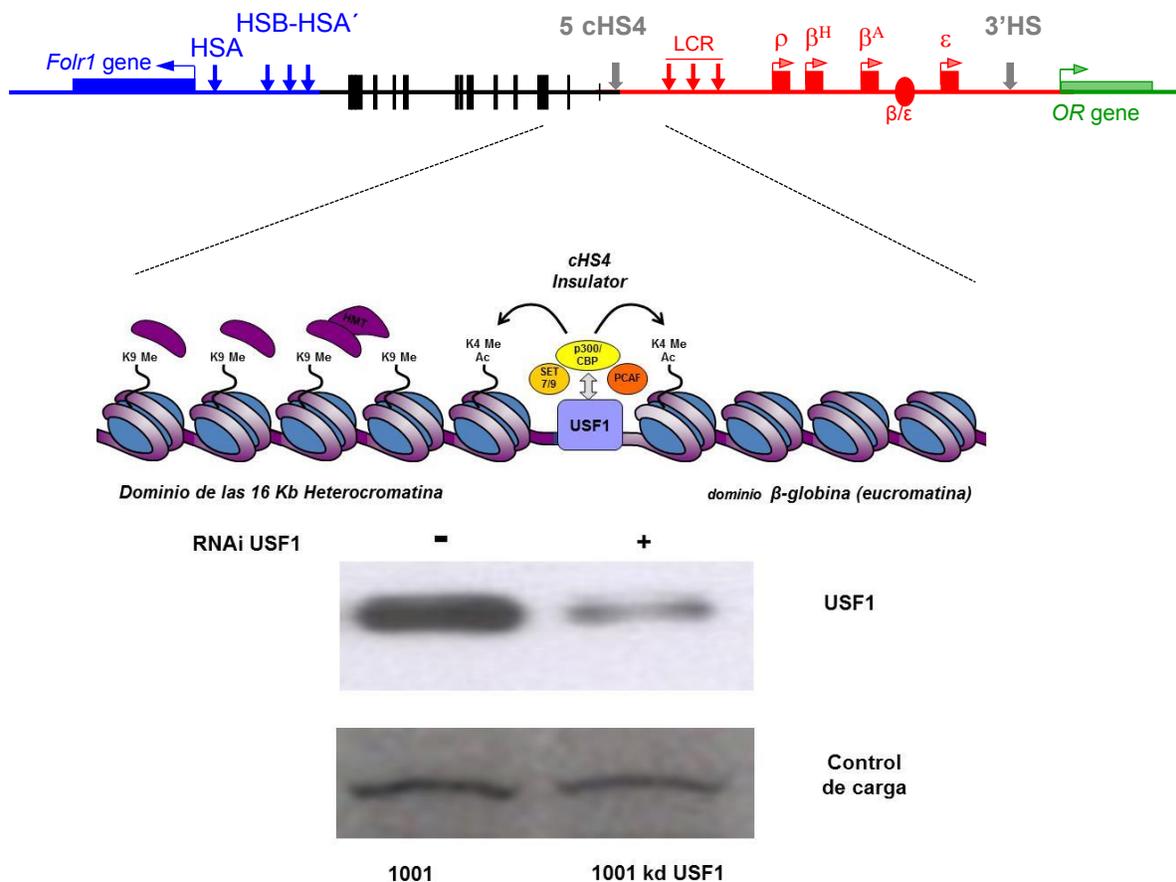
correspondiente al “insulator” 5´cHS4 en la línea HD3 (tal como lo esperábamos), y por otro lado determinamos que en ésta línea celular, la hipersensibilidad aparece a una concentración de 9 Unidades DNasa I /1x10<sup>6</sup> núcleos. Esto será útil cuando intentemos abatir la actividad de barrera del “insulator” demostrándolo mediante la pérdida de la hipersensibilidad.

*2) Contribución de USF1 a la estructura de la cromatina en el “insulator” 5´cHS4.*

Queremos conocer si *in vivo*, la función de barrera del “insulator” 5´cHS4 reside en el reclutamiento de maquinaria de remodelaje que realiza USF1/2 en el sitio FIV del “core” del “insulator” 5´cHS4 (Fig. 33). Felsenfeld y colaboradores demostraron que el heterodímero USF1/2 se une al sitio FIV del core del “insulator” 5´cHS4 y que interactúa físicamente con SET7/9 y con PCAF (metiltransferasa H3K4 y acetiltransferasa-H3, respectivamente) (108). Con estos datos Felsenfeld y colaboradores propusieron que USF1/2 es necesario para el reclutamiento de la maquinaria que mantiene las señales provenientes de la cromatina compacta en equilibrio con las señales de apertura de la cromatina en el “insulator” 5´cHS4 (108). Nos planteamos realizar un RNA de interferencia contra USF1 para evitar la formación del heterodímero USF1/USF2 y por ende el reclutamiento de las demás proteínas que contribuyen a que el “insulator” 5´cHS4 pueda contener contra las marcas represivas de la cromatina compacta(Fig.33). De esta manera, planteamos perturbar el equilibrio teórico que se guarda en éste elemento y observar consecuentemente la propagación de las 16 kb de heterocromatina en su contexto endógeno (Fig 34).

