



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE MÉXICO



---

---

**FACULTAD DE QUÍMICA**

PROGRAMA DE MAESTRÍA Y DOCTORADO  
EN CIENCIAS BIOQUÍMICAS

**ESTUDIO DE LA FORMACIÓN Y  
MANTENIMIENTO DE UN DOMINIO GÉNICO:  
EL PAPEL DE LOS *INSULATORS*  $\alpha$ -GLOBINA DE  
POLLO**

**T E S I S**

QUE PARA OBTENER EL GRADO DE:

**DOCTOR EN CIENCIAS (BIOQUÍMICAS)**

P R E S E N T A:

**MAYRA FURLAN MAGARIL**



Tutor: Dr. Félix Recillas Targa

MEXICO D.F. Diciembre 2010



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.



# AGRADECIMIENTOS

Quisiera agradecer a Félix por su invaluable asesoría durante la realización de este trabajo. Su dedicación, entusiasmo y compromiso con la ciencia y la formación de alumnos es motivo de admiración y transmite el mensaje esencial y a veces oculto para tener éxito en la investigación: las ideas brillantes no son nada sin constancia en el trabajo cotidiano.

Agradezco a Geito, con tu alegría y cariño haces del laboratorio un lugar placentero para trabajar. Gracias por tu apoyo tu escucha y tu buen ánimo. Te voy a extrañar mucho en cualquier lugar a donde vaya.

Agradezco a Eria porque cuando están juntas “las hermanas probeta” cualquier experimento es un juego y cualquier discusión científica una forma más de amistad.

A Martín, porque su buena vibra y amistad me ayudaron a transitar por el doctorado y no perder la actitud alegre en el camino.

A mis compañeros Meche y Edgar. Por las pláticas, los cafés, las confesiones y consejos. Por su cariño y disposición emocional y académica. Gracias por escucharme y rescatarme en diversas ocasiones. Los quiero mucho. Se que nos seguiremos encontrando.

A las chicas: Charis, Erandita y Andre. Porque en los últimos años ustedes vinieron a inyectarle brillo y carisma al 122Nte y me recordaron lo increíble de tener un equipo de trabajo en donde la solidaridad y el compañerismo estén a la par de la capacidad crítica. Chicas, son lo máximo. Les deseo la mejor de las suertes en todo lo que se propongan.

A todos los miembros del lab, de antes y de ahora. Héctor, Paul, Inti, Abrahan, Neto, Pedro, Estelita, Ricardo, David, Rodrigo, Gabriel, Cristian A, Christian V y Gianellita. Por ser un grupo increíble y enseñarme a trabajar en equipo.

A Lety García, porque su trabajo es excelente. Por conocer a todos y cada uno de los que integramos el postgrado de Bioquímicas y aguantar nuestras miles de preguntas siempre con una sonrisa.

A mamá, por que su amor y orgullo me hacen sentir que lo que hago tiene un valor trascendental, aunque solo sea un paso pollito para la humanidad. Por ser una mujer fuerte y decidida que malabarea miles de cosas a la vez sin perder nunca el estilo y las ganas de reír.

A papá, porque su entereza es monumental y su vitalidad me contagia de ganas de disfrutarlo todo. Porque siempre está interesado en cómo van mis experimentos y quiere saber todos los detalles para tratar de aportar algo. Por su alma punk y su sentido del humor.

A Logí, por todo su amor. Porque mi vida sin ella no sería válida. Porque juntas podemos deshacerlo todo y volver a empezar. Porque... chee, contáte algo.

A Flor y Nico, porque por suerte llegaron a esta ciudad. Porque junto a ustedes todo tiene mucho más sentido y se me han vuelto indispensables, desde donde quiera que estemos.

A Ceci, por todo tu cariño. Por coincidir en el camino y decidir encontrarnos.

A los Rosen todos: Ana, Fer, Seba, Hilda y Jorge. Gracias por su presencia constante y amorosa.

A mis amigas, mi segunda familia, mi alma toda:

A Truchita, otra vez. Gracias por compartir conmigo la vida pescadito. Tu amistad es para mi una evidencia de lo que realmente importa. Juntas hemos aprendido a reirnos de todo aunque sea entre lágrimas y ahora, es esa la única manera en la que sé y quiero vivir.

A la nena, Hele. Gracias por todo tu amor mi Hel, por tu escucha sin juicio. Gracias por compartir tu mundo conmigo y pensar que mis consejos pueden servir de algo. Gracias por siempre estar cerca y no asustarte ante la adversidad sino permitir que nos asustemos juntas. Te quiero.

A Julieta. Te agradezco porque, a pesar de habitar territorios distintos, nuestra amistad solo se ha hecho más fuerte y siempre te siento cerca. Por tu cariño, porque tu forma de ver la vida me recuerda que todo es posible, que lo que importa es lo que late y suena... lo demás, viene y va.

A Julita. Te agradezco mi Juji porque siempre contigo me encuentro en ese mundo al que me encanta pertenecer, en dónde habitan los personajes fantásticos de nuestra amistad y cariño. Porque siempre te reencuentro y vuelvo a admirarte, quererte y sorprenderme.

A Nadia. Porque eres mi hermana del alma y te quiero desde el fondo de mi corazón. Porque mi vida tiene una forma y contenido que de adentro hacia afuera llevan impresos tu amistad... más allá del tiempo-espacio que la contenga.

A Andre. Gracias por estar ahí todos estos años, por todo el cariño y por seguir compartiendo lo que pasa y nos pasa. Porque la alegría y entusiasmo con la que emprendes todos tus proyectos es brillante y se transmite.

A Ju Pozas. Porque a pesar de no coincidir geográficamente siempre estuviste presente mi Ju y tu alegría y compañía son esenciales para mi.

A Brain y a Javi. Por su amistad y cariño. Porque cuando uno dice "las chicas" ustedes dos siempre están incluidos. Los quiero mucho.

A Tucho, por su perritud alargada y encantadora.

Finalmente agradezco a Mari:

Gordino, gracias por acompañarme durante estos años. Por quererme, por aguantar mi locura, por hacerme reír. Por escuchar las mil y un historias que conformaron a ésta que se convirtió en tesis y por querer construir conmigo mil historias más. Porque a tu lado, todo se transforma en algo que disfrutar.

Gracias Totales.

# INDICE

<b><u>Resumen</u></b> .....	4
<b><u>Abstract</u></b> .....	5
<b><u>Abreviaturas</u></b> .....	6
<b>1. <u>INTRODUCCIÓN</u></b> .....	12
<b>1.1 El genoma eucarionte y la estructura de la cromatina</b> .....	12
1.1.1 Modificaciones post-traduccionales de las histonas.....	15
1.1.2 Variantes de histonas.....	19
1.1.3 Complejos remodeladores de cromatina.....	22
<b>1.2 La organización del genoma al interior del núcleo</b> .....	23
1.2.1 Territorios cromosomales.....	23
1.2.2 Fábricas Transcripcionales.....	27
<b>1.3 Dominios cromosómicos y sus elementos de regulación</b> .....	28
1.3.1 <i>Insulators</i> .....	30
1.3.2 El factor CTCF.....	36
<b>2. <u>ANTECEDENTES</u></b> .....	43
<b>2.1 El dominio <math>\beta</math>-globina</b> .....	44
2.1.1 El dominio $\beta$ -globina de pollo.....	46
<b>2.2 El dominio <math>\alpha</math>-globina</b> .....	50
2.2.1 El dominio $\alpha$ -globina de pollo.....	53
<b>3. <u>PLANTEAMIENTO DEL PROBLEMA</u></b> .....	57
<b>4. <u>HIPÓTESIS</u></b> .....	57

<b>5. <u>OBJETIVO GENERAL</u></b> .....	<b>58</b>
<b>5.1 Objetivos Particulares</b> .....	<b>58</b>
<b>6. <u>ESTRATEGIAS EXPERIMENTALES</u></b> .....	<b>59</b>
<b>6.1 Cultivo de líneas celulares</b> .....	<b>59</b>
<b>6.2 Diferenciación de células HD3</b> .....	<b>59</b>
<b>6.3 Transfecciones transitorias y estables</b> .....	<b>60</b>
<b>6.4 Ensayo de protección contra efectos de posición cromosomal (EPC)</b> .....	<b>61</b>
<b>6.5 Inmunoprecipitación de la cromatina (ChIP)</b> .....	<b>62</b>
<b>6.6 Generación del “knockdown” contra CTCF y tratamiento con 3-Aminobenzamida</b> .....	<b>64</b>
<b>6.7 Análisis de expresión mediante RT-PCR</b> .....	<b>66</b>
<b>6.8 Western blot: Electroforesis e Inmunodetección</b> .....	<b>66</b>
<b>6.9 Ensayo de Hipersensibilidad a la DNasa I</b> .....	<b>68</b>
<b>6.10 Anticuerpos</b> .....	<b>69</b>
<b>6.11 Estadística</b> .....	<b>70</b>
<b>6.12 Oligonucleótidos y plásmidos utilizados</b> .....	<b>70</b>
<b>7. <u>RESULTADOS y DISCUSIÓN</u></b> .....	<b>78</b>
<b>7.1 El <i>insulator</i> <math>\alpha</math>EHS-1.4 protege a un transgén contra EPC en células eritroblásticas de pollo</b> .....	<b>78</b>
<b>7.2 Los transgenes protegidos con <i>insulators</i> presentan un perfil permisivo de marcas post-traduccionales de histonas</b> .....	<b>85</b>
<b>7.3 Los EPC conllevan a la desacetilación de histonas y a la metilación del ADN de los transgenes</b> .....	<b>90</b>

<b>7.4 El <i>insulator</i> <math>\alpha</math>EHS-1.4 protege contra EPC en ratones transgénicos.....</b>	<b>91</b>
<b>7.5 CTCF es necesario en la actividad del <i>insulator</i> <math>\alpha</math>EHS-1.4.....</b>	<b>97</b>
7.5.1 Mutaciones puntuales sobre la secuencia de unión de CTCF en el <i>insulator</i> $\alpha$ EHS-1.4 interfieren con su actividad....	<b>98</b>
7.5.2 La asociación de CTCF al <i>insulator</i> $\alpha$ EHS-1.4 endógeno durante la diferenciación eritroide y el desarrollo del organismo es dinámica .....	<b>100</b>
<b>7.6 La actividad del <i>insulator</i> <math>\alpha</math>EHS-1.4 correlaciona con la asociación de CTCF durante la diferenciación eritroide.....</b>	<b>101</b>
<b>7.7 Alteración de la abundancia y la función de CTCF.....</b>	<b>104</b>
7.7.1 Generación de líneas celulares “knockdown” contra CTCF.....	<b>104</b>
7.7.2 La inhibición de la poli(ADPribosil)asa polimerasa (PARP-1) interfiere con la función del <i>insulator</i> $\alpha$ EHS-1.4.....	<b>109</b>
<b>7.8 La alteración de la función del <i>insulator</i> <math>\alpha</math>EHS-1.4 genera una configuración activa de la cromatina y una desregulación de la expresión de los genes <math>\alpha</math>-globina adultos.....</b>	<b>111</b>
<b>7.9 La transactivación génica por parte del <math>\alpha</math>MRE se potencia ante la disociación de CTCF del <i>insulator</i> <math>\alpha</math>EHS-1.4.....</b>	<b>115</b>
<b>8. <u>CONCLUSIONES</u>.....</b>	<b>119</b>
<b>9. <u>PERSPECTIVAS</u>.....</b>	<b>125</b>
<b>10. <u>REFERENCIAS</u>.....</b>	<b>128</b>
<b>11. <u>ANEXO 1</u>. Identificación de nuevos elementos de regulación en el dominio <math>\alpha</math>-globina de pollo.....</b>	<b>135</b>

<b>11.1 Caracterización de la región 3' no-codificante del dominio <math>\alpha</math>-globina de pollo.....</b>	<b>135</b>
11.1.1 Ensayo de hipersensibilidad al corte por la DNasa I en la región 3' no-codificante del dominio $\alpha$ -globina de pollo.....	<b>136</b>
11.1.2 Evaluación de la función de los elementos identificados en la región 3' no-codificante del dominio $\alpha$ -globina de pollo.....	<b>139</b>
<b>11.2 Ensayo de accesibilidad al corte por la DNasa I en la región 5' no-codificante del dominio <math>\alpha</math>-globina de pollo.....</b>	<b>142</b>
11.2.1 Ensayo de hipersensibilidad al corte por la DNasa I en la región 5' no-codificante del dominio $\alpha$ -globina de pollo.....	<b>143</b>
11.2.2 Evaluación de la función del elemento $\alpha$ MRE identificado en la región 5' no-codificante del dominio $\alpha$ -globina.....	<b>145</b>
<b>12. <u>ANEXO 2</u>. Publicaciones.....</b>	<b>148</b>

## RESUMEN

La organización del genoma en dominios transcripcionales activos representa uno de los primeros niveles de regulación de la expresión génica. El concepto de dominio cromatínico es aceptado de manera general, sin embargo, se conoce muy poco acerca de cómo la organización del genoma en dominios afecta la regulación de la expresión diferencial de genes. Los elementos de regulación conocidos como *insulators* o fronteras de cromatina son relevantes para delimitar a los dominios transcripcionales protegiéndolos tanto de señales externas como del estado cromatínico circundante.

En estudios previos del laboratorio se identificó un *insulator* dependiente del factor CTCF con actividad de bloqueo de *enhancer*, embebido en la región 5' no-codificante del dominio  $\alpha$ -globina de pollo. En el presente trabajo de investigación demostramos que dicho *insulator*, denominado  $\alpha$ EHS-1.4, protege un transgén contra efectos de posición cromosomales en líneas celulares transfectadas de manera estable y en ratones transgénicos. Además, encontramos que este *insulator* genera un ambiente cromatínico regulado en el dominio  $\alpha$ -globina que coincide con el inicio de la expresión de los genes  $\alpha$ -globina adultos. Por otro lado, encontramos que la actividad del *insulator*  $\alpha$ EHS-1.4 es dependiente de la unión regulada de CTCF así como de su poly(ADP-ribosil)ación. Nuestros resultados sugieren que el *insulator*  $\alpha$ EHS-1.4 contribuye en organizar la estructura de la cromatina del dominio  $\alpha$ -globina y previene la activación de los genes adultos en la etapa de eritroblastos via CTCF. El papel que hemos descrito para el *insulator*  $\alpha$ EHS-1.4 dentro del dominio  $\alpha$ -globina, sugiere que dichos elementos no sólo forman fronteras constitutivas que delimitan a un dominio génico del resto del genoma, sino que también pueden participar en la regulación transcripcional fina de genes durante procesos de diferenciación celular.

## ABSTRACT

Genome organization into transcriptionally active domains denotes one of the first levels of gene expression regulation. Although the chromatin domain concept is generally accepted, only little is known on how domain organization impacts the regulation of differential gene expression. Regulatory elements known as insulators or chromatin barriers are relevant to delimit transcriptional domains protecting them from external signals and the surrounding chromatin environment.

We have previously identified a CTCF-dependent insulator with enhancer-blocking activity embedded in the 5' non-coding region of the chicken  $\alpha$ -globin domain. Here, we demonstrate that this element, called the  $\alpha$ EHS-1.4 insulator, protects a transgene against chromosomal position effects in stably transfected cell lines and transgenic mice. We found that this insulator can create a regulated chromatin environment that coincides with the onset of adult  $\alpha$ -globin gene expression. Furthermore, such activity is in part dependent on the in vivo regulated occupancy of CTCF at the  $\alpha$ EHS-1.4 element. Insulator function is also regulated by CTCF poly(ADP-ribosyl)ation. Our results suggest that the  $\alpha$ EHS-1.4 insulator contributes in organizing the chromatin structure of the  $\alpha$ -globin gene domain and prevents activation of adult  $\alpha$ -globin gene expression at the erythroblast stage via CTCF. The role that we have described for the  $\alpha$ EHS-1.4 insulator inside the  $\alpha$ -globin gene domain suggests that these elements could act, not only as constitutive chromatin barriers, but also as active participants in the transcriptional regulation of specific genes throughout cellular differentiation.

# 1. INTRODUCCIÓN

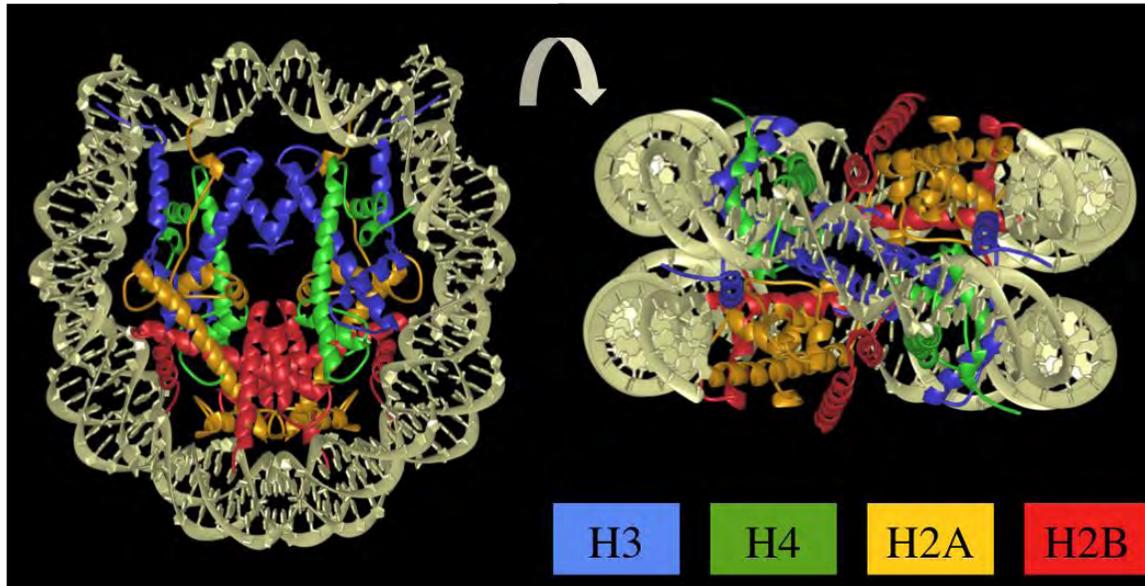
## 1.1 El genoma eucarionte y la estructura de la cromatina

El desarrollo de los organismos multicelulares conlleva la especificación de múltiples identidades celulares. Las células somáticas dentro de un organismo contienen el mismo genoma y sin embargo, poseen características distintas derivadas del perfil transcripcional específico de cada tejido. ¿Cómo es entonces que se mantienen dichos perfiles de expresión génica durante las divisiones celulares?, parte de la respuesta a dicha interrogante se encuentra en el concepto de epigenética.

Antes de la década de 1950, el término epigenética fue utilizado para categorizar todos los eventos del desarrollo que llevan del cigoto al organismo maduro, es decir, el conjunto de procesos regulados que, empezando por el material genético, dan forma al producto final (Waddington 1953, revisado en Epigenetics, 2006). Esta concepción de la epigenética tiene sus orígenes en trabajos de embriología, biología celular y genética realizados desde finales del siglo XIX y principios del siglo XX, que fundaron las bases para generar nuestro entendimiento actual de la relación entre los genes y el desarrollo del organismo. El concepto actual de epigenética comprende a toda la información estable y heredable (mitótica y/o meióticamente) que no está codificada en la secuencia del ADN (revisado en Epigenetics, 2006; Margueron y Reinberg, 2010).

La epigenética incluye diversos procesos que en su mayoría tienen como sustrato a la estructura de la cromatina. Las células eucariontes a diferencia de las procariontes, tienen una estructura nuclear que contiene al genoma. De forma lineal el genoma eucarionte tiene una longitud de 2 a 4 metros mientras que el núcleo presenta un diámetro de entre 10 y 20 micras, lo cual indica que debe existir un sistema a través del cual el genoma se compacta y empaqueta en el núcleo celular. Ahora sabemos que este sistema de empaquetamiento consiste en la asociación del ADN con una serie de proteínas llamadas histonas que en su conjunto conforman a la cromatina. Inicialmente se

pensó que la cromatina tenía sólo una función de empaquetamiento, sin embargo, muchas evidencias han mostrado que es una entidad dinámica central en una gran variedad de actividades bioquímicas y en la regulación de la expresión genética.

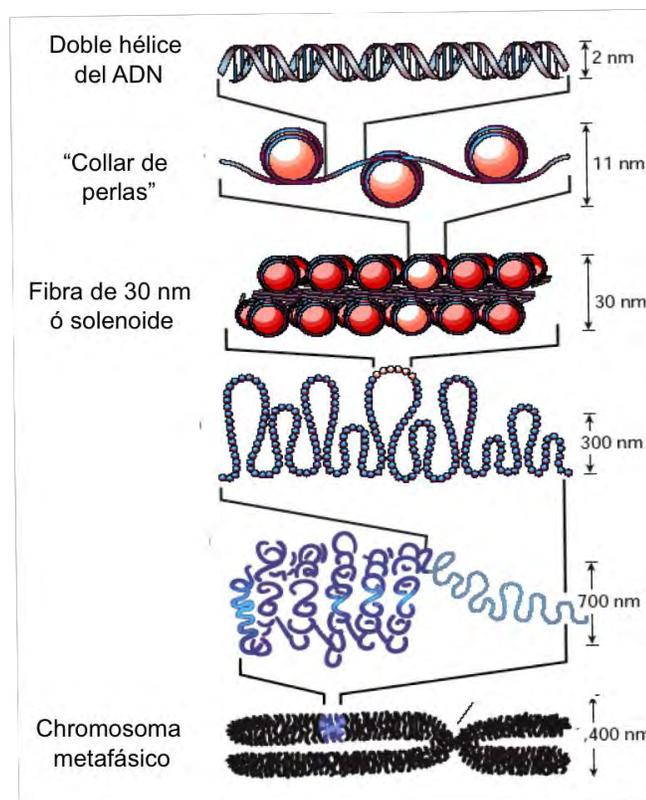


**Figura 1. La estructura del nucleosoma.** Se muestran el ADN (en blanco) y el octámero de histonas (conformado por un tetrámero de H3-H4 y 2 dímeros de H2A-H2B) (Modificado de [cmgm.stanford.edu/biochem201/Slides/Chromatin](http://cmgm.stanford.edu/biochem201/Slides/Chromatin))

La unidad fundamental de la cromatina es el nucleosoma el cual consiste de alrededor de 147 pares de bases de ADN enrollados en dos giros superhelicoidales sobre un octámero formado por dos de las histonas H2A, H2B, H3 y H4, lo cual conforma el núcleo o “core” del nucleosoma (Fig. 1). Cada nucleosoma está conectado al siguiente a través de un segmento de ADN inter-nucleosomal de 10-80 pares de bases. Éste es el nivel mínimo de organización y se conoce como el collar de perlas o fibra de 11 nm. El collar del perlas se estructura a su vez formando una fibra compacta de alrededor de 30 nm o solenoide, la cual se estabiliza por la unión de la histona H1 al ADN internucleosomal. El estado más abundante de la cromatina en la célula eucarionte en interfase es el de la fibra de 30 nm, por lo cual, toda la maquinaria metabólica del ADN se enfrenta con esta estructura y es capaz de remodelarla para generar de manera transitoria a la fibra de 11 nm o “collar de perlas”, la cual es más accesible a las diferentes maquinarias de regulación del genoma (Li *et al.*, 2010). Posteriormente, existen niveles de compactación superiores

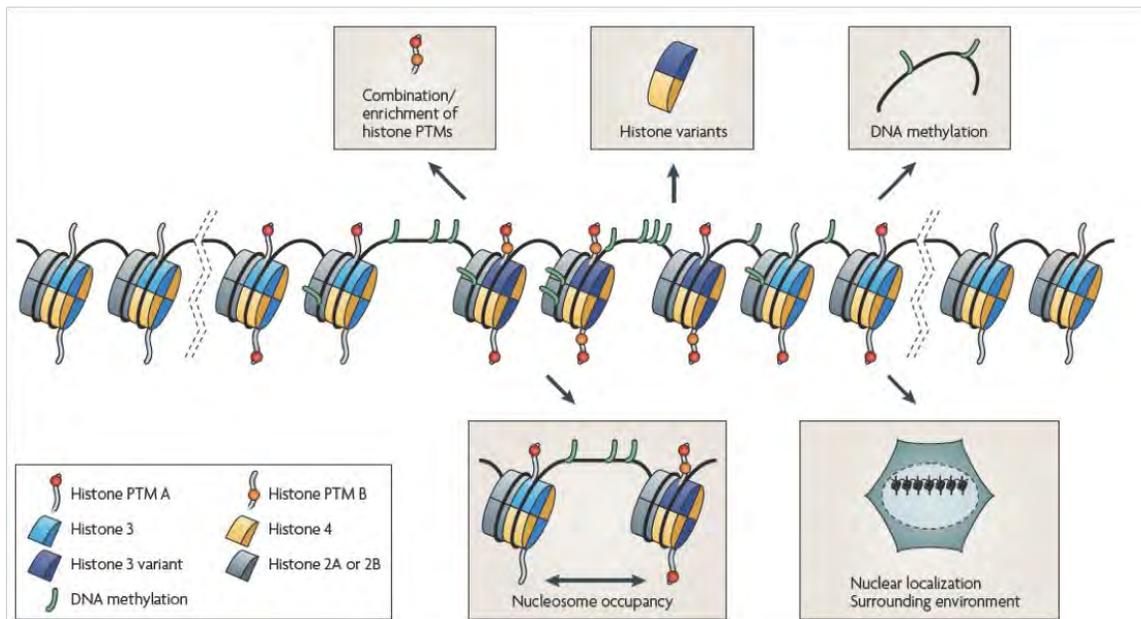
hasta alcanzar el grado máximo que se encuentra en el cromosoma metafásico de aproximadamente 1400 nm (Felsenfeld y Groudine, 2003) (Fig. 2).

El grado de compactación de la cromatina determina el acceso al ADN que tienen los factores involucrados en la transcripción y otros complejos que participan en el metabolismo del ADN. Existen regiones en donde la cromatina se encuentra en un estado muy compacto (heterocromatina) que son transcripcionalmente inactivas, mientras que las regiones transcripcionalmente activas, presentan una cromatina más laxa (eucromatina). La diferencia en los niveles de compactación de la cromatina se debe a que sus componentes están sujetos a una serie de modificaciones que pueden ocurrir a los diferentes niveles de organización de su estructura. Sin embargo, la mayoría de los estudios se han enfocado en estudiar las alteraciones que ocurren a nivel del nucleosoma.



**Figura 2. Los diferentes niveles de compactación del genoma.** El “collar de perlas” o fibra de 11 nm está compuesta del ADN enrollado en un octámero de proteínas llamadas histonas. Al incorporarse la histona H1 a la zona internucleosomal, se genera el solenoide o fibra de 30 nm. Existen niveles de compactación subsecuentes hasta llegar al cromosoma metafásico presente únicamente durante la mitosis (modificado de Felsenfeld y Groudine, 2003).