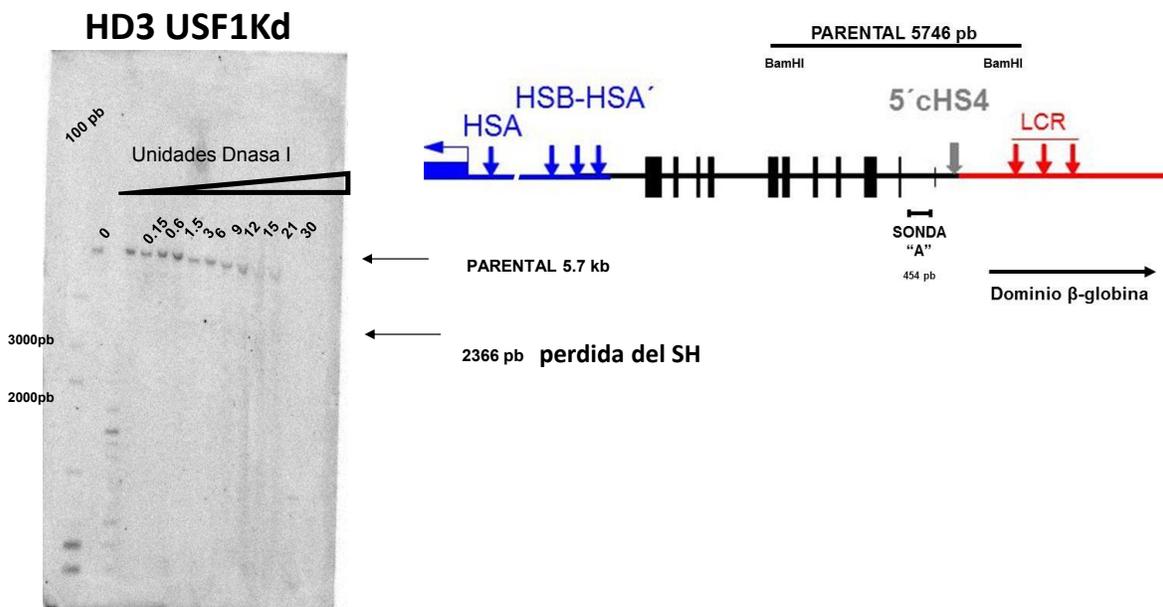
3). Generación de una línea estable deficiente de *USF1* mediante ARN interferencia y ensayo de hipersensibilidad a la *DNasa I* en la línea celular HD3-*USF1*Kd.

Se co-transfectaron 5 plásmidos independientes a células HD3 que codifican para RNAi contra *USF1*. Se aisló la clona 1001(2) y se realizó un Western-blot para corroborar la interferencia de *USF1*(Fig. 33)

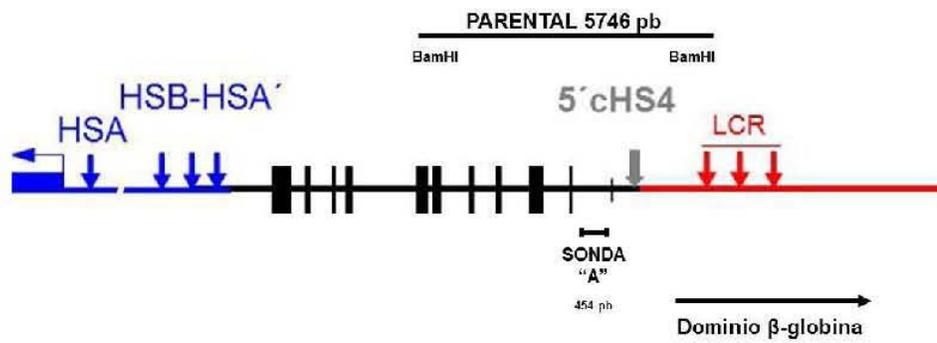
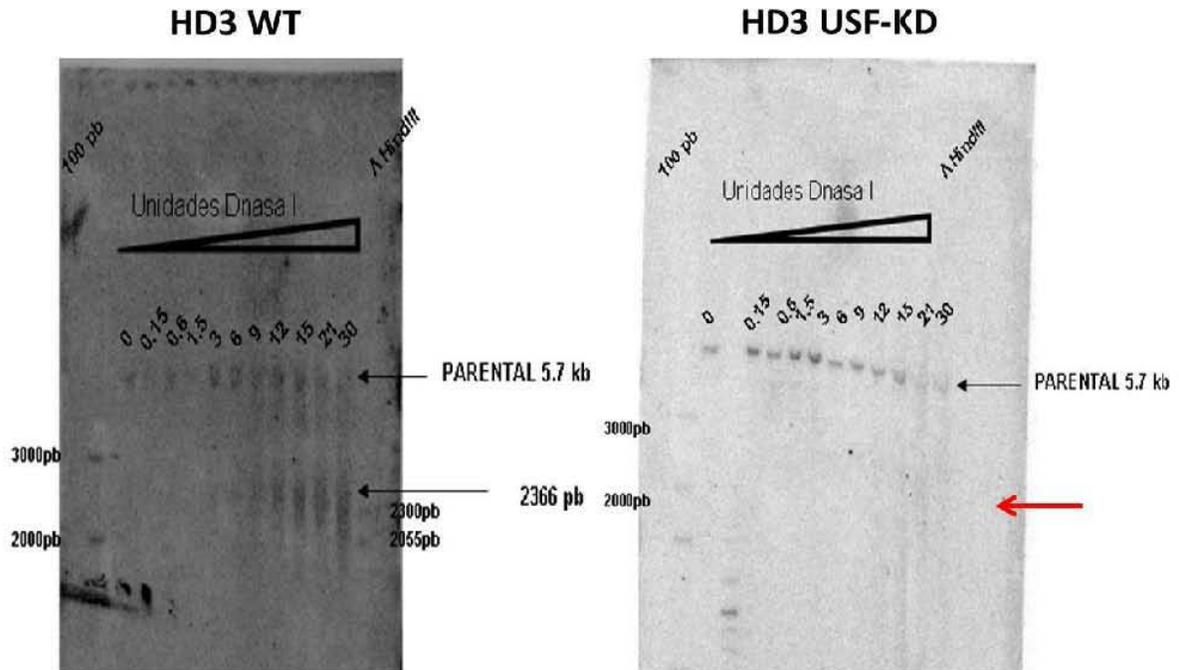