Existen diferentes tipos de alteraciones en la estructura de la cromatina y la amplia gama de los componentes involucrados en regular su función. Entre ellos se encuentran las modificaciones covalentes de las histonas, el reemplazo de histonas del núcleo nucleosomal por variantes particulares, el remodelaje de nucleosomas por complejos proteicos que utilizan ATP, las modificaciones realizadas por los grupos de proteínas represoras Polycomb y activadoras Thrtorax, la metilación del ADN, la función de transcritos no-codificantes y la organización y dinámica tridimensional de la cromatina en el espacio nuclear. Todos estos mecanismos participan y definen un dominio cromatínico en el genoma (Fig. 3). A continuación se abordarán solo los más relevantes en relación al presente trabajo de investigación.



**Figura 3. El conjunto de procesos epigenéticos determinan el estatus de un dominio cromatínico en el genoma.** En esta figura se muestra como la combinatoria de rasgos epigenéticos como las variantes de histonas, la metilación del ADN, las modificaciones post-traduccionales de las histonas (MPTs) y la localización en el espacio nuclear, determinan el estado y accesibilidad de un dominio al interior del genoma (modificado de Margueron y Reinberg, 2010).

### 1.1.1 Modificaciones post-traduccionales de las histonas

Las histonas que conforman al nucleosoma están sujetas a diversas modificaciones post-traduccionales (MPTs) tanto en el “core” como en la región amino terminal (Fig. 4). La combinatoria de MPTs presentes en un nucleosoma genera respuestas biológicas diferentes que residen al centro de la regulación de los procesos que ocurren sobre el

templado de cromatina. Dado que las MPTs de las histonas tienen influencia en el reclutamiento de efectores que modifican a la cromatina, se ha propuesto la existencia de un “código histónico” para intentar correlacionar a los distintos perfiles de MPTs con respuestas biológicas particulares (revisado en Campos y Reinberg, 2009). A la luz de esta hipótesis, se han realizado enormes esfuerzos por intentar describir la complejidad de los mensajes que encierran los perfiles de MPTs de histonas a nivel global en el genoma. Sin embargo, el dinamismo de la estructura nucleosomal y cromatínica, así como la interrelación de todos sus componentes, (remodeladores ATP-dependientes, variantes de histonas, metilación del ADN entre otros), ha dificultado el poder traducir con precisión el significado (si es que el código existe como tal), encriptado en los perfiles de MPTs en las histonas y ha derivado a proponer la existencia de un “código multivalente” ó “código nucleosomal” que incorpore la enorme complejidad observada (revisado en Campos y Reinberg, 2009).

La interpretación de los perfiles de MPTs de las histonas se lleva a cabo mediante diversos mecanismos no excluyentes. En primer lugar, existen efectos directos sobre la estructura del nucleosoma, particularmente mediante la neutralización o adición de carga, lo cual aumenta la movilidad nucleosomal abatiendo contactos particulares entre el ADN y las histonas dentro de éste. Tal es el caso de la acetilación de las lisinas en las histonas H3 y H4. De manera inversa la desacetilación promueve un nucleosoma con menor movilidad promoviendo una estructura cromatínica más compacta (revisado en Ruthenburg *et al.*, 2007).

Otro mecanismo radica en la alteración de la estructura de la cromatina a través de procesos mediados por efectores no histónicos que reconocen cierto perfil de MPTs en un nucleosoma. Las proteínas efectoras pueden mediar las propiedades de la cromatina a través de entrecruzar uno o más nucleosomas, aumentar la ocupación de la RNA polimerasa II y sus factores asociados en la fibra cromatínica, y a través del reclutamiento activo de remodeladores que modifican o alteran químicamente su estructura (revisado en Ruthenburg *et al.*, 2007). Por ejemplo, la acetilación de histonas es reconocida por proteínas efectoras con bromodominios que promueven la apertura de la cromatina. Otro

ejemplo es el caso de la metilación de lisinas como la H3K4me3 que es reconocida por factores que contienen dominios PHD, en zonas que correlacionan con cromatina transcripcionalmente activa, generalmente promotores. Por el contrario, la metilación de la H3K9me3 correlaciona con la proteína HP1 en regiones de heterocromatina constitutiva. Sin embargo, a medida que los estudios a gran escala se incrementan, dichas correlaciones parecen no ser tan estrictas. Por ejemplo, en células troncales embrionarias, la H3K4me3 coexiste en algunos sitios con la H3K27me3, marca relacionada con la compactación de la cromatina a través del grupo de proteínas Polycomb. Dicha combinación genera un estatus de cromatina “bivalente” de gran importancia en la plasticidad del epigenoma en este tipo celular pluripotente que generará múltiples tipos celulares diferentes (revisado en Ruthenburg *et al.*, 2007).

Existen otras modificaciones post-traduccionales de histonas que alteran a la estructura de la cromatina a través de mecanismos menos conocidos. Estas incluyen la fosforilación, la sumoilación, la PARilación y la ubiquitinación entre otras (Fig. 4).

# Regulation of Histone Modifications

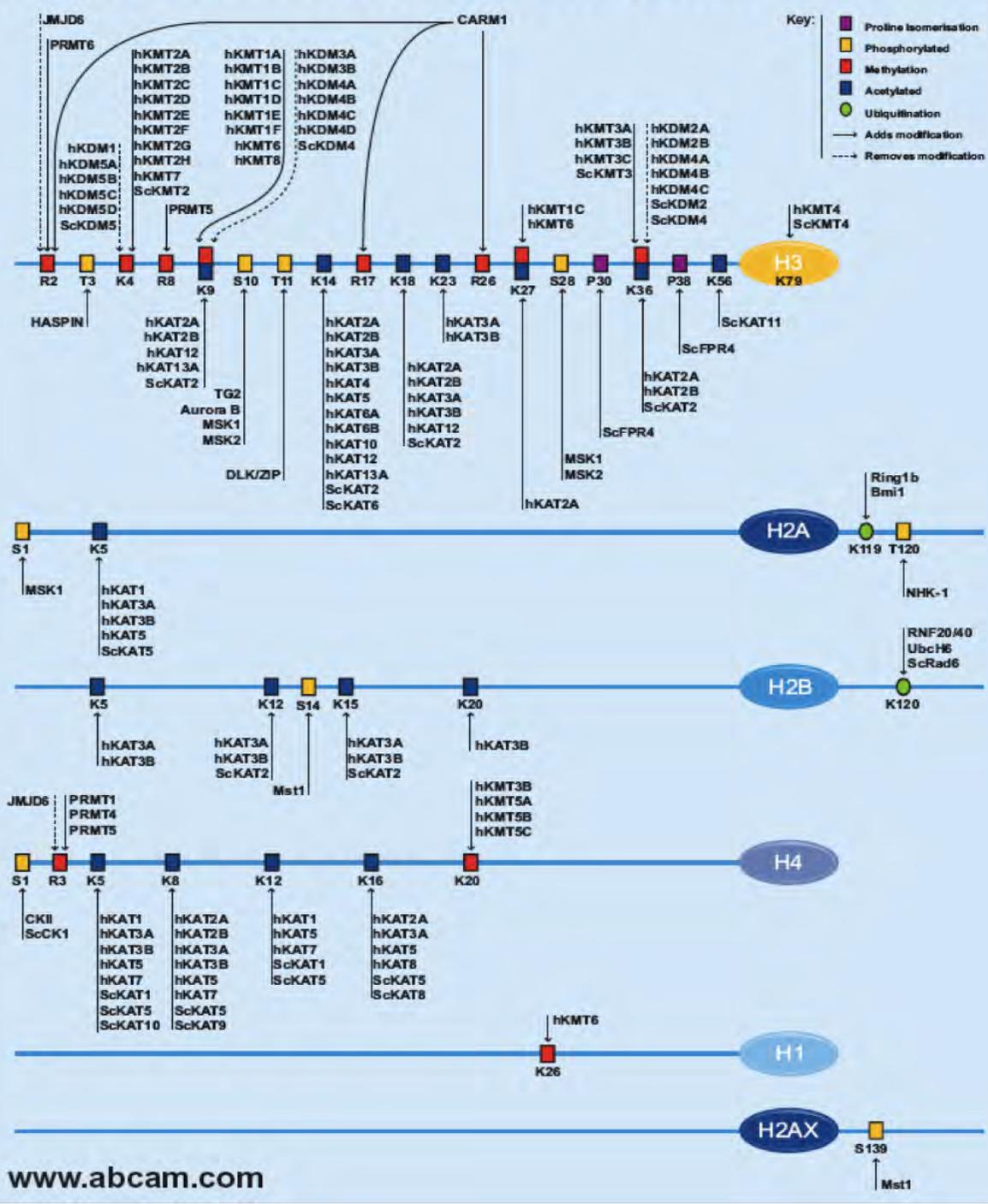


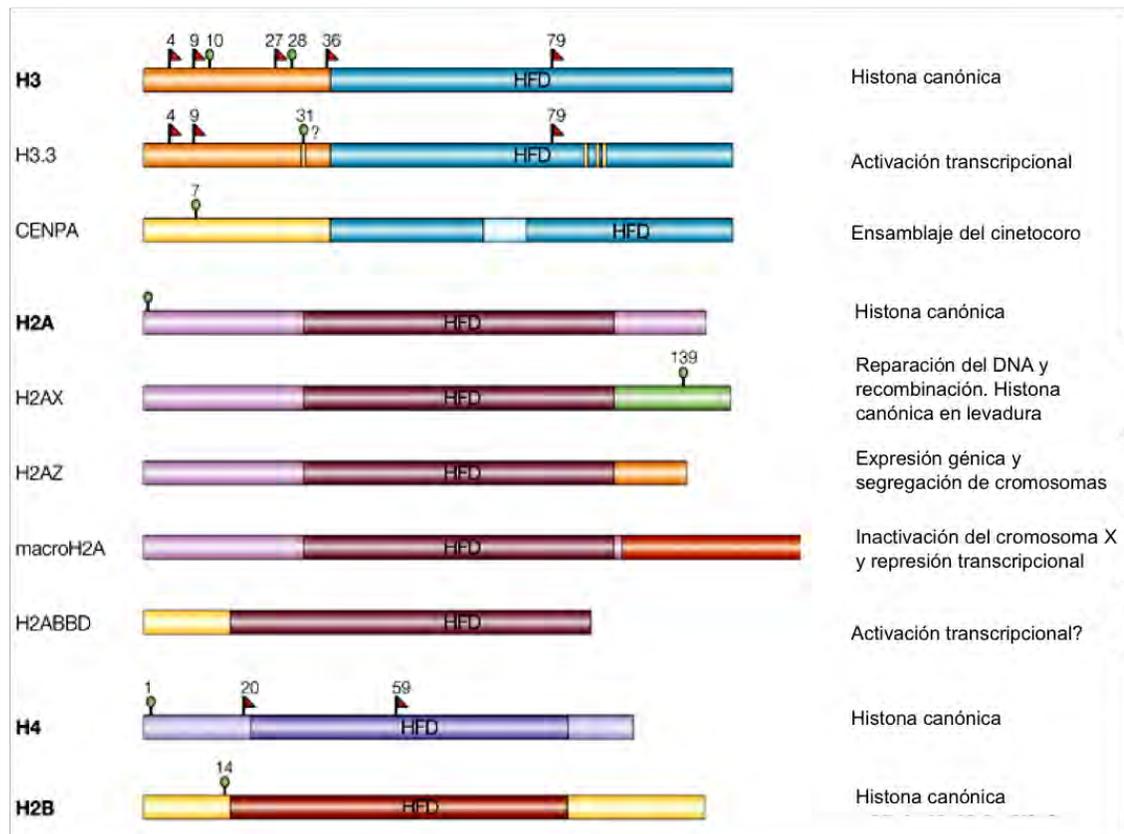
Figura 4. Esquema de las modificaciones post-traduccionales de histonas (MPTs) conocidas y las enzimas que las catalizan (tomado de www.abcam.com).

### 1.1.2 Variantes de histonas

Otro mecanismo a través del cual la estructura del nucleosoma puede ser modificada es el reemplazo de histonas. La existencia de variantes de histonas se reconoció desde hace tiempo, sin embargo, el reemplazo de las histonas canónicas como mecanismo de regulación de la estructura de la cromatina sólo recientemente ha adquirido relevancia (Henikoff *et al.*, 2004) y resulta interesante considerarlo en los distintos fenómenos asociados con la regulación epigenética.

Los genes que codifican para las histonas canónicas se organizan como grupos de arreglos repetidos en el genoma y su transcripción se encuentra estrechamente acoplada a la replicación del ADN. Las variantes de histonas poseen diferencias a nivel de la secuencia primaria de aminoácidos con respecto a su paróloga canónica y los genes que las codifican, se presentan en copia única en el genoma y se transcriben a lo largo de todo el ciclo celular (Sarma y Reinberg, 2005). Las variantes de histonas tienen distintas funciones en una diversidad de procesos como la reparación del ADN, la recombinación meiótica, la segregación de cromosomas, el silenciamiento del cromosoma X, el inicio y término de la transcripción y la accesibilidad de la cromatina, entre otros (revisado en Talbert y Henikoff, 2010).

Algunas de las variantes de histonas están sumamente conservadas desde las primeras diversificaciones dentro del dominio eucarionte. Estas incluyen las variantes de la histona 3, CENP-A y H3.3 así como las variantes de la histona H2A, H2AX y H2AZ (Fig. 5).



**Figura 5. Variantes de histonas.** Existen variantes de las histonas canónicas H3 y H2A que juegan papeles fundamentales en la regulación de la función genómica (ver texto) (modificado de Ahmad y Henikoff, 2002).

Por ejemplo, la variante centromérica de la histona 3 (CENP-A, en humanos), juega un papel relevante en el ensamblaje del cinetocoro durante la mitosis. CENP-A presenta una identidad del 60% con respecto a la histona canónica H3 con un dominio amino terminal muy diferente y se deposita exclusivamente en el centrómero en donde al parecer genera una estructura nucleosomal más rígida (Henikoff *et al.*, 2004; revisado en Talbert y Henikoff, 2010). Otra de las variantes de la histona H3 presente en la mayoría de los eucariontes es la isoforma H3.3. Ésta difiere sólo en 4 aminoácidos de la histona H3 (3 en la región nuclear de la histona y 1 en el extremo amino-terminal) y estudios recientes muestran que se enriquece preferentemente en regiones de transcripción activa (Ahmad y Henikoff, 2002; Janicki *et al.*, 2004; Mito *et al.*, 2005). A favor de esta idea, se ha observado que la H3.3 posee mayoritariamente, marcas de cromatina abierta tales como la acetilación en residuos de lisinas (McKittrik *et al.*, 2004).

La incorporación de la histona H3.3 ocurre tanto durante la replicación como de manera independiente de ella. El complejo proteico asociado a la deposición independiente de la replicación contiene a la proteína chaperona HIRA (Tagami *et al.*, 2004) y dicho proceso parece estar acoplado a la transcripción (Schwartz y Ahmad, 2005). Por otro lado, la tasa de reemplazo de la histona H3.3 es mayor que la de H3 y se encuentra enriquecida en regiones hipersensibles al corte por la endonucleasa DNasa I, lo cual sugiere que la presencia de dicha variante favorece una estructura más laxa y permisiva de la cromatina de la que forma parte (revisado en Talbert y Henikoff, 2010).

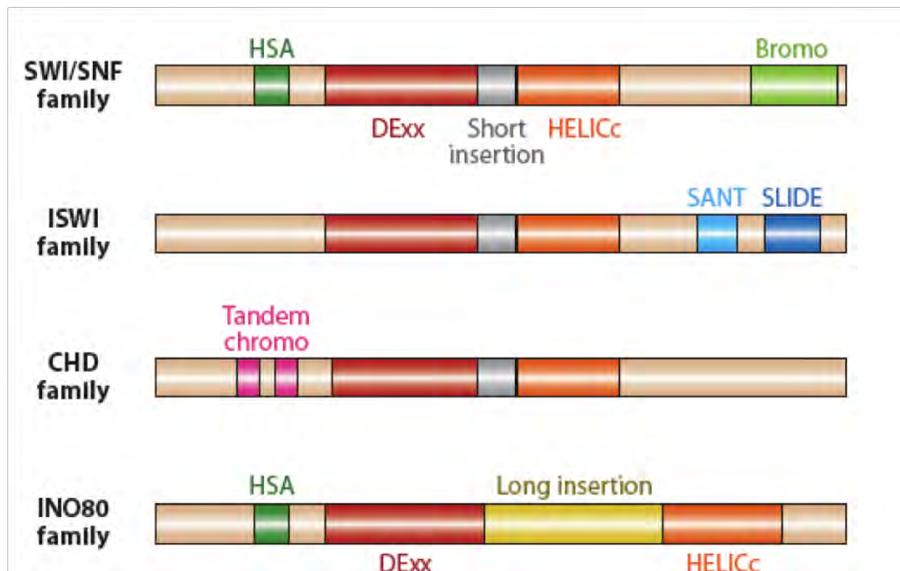
Así como la familia de H3, la histona H2A posee dos variantes universales, la H2A.Z y la H2A.X. La histona H2A.Z se localiza en regiones de inicio de la transcripción delimitando zonas libres de nucleosomas, en donde promueve el reclutamiento de la RNA polimerasa II tanto en levaduras como en células humanas. También H2A.Z participa en una serie de procesos que incluyen la apertura y cierre de la estructura de la cromatina, el recambio nucleosomal, la reparación del ADN y la contención contra heterocromatina por parte de *insulators*, entre otros. Existen algunas evidencias contradictorias con respecto a la función específica de la variante y diversos estudios serán requeridos para dilucidar de que depende el papel que ésta lleve a cabo en los diferentes contextos (revisado en Talbert y Henikoff, 2010). Por su parte, la histona H2A.X se deposita en las regiones particulares de la cromatina en respuesta a la presencia de rompimientos de doble cadena en el ADN. La fosforilación en una serina presente en el extremo carboxilo terminal de H2A.X, promueve el reclutamiento y mantenimiento de proteínas de reparación de ADN, modificadores de histonas y remodeladores de la cromatina en los sitios en donde ha ocurrido el daño. También, la variante H2A.X se ha visto involucrada en la inactivación de cromosomas sexuales durante la meiosis masculina en ratón, en donde precede a la deposición global de la variante H3.3 (revisado en Talbert y Henikoff, 2010).

La combinatoria del tipo de variantes de histonas que integran un nucleosoma y la gran diversidad de modificaciones post-traduccionales que pueden encontrarse tanto en los extremos amino terminales como en el núcleo de las histonas, representan una compleja plataforma de regulación del estado funcional de la cromatina.

### 1.1.3 Complejos remodeladores de cromatina

Otro de los mecanismos para modificar el estatus intrínsecamente represivo del templado cromatínico en donde no intervienen modificaciones covalentes, proviene del reclutamiento de complejos remodeladores que utilizan la energía generada a partir de la hidrólisis del ATP para reorganizar el posicionamiento nucleosomal en la fibra cromatínica. Las enzimas remodeladoras pueden exponer u ocultar secuencias para facilitar la interacción entre el ADN nucleosomal y otras proteínas o bien para estabilizar conformaciones represivas de la cromatina. La actividad remodeladora de estos complejos es requerida en diversos procesos como la transcripción, la replicación, la recombinación, el remplazo de histonas por variantes y la reparación del ADN (revisado en Tang *et al.*, 2010).

Los complejos remodeladores de la cromatina ATP-dependientes son complejos grandes (>1 MDa) compuestos de componentes múltiples (de 4 hasta 17 subunidades) que se encuentran conservados entre los eucariontes. Se caracterizan por contener una subunidad ATPasa perteneciente a la superfamilia II de proteínas tipo helicasas. El dominio de ATPasa está comprendido a su vez de dos partes, las regiones DExx y HELICc separadas por un linker. Dichos complejos se subdividen en 4 familias diferentes: SWI/SNF, ISWI, NURD/Mi-2/CHD e INO80 con base en la presencia de dominios adicionales únicos para cada una de ellas (Fig. 6). Las diferentes subunidades de los complejos ATP-dependientes adicionales a la ATPasa, le confieren regulación y especificidad a la actividad remodeladora (revisado en Clapier y Cairns, 2009).

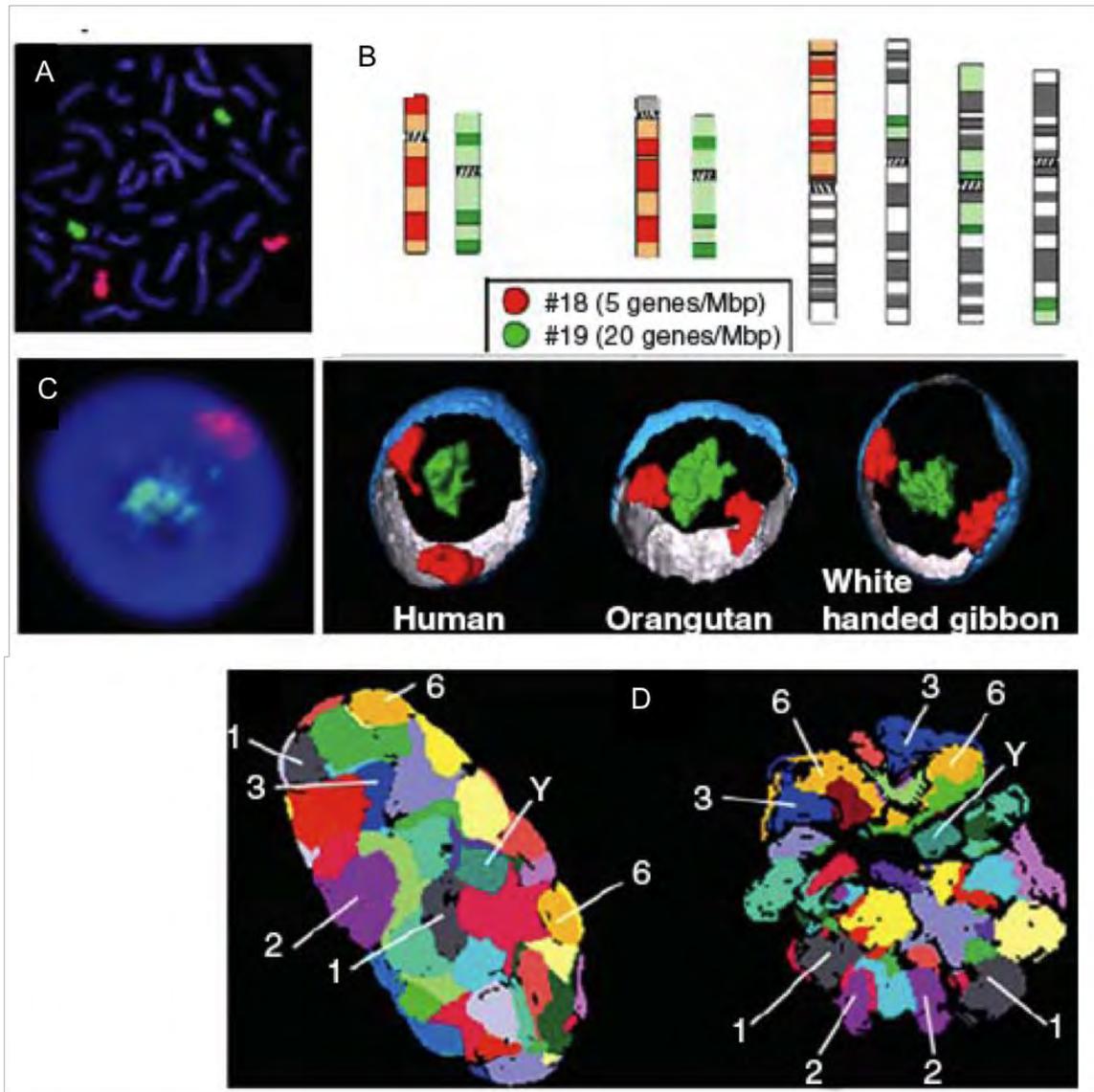


**Figura 6. Las cuatro familias de complejos remodeladores de la cromatina ATP-dependientes.** Todas las familias de remodeladores contienen una subunidad ATPasa del tipo SWI2/SNF2 que contiene un dominio ATPasa dividido en dos: DExx (caja roja) y HELICc (caja naranja). Lo que define a cada subfamilia son los dominios adicionales presentes (representados por las cajas de diferentes colores) (Tomado de Clapier y Cairns, 2009).

## 1.2 La organización del genoma al interior del núcleo

### 1.2.1 Territorios cromosomales

Por muchos años se pensó que el núcleo era un contenedor pasivo del genoma. En las últimas décadas diversas evidencias provistas por varios grupos de investigación han cambiado esta noción y han colocado a la arquitectura nuclear como un factor epigenético clave involucrado en la mayoría de los procesos de regulación del genoma. Varios experimentos que incluyeron desde la irradiación de núcleos con luz UV y la detección de los sitios de daño hasta el desarrollo de la técnica de hibridación *in situ* fluorescente (FISH) posibilitaron el surgimiento del concepto de territorio cromosomal (TC) (Cremer y Cremer, 2002). Este concepto postula que la cromatina que conforma a cada cromosoma al interior del núcleo en interfase, se encuentra estructurada ocupando sitios discretos en el espacio nuclear (Fig. 7).



**Figura 7. Territorios cromosomales y su localización al interior del núcleo eucarionte. A.** Visualización de cromosomas metafásicos individuales en células humanas, el cromosoma 18 en rojo y el 19 en verde después de la realización de la técnica de FISH utilizando sondas específicas para “pintar” cada cromosoma. **B.** Un sector óptico único de un núcleo completo utilizando las sondas específicas para los mismos cromosomas que en A (en verde el cromosoma 19 al centro y en rojo el cromosoma 18 en la periferia nuclear) y la realización del FISH. **C.** Reconstrucciones de la posición de ambos cromosomas en células humanas, de orangután y de gibón de mano blanca. **D.** Delineación simultánea de todos los cromosomas en un fibroblasto humano (izquierda) y un FISH multicolor de una prometafase en roseta. Se muestran cortes ópticos individuales con todos los cromosomas representados en colores falsos. Algunos ejemplos de cromosomas se denotan por su número de kariotipo (modificado de Cremer y Cremer, 2010).