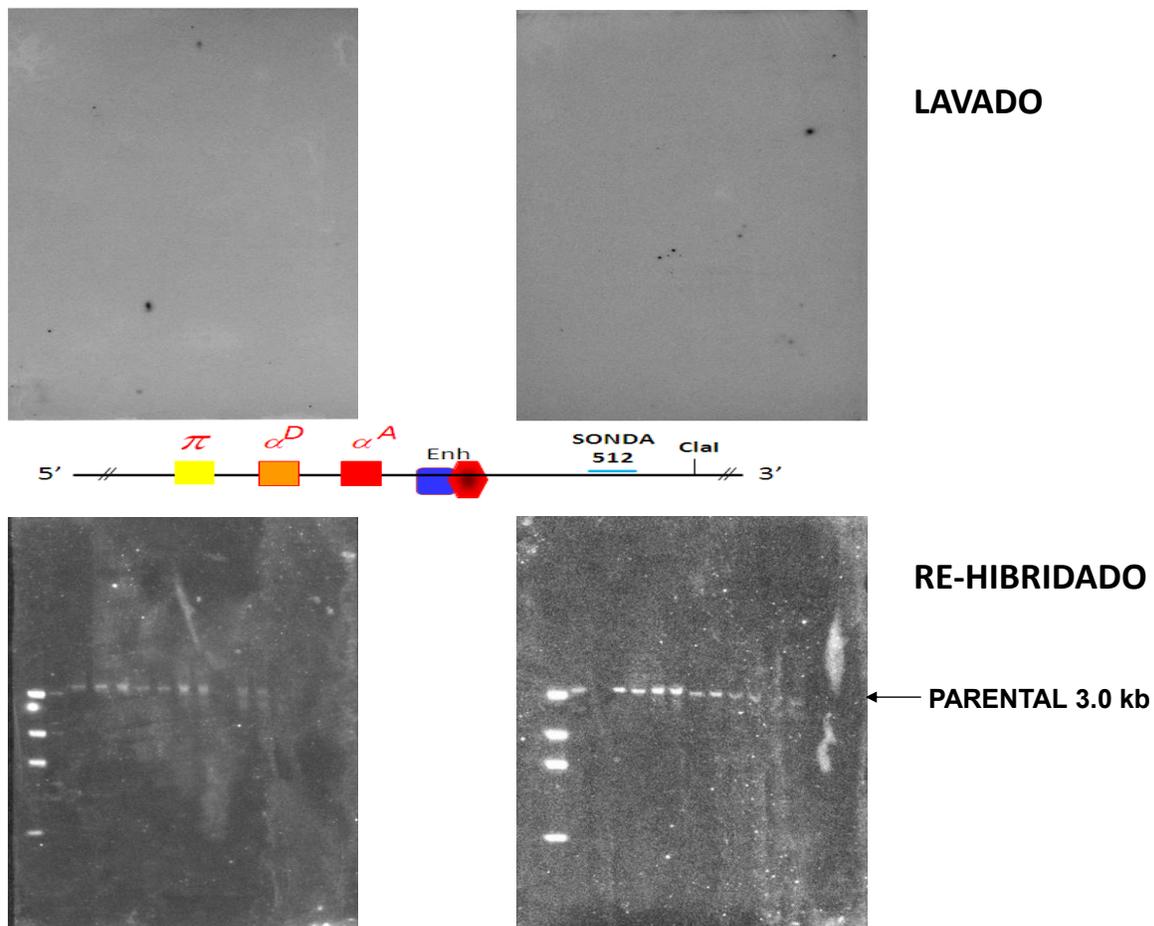
**Figura 33. Western blot.** Se utilizaron extractos totales de células HD3 wt y HD3USFkd, utilizamos un anticuerpo anti-*USF1*. Obsérvese la disminución de la cantidad de *USF1* tras la co-transfección de 5 RNAi específicos.