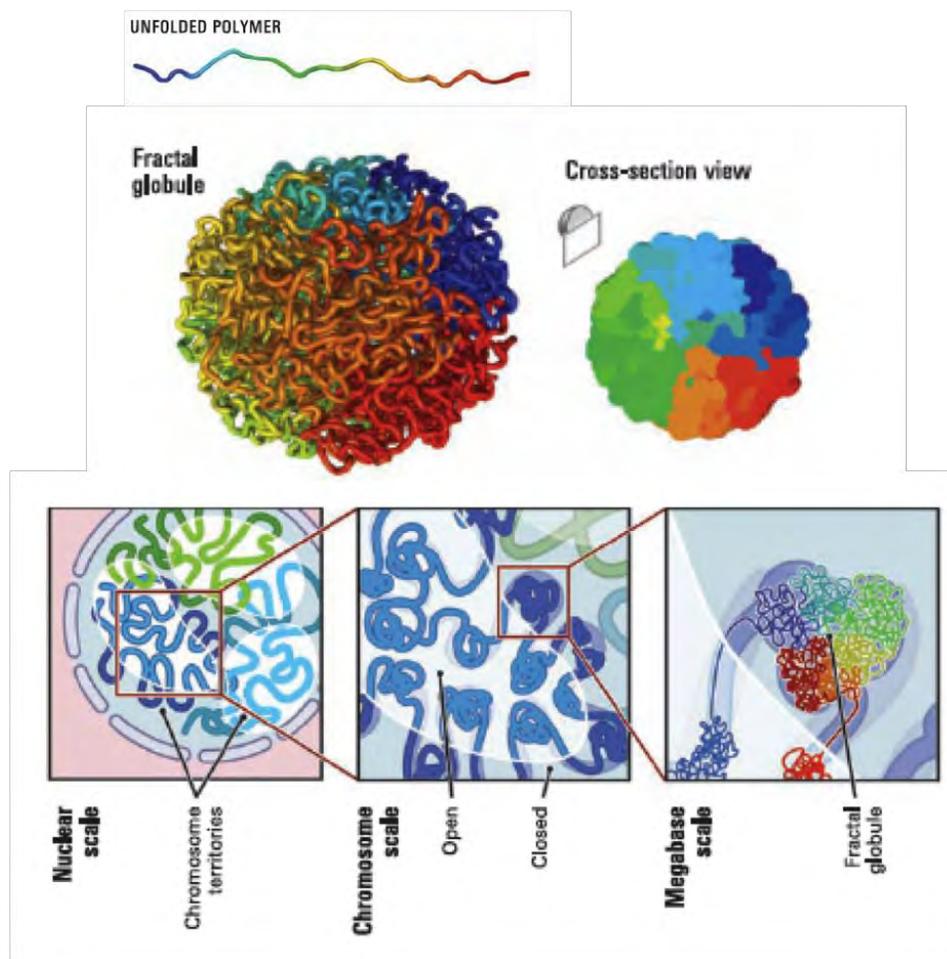
La localización de cada TC en el espacio nuclear parece no ser aleatoria. Por ejemplo, considerando la distribución nuclear radial y dependiendo de la densidad de

genes presentes en un cromosoma dado, es más frecuente encontrarlo hacia el interior del núcleo o hacia la periferia si el cromosoma es rico en genes o pobre en genes, respectivamente (Fig. 7). Asimismo, la frecuencia en la proximidad entre ciertos cromosomas ha llevado a la noción de que existen “vecinos preferentes” dependiendo del tipo celular en estudio. Las evidencias previas resultan sumamente interesantes y actualmente, existen muchos estudios interesados en los mecanismos que se pueden haber seleccionado durante la evolución para hacer de la localización espacial al interior del núcleo un nivel adicional de la regulación del genoma (revisado en Cremer y Cremer, 2010).

Además de los estudios que utilizan a la visualización directa como forma de evaluar las interacciones entre cromosomas o secuencias definidas a lo largo del genoma tanto en *cis* como en *trans*, en los últimos años han surgido otro tipo de aproximaciones experimentales sumamente poderosas basadas en la técnica conocida como Captura Conformacional de Cromosomas (3C) (Dekker *et al.*, 2002). Dicha técnica se basa en el entrecruzamiento entre secuencias de ADN y proteínas a unos Amstrongs de distancia, su fragmentación y posterior ligación para finalmente identificar las secuencias en gran proximidad por medio de diferentes métodos. La técnica de 3C se ha expandido a 4C en donde se permite la evaluación de las interacciones de una secuencia de interés con el resto del genoma por medio de el uso de microarreglos ó secuenciación masiva. Finalmente, en el 2009 Lieberman-Aiden y colaboradores reportaron una nueva técnica denominada Hi-C en donde se evalúan simultáneamente todas las interacciones dadas dentro del núcleo celular utilizando adaptadores universales para generar productos de ligación entre todas las secuencias y su posterior secuenciación masiva. En este reporte se muestra el primer mapa de proximidad de un genoma completo generado a partir de células linfoblásticas humanas con una resolución de 1 Mb (Lieberman-Aiden *et al.*, 2009). En dicho estudio se reafirman las nociones de territorios cromosomales, de la cercanía entre cromosomas de mayor y menor densidad génica respectivamente y la segregación entre regiones de heterocromatina y eucromatina. De manera adicional, los autores modelan el comportamiento de la fibra de la cromatina utilizando algoritmos utilizados

para descifrar el plegamiento de polímeros y observan que en la escala de Megabases (Mb), la cromatina se comporta como un glóbulo fractal en donde las regiones más cercanas interactúan con mayor frecuencia generando parches (o regiones) que se distinguen entre sí (Fig. 8) (Lieberman-Aiden *et al.*, 2009).

El desarrollo de estas nuevas herramientas experimentales que permiten la evaluación del arreglo espacial del genoma al interior del núcleo, proveen evidencias claras que sitúan a la organización nuclear como un componente central de la regulación del mismo.



**Figura 8.** La estructuración de la fibra de cromatina en la escala de 1 Mb se comporta como un **glóbulo fractal**. La conformación en un glóbulo fractal refleja que las regiones más próximas entre sí interactúan con mayor frecuencia que las localizadas a mayor distancia y esto coincide con lo observado para el plegamiento de la fibra de cromatina (modificado de Lieberman-Aiden *et al.*, 2009).

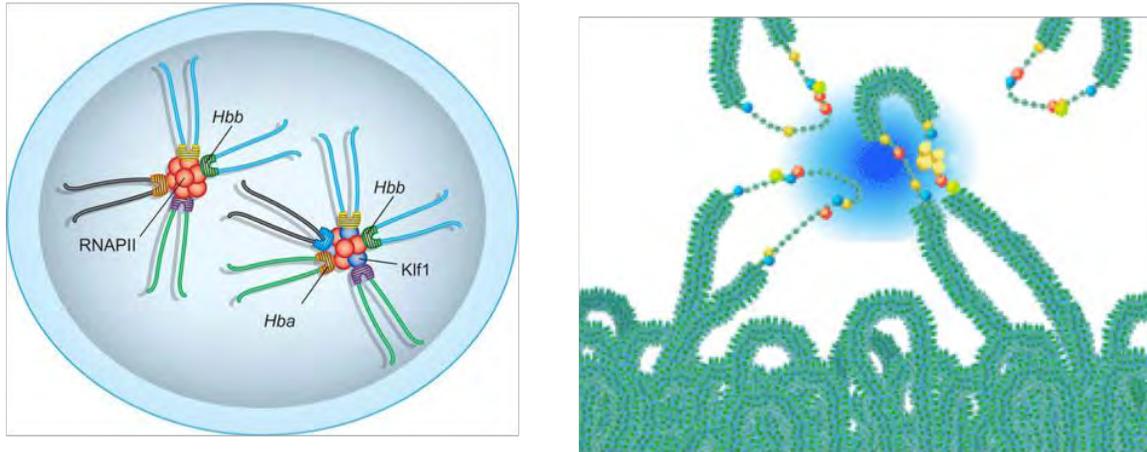
## 1.2.2 Fábricas Transcripcionales

Todos los estudios presentados en la sección anterior muestran que, aún cuando existe una estructuración del genoma eucarionte que responde a las propiedades fisicoquímicas de la fibra de cromatina y a las restricciones espaciales dentro del núcleo celular, existen interacciones no aleatorias entre regiones genómicas que parecen responder a adaptaciones seleccionadas aprovechando dicha distribución espacial. Una pregunta vigente que surge de manera natural a partir de las observaciones previas tiene que ver con cuáles son los mecanismos que subyacen a la organización del genoma? Una de las propuestas más interesantes sugiere que es la transcripción génica el proceso central que regula a dicha arquitectura.

Dentro del núcleo eucarionte se han descrito diversos compartimentos compuestos de diferentes componentes tanto nucleicos como proteicos. Ejemplos de éstos son el nucleolo, las motas o speckles, los cuerpos PML y las fábricas transcripcionales entre otros (revisado en Zhao *et al.*, 2009).

La idea de fábricas transcripcionales surge de la observación de zonas en el núcleo en donde se acumula la ARN polimerasa II activa a donde los genes “acuden” a ser transcritos (revisado en Cook 2010; Osborne *et al.*, 2004). La propuesta de la existencia de estructuras relativamente fijas en donde se acumula la polimerasa es controversial y existen evidencias tanto a favor como en contra de ella (Shuterland y Bickmore 2009). Sin embargo, más allá de la falta de una demostración definitiva de la existencia de dichas fábricas, existen datos muy interesantes en relación a la transcripción como uno de los mecanismos centrales en la generación de interacciones no aleatorias entre segmentos del genoma. Recientemente datos del grupo de Fraser y colaboradores han mostrado como la regulación transcripcional y la arquitectura tridimensional del genoma están estrechamente relacionadas. En particular, el reporte de Schoenfelder y colaboradores muestra como los locus  $\alpha$  y  $\beta$ -globina de ratón se asocian con cientos de genes activos a lo largo del genoma en fábricas transcripcionales. Muestra también que existen fábricas

especializadas que contienen al factor transcripcional Klf 1 y éste media las interacciones entre los genes a los que regula (Schoenfelder *et al.*, 2010). Estas evidencias sugieren la existencia de un “interactoma” transcripcional específico de células eritroides. En el futuro será muy interesante estudiar si existen “interactomas” tejido-específicos ó particulares de cierta etapa del desarrollo o diferenciación (Fig. 9).



**Figura 9. Fabricas Transcripcionales.** De lado izquierdo se muestra un diagrama de la interacción entre los locus  $\alpha$  y  $\beta$ -globina de ratón con otros genes activos en células eritroides mediados por la ARN pol II y Klf 1 (modificado de Ragoczy y Groudine 2010). A la derecha se muestra un esquema de asas de cromatina alejándose del centro de su territorio cromosomal para acercarse a un fábrica transcripcional representada por un círculo azul (tomado de [http://www.babraham.ac.uk/pjl\\_pages/osborne/osborne.html](http://www.babraham.ac.uk/pjl_pages/osborne/osborne.html)).

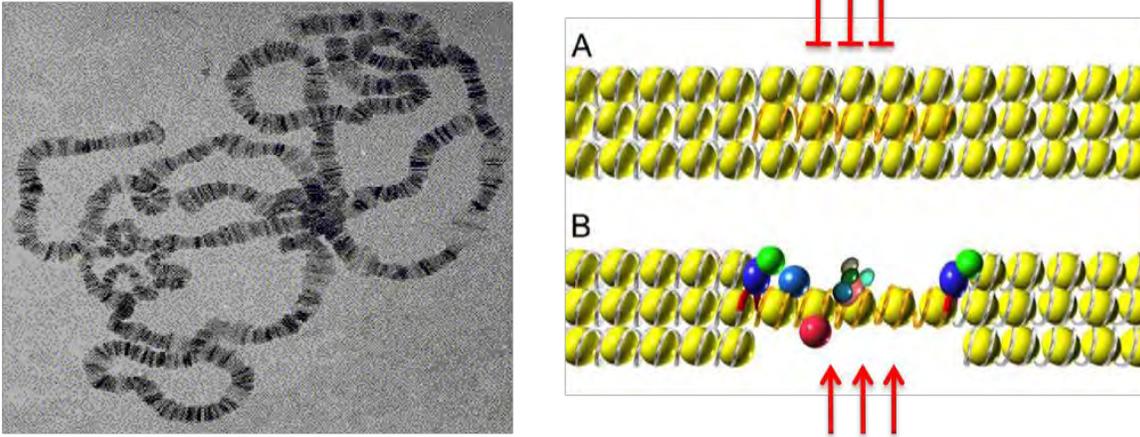
### 1.3 Dominios cromosómicos y sus elementos de regulación

Además de las evidencias topológicas presentadas en la sección previa, existen otro tipo de argumentos que postulan la existencia de dominios cromatínicos funcionales que representan unidades independientes de actividad transcripcional. Es decir que el genoma eucarionte se compartamentaliza en dominios de expresión genética con límites definidos que les permiten ser independientes de su entorno cromosómico para así expresarse en el espacio y tiempo adecuados (Capelson y Corces, 2004).

Una de las primeras observaciones realizadas que apuntaron hacia la existencia de regiones discretas en el genoma surgió del estudio de cromosomas politénicos de

glándulas salivales de *Drosophila*. En éstos se observa un patrón de bandas de cromatina compacta separadas por interbandas de cromatina más laxa que correlacionan con estados transcripcionales diferentes. Otra de las evidencias de la existencia de dominios cromatínicos proviene del uso de nucleasas como la DNasa I. Al tratar núcleos con esta enzima se pudo observar que existen regiones de mayor accesibilidad al corte por dicha enzima reflejando diferentes estados de compactación del genoma (Fig. 10). En la actualidad sabemos que las regiones de mayor accesibilidad al corte por la DNasa I corresponden a dominios cromatínicos abiertos y los sitios hipersensibles a cuerpos génicos o elementos de regulación dentro de éstos (Recillas-Targa y Razin, 2001). El estudio de dichos elementos de regulación en *cis* tales como potenciadores o *enhancers*, regiones de control de locus (LCRs) y silenciadores también han contribuido al concepto de dominios funcionales al interior del genoma. Estos elementos de regulación pueden actuar a grandes distancias de su gen blanco y sin embargo, se ha observado que no existe regulación cruzada entre elementos de dominios distintos lo cual sugiere la existencia de unidades independientes de regulación (West y Fraser, 2005). Por último, experimentos bioquímicos en donde se extraen factores asociados a la cromatina en altas concentraciones de sal en donde se muestra la presencia de asas de ADN independientes anclados a la matriz nuclear u otras estructuras nucleares como el nucleolo o lámina nuclear. La conformación en asas del genoma se ha interpretado como una forma de conferir independencia funcional entre dominios (Capelson y Corces, 2004).

Todas estas evidencias sugieren la existencia de dominios transcripcionales que se encuentran delimitados de tal forma que su regulación es independiente del entorno cromosómico.



**Figura 10. Evidencia de la existencia de dominios cromatínicos funcionales.** Del lado izquierdo se muestra una imagen de cromosomas politénicos de glándulas salivales de *Drosophila* en donde se pueden observar las bandas e interbandas que reflejan distintos estados de compactación de la cromatina. Del lado derecho se muestra una caricatura de la accesibilidad al corte por nucleasas como la DNasa I (flechas rojas) en dos regiones de la cromatina. La región accesible (abajo) coincide con un dominio en donde se asocian distintos complejos proteicos (modificado de <http://www.biology.emory.edu/research/Corces/Research2.html>)

### 1.3.1 *Insulators*

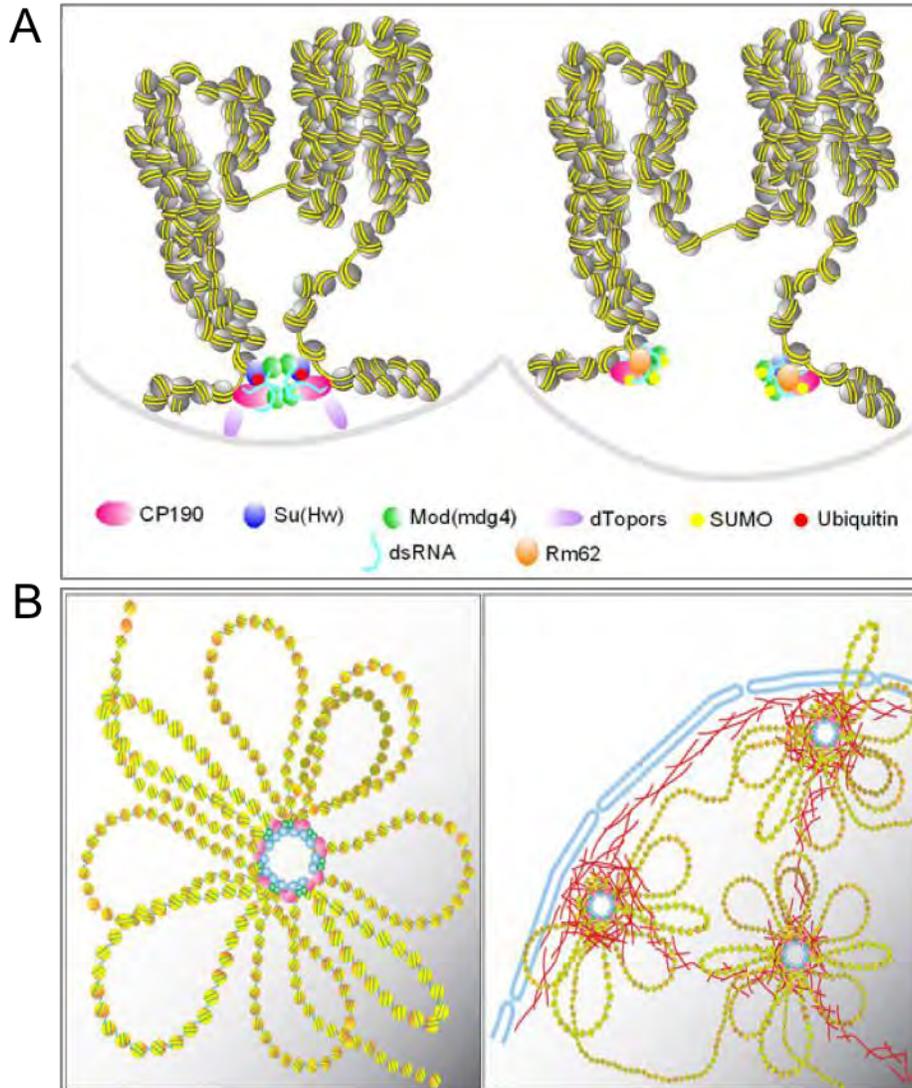
Como se ha mencionado anteriormente, los genes y grupos de genes constituyen unidades transcripcionales autónomas por medio de su organización en dominios que están protegidos de la acción tanto de reguladores externos al dominio, como del entorno cromatínico. Se ha propuesto que dicha autonomía, traducida en dominios independientes, se consigue a través de elementos localizados en los límites de los dominios llamados fronteras de cromatina o *insulators*.

Los *insulators*, descritos desde la levadura hasta el humano, son secuencias neutras que comparten la habilidad de proteger a los genes de señales inapropiadas de su entorno cromatínico y en general coinciden con sitios de hipersensibilidad al corte por la DNasa I (West y Fraser, 2005). En algunos casos, estos elementos participan en el establecimiento de regiones de cromatina abierta a nivel del dominio que delimitan (Litt *et al.*, 2001; Razin *et al.*, 2003).

A la fecha se conocen varios *insulators* de función demostrada experimentalmente. El *insulator* presente en el retrovirus *gypsy* en *Drosophila* es uno de los mejor caracterizados de donde se desprende una buena parte del conocimiento referente al funcionamiento de dichos elementos (Fig. 11).

La secuencia más importante para la actividad del *insulator gypsy* abarca 350 pb de longitud y contiene 12 copias de la secuencia de 26 pb a donde se une el factor Suppressor of Hairy-wing [Su(Hw)]. Esta proteína contiene 12 dedos de zinc involucrados en su unión al ADN y una región HLH-zip requerida para la función de *insulator*. Esta región media la interacción entre dicho factor y otro componente del *insulator gypsy* denominado Modifier of modg4 2.2 [Mod(mdg4)2.2]. Mod(mdg4)2.2 contiene un dominio BTB a través del cual dimeriza e interactúa con CP190, un tercer componente del *insulator gypsy* (Fig. 11A).

Las proteínas Su(Hw), Mod(mdg4)2.2 y CP190 colocalizan en cientos de sitios en cromosomas politénicos de *Drosophila*. Estos sitios no correlacionan con inserciones de retrotransposón *gypsy* sino que son *insulators* que regulan la expresión de genes en el genoma de *Drosophila*. La localización de estos *insulators* coincide generalmente con la frontera entre las bandas de cromatina compacta y las interbandas de cromatina más laxa en donde se encuentran genes transcripcionalmente activos (revisado en <http://www.biology.emory.edu/research/Corces/Research2.html>).

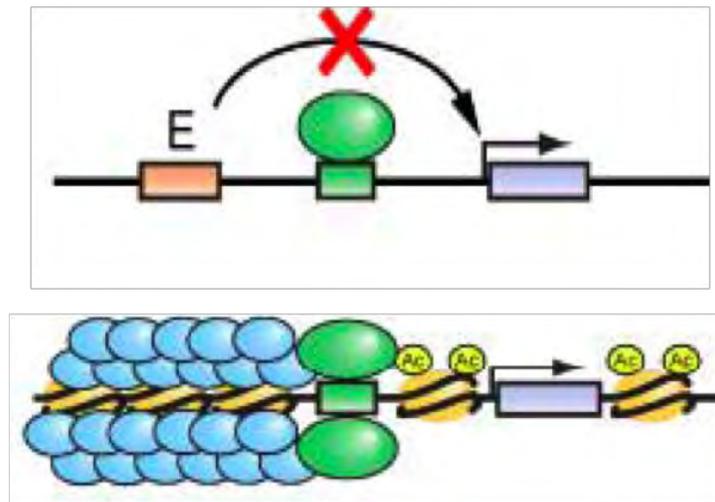


**Figura 11. Insulators tipo gypsy en *Drosophila*.** **A.** La actividad de los *insulators* pueden estar reguladas por la ubiquitinación y sumoilación de las proteínas que los comprenden. A la izquierda, dos *insulators* activos se juntan en un cuerpo de *insulators*. dTopors está presente por lo cual Mod(mdg4)2.2 y CP190 no están sumoiladas y dTopors sirve como puente de anclaje a la lámina nuclear. A la derecha, dos *insulators* inactivos en donde dTopors no está presente y estos no pueden formar parte de un cuerpo de *insulators*. **B.** Los *insulators* organizan la fibra de cromatina al interior del núcleo estableciendo estructuras de orden mayor. Los cuerpos de *insulators* interactúan con la lamina anclándose a la periferia nuclear (modificado de <http://www.biology.emory.edu/research/Corces/Research2.html>).

Al definir la distribución de Su(Hw) y Mod(mdg4)2.2 en núcleos en interfase, se observa que estos se acumulan en sitios particulares formando estructuras denominadas cuerpos de *insulators*. Se ha demostrado que estas acumulaciones generan estructuras de tipo roseta en donde asas de cromatina se anclan con la ayuda de las proteínas que

conforman los *insulators*. Los cuerpos de *insulators* se posicionan en relación a la envoltura nuclear a través de interacciones con lamina y la ligasa de ubiquitina E3 dTopors. dTopors puede estar ubiquitinando proteínas del *insulator* como Su(Hw) y de esta forma regular su función en conjunto con consecuente sumoilación de otros componentes como CP190 y Mod(mdg4)2.2. Los *insulators* que contienen Su(Hw) poseen además otros componentes como moléculas de ARN generadas por la maquinaria de ARN de interferencia necesarios para mantener la integridad de los cuerpos de *insulators*. Por otra parte, la proteína RM62 interactúa con el ARN y regula de manera negativa la función de los *insulators* irrumpiendo la formación de cuerpos de *insulators* (Fig. 11) (revisado en <http://www.biology.emory.edu/research/Corces/Research2.html>).

La identificación y caracterización de elementos tipo *insulators* se consigue por medio de dos ensayos funcionales. El primero es el de bloqueo de *enhancer*, en donde se evalúa la capacidad de un *insulator* de bloquear la acción de un *enhancer* sobre un promotor al ser colocado entre ambos (Recillas-Targa *et al.*, 2002). El segundo se refiere a la propiedad de los *insulators* de proteger a un transgene del efecto de posición cromosomal al ser colocados a cada lado de éste (Recillas-Targa *et al.*, 2004) (Fig. 12).



**Figura 12. Las propiedades de los *insulators*.** Bloqueo de *enhancer* (arriba) y protección contra los efectos de posición cromosomal (abajo) (ver texto) (modificado de Kimura y Horikoshi, 2004).

Éstas funciones en algunos *insulators* como el 5' cHS4 del dominio  $\beta$ -globina de pollo son separables (Recillas-Targa *et al.*, 2002) mientras que en otros casos la separación

no es tan clara ya que el factor asociado al bloqueo de *enhancer* parece tener funciones de barrera (Filippova *et al.*, 2005 y Cho *et al.*, 2005) (ver sección de CTCF más adelante). Los mecanismos precisos a través de los cuales los *insulators* actúan se desconocen pero existen diversas propuestas que se resumen a continuación.

### *Bloqueo de enhancer*

Uno de los modelos propuestos recientemente para la función de bloqueo de *enhancer* involucra el anclaje de componentes del *insulator* en estructuras nucleares específicas. Se ha observado que el *insulator gypsy* de *Drosophila* se ancla cerca de la periferia nuclear (Gerasimova *et al.*, 2000) mientras el *insulator 5' cHS4*  $\beta$ -globina de pollo parece anclarse a la periferia nucleolar en células humanas (Yusufzai *et al.*, 2004). Por otro lado, varios elementos caracterizados como *insulators* presentan secuencias de unión a matriz nuclear (Yusufzai y Felsenfeld, 2004). El bloqueo de *enhancer* a través del anclaje a subestructuras nucleares, se ha interpretado como el resultado de la formación de asas que mantienen separados a los elementos de regulación impidiendo así el intercambio erróneo de señales entre dominios (West *et al.*, 2002).

Resulta interesante notar que, como se ha mencionado, en *Drosophila* se han identificado varias proteínas asociadas a la acción de bloqueo de *enhancer* mientras que en vertebrados, sólo una proteína con once dedos de zinc de nombre CTCF (CCCTC-binding factor) se ha vinculado con dicha actividad (ver más adelante).

### *Protección contra efecto de posición cromosomal*

La expresión de transgenes en la célula eucariota está sujeta a su sitio de integración en el genoma y se ha observado que frecuentemente ésta va disminuyendo con el paso del tiempo. Este silenciamiento por efectos de posición cromosomal (EPC) está acompañado de cambios epigenéticos como la desacetilación de histonas y la metilación del ADN en las secuencias del transgén. Al enmarcar a un gen reportero con dos copias de un *insulator* a cada lado y evaluar su expresión en el tiempo, se observa que se previene el silenciamiento y los cambios que conlleva (Pikaart *et al.*, 1998; Recillas-Targa *et al.*, 2000).