Con ésta clona “knockdown” contra USF1 proseguimos con el ensayo de la DNasa I para observar posibles cambios en la formación de la hipersensibilidad 5'cHS4 (Fig. 34). El ensayo se realizó en las mismas condiciones que el experimento anterior. Tras el revelado con la sonda “A454” observamos únicamente la presencia de la banda parental de 5.7 kb, lo que significa que el “insulator” 5'cHS4 dejó de ser hipersensible al corte por la DNasa I. Resulta importante recalcar que inclusive en el rango de concentraciones más elevadas no fue posible observar el sitio de hipersensibilidad (Fig. 34 y 35).



**Figura 34. Ensayo de Hipersensibilidad al corte por la DNasa I en células HD3 kdUSF1.** En el ensayo anterior la hipersensibilidad a la DNasa I aparece a  $9 \text{ U}/1 \times 10^6$  núcleos, carril 7 (Fig. 4). Este ensayo es representativo de dos experimentos independientes.

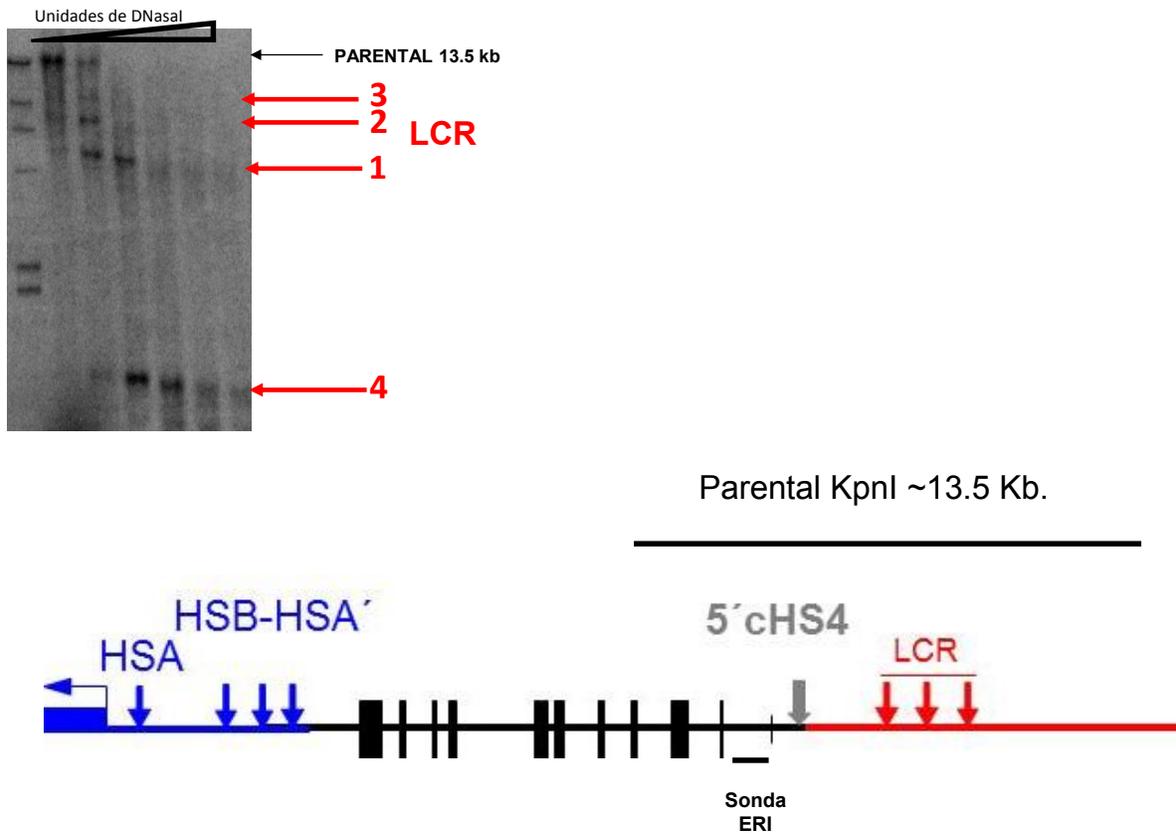




**Figura 35. Comparación del Ensayo de Hipersensibilidad al corte por la DNasa I en células HD3 y HD3USF1kd.** Observación de la pérdida de 5'cHS4. Se muestra el control de corte por la DNasa I utilizando la sonda 512 del dominio  $\alpha$ -globina de pollo que revela un parental de 3 kb. Se muestra un esquema de la región genómica analizada.