Esta propiedad de los *insulators* se ha explicado de distintas maneras, en algunas de ellas se propone que el silenciamiento puede deberse a la invasión de heterocromatina circundante y que los *insulators* actúan previniéndola. Uno de los modelos propuestos tiene que ver, al igual que el modelo planteado para el bloqueo de *enhancer*, con la capacidad de los factores nucleares asociados al *insulator* de anclarse a estructuras nucleares generando un impedimento estérico para la expansión de la heterocromatina a través de la formación de asas. Este fenómeno se ha observado principalmente en levadura (Ishii *et al.*, 2002). Otro modelo propone que los elementos asociados al *insulator* pueden promover la disociación de uno o varios nucleosomas, de tal forma que el molde para la expansión de la heterocromatina se pierde. En un modelo similar al anterior se propone que los elementos en el *insulator* no solo interrumpen pasivamente a la formación de la heterocromatina, sino que interactúan con las histonas de nucleosomas adyacentes compitiendo con los factores de silenciamiento (Ferrari *et al.*, 2004). Por último, existe otro modelo en el cual los elementos del *insulator* reclutan activamente enzimas modificadoras de histonas promoviendo la formación de cromatina abierta. Un estudio que favorece este modelo identifica a la proteína USF, que se une al elemento 5' cHS4 del dominio  $\beta$ -globina de pollo, como la responsable de reclutar acetilasas de histonas y otras marcas de cromatina abierta (Litt *et al.*, 2001; West *et al.*, 2004). Adicionalmente, un reporte reciente revela la asociación de una helicasa con cromodominio (CHD8) con CTCF. La pérdida de dicha proteína afecta el estado de acetilación de las histonas y la metilación del ADN en la región circundante a los sitios de unión a CTCF (Ishihara *et al.*, 2006). Esto sugiere que la función de CTCF puede estar asociada a la remodelación epigenética activa como parte del mecanismo de ciertos *insulators*.

Las propiedades descritas para los *insulators* los proponen como uno de los elementos importantes que organizan al genoma en dominios funcionales independientes.

### 1.3.2 El factor CTCF

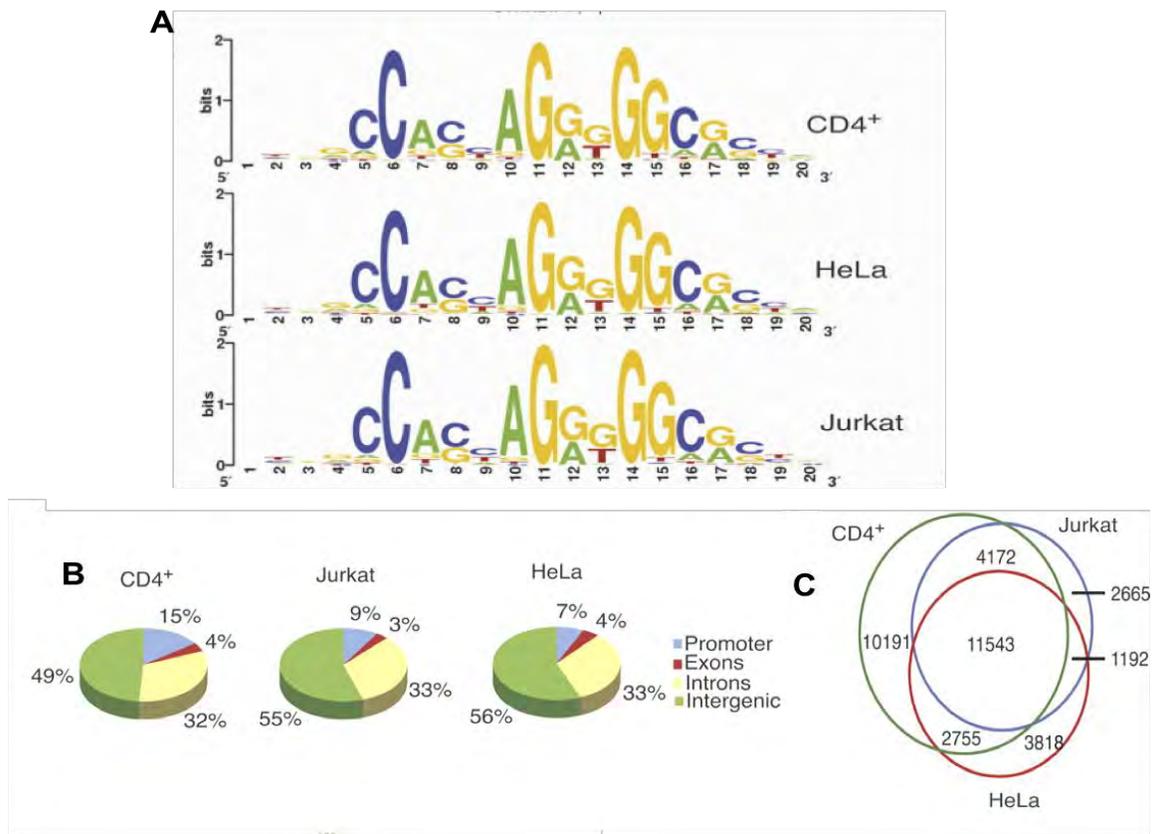
CTCF representa el factor por excelencia asociado a los elementos *insulator* y sus propiedades. En los últimos años han surgido una gran cantidad de reportes que profundizan con respecto a la actividad y distribución de dicho factor y lo han propuesto como pieza fundamental para generar el “tejido del genoma”, es decir el conjunto de interacciones que dan andamiaje y organización a la arquitectura de la cromatina y distribución del genoma al interior del núcleo (revisado en Phillips y Corces, 2009).

El factor CTCF se encuentra altamente conservado en los eucariontes más complejos. La proteína posee un dominio central de unión al ADN que comprende 11 dedos de zinc. Éste presenta un índice de identidad de casi el 100% entre el pollo, el ratón y el humano embebido en extremos amino y carboxilo terminales más divergentes. La capacidad de unirse a secuencias diferentes en el ADN y a interactuar con factores diversos a través de este dominio de dedos de zinc, llevaron a proponer a CTCF como un factor multivalente de funciones versátiles. CTCF está implicado en procesos tales como la activación o represión transcripcional, la impronta genómica, la inactivación del cromosoma X y la actividad de los *insulators*, entre otros (revisado en Phillips y Corces, 2009).

Estudios recientes utilizando la técnica de inmunoprecipitación de cromatina (ChIP) (ver métodos) combinada con estrategias de microarreglos o de secuenciación masiva del genoma (ChIP-Seq) han identificado la ocupación de CTCF a escala global en el genoma en diferentes tipos celulares, que incluyen fibroblastos humanos, células T CD4+ humanas, células HeLa, Jurkat, células troncales embrionarias de ratón y más recientemente, células eritroides de pollo (ver Anexo 2, Martín *et al.*, sometido). La distribución global de CTCF en el genoma es de alrededor del 46% en regiones intergénicas, 22% intrónica, 12% exónicas y 20% a una distancia de 2.5 kb o menor de promotores y el número de sitios de unión varía desde 13,800 en fibroblastos humanos, alrededor de 19,500 en células HeLa y Jurkat, 21,500 en eritrocitos de pollo de 10 días de desarrollo, hasta 39,600 en células troncales embrionarias de ratón. Actualmente se

desconoce si estas diferencias en la ocupación de CTCF dependiendo del tipo celular refleja alguna particularidad funcional o se debe a los diferentes métodos tanto experimentales como computacionales utilizados por los investigadores (revisado en Phillips y Corces, 2009).

A partir de estos estudios se ha identificado un consenso de alrededor de 20 pb al cual se une CTCF (Fig. 13).

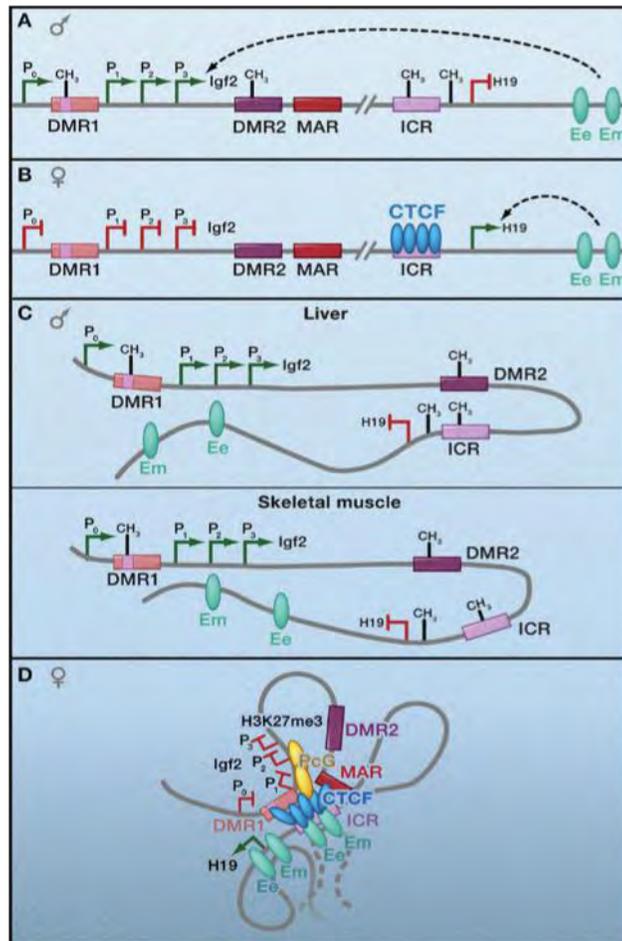


**Figura 13. Algunos sitios de unión al factor CTCF coinciden entre distintas líneas celulares. A.** El motivo consenso de unión de CTCF en células T CD4<sup>+</sup>, HeLa y Jurkat. **B.** La distribución global de CTCF en el genoma en los diferentes tipos celulares. **C.** Diagramas de Benn en donde se muestra el traslape de los sitios entre los tipos celulares (modificado de Cuddapah *et al.*, 2009).

Sin embargo, es importante mencionar que existen sitios corroborados experimentalmente que no reflejan el consenso obtenido (tales como los que se encuentran en hTERT y ALF) lo cual sugiere que existen sitios de unión a CTCF no detectables al utilizar las metodologías previamente descritas (revisado en Ohlsson *et al.*, 2010).

La distribución global de CTCF en comparación con la distribución de otros factores transcripcionales tales como Nanog, Sox-2, Oct4, Smad1 y c-Myc entre otros, sugiere que la función preponderante de CTCF no es la de un factor transcripcional convencional sino más bien es la de *insulator*, delimitando dominios génicos de alrededor de 50 kb y posiblemente generando una red de interacciones entre fibras de cromatina confiriendo estructura y organización al genoma (Chen *et al.*, 2008).

La función particular y más detallada de CTCF se conoce en el caso de solo algunos dominios. Uno de ellos es el caso del locus *Igf2/H19*, en donde CTCF se une a una región denominada ICR (por sus siglas en inglés Imprinting Control Region) en un estado no metilado asociado al locus materno. Cuando CTCF está unido al ICR, el *enhancer* localizado río abajo de éste es bloqueado y no es capaz *trans*-activar al gene *Igf2* localizado río arriba del ICR mientras que *trans*-activa al gene *H19*. En el alelo paterno en cambio, el ICR se encuentra metilado, CTCF no está unido y ahora el *enhancer* es capaz de *trans*-activar la expresión del gen *Igf2* (Fig. 14). Este es el ejemplo más claro de la participación de CTCF en la regulación de un locus improntado y de la incapacidad de éste a unirse a su sitio blanco en el ADN cuando éste se encuentra metilado. El mecanismo a través del cual CTCF aísla al gene *Igf2* está dado por la formación de asas cromatínicas excluyentes como se muestra en Figura 14D (revisado en Phillips y Corces, 2009).



**Figura 14. CTCF organiza contactos entre la cromatina del locus improntado *Igf2/H19*.** A y B muestran al locus de manera lineal y a sus elementos de regulación que incluye al ICR que se metila en el alelo paterno y no en el materno en donde se une CTCF bloqueando la acción de los *enhancers* localizados río abajo del gen *H19*. C y D muestran la estructuración en asas de cromatina que se genera para cada uno de los alelos (tomado de Phillips y Corces, 2009).

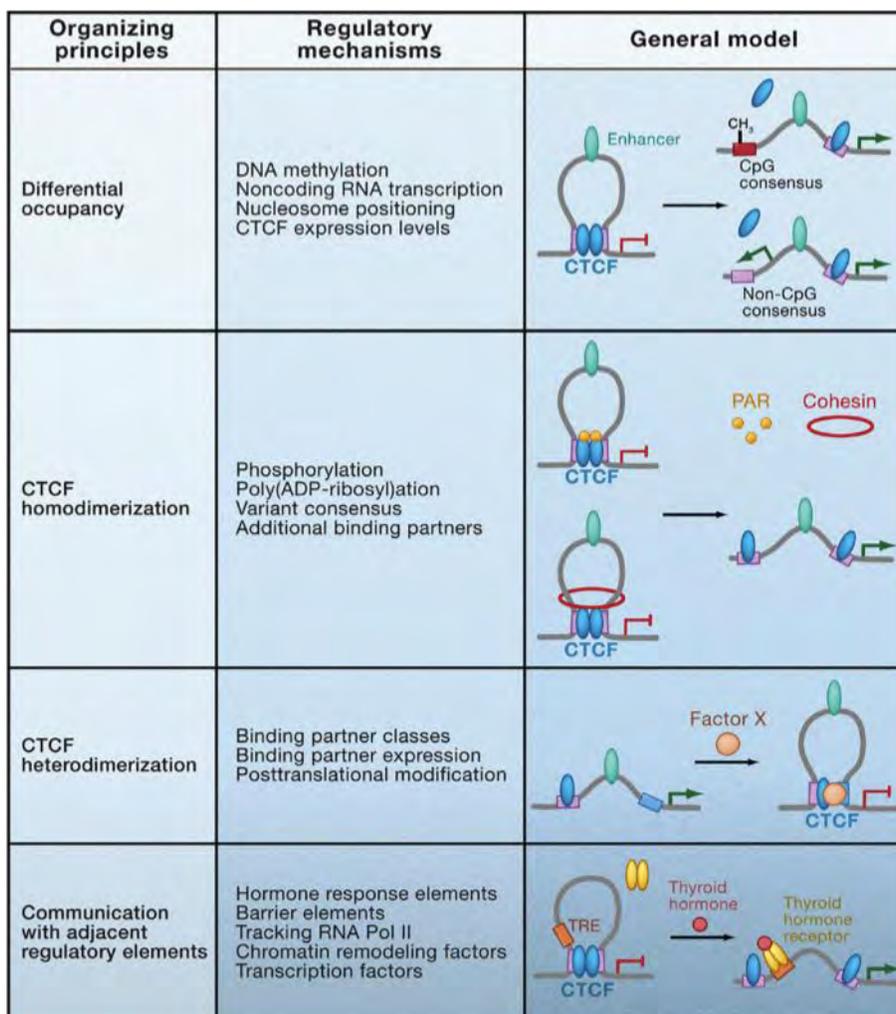
Otro locus en donde la función de un *insulator* dependiente de CTCF y su regulación se conoce con mayor detalle es el locus de la lisozima de pollo. Este es un caso interesante ya que la regulación de la unión de CTCF parece estar mediada por el paso de la ARN Polimerasa II a través de su sitio blanco. Sin embargo, no es claro si es el recorrido de la maquinaria transcripcional es lo que genera directamente el desprendimiento de CTCF o si existe la regulación se da a través de una interacción entre el ARN no-codificante y CTCF en *trans* (Lefevre *et al.*, 2008). Otro locus en donde la función de CTCF se conoce a detalle es el locus  $\beta$ -globina del ratón, este se abordará más adelante (ver sección de antecedentes).

Finalmente, CTCF puede modificarse post-traduccionalmente y tiene varias parejas de interacción que han mostrado ser importantes para su función. CTCF puede ser modificado post-traduccionalmente mediante fosforilación y poli(ADP-ribosil)ación (PARilación). La marca de poli(ADP-ribosa) (PAR) se ha detectado en más de 100 secuencias tipo *insulators* en todo el genoma incluyendo al *insulator* en el locus *Igf2/H19*. La inhibición de las enzimas PAR polimerasas (PARPs) causa la pérdida de bloqueo por parte de diversos *insulators* dependientes de CTCF en ensayos *in vitro* lo cual indica que la PARilación puede ser un mecanismo de regulación de la función de los *insulators* más generalizado. De manera interesante, la inhibición de las PARPs no genera la disociación de CTCF de su secuencia blanco por lo cual, se sugiere que la marca de PAR puede conferir estabilidad a las interacciones generadas por CTCF con otros factores o consigo mismo (Yu *et al.*, 2004). De manera similar, la fosforilación de CTCF parece abatir sus propiedades de bloqueador de un *enhancer* sin promover su disociación del ADN sin embargo, existe menos información con respecto al efecto de la fosforilación sobre la regulación de CTCF a una escala más general.

Además de las modificaciones post-traduccionales, se han caracterizado una serie de factores que interactúan con CTCF, vinculados con la función de *insulator* o su regulación (Fig. 15). Uno de ellos es la proteína CHD8, una helicasa con cromodominio de la familia SNF2-like. CHD8 se ha asociado a la actividad de varios *insulators* incluyendo al ubicado en el locus *Igf2/H19*. Sin embargo, la relación entre la función de helicasa y la de CTCF en la actividad de *insulator* se desconoce (Ishihaha *et al.*, 2006). Las proteínas cohesinas son otros de los factores que colocalizan con CTCF en un 60 a 90% de sitios en el genoma dependiendo del tipo celular. Existen sitios en donde sólo se une CTCF, sitios en donde sólo hay cohesinas y sitios en donde se encuentran ambos. En este caso, CTCF parece ser necesario para el reclutamiento de las cohesinas a dichos sitios (revisado en Phillips y Corces, 2009).

Las cohesinas conforman una estructura de tipo anillo que se ha visto involucrada en la cohesión de las cromátidas hermanas durante la mitosis. Es interesante pensar que dicha estructura pudiera estabilizar las asas de cromatina generadas por medio de CTCF.

Finalmente, otra de las parejas de interacción reportadas para CTCF es la ARN Polimerasa II (Chernukhin *et al.*, 2007). Sin embargo, aun cuando es posible que CTCF reclute a la polimerasa, probablemente la coexistencia de ambos factores sea el reflejo de la capacidad de CTCF de frenar el paso de ésta como parte de su mecanismo de bloqueo de *enhancer*, la cual se ha propuesto para el locus de *c-myc* en mamíferos (revisado en Phillips y Corces, 2009).



**Figura 15. Posibles mecanismos que regulan la formación de asas dependientes de CTCF.** (Ver texto) (tomado de Phillips y Corces, 2009).

En resumen, los *insulators* dependientes de CTCF y sus cofactores son elementos que juegan tanto un papel estructural en el genoma organizando a la fibra de cromatina

en asas, como regulatorio, impidiendo la regulación cruzada entre elementos de dominios diferentes, o regulando la acción de *enhancers* particulares sobre sus genes blanco.

En el presente trabajo de investigación estamos interesados en estudiar la función de *insulators* localizados al interior de un dominio génico, el dominio  $\alpha$ -globina de pollo. A continuación se detalla información importante respecto a los dominios  $\alpha$  y  $\beta$ -globina de vertebrados.

## 2. ANTECEDENTES

La diferenciación eritroide en vertebrados es un proceso complejo y dinámico que involucra varias etapas y depende de una gran diversidad de factores tanto extrínsecos como intrínsecos. Brevemente, la eritropoyesis surge a partir de células troncales hematopoyéticas seguidas de progenitores multipotentes tempranos que dan origen a células precursoras comprometidas al linaje eritroide (eritroblastos) que finalmente generan eritrocitos maduros (RBC, por sus siglas en inglés Red Blood Cells). Para que este proceso se lleve a cabo, se requieren de una serie de reguladores clave que incluyen a la homeostasis del hierro, la respuesta a la tensión de oxígeno, factores de crecimiento, factores transcripcionales y micro ARNs (revisado en Tsiftoglou *et al.*, 2009).

Desde el enfoque de la regulación transcripcional, la diferenciación eritroide requiere de factores transcripcionales específicos que determinan los perfiles de expresión génica necesarios para dar identidad a las células eritroides. La mayoría de estos han sido identificados por unirse a los promotores de genes eritroides, en particular los genes  $\beta$ -globina. Entre los factores transcripcionales que regulan la diferenciación eritroide se encuentran en orden de importancia: GATA-1, FOG, el complejo TAL-1/SCL/LMO2/Ldb1/E2A y EKLF, entre otros. GATA-1 es un factor transcripcional crítico para la diferenciación eritroide ya que su ausencia resulta en letalidad embrionaria en estadios de E10.5-E11.5 en el ratón a causa de una anemia severa. Todas las funciones que lleva a cabo el factor GATA-1 se consiguen a través de la orquestación de la activación o represión de la expresión de genes involucrados en todas las facetas de la diferenciación eritroide. Entonces, además de promover la expresión del programa transcripcional eritroide, GATA-1 reprime la expresión de genes que confieren multipotencia a progenitores hematopoyéticos tempranos así como de genes que promueven identidades celulares distintas a la eritroide. Además de lo anterior, GATA-1 tiene también un papel antiapoptótico fundamental que evita la muerte celular de los eritrocitos maduros cuando la proliferación celular se detiene (revisado en Tsiftoglou *et al.*, 2009). La función de FOG-1 y el complejo TAL-1/SCL/LMO2/Ldb1/E2A se conoce con

menor detalle, sin embargo, la mayor parte de su participación tiene que ver con formar parte de complejos que incluyen al factor GATA-1 para facilitar su función. De manera interesante, el factor EKLF parece jugar un papel en la estructuración tridimensional de los dominios que contienen a los genes eritroides al interior del núcleo así como participar en la remodelación local de la cromatina que los conforma (revisado en Tsiftoglou *et al.*, 2009).

Los dominios que contienen a los genes que codifican para las globinas  $\alpha$  y  $\beta$  representan uno de los blancos principales de regulación durante la eritropoyesis a nivel transcripcional.

## 2.1 El dominio $\beta$ -globina

El dominio de los genes  $\beta$ -globina representa uno de los grupos génicos regulados durante el desarrollo, que se han estudiado con gran detalle. El inicio de la eritropoyesis en células precursoras y la expresión progresiva de los diferentes miembros de la familia de las  $\beta$ -globinas durante el desarrollo y la diferenciación eritroides, depende en gran medida de cambios epigenéticos a nivel del locus.

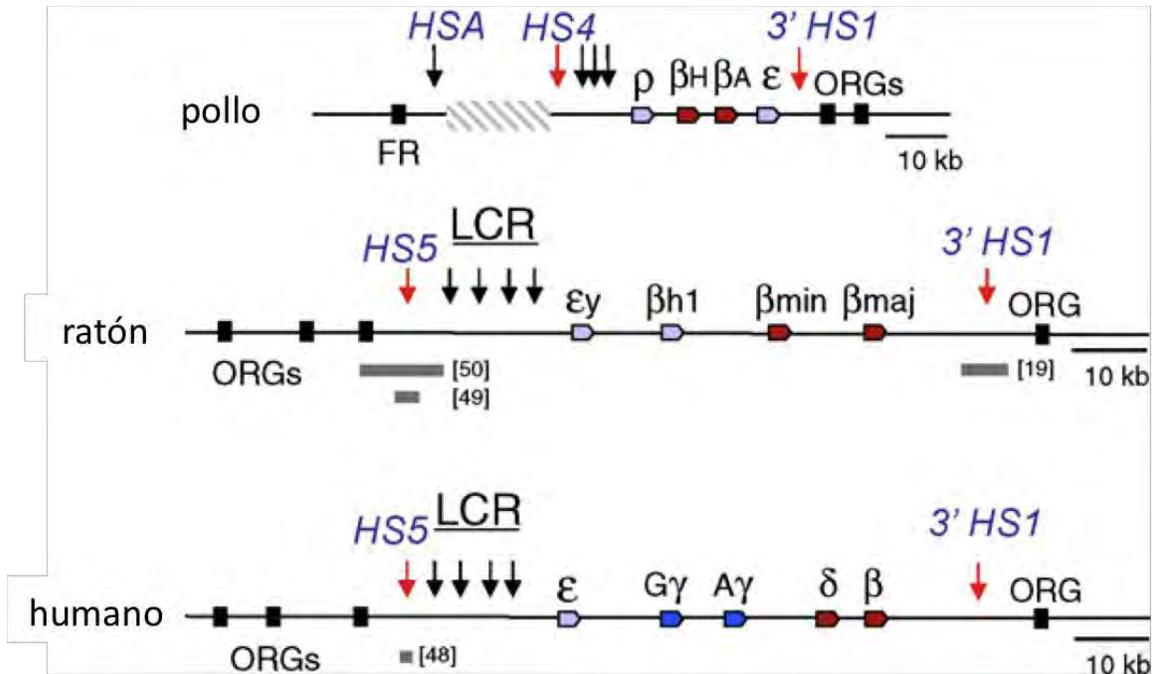
El dominio  $\beta$ -globina de humano localizado en el cromosoma once, comprende cinco genes ( $\epsilon$ ,  $G\gamma$ ,  $A\gamma$ ,  $\delta$  y  $\beta$ ) que se expresan de manera secuencial durante el desarrollo. Los polipéptidos de tipo  $\beta$  se combinan con los codificados en el dominio  $\alpha$ -globina localizado en el cromosoma 16 para formar tetrámeros de hemoglobina, la molécula transportadora de oxígeno dentro de los eritrocitos sanguíneos.

Las etapas durante la eritropoyesis en las cuales se expresan los diferentes genes que codifican para las  $\beta$ -globinas, corresponden de manera gruesa con los cambios en los sitios de eritropoyesis durante el desarrollo de los mamíferos. El gen  $\epsilon$ -globina se expresa en las islas sanguíneas del saco vitelino mientras que los genes  $\gamma$ -globinas fetales se expresan durante la eritropoyesis en el hígado fetal. Al momento del nacimiento y de

manera post natal, la eritropoyesis se traslada a la médula ósea en donde se expresa el gen  $\beta$ -globina adulto en cantidades elevadas y el gen  $\delta$ -globina en menor proporción (revisado en Kiefer *et al.*, 2008).

La estructura general de los dominios  $\beta$ -globina así como algunos de los elementos más sobresalientes, están muy conservados en vertebrados (humano, ratón y pollo) (Fig. 16), sin embargo existen algunas diferencias.

El arreglo de los genes  $\beta$ -globina difiere entre los mamíferos y el pollo. Tanto en el humano como en el ratón, los genes que se expresan de manera embrionaria durante la eritropoyesis primitiva se localizan río arriba de los genes adultos, mientras que en el pollo, los genes adultos se sitúan entre los dos genes embrionarios (Fig. 16). El elemento LCR (por Locus Control Region) en todos los dominios es el regulador eritroide principal para lograr una expresión robusta y sostenida de los genes globina. Sin embargo, en el pollo existe también un elemento *enhancer* entre el gen adulto  $\beta^A$  y el gen embrionario  $\varepsilon$  requerido para la expresión de los genes. Tanto en el ratón como en el humano existen dos sitios de hipersensibilidad que delimitan al dominio, ortólogos de los *insulators* 5' cHS4 y 3'cHS1 del pollo. En los dominios de mamíferos no existen 16 kb de heterocromatina al costado 5' , sin embargo, ambos loci están embebidos en un grupo de genes que codifican para receptores olfatorios que no son transcritos en el contexto eritroide (revisado en Kiefer *et al.*, 2008).



**Figura 16. Los dominios  $\beta$ -globina del pollo, el ratón y el humano.** Se muestran los genes (en rojo los genes globina adultos, en azul claro los embrionarios, en azul oscuro los fetales y en negros genes que codifican para receptores olfatorios) y los sitios de hipersensibilidad al corte por la DNasa I que corresponden a los elementos de regulación (flechas negras) incluyendo los elementos tipo *insulators* que los delimitan (flechas rojas). En este esquema el insulador 5' cHS4 del pollo se denomina únicamente HS4 (modificado de Kiefer *et al.*, 2008)

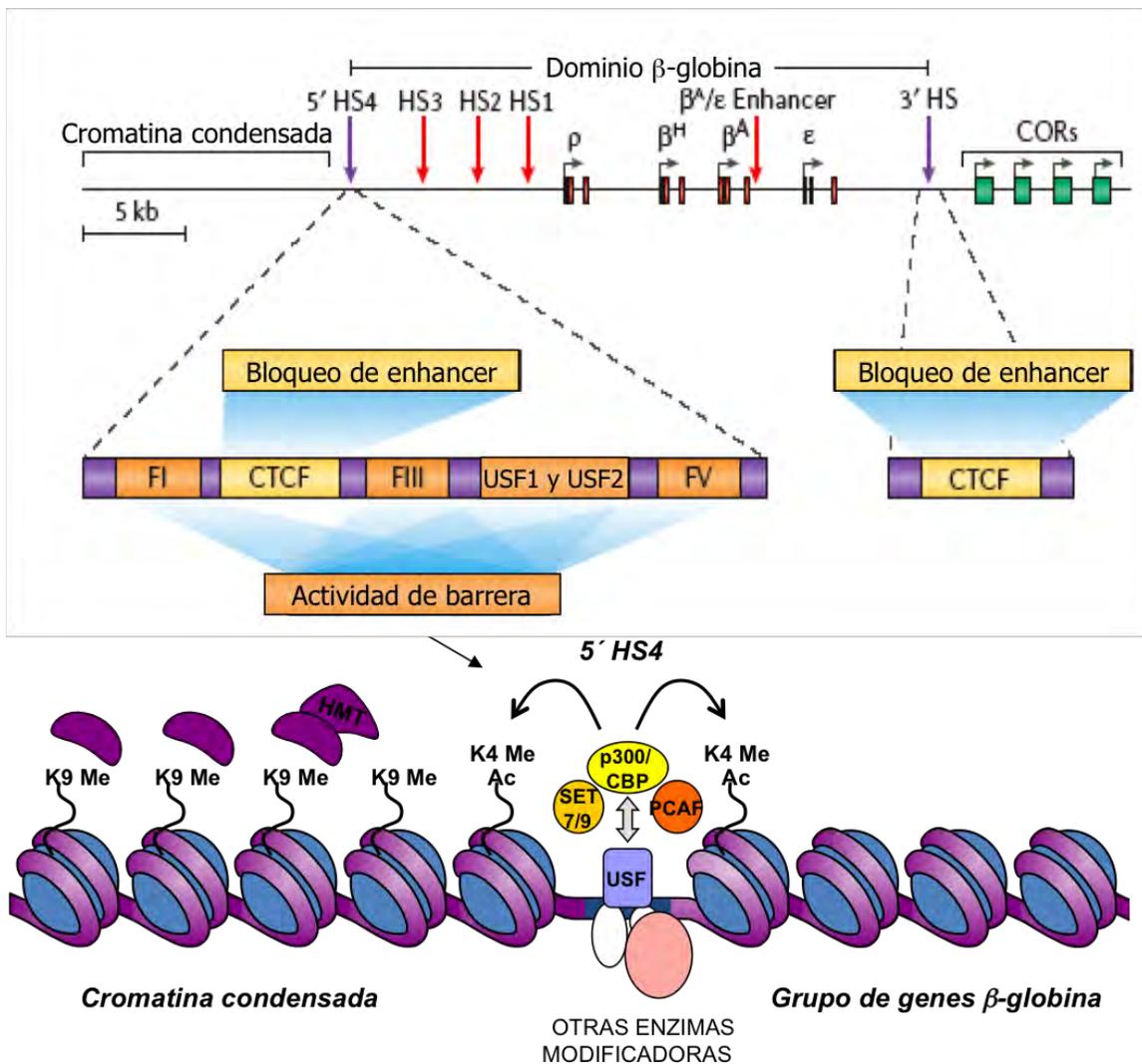
Ya que el presente proyecto se centra en los dominios globina del pollo, a continuación se profundizará en la descripción de éstos.

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

El dominio  $\beta$ -globina comprende 30 kb y contiene un gen embrionario ( $\epsilon$ ), uno fetal ( $\rho$ ) y dos genes adultos ( $\beta^H$  y  $\beta^A$ ) (Fig. 16 y 17). Entre el gen  $\beta^A$  y el gen  $\epsilon$  se localiza un *enhancer* eritroide específico similar al descrito en el dominio  $\alpha$ -globina, que regula a ambos genes y determina su expresión dependiendo de la etapa del desarrollo. Además, presenta un LCR que confiere de niveles elevados de expresión y favorece la apertura regulada del dominio a nivel cromatínico.

En la región 5' no-codificante de este dominio se han caracterizado cuatro sitios de hipersensibilidad. Tres de estos sitios que constituyen al LCR del dominio, son eritroide-

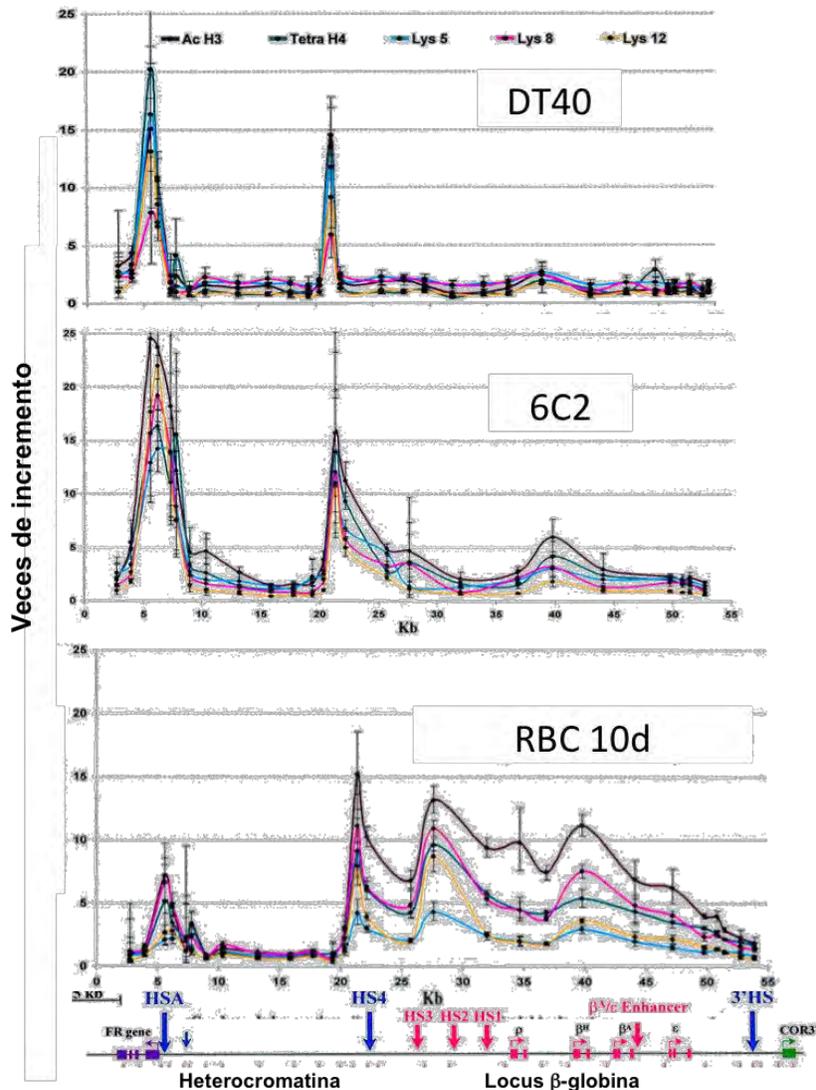
específicos y se ubican cerca de los genes  $\beta$ -globina. El cuarto sitio localizado hacia el extremo 5' es constitutivo (5' cHS4) y define la frontera cromatínica del dominio. Róo arriba del sitio 5' cHS4 existe una región de 16 kb de heterocromatina rica en secuencias repetidas y ADN metilado seguida del gen que codifica para el receptor del folato. Hacia el extremo 3' del dominio se ubica otro sitio de hipersensibilidad (3' cHS1) que representa el límite del dominio en este costado. Posterior a este sitio se encuentra un gen que codifica al receptor olfatorio COR3 $\beta$  que se expresa de manera cerebro-específica (Recillas-Targa y Razin, 2001; Gazner y Felsenfeld, 2006) (Fig. 17).



**Figura 17. El dominio  $\beta$ -globina de pollo.** Arriba, el dominio  $\beta$ -globina con los genes (cajas rojas), los elementos de regulación representados por sitios de hipersensibilidad (flechas rojas) y los dos *insulators* que delimitan al dominio (flechas moradas). Se muestran las cajas de unión de

diferentes factores proteicos incluyendo a CTCF, USF1 y USF2. Abajo, se muestra el modelo propuesto para la actividad de barrera por parte del *insulator* 5' cHS4 en donde están involucrados los factores USF, las acetilasa de histonas p300/CBP, PCAF y la metiltransferasa de histonas SET 7/9 (modificado de Gazner y Felsenfeld 2006).

El *insulator* 5' cHS4 posee tanto la actividad de bloqueo de *enhancer* (Bell *et al.*, 1999) como la de proteger a transgenes contra efecto de posición cromosomal (EPC) (Pikaart *et al.*, 1998, West *et al.*, 2004), y ambas actividades son separables (Recillas-Targa *et al.*, 1999). La capacidad de bloqueo de *enhancer* por parte de éste elemento se adjudica principalmente a la proteína CTCF mientras que la protección contra el efecto de posición cromosomal se asocia con la proteína USF y remodeladores de la cromatina (West *et al.*, 2004). De manera interesante, el *insulator* 3'HS1 no tiene sitios de unión al factor USF y sólo posee actividad de bloqueo de *enhancer* dependiente de CTCF. El modelo que se ha propuesto para explicar la capacidad de protección contra EPC por parte del *insulator* 5' cHS4 se muestra en la figura 17. Éste propone que los factores USF1-2 pueden reclutar actividades modificadoras de la cromatina como a p300/CBP y a Set7/9 que fomentan una conformación permisiva de la cromatina y así evitan la expansión de la heterocromatina adyacente. De manera consistente con esta idea, las histonas que conforman la cromatina en el elemento 5' cHS4 se encuentra permanentemente acetiladas mientras que el resto del dominio se acetila paulatinamente a medida que avanza la diferenciación eritroide (Fig. 18) (Litt *et al.*, 2001).



**Figura 18.** La estructuración de la cromatina del dominio  $\beta$ -globina de pollo durante la diferenciación eritroide. Gráficas que muestran el enriquecimiento evaluado por ChIP de diversas acetilaciones de lisinas en las histonas a lo largo del dominio  $\beta$ -globina de pollo y regiones adyacentes. De arriba a abajo se muestran los resultados para la línea linfóide DT40, la línea pre-eritroblástica 6C2 y eritrocitos terminalmente diferenciados de 10 días de desarrollo RBC 10d. En la parte inferior se muestra el esquema del dominio  $\beta$ -globina de pollo y regiones aledañas a lo largo del cual se evaluó el enriquecimiento de las acetilaciones de histonas mostrado en las gráficas superiores. Se observa un pico de enriquecimiento de histonas acetiladas sobre la región del *insulator* 5' cHS4 (flecha azul central) en todas las líneas celulares mientras que el resto del dominio (mostrado en rojo en el esquema), se enriquece paulatinamente de acetilaciones de histonas a medida que la diferenciación eritroide progresa (modificado de Litt *et al.*, 2001).

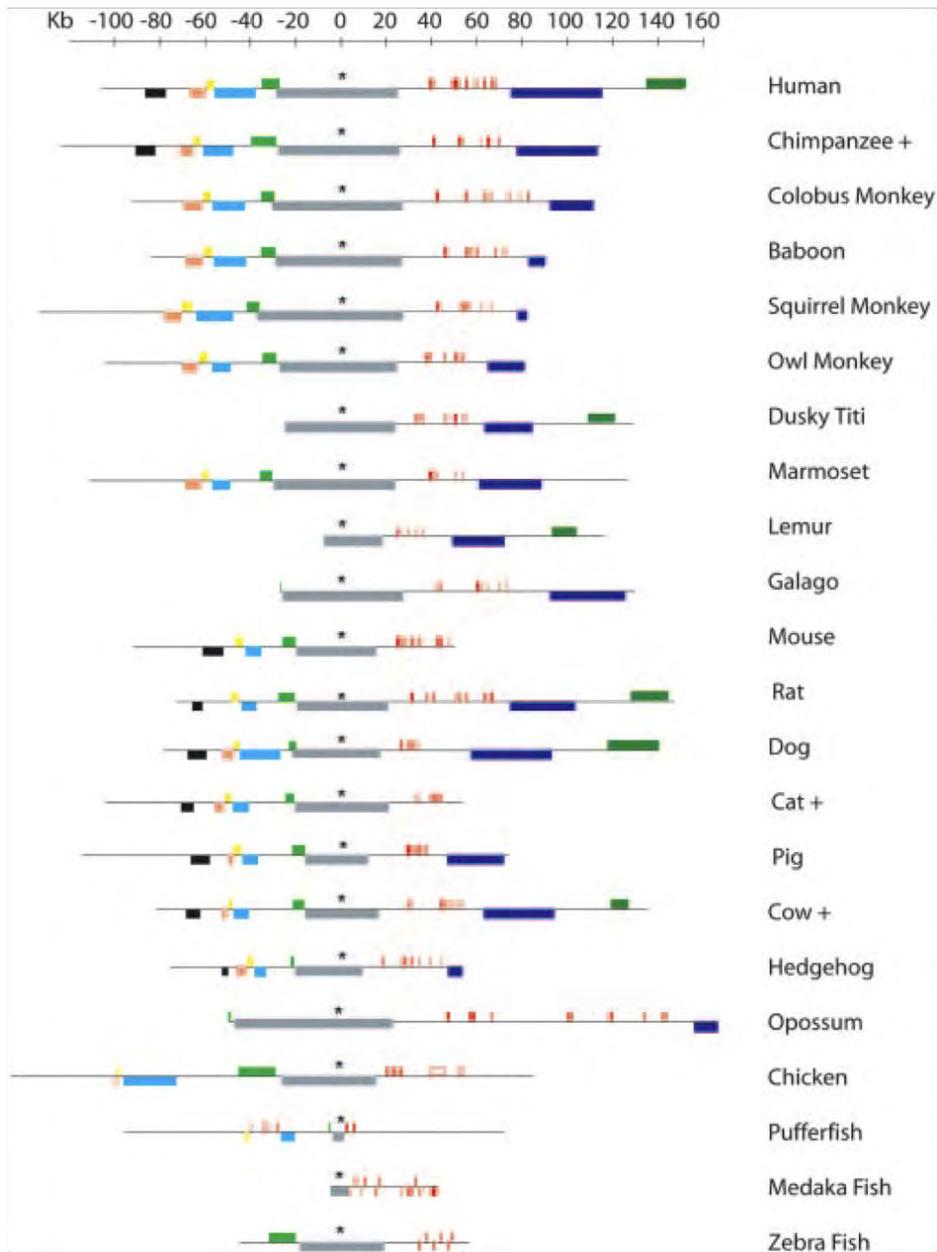
El dominio  $\beta$ -globina y todos sus componentes es uno de los locus mejor estudiados en vertebrados. En particular, la función del *insulator* 5' cHS4 se ha caracterizado a profundidad por lo cual, fue utilizado como control a lo largo de los

experimentos planteados en el presente trabajo en donde se evaluaron elementos todavía no definidos que pertenecen al dominio  $\alpha$ -globina de pollo.

## 2.2 El dominio $\alpha$ -globina

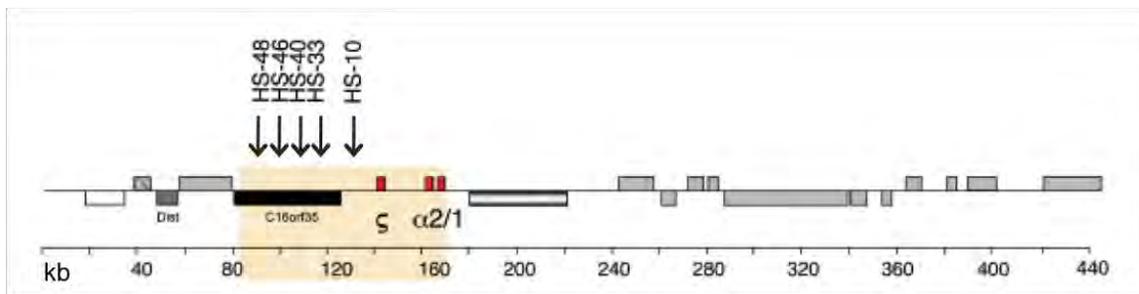
Al igual que el dominio  $\beta$ -globina, el dominio  $\alpha$ -globina comprende genes embrionarios y adultos que codifican para las cadenas  $\alpha$  de la hemoglobina y su expresión se regula a lo largo del desarrollo y durante la eritropoyesis. Sin embargo, el contexto cromosómico en donde se localiza el locus  $\alpha$ -globina difiere notoriamente del  $\beta$ -globina, ya que éste se encuentra embebido en una zona rica en genes (algunos constitutivos) que se mantiene en una conformación eucromática aún en células no-eritroides en donde los genes globina no se expresan (Hughes *et al.*, 2005).

El locus  $\alpha$ -globina y la sintenia de la región genómica en donde se ubica, se encuentra muy conservada entre los vertebrados. De igual forma, en la mayoría de los organismos (a excepción del ratón), el locus  $\alpha$ -globina se localiza cerca del telómero (Hughes *et al.*, 2005) (Fig. 19).



**Figura 19. El dominio  $\alpha$ -globina se localiza embebido en una región genómica rica en genes de sintenia conservada en la evolución de vertebrados.** Las cajas rojas representan a los genes  $\alpha$ -globina, la caja gris representa al gen de nombre *C16orf35* y el resto de las cajas corresponden a los diferentes genes que circundan al dominio  $\alpha$ -globina. Dentro del marco de lectura abierta del gen *C16orf35*, se localizan una serie de elementos de regulación del dominio  $\alpha$ -globina (ver más adelante). En particular se localiza un elemento eritroide de regulación muy conservado (representado por un asterisco en el esquema), que corresponde al LCR del dominio  $\alpha$ -globina denominado  $\alpha$ MRE (ver más adelante). Dicho elemento fue utilizado como punto central para realizar el alineamiento presentado (Hughes *et al.*, 2005).

En el dominio  $\alpha$ -globina de humano existen una serie de elementos de regulación incluyendo al  $\alpha$ MRE representado por el sitio de hipersensibilidad -40 (HS -40) (Fig. 20). Este elemento de regulación parece ser el único indispensable para conseguir una expresión robusta y sostenida de los genes globina humanos. Favoreciendo lo anterior, un estudio reciente que utiliza a la técnica de 3C ha demostrado que existe una interacción directa de todos los elementos de regulación (HS-48 a HS-10) con los genes globina en células eritroides, sin embargo, en ausencia del HS-40 estas interacciones se pierden y la expresión de los genes decae dramáticamente. Además, la reinsertión de dicho elemento en el costado 3' del dominio restaura las interacciones y consecuentemente los niveles de expresión (Vernimmen *et al.*, 2009). Con base en las evidencias anteriores se sugiere que el dominio  $\alpha$ -globina humano tiene una conformación más sencilla en comparación con el de ratón o el de pollo, en donde existen elementos adicionales necesarios para la regulación fina y diferencial de los genes al interior de cada uno de los dominios (ver detalles del dominio en el pollo más adelante (Gobbi, *et al.*, 2007).



**Figura 20. El dominio  $\alpha$ -globina de humano.** Se muestra una región genómica de 440 kb en la cual se ubica el dominio. Los genes globina se muestran en cajas rojas y los sitios de hipersensibilidad (HS-48, 46, 40, 33, 10) correspondientes a elementos de regulación eritroides se muestran con flechas (modificado de Vernimmen *et al.*, 2009).

Estudios de la localización nuclear del dominio  $\alpha$ -globina en células humanas han mostrado que existen diferencias entre éste y el dominio  $\beta$ -globina. El dominio  $\beta$ -globina se localiza en cromatina pericentromérica en células no-eritroides y se re-localiza hacia zonas de eucromatina en células eritroides. El dominio  $\alpha$ -globina en cambio, se sitúa siempre en regiones de eucromatina independientemente de su estado transcripcional, lo

cual concuerda con que la región del genoma en donde se ubica es rica en genes de expresión constitutiva (Brown *et al.*, 2001).

El dominio  $\alpha$ -globina tanto en el ratón como en el humano poseen una cromatina con histonas ricamente acetiladas en líneas celulares eritroides, con un mayor enriquecimiento de la acetilación en eritrocitos maduros, en donde los genes globina alcanzan su máxima expresión. En particular la acetilación de histonas y la metilación de la lisina 4 de la histona 3 (H3K4me1,2 y 3) se enriquece en los elementos de regulación y en los promotores de los genes globina a lo largo de la diferenciación eritroide (Anguita *et al.*, 2001, Vernimmen *et al.*, 2009).

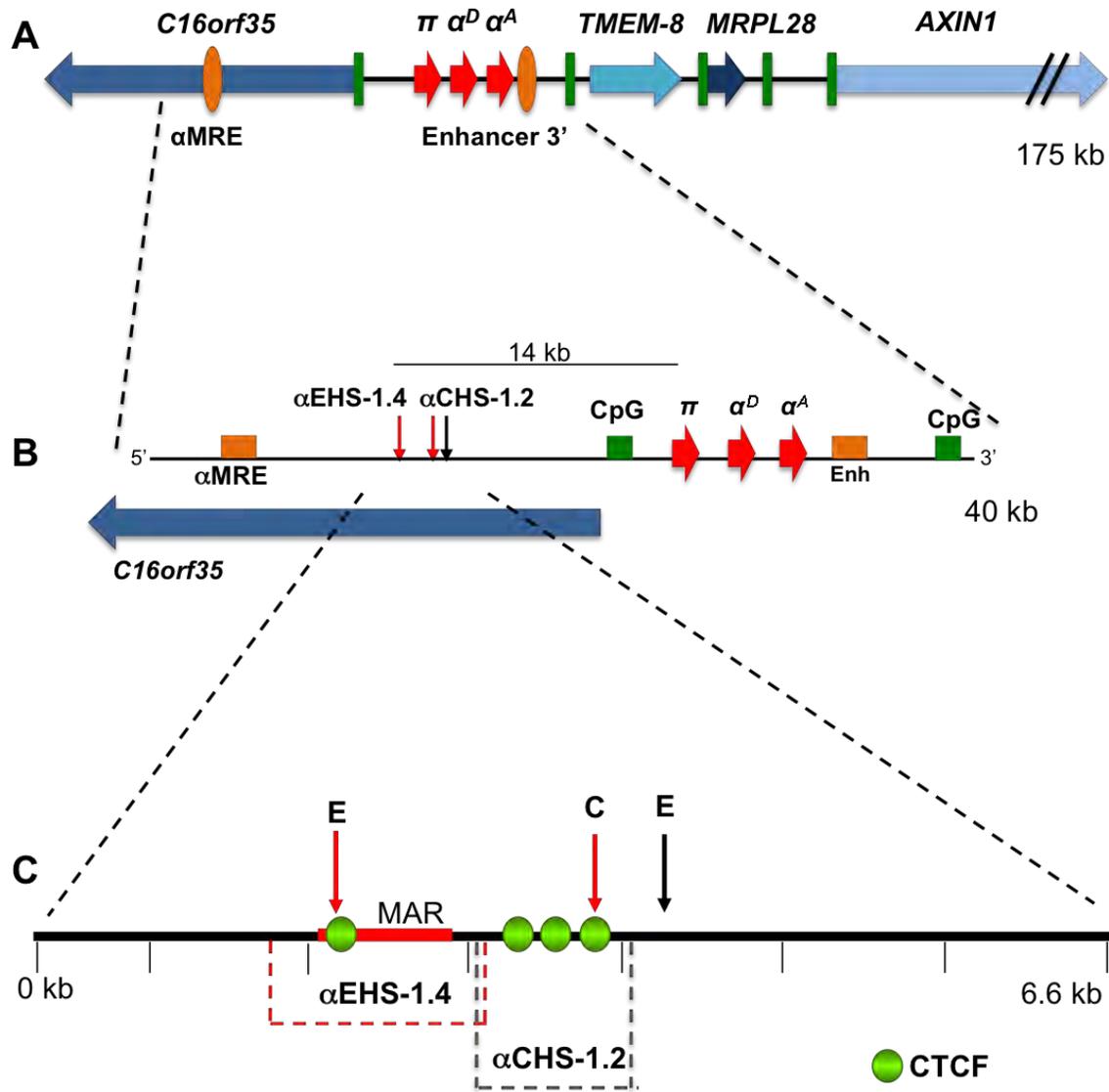
Otro rasgo interesante del dominio  $\alpha$ -globina de humano es que, en células de linaje no-eritroide en donde los genes globina no se expresan, éste se encuentra enriquecido en proteínas represoras del grupo Polycomb y en la marca represiva H3K27me3 asociada al silenciamiento. Esto contrasta con el dominio  $\beta$ -globina en donde no hay una relación entre el silenciamiento de los genes globina en linajes no-eritroides y el grupo de proteínas Polycomb (Garrick *et al.*, 2008). La presencia y ausencia respectiva de Polycomb en los dominios  $\alpha$  y  $\beta$ -globina, sugiere que el contexto cromosómico en donde se encuentran cada uno tiene influencia sobre el mecanismo de silenciamiento utilizado en células en donde los genes globina no se expresan. El locus  $\alpha$ -globina se localiza en una región rica en genes que se expresan de manera constitutiva, por lo tanto, probablemente requiere de un mecanismo activo de silenciamiento como el reclutamiento de factores Polycomb. En cambio, el locus  $\beta$ -globina, ubicado en una región pobre en genes y cercana a heterocromatina (como en el caso del pollo), posiblemente utilice un mecanismo que aproveche el silenciamiento que pudiera originarse y propagarse desde dicha heterocromatina.