Con base en este resultado decidimos realizar una digestión genómica distinta utilizando la enzima KpnI que permite evidenciar una banda parental de 13.5 kb en donde se incluyen (Fig. 36) las hipersensibilidades del LCR (HS3-1), las cuales se encuentran alejadas en relación al “insulator”, para observar si éstas fueron afectadas también (Fig. 36). Hemos establecido las condiciones experimentales

que nos permitirán demostrar la presencia o ausencia de las hipersensibilidades correspondientes al “insulator” 5'cHS4 y al LCR (HS3-1) del ensayo con la digestión genómica KpnI. Este sistema experimental nos permitirá abordar nuevas e interesantes preguntas. Por ejemplo, el proceso de propagación de la heterocromatina aledaña al dominio  $\beta$ -globina de pollo.



**Figura 36.** Ensayo de Hipersensibilidad al corte por la DNasa I en células HD3 utilizando una digestión genómica con la endonucleasa KpnI que abarca una región de 13.5 kb. Observación de 5'cHS4, y tres hipersensibilidades correspondientes a la región controladora del locus (LCR) del dominio  $\beta$ -globina de pollo. Se presenta el esquema de la región genómica analizada.

## Publicaciones

1. Furlan-Magaril M, Rebollar E, Guerrero G, Fernández A, Moltó E, González-Buendía E, Cantero M, Montoliu L, Recillas-Targa F. An insulator embedded in the chicken  $\alpha$ -globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic Acids Res.* (2011) Jan;39(1):89-103.
2. Dávalos-Salas M, Furlan-Magaril M, González-Buendía E, Valdes-Quezada C, Ayala-Ortega E, Recillas-Targa F. Gain of DNA methylation is enhanced in the absence of CTCF at the human retinoblastoma gene promoter. *BMC Cancer.* (2011) Jun 10;11:232.
3. Saldaña-Meyer R, González-Buendía E, Guerrero G, Narendra V, Bonasio R, Recillas-Targa F, Reinberg D. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* (2014) Apr 1;28(7):723-34.
4. González-Buendía E, Pérez-Molina R, Ayala-Ortega E, Guerrero G, Recillas-Targa F. Experimental strategies to manipulate the cellular levels of the multifunctional factor CTCF. *Methods Mol Biol.* (2014) 1165:53-69.
5. González-Buendía E, Saldaña-Meyer R, Meier K, Recillas-Targa F. Transcriptome-Wide Identification of In Vivo Interactions Between RNAs and RNA-Binding Proteins by RIP and PAR-CLIP Assays. *Methods Mol Biol.* (2015) 1288:413-28.
6. González-Buendía E, Escamilla-Del-Arenal M, Pérez-Molina R, Tena J, Guerrero G, Suaste-Olmos F, Ayala-Ortega E, Gómez-Skarmeta JL, Recillas-Targa F. A novel chromatin insulator regulates the chicken folate receptor gene from the influence of nearby constitutive heterochromatin and the  $\beta$ -globin locus. *Biochim Biophys Acta.*(2015) Aug;1849(8):955-65.
7. Ayala-Ortega E, Arzate-Mejía, Pérez-Molina R, González-Buendía E, Guerrero G, and Recillas-Targa F. Epigenetic silencing of miR-181c by DNA methylation in glioblastoma cell lines. *Carcinogenesis* (2015).

## Bibliografía.

1. Felsenfeld G. A brief history of epigenetics. Cold Spring Harbor perspectives in biology. 2014;6(1). Epub 2014/01/05.
2. Riggs AD, Martienssen RA, Russo VEA. Introduction 1996.
3. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.
4. Consortium TEP. The ENCODE (ENCyclopedia Of DNA Elements) Project. Science (New York, NY). 2004;306(5696):636-40.
5. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
6. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes & development. 2011;25(21):2227-41. Epub 2011/11/08.
7. Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014;509(7502):575-81.
8. Roy AL, Singer DS. Core promoters in transcription: old problem, new insights. Trends in Biochemical Sciences. 2015;40(3):165-71.
9. Ohno S, editor. So much "junk" DNA in our genome. Brookhaven Symp Biol; 1972.
10. Kadonaga JT. Regulation of RNA Polymerase II Transcription by Sequence-Specific DNA Binding Factors. Cell. 2004;116(2):247-57.
11. Butler JE, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes & development. 2002;16(20):2583-92. Epub 2002/10/17.
12. Reina JH, Hernandez N. On a roll for new TRF targets. Genes & development. 2007;21(22):2855-60. Epub 2007/11/17.
13. Carlberg C, Molnár F. Overview: What Is Gene Expression? Mechanisms of Gene Regulation: Springer Netherlands; 2014. p. 3-15.
14. Lin Z, Wu W-S, Liang H, Woo Y, Li W-H. The spatial distribution of cis regulatory elements in yeast promoters and its implications for transcriptional regulation. BMC Genomics. 2010;11(1):581.
15. Flemming W. Beitrage zur Kenntniss der Zelle und ihrer Lebenserscheinungen. Archiv für mikroskopische Anatomie. 1879;16(1):302-436.
16. Flemming W. Zellsubstanz, kern und zelltheilung: Vogel; 1882.
17. Strickfaden H, Zunhammer A, van Koningsbruggen S, Köhler D, Cremer T. 4D Chromatin dynamics in cycling cells. Nucleus. 2010;1(3):284-97.
18. Felsenfeld G, Groudine M. Controlling the double helix. Nature. 2003;421(6921):448-53. Epub 2003/01/24.
19. Oberg C, Izzo A, Schneider R, Wrangé O, Belikov S. Linker histone subtypes differ in their effect on nucleosomal spacing in vivo. Journal of molecular biology. 2012;419(3-4):183-97. Epub 2012/03/27.
20. Robinson PJ, Rhodes D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Current opinion in structural biology. 2006;16(3):336-43. Epub 2006/05/23.