### **2.2.1 El dominio $\alpha$ -globina de pollo**

El dominio  $\alpha$ -globina de pollo también se localiza en una zona rica en genes. La región que abarca a los genes globina y a sus elementos de regulación conocidos hasta el momento comprende alrededor de 40 kb. El dominio incluye a un gen embrionario  $\pi$  y dos

genes adultos  $\alpha^D$  y  $\alpha^A$  que se expresan en forma diferencial durante el desarrollo. El gen  $\pi$  se expresa en eritrocitos embrionarios hasta los 6-7 días del desarrollo en donde ocurre una transición a la expresión de los genes adultos  $\alpha^D$  y  $\alpha^A$  (Rincón-Arano *et al.*, 2009). En la región 3' del dominio nuestro grupo ha descrito un elemento *enhancer* ubicado a 1 kb del gen  $\alpha^A$  (Escamilla-del-Arenal y Recillas-Targa, 2008; Rincón-Arano *et al.*, 2009). El gen más cercano río abajo del gen  $\alpha^A$  está aproximadamente a 10 kb del elemento *enhancer* y codifica para una proteína transmembranal (*TMEM8*) (Flint *et al.*, 2001) (Fig. 21). En esta región no-codificante se han identificado secuencias repetidas tipo CR1 características del genoma del pollo, y secuencias ricas en AT que parecen coincidir con regiones de unión a matriz nuclear (MAR) (Recillas-Targa y Razin, 2001). La estructura de la cromatina de la región 3' no-codificante del dominio  $\alpha$ -globina del pollo se desconoce y tampoco se sabe si existen elementos que pudieran actuar como el límite 3' del dominio.

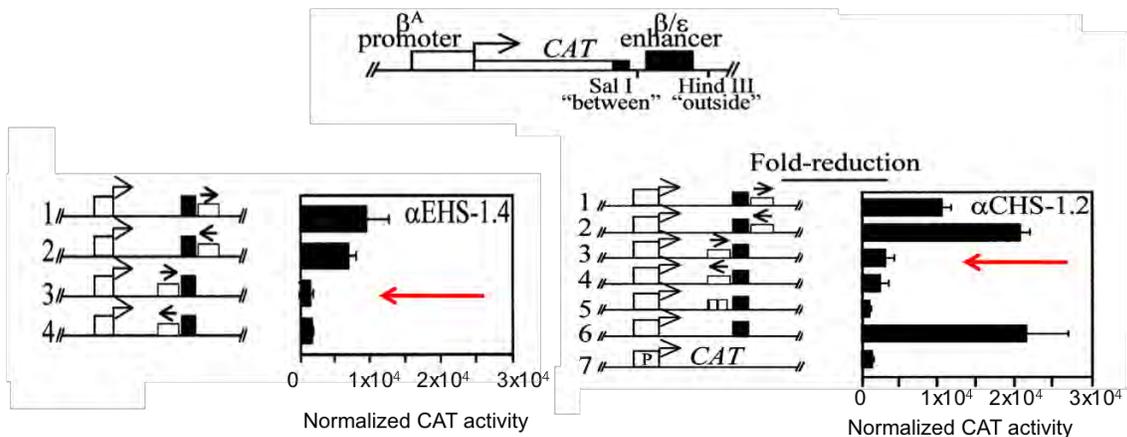
En la region 5' del dominio existe un gen constitutivo denominado *ggPRX* (*C16orf35*) que se transcribe en dirección opuesta a los genes globina y colocaliza con varios elementos de regulación eritroides (Recillas-Targa y Razin, 2001). En particular, se ha identificado un elemento de regulación conservado que representa al LCR del dominio ( $\alpha$ MRE, descrito previamente), y a tres sitios de hipersensibilidad que corresponden a elementos de función prácticamente desconocida y que son objeto del presente trabajo de investigación.



**Figura 21. Esquema del dominio  $\alpha$ -globina del pollo.** A. la región genómica (~175 kb) en donde se ubica el dominio  $\alpha$ -globina de pollo. Se muestran los genes codificantes (flechas horizontales), las islas CpG (bloques verdes), los elementos de regulación principales del dominio (óvalos o cuadros naranjas). B. El dominio  $\alpha$ -globina con los tres sitios de hipersensibilidad localizados a 14 kb del gen embrionario  $\pi$  (flechas verticales rojas para los elementos en estudio en este trabajo y flecha negra para un elemento de función desconocida). C. Acercamiento a la región de 6.6 kb que comprende a los sitios de hipersensibilidad de interés en el presente trabajo. El elemento  $\alpha$ EHS-1.4 y el elemento  $\alpha$ CHS-1.2 son secuencias con capacidad de bloqueo de *enhancer*. CTCF se une *in vitro* en ambos elementos (Valadez-Graham *et al.*, 2004).

Los sitios de hipersensibilidad ubicados río abajo del  $\alpha$ MRE se localizan a 14 kb del gen embrionario  $\pi$ . Éstos constituyen un sitio constitutivo rodeado de dos sitios eritroide-específicos (Fig. 21). El sitio eritroide-específico de 1.4 kb que se encuentra en el extremo

5' coincide con una secuencia de anclaje a la matriz nuclear (MAR) (Razin *et al.*, 1999). Un estudio realizado en nuestro laboratorio, permitió caracterizar a estos tres sitios de hipersensibilidad y determinó que tanto el sitio constitutivo de 1.2 kb como el sitio eritroide específico de 1.4 kb tienen la propiedad de bloqueo de *enhancer* y en ambos casos el factor CTCF puede unirse *in vitro* (Fig. 22) (Valadez-Graham *et al.*, 2004).



**Figura 22. Los elementos  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.4 poseen capacidad de bloqueo de *enhancer*.** Resultados de transfecciones transitorias utilizando al gen *CAT* como reportero. Al colocar tanto al elemento  $\alpha$ EHS-1.4 como al  $\alpha$ CHS-1.2 entre el *enhancer*  $\beta/\epsilon$  del dominio  $\beta$ -globina de pollo y el gen reportero *CAT* bajo el promotor  $\beta^A$  del mismo dominio, se observa una reducción en la activación del reportero debida al bloqueo del *enhancer* (indicado por las flechas). No así cuando los elementos se colocan al costado 3' del *enhancer*. En el panel izquierdo, además se muestran los resultados utilizando al *insulador* 5' *chs4* como control de bloqueo (caja rallada) y a las construcciones con el *enhancer* solo y el promotor solo con las que se normalizaron los datos (modificado de Valadés-Graham *et al.*, 2004).

La función específica de los elementos  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2 al interior del dominio  $\alpha$ -globina de pollo se desconoce. Es interesante notar que ambos elementos están inmersos en el dominio globina y tienen la propiedad de bloqueo de *enhancer in vitro*. No existen muchos ejemplos en la literatura en donde elementos tipo *insulator* formen parte integral de la regulación fina de la expresión de genes ya que en general, éstos se localizan delimitando al dominio en cuestión, como en el caso el dominio  $\beta$ -globina de pollo. El dominio  $\alpha$ -globina de pollo provee de un escenario en donde estudiar la combinación de elementos necesarios para conseguir la expresión correcta de genes regulados en la diferenciación y el desarrollo y en particular, la participación de elementos *insulator* en dicha regulación.

### 3. PLANTEAMIENTO DEL PROBLEMA

El locus  $\alpha$ -globina de pollo representa un sistema interesante para abordar el estudio de la conformación del genoma eucarionte en dominios funcionales regulados a lo largo de la diferenciación y el desarrollo. Como se ha mencionado, en la literatura se ha postulado que los dominios  $\alpha$ -globina están inmersos en regiones constitutivamente abiertas del genoma en una gama de organismos a lo largo de la evolución. Sin embargo, datos preliminares de nuestro laboratorio sugieren que la cromatina del dominio  $\alpha$ -globina de pollo se regula a través del desarrollo y la diferenciación eritroide, lo cual sugiere que el estado epigenético del locus se modula dependiendo de las necesidades transcripcionales. En este contexto resulta interesante comprender la contribución de los *insulators*  $\alpha$ EHS -1.4 y  $\alpha$ CHS-1.2 identificados en la región 5' del dominio en la regulación y definición del mismo. Consideramos que profundizar en la función de estos elementos nos permitirá concebir un panorama más amplio de la estructuración epigenética del dominio y su regulación. Así mismo, al comprender los mecanismos de acción de estos *insulators*, esperamos aportar información relevante que contribuya a entender la diversidad de funciones que pueden llevar a cabo este tipo de elementos en la regulación epigenética del genoma al interior del núcleo celular.

## 4. HIPÓTESIS

Los *insulators*  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2 localizados en el costado 5' no-codificante del locus contribuyen a la correcta estructuración cromatínica del dominio y a la regulación de la expresión diferencial de los genes  $\alpha$ -globina a lo largo de la diferenciación eritroide y del desarrollo del organismo.

## 5. OBJETIVO GENERAL

Entender la contribución de los *insulators*  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2 en la formación y regulación de la estructura de la cromatina del dominio y su influencia en la regulación de la expresión diferencial de los genes  $\alpha$ -globina durante la diferenciación eritroide y el desarrollo del organismo.

### 5.1 Objetivos Particulares

1. Estudiar las propiedades de los *insulators*  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2 del dominio  $\alpha$ -globina de pollo.
2. Estudiar la contribución del factor CTCF en la función de los *insulators*  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2.

## 6. ESTRATEGIAS EXPERIMENTALES

### 6.1 Cultivo de líneas celulares

Todas las líneas celulares utilizadas se mantuvieron a 37 °C, 5% CO<sub>2</sub> con el medio de cultivo correspondiente. La línea celular 6C2 es una línea celular obtenida a partir de médula ósea de pollo y transformada con el virus de la eritroblastosis aviar (AEV), dichas células son consideradas pre-eritroblastos detenidos en estado CFU-E (Unidades Formadoras de colonias Eritroides). Esta línea celular se cultiva en medio de cultivo  $\alpha$ -MEM (GIBCO) suplementada con 10% de Suero Fetal Bovino (SFB), 2% de suero de pollo (ChS), 1 mM de Hepes (pH=7.2) y 50 mM de  $\beta$ -mercaptoetanol. La línea celular HD3 arrestada mediante la infección con el virus AEV, corresponde al estadio de eritroblastos. Esta línea celular se mantiene con medio DMEM suplementado con 8% de SFB y 2% de ChS. La línea celular linfóide DT40 (célula hematopoyética no eritroide) fue mantenida en DMEM suplementado con 8% SFB y 2% de ChS. Todos los medios de las líneas celulares fueron complementados con penicilina y estreptomina en una concentración de 500 U/ml. Los cultivos primarios de eritrocitos terminalmente diferenciados de pollo fueron obtenidos mediante punción de la vena apical de embriones de pollo incubados durante distinto número de días (embriones de *Gallus gallus domesticus*, Alpes, Puebla México). El inhibidor de desacetilasas de histonas Tricostatina-A (TSA) fue utilizado a una concentración de 2.5 ng/ml en medio de cultivo por 24 h. El inhibidor de la mutación del ADN 5-aza-2-deoxicitidina (5-azadC) fue utilizado a una concentración de 3  $\mu$ M por tres días realizando cambios diarios de dicho reactivo.

### 6.2 Diferenciación de células HD3

Las células HD3 correspondientes a eritroblastos de pollo fueron cultivadas en medio de diferenciación (DMEM suplementado con 10% SFB, 10% ChS, 1 mg/ml de albúmina bovina (BSA) y 100 mM  $\beta$ -mercaptoetanol) y en presencia de 5% de suero de pollo anémico a 42°C por 5, 7 ó 12 días. El suero de pollo anémico se obtuvo a través de

desangrar por dos días consecutivos gallinas de primera puesta (12 semanas). Al tercer día los animales se sangraron a blanco. El suero se obtuvo a través de incubar la sangre por 1 h a temperatura ambiente y 1 h a 4° C. La sangre se centrifugó a 3000 rpm por 10 min a 4° C en la centrífuga Labofuge 400R Heraeus. Alrededor de 8 ml de suero de cada gallina se colectaron bajo las circunstancias ya descritas. Para determinar la diferenciación de las células HD3 se tiñeron con benzidina en medio ácido (0.5% benzidina en 0.3% de H<sub>2</sub>O<sub>2</sub> y 1M de ácido acético) para su posterior análisis a microscopio de luz obteniendo alrededor de un 70% de células diferenciadas. En algunos experimentos la diferenciación se llevó a cabo en ausencia de suero anémico y se corroboró la expresión de los genes  $\alpha$ -globina por medio de RT-PCR dúplex. Asimismo, se obtuvieron pellets de las células diferencias y se comparó la coloración de éstos con las células HD3 control. Las células diferenciadas presentan una clara coloración rojiza debido a las presencia de hemoglobinas.

### 6.3 Transfecciones transitorias y estables

Las transfecciones transitorias se realizaron por medio de electroporación para los eritrocitos primarios. Se extrajo sangre de embriones de pollo de 10 días de desarrollo y se fue acumulando en la solución 1 (PBS 1X, EDTA al 10%) en hielo. Se centrifugó la sangre por 10 min a 1000 rpm en la centrífuga Labofuge 400R Heraeus. Se lavaron dos veces con PBS y se resuspendieron en 10 ml de solución 1. Se cuantificaron las células (mediante leer la absorbancia a 412 nm en donde 1 OD=  $8.5 \times 10^6$  eritrocitos de 10 días de desarrollo) y se resuspendieron en medio L15 sin complementar dividiéndolas en el número de tubos necesario según el número de transfecciones a realizar tomando en cuenta que se requiere de  $1 \times 10^8$  RBC por transfección. Las transfecciones se realizaron utilizando el sistema “Dual-Luciferase Reporter Assay System” (DLR™ Promega) que consiste en la co-transfección del plásmido experimental que contiene al gen de la luciferasa de luciérnaga (*Photinus pyralis*) y un plásmido control que codifica para la el gen de la luciferasa de pólipo coralino (*Renilla reniformis*). Se añadieron a cada tubo ambos plásmidos (renilla + luciferasa) y se procedió a electroporar. Se agregó 1 ml de medio L15 complementado a

cada transfección y se incubaron a 37°C por 36 hrs. Se centrifugaron los tubos a 1300 rpm por 2 min en la centrífuga Labofuge 400R Heraeus. Se aspiró el sobrenadante y se agregó 50 µl de amortiguador de lisis a cada tubo hasta resuspender. Se congelaron las células a -70°C dos veces. Se centrifugaron las células 2 min a 3200 rpm en la centrífuga Labofuge 400R Heraeus. Se tomó el sobrenadante y se agregó el sustrato de la luciferasa, se tomó la lectura experimental en el luminómetro. Se agregó el sustrato de la renilla y se midió la señal control. Cada medida se realizó por triplicado. Posteriormente se normalizaron los datos experimentales con la señal obtenida para el plásmido control.

Las transfecciones estables de la línea celular HD3 se llevaron a cabo por medio de lipofección (Invitrogen). Los plásmidos a transfectar (ver cada caso) se linearizaron por medio de digestiones enzimáticas específicas. Se transfectaron diferentes cantidades (ver cada caso) ajustando la cantidad de Lipofectamina según las instrucciones del producto (3 µl de Lipofectamina por cada 1 µg de ADN). Se incubó la lipofectamina con el ADN en DMEM sin suero (200 µl) por 20-30 min a temperatura ambiente. Se cuantificaron las células utilizando  $1 \times 10^6$  células por transfección en 0.8 ml de DMEM sin suero. Se colocaron en placas de 6 pozos y se agregó el plásmido + Lipofectamina por goteo lento. Se dejó incubar de 4-6 horas a 37°C, 5% CO<sub>2</sub> y se agregó 3 ml de medio HD3. Se dejó incubar de 48 a 72hrs para proceder a la selección con el antibiótico correspondiente en cada caso.

## **6.4 Ensayo de protección contra efectos de posición cromosomal (EPC)**

La transfección estable de los diferentes plásmidos utilizados se realizó por lipofección como descrito anteriormente. Los plásmidos transfectados fueron el pGα<sup>D</sup> que contiene al gen reportero *EGFP* bajo la expresión del promotor α<sup>D</sup> del dominio α-globina de pollo, el pGα<sup>D</sup>HS4 que contiene al “core” del *insulator* 5' cHS4 del dominio β-globina de pollo a cada lado del gen reportero, el pGα<sup>D</sup>1.4 que contiene al elemento αEHS-1.4 del dominio α-globina de pollo y el pGα<sup>D</sup>1.2 que incluye al elemento αCHS1-2 también

perteneciente al dominio  $\alpha$ -globina de pollo. Una vez realizada la transfección, clonas con un solo evento de integración fueron seleccionadas utilizando medio semisólido (Methocel, Fluka) con 0.9 mg/ml de geneticina (Calbiochem). Después de 14 días de selección, clonas individuales fueron aisladas y expandidas. El nivel de fluorescencia se midió por citometría de flujo (FACS) (día 0 de cultivo). La integridad y número de copias del transgene (uni o multicopia) fue evaluada por southern blot utilizando una sonda de ADN contra un fragmento del gene de *EGFP*. El nivel de fluorescencia de las distintas clonas fue evaluado por FACS cada quince días por un total de 100 días de cultivo continuo sin geneticina.

Las versiones mutantes  $\alpha$ EHS-1.4 $\Delta\Delta$  del *insulator* fueron generadas a partir de PCRs anidadas utilizando oligonucleótidos que contienen mutaciones en las cajas  $\Delta$ CAG y  $\Delta$ CTAG (Valadéz-Graham *et al.*, 2004) (ver Tabla 1 en p. 64-65). Primero se realizó una PCR utilizando un oligonucleótido silvestre en el extremo 5' del  $\alpha$ EHS-1.4 y un reverso 3' mutante. De la misma forma se amplificó el fragmento de ADN restante del  $\alpha$ EHS-1.4 utilizando un oligonucleótido forward mutante y un reverso silvestre en el extremo 3' del  $\alpha$ EHS-1.4. Se purificaron ambos fragmentos de ADN y se realizó una nueva PCR en donde ambos fragmentos se agregaron de manera equimolar (sin templado). Después de 4 ciclos de amplificación se detuvo la PCR y se agregaron los oligonucleótidos silvestres de los extremos 5' y 3' del  $\alpha$ EHS-1.4 obteniendo así la versión del *insulator* con las dos cajas mutadas  $\alpha$ EHS-1.4 $\Delta\Delta$ . Éste se clonó en el pG $\alpha^D$ 3 para generar el plásmido pG $\alpha^D$ EHS-1.4 $\Delta\Delta$ . El ensayo de protección contra efectos de posición cromosomal se llevó a cabo como descrito anteriormente.

## 6.5 Inmunoprecipitación de la cromatina (ChIP)

Se utilizaron de 1 a  $2 \times 10^7$  células por cada inmunoprecipitación ó 400  $\mu$ g de cromatina. Las células se lavan dos veces con PBS-SFB 2% y se resuspendieron en 30 ml de PBS-Formaldehido 1% durante 10 min a temperatura ambiente con agitación suave constante. El entrecruzamiento se detuvo con glicina a una concentración final de 125

mM. A partir de este momento todo se lleva a cabo en hielo con las soluciones a 4°C a excepción del buffer de lisis. Se lavaron las células dos veces con PBS frío, se resuspendieron en buffer de lisis (1% SDS, 5 mM EDTA, 50 mM Tris-HCl, pH=8.1 más inhibidores de proteasas) y se incubaron en hielo por 10 min. Se precedió a sonicarlas para obtener fragmentos de cromatina entre 300 y 600 pb (rango corroborado por electroforesis). Se centrifugó la muestra a máxima velocidad a 4°C en una centrífuga Eppendorf 5417-R y el sobrenadante se colocó en un tubo limpio. Se tomó una alícuota de 50 µl como INPUT. El resto se llevó a 10 ml con buffer de dilución (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl, pH=8.1 más inhibidores de proteasas). En el caso de haber partido de un número de células distinto, entonces se cuantificaron las proteínas para asegurar tener 400 µg de cromatina por inmunoprecipitación en 1 ml de buffer dilución:lisis (9:1). Se incubó la cromatina con 50 µl de perlas (proteína A/G sefarosa, Amersham) prebloqueadas para reducir el ruido de fondo por 2 h a 4° C en agitación giratoria. Las perlas se removieron por centrifugación (3000 rpm, 5 min, 4°C en una centrífuga Labofuge 400R Heraeus). Se dividió la cromatina en alícuotas de 1 ml para su posterior inmunoprecipitación.

Se agregó el anticuerpo de interés a cada uno de los tubos y se incubó con agitación giratoria a 4°C toda la noche. Se agregaron 30 µl de perlas (proteína A/G sefarosa) prebloqueadas a cada tubo y se incubaron 2-4 h a 4°C en agitación giratoria. Una vez incubadas las perlas se recuperaron por centrifugación (2000 rpm por 1 min a 4°C en una centrífuga Eppendorf 5417-R). Se lavaron una vez incubando en agitación giratoria por 10 min a 4°C con cada uno de los siguientes buffers: Paro I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl, pH=8.1 más inhibidores de proteasas). Paro II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl, pH=8.1 más inhibidores de proteasas). Paro III (0.25 M LiCl, 1% NP40, 1% deoxicolato, 1mM EDTA, 10 mM Tris-HCl, pH=8.1 más inhibidores de proteasas), TE (10 mM Tris-HCl, 2 mM EDTA, pH=8 más inhibidores de proteasas). Todos los buffers se mantuvieron en hielo durante el proceso y entre cada uno las perlas se recuperaron por centrifugación (2000 rpm, 1 min, 4°C en centrífuga eppendorf 5417-R).

Finalmente se añadió 150 µl de solución de elución a cada inmunoprecipitación (1 % SDS, 0.1 M NaHCO<sub>3</sub>) se incubaron con Proteinasa K (Sigma) y RNasa por 1-2 h a 37°C y se revirtió el entrecruzamiento incubando a 65°C toda la noche. La fracción soluble fue purificada utilizando en sistema Minielute (Qiagen). El ADN recuperado en cada inmunoprecipitación se utilizó como templado para realizar reacciones de PCR dúplex en presencia de [ $\alpha$ -<sup>32</sup>P]dCTP para su posterior cuantificación. Los fragmentos amplificados se resolvieron en geles de acrilamida verticales al 6%.

La PCR dúplex se realizó con dos pares de oligonucleótidos; uno que amplifica la región de interés y otro que amplifica una región control en la cual no hay asociación de la proteína evaluada. Para cada PCR dúplex se estandarizaron las condiciones para obtener una amplificación similar de ambos fragmentos (relación cercana a 1) en el rango lineal de la amplificación (ver Tabla 2 p. 69). Cada gel se cuantificó utilizando el programa ImageQuant3 utilizando la siguiente ecuación:

**Enriquecimiento=**

Fragmento experimental / Fragmento control) anticuerpo experimental

(Fragmento experimental/ Fragmento control) IgG inespecífica

siendo un valor > 1 considerado enriquecimiento significativo.

## **6.6 Generación del *knockdown* contra CTCF y tratamiento con 3-Aminobenzamida (3-ABA)**

Para generar el *knockdown* contra CTCF se utilizaron una colección de RNAs interferentes con una estructura tallo-asa (shRNAs) provenientes de distintas fuentes; 3 de ellos fueron diseñados por nosotros con la ayuda del Dr. Fabián Flores (pSiLCTCF Hygro1,2 y 3), 3 de ellos fueron diseñados con base en las secuencias reportadas por Lefevre *et al.*, 2008 (pSiLefevre 1, 2 y 3) y uno de ellos fue amablemente proporcionado

por el Dr. Ishihara (Kumamoto University, Kumamoto, Japón) (pCT1) (Para ver las secuencias ir a Tabla 1). También se diseñaron oligonucleótidos para generar un shRNAi control (nombrado scramble) que corresponde a los nucleótidos en el shRNAi del plásmido pCT1 distribuidos aleatoriamente. Todos ellos (a excepción del pCT1 que recibimos como vector), fueron sintetizados por SIGMA, purificados por PAGE e hibridados siguiendo las indicaciones del fabricante (Ambion) para ser clonados en el plásmido correspondiente. Una serie de combinaciones de estos shRNAis fueron utilizados para obtener reducciones adecuadas de los niveles de CTCF. Inicialmente se utilizaron los 3 shRNAis diseñados por nosotros (pSiLCTCF Hygro 1, 2 y 3) dando como resultado una disminución ineficiente de CTCF (alrededor del 30%). Los plásmidos fueron transfectados de manera estable en células HD3 en distintas cantidades (250 ng, 500 ng y 1  $\mu$ g) y fueron seleccionados con higromicina 0.8mg/ml (Calbiochem). El análisis de la cantidad de CTCF fue realizado por medio de western blot (ver abajo) en cuanto el cultivo se recuperó de la selección (2 semanas aproximadamente). Posteriormente se transfectó el plásmido pCT1 (500 ng y 1  $\mu$ g) obteniendo una mejor reducción de los niveles de CTCF (alrededor del 50 %). La transfección se realizó de manera estable en células HD3 y las células transfectantes fueron seleccionadas con puromicina (1  $\mu$ g/ml) (SIGMA). Posteriormente se transfectaron los 3 shRNAis (pSiLefevre 1, 2 y 3) en distintas cantidades (250 ng y 500 ng) de la manera descrita anteriormente. Las células transfectantes se seleccionaron con puromicina (1  $\mu$ g/ml) y el análisis de la disminución de los niveles de CTCF se realizó en cuanto se recuperó el cultivo (2 semanas aproximadamente). La disminución observada fue de alrededor del 50% al igual que para el pCT1. Finalmente se intentaron diferentes combinaciones con los 7 shRNAis. La combinación más eficiente fue combinar al pCT1 con los plásmidos pSiLefevre 1, 2 y 3 (500 ng de cada uno) generando una disminución de alrededor del 55-60% (los detalles se describen en la sección de resultados).

Para evaluar el efecto del inhibidor de la PARP-1 (3-ABA) en la capacidad de protección contra EPC por parte del *insulator*  $\alpha$ EHS-1.4 las líneas celulares 506 y 528 (que contienen al *insulator*  $\alpha$ EHS-1.4) fueron tratadas con 4 y 8 mM 3-ABA. Se analizó el nivel de fluorescencia de dichas líneas por citometría de flujo durante 30 días de cultivo. Para

llevar a cabo el análisis de ChIP, western blot y RT-PCR las células HD3 fueron incubadas de 3-6 días con 3-ABA a una concentración de 8 mM.

## **6.7 Análisis de expresión mediante RT-PCR**

Las células fueron resuspendidas en Trizol (Invitrogen) y el RNA fue aislado siguiendo las instrucciones del fabricante. El RNA obtenido fue resuspendido en agua libre de RNasas y cuantificado. La reacción de transcripción reversa se realizó utilizando el sistema de Promega. 5 µg de RNA fueron mezclados con oligo-dT en un volumen de reacción de 50 µl. Las amplificaciones por PCR fueron realizadas utilizando la ADN polimerasa en presencia de [ $\alpha$ -<sup>32</sup>P]dCTP. Las condiciones de amplificación fueron las siguientes: desnaturalización a 95°C por 5 min (1 ciclo), desnaturarización a 95°C por 30 s, apareamiento a 65°C por 30 s y una temperatura de extensión de 72°C por 50 s durante 25 ciclos para concluir con una extensión de 5 min a 72°C. La reacción de PCR fue diseñada para obtener de manera reproducible las condiciones lineales de amplificación. Los fragmentos fueron separados en un gel vertical de acrilamida al 6% en TBE 0.025X. El gel fue secado y expuesto en una pantalla sensible a radioactividad y el registro se llevó a cabo en un sistema Typhoon (Amersham). Los datos se analizaron en el programa ImageQuant3. Con el fin de hacer un cálculo semicuantitativo, se realizó una PCR dúplex considerando dos pares de oligonucleótidos, un par que corresponde al gen de interés y otro para el gen control constitutivamente expresado ( $\beta$ -actina). Los oligonucleótidos fueron diseñados sobre la zona de empalme entre exones para amplificar el mRNA maduro y de esta forma evitar la amplificación de ADN genómico (ver Tabla 1).

## **6.8 Western blot: Electroforesis e Inmunodetección**

### *Electroforesis.*

Se utilizó la cámara de BioRad como sugiere el fabricante. Se prepararon geles de resolución de bisacrilamida (al 7.5% para CTCF de 130 kD y al 6.5% para resolver las dos formas de CTCF, una de 130 kD y otra de 180 kD que se encuentra poli-ADP-ribosilada).

Los geles se prepararon mediante una mezcla de 1.25 ml de buffer resolving (48 mM de HCl, 3 M Tris base, pH=8.9), 1.875 ml ó 1.625 ml de solución de bisacrilamida al 40% (para gel al 7.5 ó 6.5 %, respectivamente) (38.96 g de acrilamida (Invitrogen), 1.04 g de bis-acrilamida (Invitrogen) en agua desionizada para 100 ml), 100 µl de 10% dodecil sulfato de sodio (SDS), 100 µl de 10% persulfato de amonio (APS), 10 µl de TEMED ajustado a 10 ml con agua desionizada. Dicha mezcla se deja polimerizar en la cámara de electroforesis agregando una capa de isopropanol en la parte superior para obtener un borde del gel uniforme. Posteriormente se retira el isopropanol, se enjuaga y se vierte al gel concentrador al 3.5% de acrilamida. El gen concentrador se genera mezclando 1.2 ml del buffer concentrador (500 mM Tris base pH=6.7 en agua desionizada), 0.75 ml de acrilamida al 40%, 100 µl de SDS 10 %, 100 µl de APS 10% y 10 µl de TEMED llevado a 10 ml con agua desionizada. Se agregó la mezcla del gel concentrador sobre el gel resolvidor y se colocaron los peines para depositar las muestras.

Una vez polimerizado el gel preparamos las muestras protéicas hirviéndolas a 98°C durante 5 minutos e incubándolas en hielo rápidamente. Después agregamos el buffer de carga (que contiene 200 mM DTT ó 5% β-mercaptoetanol) y separamos las muestras por el tiempo conveniente para resolver los tamaños de las proteínas de interés a 100 V hasta el gel resolvidor y luego a 200 V en buffer de corrida (25 mM Tris base, 190 mM Glicina, 1 % SDS en agua desionizada).

### *Inmunodetección*

La membrana de bifluoruro de polivinildeno (PVDF) se humedeció en metanol por 1 min, se enjuagó en agua destilada y se incubó por 15 min en buffer de transferencia semi-seca (48 mM Tris, 39 mM Glicina, 20% metanol, 1.3 mM SDS, pH=9.2) a 4°C. Por otro lado el gel donde se resolvieron las proteínas se colocó en la cámara de transferencia semi-seca siguiendo las instrucciones del fabricante (BioRad). La transferencia se llevó a cabo a 14 V por 1 h. Una vez transferidas las proteínas la membrana de PVDF se incubó en solución de bloqueo (5 % de leche carnation en polvo, 0.05 % Tween en TBS [0.05 M Tris, 0.15 M NaCl, pH=7.5]) por dos horas a temperatura ambiente en agitación constante.

Posteriormente se diluyó el anticuerpo a utilizar en la solución de bloqueo en la dilución correspondiente (para el anticuerpo  $\alpha$ CTCF generado en el laboratorio 1:2000, para  $\alpha$ -lamina 1:5000, para  $\alpha$ -actina 1:500) y se incubó a la membrana toda la noche con agitación constante a 4°C. Posteriormente se retiró la solución del anticuerpo y se lavó la membrana 5 veces con buffer TBS-0.05% Tween. Posteriormente se incubó el anticuerpo secundario correspondiente en dilución 1:3000 (anti-IgG de conejo ó anti-IgG de ratón acoplada a la peroxidasa obtenida de rábano) en solución de bloqueo por 1 h a temperatura ambiente en agitación constante. La membrana se lava cinco veces con solución TBS-0.05% Tween por 5 min cada vez y posteriormente se incubó con el substrato quimioluminiscente (SuperSignal, PIERCE) por 5 min a temperatura ambiente. Finalmente la membrana se expuso en placas fotográficas para rayos X revelando la señal. El mismo procedimiento se realizó para los anticuerpos control como  $\alpha$ -lamina ó  $\alpha$ -actina.

## **6.9 Ensayo de Hipersensibilidad a la DNasa I**

Se obtuvo sangre de 20 embriones en 30 ml de PBS-EDTA 2 mM. Se centrifugó la sangre por 5 min a 1400 rpm a 4°C en la centrífuga Labofuge 400R Heraeus. Se lavó con PBS-EDTA frío y se resuspendió en 10 ml de PBS sólo. Después de centrifugar por 5 min a 2000 rpm a 4°C se tomaron 3 alícuotas de 10  $\mu$ l para cuantificar el número de células. Se resuspendieron las células en un volumen de buffer NI (Tris-HCl 15 mM, pH=7.5, sacarosa 300 mM, HEPES 10 mM, KCl 60 mM, MgCl 4 mM, DTT 1 mM y NaCl 5 mM) más un volumen de buffer NI con NP-40 (Igepal) al 1% frío. Se pipetearon las células lentamente tres veces en hielo (un minuto aproximadamente). Se centrifugó 5 min a 3000 rpm a 4°C y se lavaron dos veces con 10 ml de buffer NI y se evaluó la integridad de los núcleos bajo el microscopio de luz. Se centrifugó la muestra por 5 min a 3000 rpm a 4°C y se resuspendió en buffer NI según el número de ensayos a realizar, siempre manteniendo a las muestras en hielo. Se agregó 2  $\mu$ l de CaCl<sub>2</sub> 0.3 M a cada ensayo y se llevó a 300  $\mu$ l con buffer NI (considerando el volumen que se añadirá de la enzima DNasa I). Se agregó la DNasa I (las unidades indicadas en la figuras 46 y 53) y se digirió por 5 minutos a 25°C. Se detuvo la reacción con 300  $\mu$ l de solución STOP (Tris-HCl 10 mM pH=7.5, EDTA 20 mM, SDS 0.5%) y

se añadió 10 µl de proteinasa K a cada muestra. Se incubaron los tubos por 3-6 hrs a 50°C. Posteriormente se realizaron dos extracciones fenol/cloroformo y una extracción cloroformo. Para precipitar el ADN se agregaron 0.33 volúmenes de Acetato de Amonio 10 M y un volumen de isopropanol (mezclando muy bien con vortex en cada caso, hasta conseguir fases homogéneas con el fenol). Se incubaron las muestras durante toda la noche a una temperatura de -20°C. Al otro día se centrifugaron las muestras 30 minutos a 4 °C, se lavaron con etanol al 70% dos veces y se resuspendieron en 100 µl de agua Milli Q (Millipore).

Las muestras se cuantificaron y 40 µg de ADN se digirieron con las enzimas BamHI o EcoRI-ClaI (85 Unidades) toda la noche a 37°C, se purificaron nuevamente por fenol/cloroformo como describió anteriormente y se corrieron en geles de agarosa. Se procedió a realizar un Southern blot utilizando la sonda de ADN de interés (en este caso una sonda que hibridan en el costado 3' del dominio  $\alpha$ -globina de pollo EcoRI – ClaI de 850 pb , ver Anexo I).

## 6.10 Anticuerpos

Los anticuerpos contra las histonas acetiladas ( $\alpha$ -H3ac y  $\alpha$ -H4ac) fueron obtenidos de Upstate. Dos anticuerpos contra la dimetilación de la lisina 4 de la histona 3 ( $\alpha$ -H3K4me2) fueron utilizados, uno de Upstate con el cual se realizaron la mayoría de los experimentos y uno de Abcam con el cual se realizaron los últimos ensayos de ChIPs en las células tratadas con el inhibidor 3-ABA. Los anticuerpos contra la lisina 9 trimetilada de la histona 3 y la lisina 20 trimetilada de la histona 4 ( $\alpha$ -H3K9me3 y  $\alpha$ -H4K20me3) fueron amablemente proporcionados por el Dr. Thomas Jenuwein (Instituto de Inmunología Max Plank, Friburgo, Alemania). El anticuerpo contra la lisina  $\alpha$ -79 dimetilada de la histona 3 ( $\alpha$ -H3K79me2) fue amablemente proporcionado por el Dr. Dan Gottschling (Fred Hutchinson Cancer Research Center, Seattle, EEUU). El anticuerpo de pollo contra  $\alpha$ -CTCF fue generado en el laboratorio (Valadéz-Graham *et al.*, 2004). Éste reconoce el extremo N-terminal de CTCF de pollo. El anticuerpo inespecífico utilizado para el ensayo de ChIP es

contra  $\alpha$ -IgG de conejo (Upstate). Los anticuerpos contra actina fueron amablemente proporcionados por el Dr. Alejandro Zentella (Instituto Nacional de Nutrición, México DF, México). El anticuerpo contra  $\alpha$ -lamina fue obtenido de Zymed.

## 6.11 Estadística

Análisis de ANOVA y MANOVA se llevaron a cabo según cada caso, considerando como significativo  $* = P \leq 0.05$ ,  $** = P \leq 0.01$  y  $*** = P \leq 0.001$ . Las pruebas post hoc realizadas para distinguir que componentes contribuyen a generar la varianza entre condiciones fueron Scheffe o Games-Howell para pruebas con varianzas homogéneas y no homogéneas, respectivamente. En análisis estadístico se realizó en el programa SPSS (versión 18).

## 6.12 Oligonucleótidos y plásmidos utilizados

**Tabla 1.** Todos los oligonucleótidos fueron diseñados para este proyecto a excepción de los que presentan un asterisco (\*). <sup>\*1</sup>Eria Rebollar, <sup>\*2</sup> Valadez-Graham 2004. <sup>\*3</sup> Litt *et al.*, 2001.

Nombre	Secuencia 5' a 3'	Tm	bases	inicio	Amplicón
Región de los "insulators" de 6.6.kb en el extremo 5' del dominio $\alpha$ -globina de pollo (AY016020)					
P1F <sup>*1</sup>	GGATCCGTTTTGGATACCTG	60	20	1	291 pb
P1R <sup>*1</sup>	CCTATTGTTCTGACCGAAATG	60	21	291	
P2VF <sup>*2</sup>	CAGTCTGTCAGGAAGAAAGA	58	20	2611	140pb
P2VR <sup>*2</sup>	TGTGGCCTGGGTTTTAGCTG	62	20	2731	
P3F <sup>*1</sup>	GCACCCTTCTGAGAGCTGC	62	19	3880	260pb
P3R <sup>*1</sup>	GGAAAGTGGTTGCAGAATCC	60	20	4139	
P4F <sup>*1</sup>	GCAGTCAAAGCCTCTGATGC	62	20	4739	306pb

P4R* <sup>1</sup>	CCTATTTAGATGTGATCACC	60	21	5044	
P5F* <sup>1</sup>	CCCCAGCATTAGCCTGAGC	62	19	5932	281pb
P5R* <sup>1</sup>	CTCTTGCTGCATGCTAAGG	60	20	6212	
Región del "insulator" 5' cHS4 del dominio $\beta$ -globina de pollo (L174342)					
21.505F* <sup>1</sup>	GGTTGAAGAAAAGAAGCAGGC	62	20		243pb
21.505R* <sup>1</sup>	CGTTCAGAGGAAAGCGATCC	62	20		
21.365F* <sup>3</sup>	CTCTGTGCTCAGCATCCTTCAAT	68	23		140pb
21.365R* <sup>3</sup>	CCTTTCGGCACTTTCTTCCTT	64	22		
Región de 16 kb de heterocromatina 5' del dominio $\beta$ -globina de pollo.					
H10.35F* <sup>3</sup>	GGAACAAGTTGGCAAGGTCCTAT	66	23	10338	145pb
H10.35R* <sup>3</sup>	TCTTCTGCCCTGCCCGTAT	60	19	10482	
H15aF* <sup>1</sup>	GGCTCCCTGTGGGCTGC	60	17	15813	157pb
H15aR* <sup>1</sup>	CCTGGGGCTGCACTGGG	60	17	15970	
H15bF* <sup>1</sup>	GCAAAGTCATTGCCTGGTGC	62	20	16752	185pb
H15bR* <sup>1</sup>	GCATTTGGTTAAAATGTTTATGC	60	23	16936	
EGFPF1	ACATGAAGCAGCACGACTTC				360 pb
EGFPR1	TGCTCAGGTAGTGGTGTGC				
M9 $\Delta$ CAG F	CCTAGAGTTAATCTGGGCTTTTTCATAAGC	55	30		
M9 $\Delta$ CAG R	GCTTATGAAAAAGCCAGATTAAGTCTAGG	55	30		
M9 $\Delta$ CTAG F	GTAAATCTTTCACTTAGTTAAGCAGGGC	55	29		
M9 $\Delta$ CTAG R	GCCCTGCTTAATAAGTGAAGAATTTAC	55	29		
1.4 Xba I F	CCGGGGtCTAGATGTTTGCACCTCTAA	55	27		1400pb
1.4 Xba I R	CCGGGGTCTAGATTTTCATGAAGTT	55	24		
Afl II FOR (M9 T 3')	GCATTCTAGTTGTGGTTTGTCC	61.5	22		~380pb
FtrCTCF1.4 F (M9 t 5')	GCAGCCCACCTCAAAAAACG	61.5	20		

FtrCTCF1.4 R (M9 t 5')	GCTCTGCCAGTTGAGTTGAC	60	20		
USF en 1.4 F	GTCTCAACACAAATGGCAGC	60	20		342pb
USF en 1.4 R	CTGCTCCTTCCTGATTCTG	60	20		
$\alpha$ 1.4-730 H-S F	CCGGGGAAGCTTGTCGACCAGAAATCAGGAAGG	58	33		730 pb
$\alpha$ 1.4-730 Afl II F	CCGGGTGCTTAAGTCAGAAATCAGGAAGG	58	29		730 pb
$\alpha$ 1.4-650 H-S R	CCGGGGAAGCTTGTCGAC CCTTCCTGATTCTG	58	33		650 pb
$\alpha$ 1.4-650 Afl II R	CCGGGTGCTTAAGT CCTTCCTGATTCTG	58	29		650 pb
$\alpha$ MRE Xho F	ACTGACTGATCGCTCGAGCGGTTCCCTTAC ATGCTGA	58	37	25387	450 pb
$\alpha$ MRE Xho R	ACTGACTGATCGCTCGAGAGCAGGCAAAGTCCCAA AAGC	58	39	25813	
Oligonucleótidos para generar shRNAi contra CTCF					
HpcCTCF1	CCGAGAGACGAGGCTTGCCATATTTCA AGAGAATATGGCAAGCCTCGTCCTCC TTTTTTAAGCTTCTCGAG	90	71		
HpcCTCF2	CCCAGCGGTGTCCAGATGGTGATTCAA GAGAATCACCATCTGGACACCGCTG TTTTTTAAGCTTCTCGAG	92	71		
HpcCTCF3 F	CCGACGTGGATGTCTCTGTGTATTTT AAGAGAATACACAGAGACATCCACGTC TTTTTTAAGCTTCTCGAG	87	71		
FillcCTCF	TCGATCGATCGACTCGAGAAGCTTAAA	72	27		
pSilLefevre 1 F	GATCCGCCAGTGTGGAGGTTAGCAAATTG ATTCAAGAGACGGTCACACCTCCAATCGTTT	89.5	75		

	AACTTTTTTTGGAAA				
pSilLefevre 1 R	AGCTTTTCCAAAAAAGTTAAACGATTGGAG GTGTGACCGTCTCTTGAATCAATTTGCTAACC TCCACACTGGCG	89.5	79		
pSilLefevre 2 F	GATCCGCAGTGTGGAGGTTAGCAAATTG AAATTCAAGAGAGTCACACCTCCAATCG TTAACTTTTTTTTTGGAAA	87	76		
pSilLefevre 2 R	AGCTTTTCCAAAAAAGTTAAACGATT GGAGGTGTGACTCTCTTGAATTTCAATTT GCTAACCTCCACACTGCG	87	76		
pSilLefevre 3 F	GATCCGCCAGCAGGGATACTTACAACTGA ATTCAGAGAGGTCGTCCCTATGAATGTTTG ACTTTTTTTTTGGAAA	87	76		
pSilLefevre 3 R	AGCTTTTCCAAAAAAGTCAAACATTCATAG GGACGACCTCTCTTGAATTCAGTTTGTAAGTA TCCCTGCTGGCG	87	76		
pCT1control a (scramble)	GATCC GCGGTAGGACTCTGATTGA TTCAAGAGA TCAATCAGAGTCCTACCGCTTTTTTTGGAAA	87	63		
pCT1control b (scramble)	AGCTTTTCCAAAAA GCGGTAGGACTCTGATTGA TCTCTTGAA TCAATCAGAGTCCTACCGCG	87	63		
Oligonucleótidos para secuenciar los shRNAs in plásmidos pSil 3.1 y 2.1 Hygro y Puro					
M13F (-40)	GTTTTCCAGTCACGAC	56	17		
3.1 Hygro	GAGTTAGCTCACTCATTAGGC	57	21		
2.1 Hygro	AGGCGATTAAGTTGG	56	16		
Región 3' del dominio $\alpha$ -globina de pollo (AY016020)					
2089 Hindi III F	GCATGAAGCTTCCAGTTGCTCTCACACCTTGC	66	32	55245	2089pb
2089 Hindi III R	GCATGAAGCTTGGTTCTGCCATGGGTATCTCC	66	32	57312	

2089 BamHI F	GCGGATCCCAGTTGCTCTCACACCTTGC	66	28	55245	2089 pb
2089 BamHI R	GCGGATCCGGTTCTGCCATGGGTATCTCC	66	28	57312	
FRAG1 F (del 2089pb)	GCGTGGGATCCCAGTTGCTCTCACACCTTGC	66	31	55245	1036 pb
FRAG1 R	GCGTGGGATCCGCAGTTCAGGCTGACATTTAGG	66	36	56276	
FRAG2 F (del 2089pb)	GCGTGGGATCCTAAATGTCAGCCTGAACTGC	66	31	56276	1053 pb
FRAG2 R	GCGTGGGATCCGGTTCTGCCATGGGTATCTCC	66	32	57312	
CPG 3' F	GCTCTCAGATTTTGGATCCAT	60	21	58134	2211 pb
CPG 3' R	GGAAGTGGATCCGCACAAAGCTCAGCAGCTC	60	19	60345	
TMEM-6 (422) F	CGTGTGTTTCCATCCCTTGG	62	20	65601	422 pb
TMEM-6 (422) R	CCAGAACAGGCATATCAGTGA	62	21	65092	
Oligonucleótidos para amplificar la Histona H3.3					
H3.3-YFP F (BamHI)	CGCGGATCCACTTAATTACCAGGTAAACG	62	22		1200pb
H3.3-YFP R (Sal I)	GGGGTCGACCTTAATCAGATCTTGTGCG	62	21		
H3.3 R (Sal I)	GGGGTCGACTAGGCCCGCTGCCACGG	68	19		450pb
H3.3 FOR (Hind III)	CGAAGCTTGTCCACTTAATTACCAGGTAAACG	62	22		450pb
H3.3 REV	GGGGATCCTGTGAGCCCGCTGCCACGG	68	19		

**Tabla 2. Pares de oligonucleótidos utilizados en las PCRs dúplex**

Oligonucleotido experimental/ oligonucleótido control
P1 / H10.35
P2V / H15b
P3 / H10.35
P4 / H10.35
P5 / H10.35
EGFP/H10.35
EGFP/21.505
(Afl II for + P2V R) M9t 3'/H10.35
M9t 5'/ H10.35

**Tabla 3. Plásmidos**

Plásmidos	Descripción
pG $\alpha^D$ 3	pEGFP-1 con el promotor $\alpha$ D en el sitio BamHI y una región de clonación múltiple en Afl II (Héctor Rincón)
pGC <sub>2</sub> GC <sub>2</sub>	pG $\alpha^D$ con dos copias del <i>core</i> del insulator 5' cHS4 a cada lado del gen reportero (Héctor Rincón)
pG $\alpha^D$ 1.2+1.2	pG $\alpha^D$ con el elemento $\alpha$ CHS-1.2 a cada lado del gen reportero (Viviana Valadez)
pG $\alpha^D$ 1.4 en Afl II	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 en Afl II
pG $\alpha^D$ 1.4 en Hind III	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 en Hind III
pG $\alpha^D$ 1.4 en Kpn-Hind III	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 en Kpn I- Hind III

pG $\alpha^D$ 1.4+1.4	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 en ambos lados del gen reportero
pG $\alpha^D$ 1.4 $\Delta$ CAG	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 mutante $\Delta$ CAG
pG $\alpha^D$ 1.4 $\Delta$ CTAG	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 mutante $\Delta$ CTAG
pG $\alpha^D$ 1.4 $\Delta\Delta$	pG $\alpha^D$ con el elemento $\alpha$ EHS-1.4 con la doble mutación $\Delta$ CAG $\Delta$ CTAG
pG $\alpha^D$ MRE	pG $\alpha^D$ con el elemento $\alpha$ MRE
pG $\alpha^D$ 1.4MRE	pG $\alpha^D$ con el elemento $\alpha$ MRE y el elemento $\alpha$ EHS-1.4
pGEM-Teasy 1.4 Xbal	pGEM-Teasy con el elemento $\alpha$ EHS-1.4 en Xbal
pTOPO 1.4 $\Delta\Delta$ Afl II	TOPO con el elemento $\alpha$ EHS-1.4 $\Delta\Delta$ en Afl II
pDrive 1.4	pDrive con el elemento $\alpha$ EHS-1.4
pDrive 1.4 $\Delta\Delta$	pDrive con el elemento $\alpha$ EHS-1.4 $\Delta\Delta$
pTyr5	Plásmido con el mini-gen que codifica para la <i>tirosinasa</i> (TyR) (Dr. Montoliu)
pTyr5- $\alpha$ EHS-1.4	Plásmido con el mini-gen que codifica para la <i>tirosinasa</i> y el elemento $\alpha$ EHS-1.4
pL $\alpha^D$	pGL3Basic (plásmido que contiene al gen de la <i>luciferasa</i> ) con el promotor $\alpha^D$
pL $\alpha^D$ 2089	pL $\alpha^D$ con la región de 2089pb que comprende a los 3 sitios de hipersensibilidad de de la región 3' del dominio
pL $\alpha^D$ F1	pL $\alpha^D$ con el fragmento F1 de la región de 2089pb
pL $\alpha^D$ F2	pL $\alpha^D$ con el fragmento F2 de la región de 2089pb
pL $\alpha^D$ 4232	pL $\alpha^D$ con el fragmento de 4232pb que comprende al $\alpha$ MRE
pL $\alpha^D$ 1.4	pL $\alpha^D$ con el elemento $\alpha$ EHS-1.4
pL3B 2089	pGL3B sin promotor y con la región de 2089pb
pL3B F1	pGL3B sin promotor y con el fragmento F1 de la región de 2089pb
pL3B F2	pGL3B sin promotor y con el fragmento F2 de la región de 2089pb
pSilCTCF 1	pSilencer-3.1 Hygro con el hairpin # 1 contra CTCF de pollo
pSilCTCF 2	pSilencer-3.1 Hygro con el hairpin # 2 contra CTCF de pollo
pSilCTCF 3	pSilencer-3.1 Hygro con el hairpin # 3 contra CTCF de pollo

pSiLefevre 1	pSilencerPURO-old con el hairpin # 1 diseñado en base a Lefevre <i>et al.</i> contra CTCF de pollo
pSiLefevre 2	pSilencerPURO-old con el hairpin # 2 diseñado en base a Lefevre <i>et al.</i> contra CTCF de pollo
pSiLefevre 3	pSilencerPURO-old con el hairpin # 3 diseñado en base a Lefevre <i>et al.</i> contra CTCF de pollo
pCT1	pSilencerPURO-old con un hairpin contra CTCF de humano (Dr. Ishihara)
pScramble	pSilencerPURO-old con un hairpin diseñado con el contenido nucleotídico revuelto del pCT1
pSilPURO vacío	pSilencerPURO-old con la región BamHI-HindIII del sitio de clonación múltiple del pEGFP-1
pSilHygro vacío	pSilencer-3.1 Hygro con la región BamHI-HindIII del sitio de clonación múltiple del pEGFP-1
pSilencer-USF 1 (136)	pSilencer-3.1 Hygro con hairpins contra USF-1 # 1 (Dr. West)
pSilencer-USF 1 (142)	pSilencer-3.1 Hygro con hairpins contra USF-1 # 2 (Dr. West)
pSilencer-USF 1 (472)	pSilencer-3.1 Hygro con hairpins contra USF-1 # 3 (Dr. West)
pSilencer-USF 1 (639)	pSilencer-3.1 Hygro con hairpins contra USF-1 # 4 (Dr. West)
pSilencer-USF 1 (853)	pSilencer-3.1 Hygro con hairpins contra USF-1 # 5 (Dr. West)
pcDNA-hH3.1	pcDNA con la histona H3.1 humana (Dr. Henikoff)
pcDNA-hH3.3	pcDNA con la histona H3.3 humana (Dr. Henikoff)
pcDNA-hH3.1-FLAG	pcDNA con la histona H3.1 humana fusionada a FLAG-HA (Dr. Henikoff)
pcDNA-hH3.3-FLAG	pcDNA con la histona H3.3 humana fusionada a FLAG-HA (Dr. Henikoff)
pK399	Plásmido con la histona H3.3 de <i>Drosophila</i> fusionada a YFP (Dr. Ahmad)
pGEMH3.3 stop	pGEM-11Z con el ORF de la histona H3.3 de <i>Drosophila</i> con codón de STOP

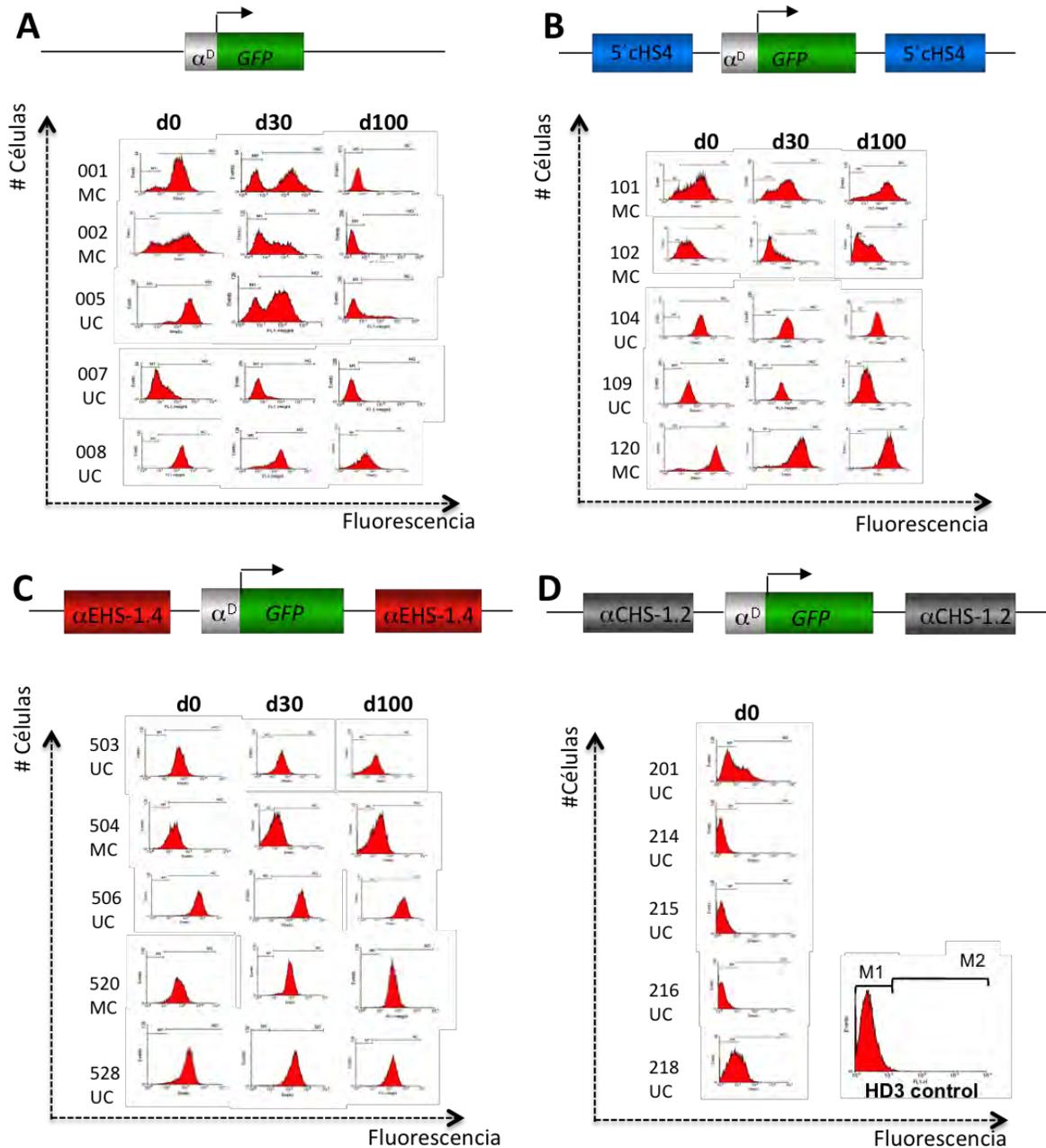
## 7. RESULTADOS y DISCUSIÓN

Para llevar a cabo la caracterización de la función de los elementos  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2 se comenzó por estudiar si éstos son capaces de proteger a un transgen de los efectos de posición cromosomal (EPC).