21. Daban J-R. Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure. *Micron*. 2011;42(8):733-50.
22. Ricci MA, Manzo C, Garcia-Parajo MF, Lakadamyali M, Cosma MP. Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*. 2015;160(6):1145-58. Epub 2015/03/15.
23. Passarge E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *American Journal of Human Genetics*. 1979;31(2):106-15.
24. Brown SW. Heterochromatin. *Science (New York, NY)*. 1966;151(3709):417-25. Epub 1966/01/28.
25. Trojer P, Reinberg D. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Molecular Cell*. 2007;28(1):1-13.
26. Muller HJ. Types of visible variations induced by X-rays in *Drosophila*. *Journal of Genetics*. 1930;22(3):299-334.
27. B M, editor. Control of gene action in maize. Brookhaven Symposia in Biology; 1965: NAT TECHNICAL INFORM SERVICE UP DEPT OF COMMERCE, SPRINGFIELD, VA 22161.
28. Sims RJ, 3rd, Nishioka K, Reinberg D. Histone lysine methylation: a signature for chromatin function. *Trends Genet*. 2003;19(11):629-39. Epub 2003/10/31.
29. Haber JE. Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics*. 2012;191(1):33-64. Epub 2012/05/05.
30. Grunstein M, Gasser SM. Epigenetics in *Saccharomyces cerevisiae*. *Cold Spring Harbor perspectives in biology*. 2013;5(7). Epub 2013/07/03.
31. Li J, Coic E, Lee K, Lee CS, Kim JA, Wu Q, et al. Regulation of budding yeast mating-type switching donor preference by the FHA domain of Fkh1. *PLoS genetics*. 2012;8(4):e1002630. Epub 2012/04/13.
32. Kueng S, Oppikofer M, Gasser SM. SIR proteins and the assembly of silent chromatin in budding yeast. *Annu Rev Genet*. 2013;47:275-306. Epub 2013/09/11.
33. Blander G, Guarente L. The Sir2 family of protein deacetylases. *Annu Rev Biochem*. 2004;73:417-35. Epub 2004/06/11.
34. Ottaviani A, Gilson E, Magdinier F. Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie*. 2008;90(1):93-107. Epub 2007/09/18.
35. Holloch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet*. 2015;16(2):71-84.
36. Yu R, Jih G, Iglesias N, Moazed D. Determinants of Heterochromatic siRNA Biogenesis and Function. *Molecular Cell*. 2014;53(2):262-76.
37. Giles KE, Ghirlando R, Felsenfeld G. Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. *Nature cell biology*. 2010;12(1):94-9; sup pp 1-6. Epub 2009/12/17.
38. Cohen AL, Jia S. Noncoding RNAs and the borders of heterochromatin. *Wiley Interdisciplinary Reviews: RNA*. 2014;5(6):835-47.
39. Cavalli G, Misteli T. Functional implications of genome topology. *Nature structural & molecular biology*. 2013;20(3):290-9. Epub 2013/03/07.
40. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*. 2014;15(4):234-46. Epub 2014/03/13.
41. Gualdi R, Bossard P, Zheng M, Hamada Y, Coleman JR, Zaret KS. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes & development*. 1996;10(13):1670-82. Epub 1996/07/01.

42. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell*. 2002;9(2):279-89. Epub 2002/02/28.
43. Tuteja G, Jensen ST, White P, Kaestner KH. Cis-regulatory modules in the mammalian liver: composition depends on strength of Foxa2 consensus site. *Nucleic acids research*. 2008;36(12):4149-57. Epub 2008/06/17.
44. Sekinger EA, Moqtaderi Z, Struhl K. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol Cell*. 2005;18(6):735-48. Epub 2005/06/14.
45. Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, et al. A genomic code for nucleosome positioning. *Nature*. 2006;442(7104):772-8. Epub 2006/07/25.
46. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes & development*. 2007;21(12):1519-29. Epub 2007/06/19.
47. Sherwood RI, Hashimoto T, O'Donnell CW, Lewis S, Barkal AA, van Hoff JP, et al. Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat Biotech*. 2014;32(2):171-8.
48. Zhang Y, Liang J, Li Y, Xuan C, Wang F, Wang D, et al. CCCTC-binding factor acts upstream of FOXA1 and demarcates the genomic response to estrogen. *Journal of Biological Chemistry*. 2010;285(37):28604-13.
49. Lobanenko VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*. 1990;5(12):1743-53. Epub 1990/12/01.
50. Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH, et al. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Molecular and cellular biology*. 1993;13(12):7612-24. Epub 1993/12/01.
51. Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and cellular biology*. 1996;16(6):2802-13. Epub 1996/06/01.
52. Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell*. 2009;137(7):1194-211. Epub 2009/07/01.
53. Kim S, Yu N-K, Kaang B-K. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med*. 2015;47:e166.
54. Klenova EM, Chernukhin IV, El-Kady A, Lee RE, Pugacheva EM, Loukinov DI, et al. Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. *Molecular and cellular biology*. 2001;21(6):2221-34. Epub 2001/03/10.
55. Ong CT, Van Bortle K, Ramos E, Corces VG. Poly(ADP-ribosyl)ation regulates insulator function and intrachromosomal interactions in *Drosophila*. *Cell*. 2013;155(1):148-59. Epub 2013/09/24.
56. Moore JM, Rabaia NA, Smith LE, Fagerlie S, Gurley K, Loukinov D, et al. Loss of maternal CTCF is associated with peri-implantation lethality of Ctfc null embryos. *PloS one*. 2012;7(4):e34915. Epub 2012/04/26.
57. Saldana-Meyer R, Gonzalez-Buendia E, Guerrero G, Narendra V, Bonasio R, Recillas-Targa F, et al. CTCF regulates the human p53 gene through direct interaction with its natural

- antisense transcript, Wrap53. *Genes & development*. 2014;28(7):723-34. Epub 2014/04/04.
58. Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet*. 2001;17(9):520-7. Epub 2001/08/30.
  59. Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet*. 2007;8(2):104-15. Epub 2007/01/19.
  60. Cremer T, Cremer C. Rise, fall and resurrection of chromosome territories: a historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. *European journal of histochemistry : EJH*. 2006;50(4):223-72. Epub 2007/01/11.
  61. Bonifer C, Hecht A, Saueressig H, Winter DM, Sippel AE. Dynamic chromatin: the regulatory domain organization of eukaryotic gene loci. *Journal of cellular biochemistry*. 1991;47(2):99-108. Epub 1991/10/01.
  62. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. *Nature*. 2007;447(7143):413-7. Epub 2007/05/25.
  63. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008;453(7197):948-51. Epub 2008/05/09.
  64. Burgess-Beusse B, Farrell C, Gaszner M, Litt M, Mutskov V, Recillas-Targa F, et al. The insulation of genes from external enhancers and silencing chromatin. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99 Suppl 4:16433-7. Epub 2002/08/03.
  65. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. 2015;518(7539):331-6. Epub 2015/02/20.
  66. Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*. 1993;74(3):505-14. Epub 1993/08/13.
  67. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*. 2000;405(6785):482-5. Epub 2000/06/06.
  68. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, et al. CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science (New York, NY)*. 2006;312(5771):269-72. Epub 2006/04/15.
  69. Chen H, Tian Y, Shu W, Bo X, Wang S. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PloS one*. 2012;7(7):e41374. Epub 2012/07/26.
  70. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*. 2008;133(6):1106-17. Epub 2008/06/17.
  71. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, et al. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome research*. 2012;22(9):1680-8. Epub 2012/09/08.
  72. Martin D, Pantoja C, Fernandez Minan A, Valdes-Quezada C, Molto E, Matesanz F, et al. Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nature structural & molecular biology*. 2011;18(6):708-14. Epub 2011/05/24.

73. de Wit E, de Laat W. A decade of 3C technologies: insights into nuclear organization. *Genes & development*. 2012;26(1):11-24. Epub 2012/01/05.
74. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 2012;485(7398):381-5. Epub 2012/04/13.
75. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *The EMBO journal*. 2008;27(4):654-66. Epub 2008/01/26.
76. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008;451(7180):796-801. Epub 2008/02/01.
77. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665-80. Epub 2014/12/17.
78. Downen Jill M, Fan Zi P, Hnisz D, Ren G, Abraham Brian J, Zhang Lyndon N, et al. Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. *Cell*. 159(2):374-87.
79. Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, Schwarzer W, et al. Functional and topological characteristics of mammalian regulatory domains. *Genome research*. 2014;24(3):390-400. Epub 2014/01/09.
80. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485(7398):376-80. Epub 2012/04/13.
81. Prioleau MN, Nony P, Simpson M, Felsenfeld G. An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. *The EMBO journal*. 1999;18(14):4035-48. Epub 1999/07/16.
82. Ma MK, Heath C, Hair A, West AG. Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity. *PLoS genetics*. 2011;7(7):e1002175. Epub 2011/08/04.
83. Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *The EMBO journal*. 2001;20(9):2224-35. Epub 2001/05/02.
84. Ghirlando R, Litt MD, Prioleau MN, Recillas-Targa F, Felsenfeld G. Physical properties of a genomic condensed chromatin fragment. *Journal of molecular biology*. 2004;336(3):597-605. Epub 2004/04/21.
85. Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science (New York, NY)*. 2001;293(5539):2453-5. Epub 2001/08/11.
86. Chung JH, Bell AC, Felsenfeld G. Characterization of the chicken beta-globin insulator. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(2):575-80. Epub 1997/01/21.
87. Metz T, Graf T. v-myb and v-ets transform chicken erythroid cells and cooperate both in trans and in cis to induce distinct differentiation phenotypes. *Genes & development*. 1991;5(3):369-80. Epub 1991/03/01.
88. Davalos-Salas M, Furlan-Magaril M, Gonzalez-Buendia E, Valdes-Quezada C, Ayala-Ortega E, Recillas-Targa F. Gain of DNA methylation is enhanced in the absence of CTCF at the human retinoblastoma gene promoter. *BMC cancer*. 2011;11:232. Epub 2011/06/15.

89. Valadez-Graham V, Razin SV, Recillas-Targa F. CTCF-dependent enhancer blockers at the upstream region of the chicken alpha-globin gene domain. *Nucleic acids research*. 2004;32(4):1354-62. Epub 2004/02/26.
90. Yuan B, Latek R, Hossbach M, Tuschl T, Lewitter F. siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic acids research*. 2004;32(Web Server issue):W130-4. Epub 2004/06/25.
91. Gomes NP, Espinosa JM. Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes & development*. 2010;24(10):1022-34. Epub 2010/05/19.
92. Gonzalez-Buendia E, Perez-Molina R, Ayala-Ortega E, Guerrero G, Recillas-Targa F. Experimental strategies to manipulate the cellular levels of the multifunctional factor CTCF. *Methods in molecular biology (Clifton, NJ)*. 2014;1165:53-69. Epub 2014/05/20.
93. Escamilla-Del-Arenal M, Recillas-Targa F. GATA-1 modulates the chromatin structure and activity of the chicken alpha-globin 3' enhancer. *Molecular and cellular biology*. 2008;28(2):575-86. Epub 2007/11/07.
94. Furlan-Magaril M, Rebollar E, Guerrero G, Fernandez A, Molto E, Gonzalez-Buendia E, et al. An insulator embedded in the chicken alpha-globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic acids research*. 2011;39(1):89-103. Epub 2010/09/04.
95. Bessa J, Tena JJ, de la Calle-Mustienes E, Fernandez-Minan A, Naranjo S, Fernandez A, et al. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*. 2009;238(9):2409-17. Epub 2009/08/05.
96. Recillas-Targa F, Valadez-Graham V, Farrell CM. Prospects and implications of using chromatin insulators in gene therapy and transgenesis. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2004;26(7):796-807. Epub 2004/06/29.
97. Mutskov V, Felsenfeld G. Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *The EMBO journal*. 2004;23(1):138-49. Epub 2003/12/20.
98. Grewal SI. RNAi-dependent formation of heterochromatin and its diverse functions. *Current opinion in genetics & development*. 2010;20(2):134-41. Epub 2010/03/09.
99. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature*. 2012;489(7414):101-8. Epub 2012/09/08.
100. Roy AL, Singer DS. Core promoters in transcription: old problem, new insights. *Trends in biochemical sciences*. 2015;40(3):165-71. Epub 2015/02/15.
101. Taft RJ, Hawkins PG, Mattick JS, Morris KV. The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. *Epigenetics & chromatin*. 2011;4:13. Epub 2011/08/05.
102. Taft RJ, Glazov EA, Cloonan N, Simons C, Stephen S, Faulkner GJ, et al. Tiny RNAs associated with transcription start sites in animals. *Nature genetics*. 2009;41(5):572-8. Epub 2009/04/21.
103. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, et al. Divergent transcription from active promoters. *Science (New York, NY)*. 2008;322(5909):1849-51. Epub 2008/12/06.
104. Yan BX, Ma JX. Promoter-associated RNAs and promoter-targeted RNAs. *Cellular and molecular life sciences : CMLS*. 2012;69(17):2833-42. Epub 2012/03/15.

105. Caudron-Herger M, Rippe K. Nuclear architecture by RNA. *Current opinion in genetics & development*. 2012;22(2):179-87. Epub 2012/01/28.
106. Lim SJ, Boyle PJ, Chinen M, Dale RK, Lei EP. Genome-wide localization of exosome components to active promoters and chromatin insulators in *Drosophila*. *Nucleic acids research*. 2013;41(5):2963-80. Epub 2013/01/30.
107. Lefevre P, Witham J, Lacroix CE, Cockerill PN, Bonifer C. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol Cell*. 2008;32(1):129-39. Epub 2008/10/15.
108. West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell*. 2004;16(3):453-63. Epub 2004/11/05.