### 7.1 El *insulator* $\alpha$ EHS-1.4 protege a un transgén contra EPC en células eritroblásticas de pollo

Para realizar el ensayo de protección contra EPC se transfectaron establemente diferentes plásmidos descritos a continuación y enlistados en la Tabla 3 (ver métodos). Todos los plásmidos contienen al gen reportero de la *EGFP* bajo el promotor  $\alpha^D$  (gen adulto del dominio  $\alpha$ -globina de pollo, ver Fig. 21). Éste promotor tiene una actividad moderada en las células HD3 lo cual posibilita medir los efectos de posición cromosomal de manera indirecta a través de cuantificar los niveles de fluorescencia emitidos por el gen reportero *EGFP*. El plásmido pG $\alpha^D$ 3 es el control sin *insulators*, es decir, que las clonas estables obtenidas con este transgén se encuentran sujetas a efectos de posición. El plásmido pC2GC2 posee al “core” del *insulator* 5' CHS4 a cada lado del gen reportero. Como se ha mencionado previamente, sabemos que este *insulator* protege contra los efectos de posición cromosomal por lo cual lo utilizamos como un control positivo. Finalmente, los plásmidos pG $\alpha^D$ 1.4+1.4 y pG $\alpha^D$ 1.2+1.2 contienen a los elementos a evaluar,  $\alpha$ EHS-1.4 y  $\alpha$ CHS-1.2, enmarcando al gen reportero respectivamente.

Una vez realizadas las transfecciones estables, se aislaron clonas individuales y la integridad del transgén y el número de copias se corroboró por southern blot utilizando una sonda contra el gen de *EGFP* (no mostrado). Las transfecciones se realizaron por triplicado para cada uno de los casos y todas las líneas con el transgén íntegro fueron consideradas para el análisis.



**Figura 23. Evaluación de la protección contra EPC.** Se muestra el nivel de fluorescencia medido por citometría de flujo (FACS) al día 0, 30 y 100 de cultivo continuo para clonas representativas de las líneas estables para los diferentes plásmidos. **A.** Líneas estables que contienen al plásmido p $\alpha^D$ 3 sin *insulators*. A los 100 días en cultivo continuo en ausencia de selección, la expresión de la EGFP se ha silenciado prácticamente por completo en la mayoría de las líneas celulares. **B.** Líneas estables que contienen al plásmido pC2GC2 con el *insulator* 5' cHS4 delimitando al gen reportero. La fluorescencia en éstas líneas se mantiene relativamente estable durante el tiempo de cultivo. **C.** Líneas celulares con el plásmido p $\alpha$ D1.4+1.4 que contienen al elemento  $\alpha$ EHS-1.4 del dominio  $\alpha$ -globina de pollo. Los niveles de fluorescencia en las diferentes líneas estables generadas se mantienen muy estables en el tiempo de cultivo. **D.** Líneas celulares que contienen plásmido p $\alpha$ D1.2+1.2 en donde el gen reportero se encuentra delimitado por el elemento  $\alpha$ CHS-1.2 del

dominio  $\alpha$ -globina. Al día 0 de cultivo la mayoría de las líneas analizadas con este plásmido presentan un perfil en donde la fluorescencia es nula y muy semejante al control de células HD3 sin transfectar. En el costado se muestra el nombre de la línea y si contiene al transgén en copia única (UC) o en multicopia (MC). Se muestran 5 clonas representativas en cada caso. M1 marca la población GFP<sup>-</sup> y M2 la GFP<sup>+</sup> (ejemplificado en el histograma de las células HD3 control, panel D).

El seguimiento del nivel de fluorescencia por citometría de flujo de las líneas estables del plásmido pG $\alpha^D3$  muestra como la expresión del reportero *EGFP* se ha extinguido en la mayoría de las líneas evaluadas para los 80-100 días de cultivo continuo (Fig. 23A). Existen algunas clonas como la 008 que a los 100 días de cultivo todavía muestran un perfil positivo de expresión. Esto nos muestra como este transgén se encuentra sujeto a los efectos de EPC de tal manera que, existen sitios de integración favorables en donde la expresión se mantiene y sitios en donde, aún cuando inicialmente la expresión es elevada, el entorno cromatínico influye y afecta la expresión del gen reportero comportamiento que ocurre en la mayoría de las clonas estables generadas.

Si analizamos la expresión de las clonas estables que contienen al plásmido pC2GC2 durante los días de cultivo continuo, podemos observar como el perfil de expresión al día 0 del cultivo se mantiene relativamente homogénea después de 100 días de cultivo continuo (Fig. 23B). Al comparar la expresión de las líneas anteriores podemos notar como el nivel de fluorescencia es más estable en las líneas que tienen al *insulator* 5' *CHS4* enmarcando al gen reportero que las que solo poseen al promotor  $\alpha^D$ , lo cual, es el comportamiento esperado para estas líneas que son las controles. Así mismo, si observamos en la cuantificación de los datos la desviación estándar, podemos concluir que las líneas del plásmido pG $\alpha^D3$  presentan mayor variabilidad con respecto a las líneas del pC2GC2 (Fig. 24). Este fenómeno es indicativo de la susceptibilidad que presenta la expresión de la *GFP* dependiendo del contexto cromosómico de integración y como ésta disminuye al incluir al "core" del *insulator* 5' *CHS4*.

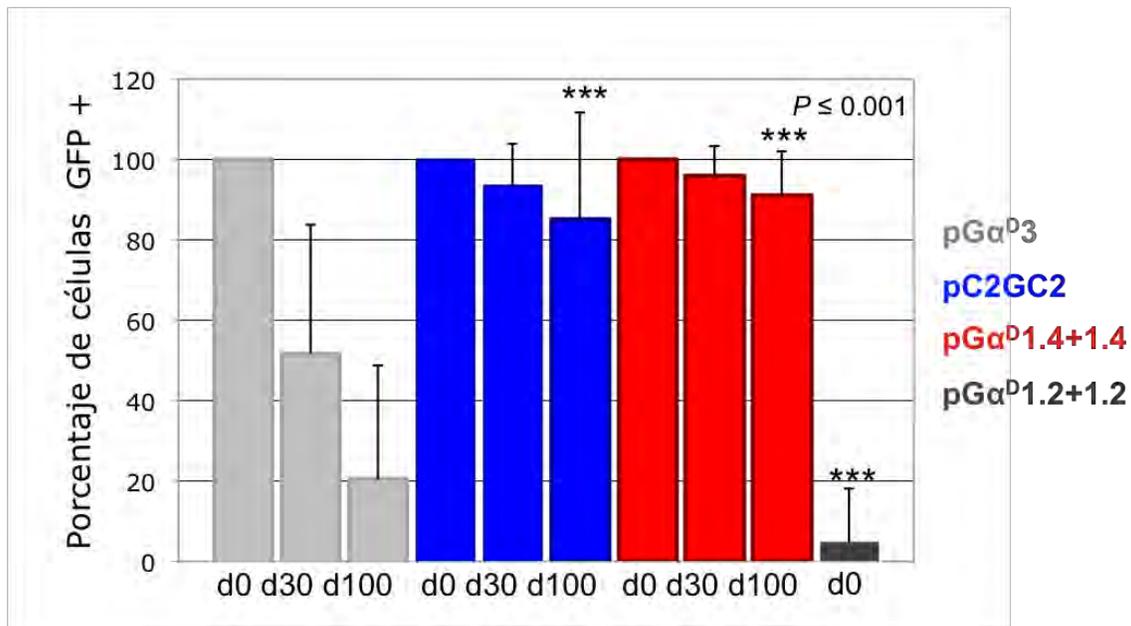
Resulta interesante notar que estas líneas además de representar un sistema control para evaluar a los *insulators* de la región 5' del dominio  $\alpha$ -globina de pollo, muestran cómo el *insulator* 5' *CHS4* del dominio  $\beta$ -globina funciona protegiendo contra EPC en un sistema diferente al utilizado originalmente y bajo un promotor endógeno el

cual corresponde al gen adulto  $\alpha^D$  (Pikaart *et al.*, 1998). Por otro lado, los perfiles de fluorescencia en cada línea son diferentes siendo unos más variegados que otros dependiendo del sitio de inserción en el genoma y del número de copias insertadas. Sin embargo, dicha heterogeneidad no impide observar la acción del *insulator* 5' cHS4 el cual consigue mantener en buena medida el perfil inicial de expresión en cada línea.

Al analizar los histogramas de citometría obtenidos para las clonas con el plásmido pG $\alpha^D$ 1.4+1.4 podemos observar que las líneas muestran un perfil homogéneo de expresión al día cero con una curva poblacional bien definida y con un nivel de fluorescencia elevado (Fig. 23C). Por otro lado, notamos que la expresión del gen reportero se mantiene prácticamente invariable a los 100 días de cultivo continuo lo cual muestra que el elemento  $\alpha$ EHS-1.4 es un *insulator* con la capacidad de proteger a un transgén contra los efectos de EPC.

Finalmente, si observamos los resultados obtenidos con el plásmido pG $\alpha$ D1.2+1.2 podemos constatar que al día cero, las clonas presentan un perfil de fluorescencia muy reducido y similar al observado para el control de fluorescencia (células HD3 sin transfectar) (Fig. 23D). Dicho resultado nos muestra que el elemento  $\alpha$ CHS-1.2 tiene una actividad represiva en un contexto integrado sobre el promotor  $\alpha^D$  ya que, a diferencia del comportamiento observado sin *insulators* (comparar Fig. 23A y D), al día cero la gran mayoría de las clonas se encuentran silenciadas.

Al obtener los promedios del porcentaje de células GFP<sup>+</sup> en distintos días de cultivo continuo, las líneas correspondientes al plásmido pG $\alpha^D$ 1.4+1.4 después de 100 días de cultivo resulta de 91%, mientras que las líneas del plásmido pG $\alpha^D$ 1.2+1.2 se encuentra por debajo del 10% desde el día 0 de cultivo (Fig. 24). Además, pudimos observar como la desviación estándar en el caso de las líneas del pG $\alpha^D$ 1.4+1.4 es pequeña reflejando la homogeneidad en el comportamiento de cada clona.

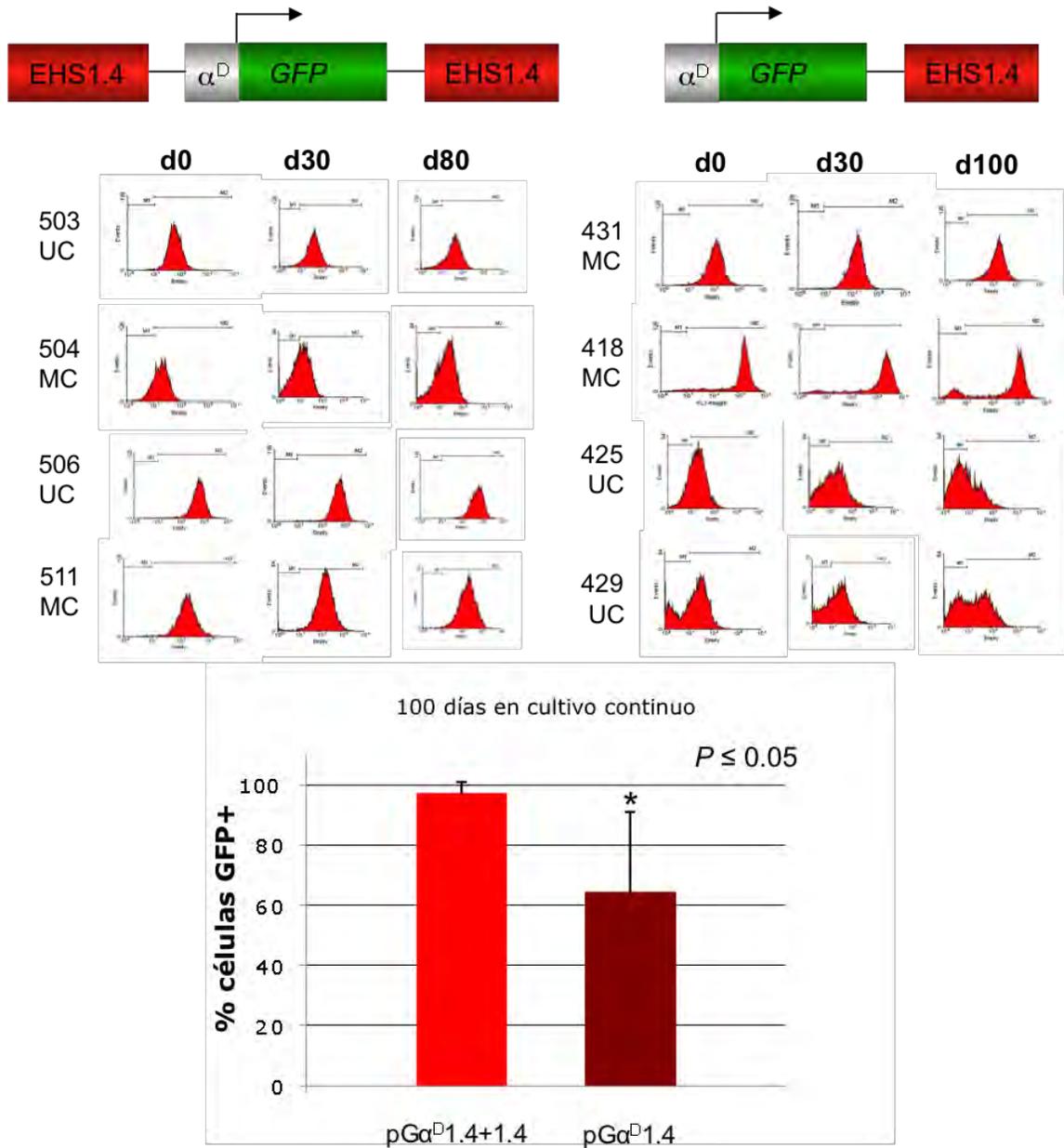


**Figura 24. Cuantificación del porcentaje de células GFP<sup>+</sup> a lo largo del cultivo continuo sin selección.** Las medidas representan el promedio +/- la desviación estándar del número de clones pGα<sup>D3</sup> n=15, pC2GC2 n=10, pGα<sup>D1.4+1.4</sup> n=10 y pGα<sup>D1.2+1.2</sup> n=17 para cada plásmido. El *insulator* αEHS-1.4 protege de manera muy eficiente contra EPC mientras que el elemento αCHS-1.2 tiene un efecto negativo sobre la expresión del transgén desde el día 0 del cultivo.

La homogeneidad en el perfil de expresión inicial en las líneas del pGα<sup>D1.4+1.4</sup> es mayor con respecto a lo que se observa para las líneas del “core” del *insulator* 5' cHS4 (comparar histogramas B y C en Fig. 23). Dicho fenómeno puede deberse a diversos factores. Por un lado, el “core” del *insulator* 5' cHS4 comprende alrededor de 250 pb que se definieron como el elemento mínimo funcional de dicho *insulator* en el ensayo de protección (Recillas-Targa *et al.*, 1998; Pikaart *et al.*, 1999; Bell *et al.*, 1999). Sin embargo es posible que en el contexto utilizado, este elemento no sea suficiente para alcanzar una expresión inicial homogénea y robusta del transgén, es decir, que el “core” del *insulator*, aun cuando logra mantener la expresión en el tiempo, no consigue contender con el contexto cromatínico asociado al sitio de integración en una primera instancia dada la eliminación de secuencias de unión a factores que podrían ser importantes para dicha función. Otra posibilidad tiene que ver con que el “core” del *insulator* 5' cHS4 fue evaluado inicialmente en presencia de un *enhancer* y promotor correspondientes al dominio β-globina al cual pertenece dicho *insulator*, a diferencia de nuestro sistema en donde el promotor utilizado forma parte del dominio α-globina (Pikaart *et al.*, 1998;

Recillas-Targa *et al.*, 2002). Esto hablaría de una actividad más eficiente del *insulator* al interactuar con elementos de su contexto genómico endógeno. Por último, es factible que el elemento  $\alpha$ EHS-1.4 incluya sitios de unión a factores nucleares diferentes a los del “core” del *insulator* 5' cHS4 y que de esta forma tenga mejor capacidad de contender con los efectos de posición cromosomal.

Además, realizamos un análisis comparando la capacidad de protección contra EPC por parte del *insulator*  $\alpha$ EHS-1.4 cuando se localiza sólo a un lado del gen reportero *EGFP* para lo cual utilizamos el plásmido pG $\alpha^D$ 1.4(Afl II) que contiene al *insulator* únicamente en el costado 3' del gen reportero. Posteriormente se realizó la transfección con un plásmido que incorpora al *insulator* sólo en el costado 5' obteniendo resultados similares (ver más adelante, figura 32). Como puede observarse en la figura 25 , al comparar la habilidad de proteger contra EPC por parte del  $\alpha$ EHS-1.4 concluimos que éste es más eficiente cuando se sitúa enmarcando al gen reportero. Este resultado sugiere que la presencia de *insulators* a cada lado favorece una conformación de la cromatina más permisiva que cuando sólo se encuentra en un costado. Esto se hace más evidente cuando se compara la integración múltiple (MC) en relación a las integraciones sencillas (UC) del transgén pG $\alpha^D$ 1.4 (solo en un costado). Al tener varias integraciones (que normalmente ocurren en tándem), el efecto de delimitar al gen reportero se reproduce y la protección es más eficiente (Fig. 24 y 25).



**Figura 25. Evaluación de la capacidad de contener contra EPC por parte del  $\alpha$ -EHS-1.4 colocado delimitando al gen reportero *EGFP* por uno o ambos costados.** En el panel de la izquierda se muestran las líneas celulares establecidas con el plásmido pGa<sup>D</sup>1.4+1.4 y en el de la derecha las líneas generadas con el plásmido pGa<sup>D</sup>1.4 (Afl II). Abajo se muestra la cuantificación del promedio +/- la desviación estándar del porcentaje de células GFP<sup>+</sup> en cada caso. La protección es más eficiente cuando el *insulator* delimita al gen reportero. pGa<sup>D</sup>1.4+1.4 (n=10) y pGa<sup>D</sup>1.4 (Afl II) (n=16).

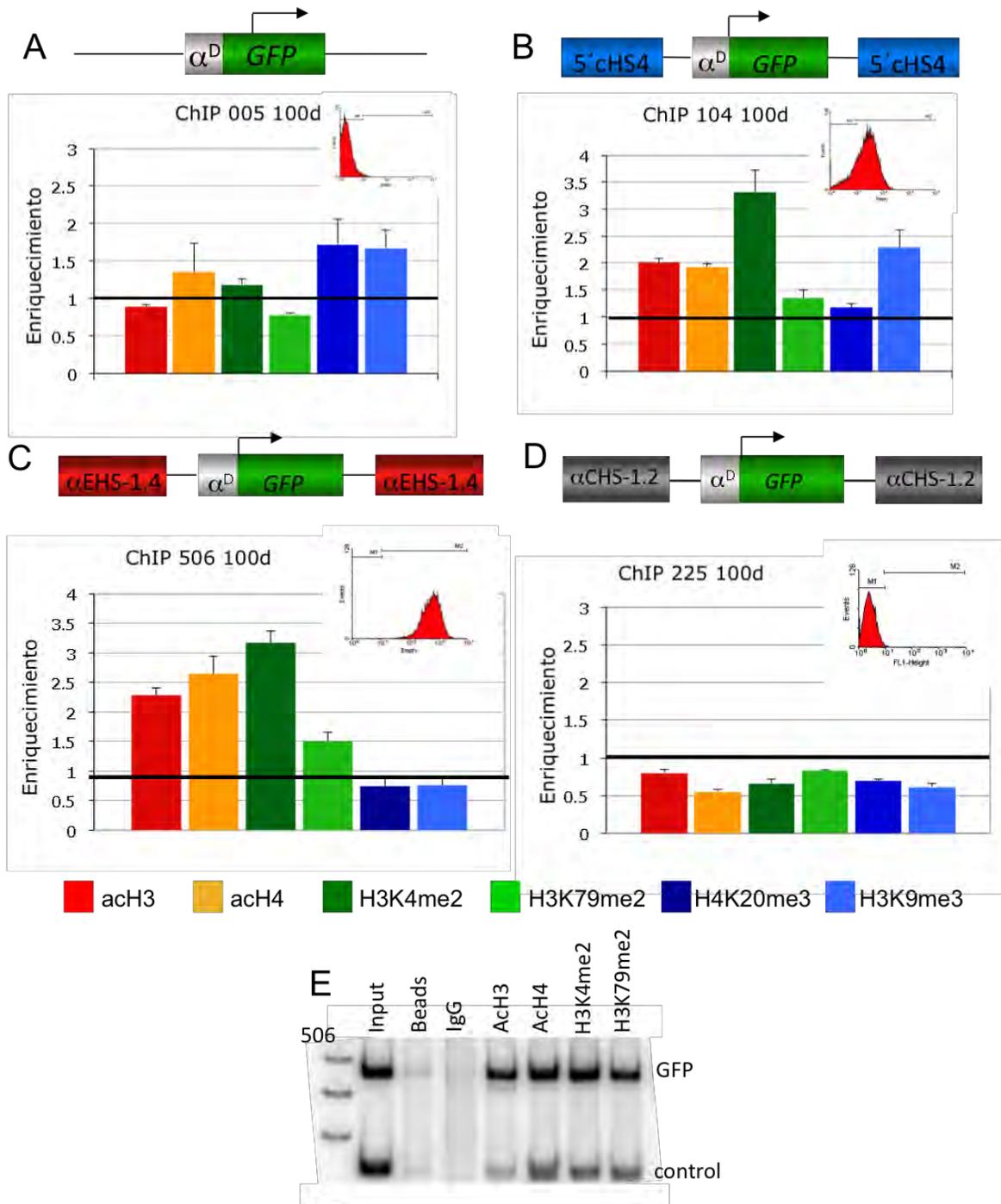
## 7.2 Los transgenes protegidos con *insulators* presentan un perfil permisivo de marcas post-traduccionales de histonas

A continuación, decidimos analizar por medio de inmunoprecipitación de cromatina (ChIP) el estado de la cromatina sobre el transgén en las clonas estables generadas, para lo cual, utilizamos anticuerpos específicos para modificaciones post-traduccionales de histonas. Como marcas de cromatina relajada y permisiva utilizamos acH3, acH4, H3K4me2 y H3K79me2. Como marcas de cromatina no permisiva utilizamos H4K20me3 y H3K9me3 (ver métodos, sección 6.5).

Se seleccionaron líneas con integración única representativas de cada uno de los plásmidos a los 100 días de cultivo para llevar a cabo el ensayo de ChIP. La cuantificación de los datos obtenidos se realizó a través de PCR dúplex utilizando dos regiones control; una región de heterocromatina ubicada en el costado 5' del dominio  $\beta$ -globina de pollo para las marcas de cromatina permisiva, y una región abierta en la región de unión a CTCF del *insulator* 5' cHS4 para las marcas correspondientes a cromatina compacta (marcas inversas a las que se desean evaluar). La obtención del valor de enriquecimiento de cierta marca se llevó a cabo haciendo una razón entre la señal del anticuerpo específico en la región de interés versus la control y la señal obtenida con el anticuerpo inespecífico en la región de interés versus la región control (ver métodos, sección 6.5).

Al realizar los ensayos de ChIP sobre el transgén en la línea 005 pudimos detectar la ausencia de marcas de cromatina abierta, así como un enriquecimiento moderado en las marcas de cromatina cerrada (Fig. 26A). Este comportamiento coincide con el bajo perfil de expresión de dicha clona a los 100 días de cultivo. La pérdida de marcas post-traduccionales de las histonas que denotan una cromatina relajada y la adquisición de marcas de cromatina más compacta, demuestran como el transgén ha sido afectado por los efectos de posición cromosomales y de esta forma, el silenciamiento de la expresión del gen reportero está acompañada por cambios en la conformación local de la cromatina.

Dicho fenómeno es el esperado para un transgén desprotegido y es consistente con los reportes previos (Pikaart *et al.*, 1998).



**Figura 26. Ensayo de CHIP evaluando al transgén en las distintas líneas estables. A.** Gráfica de enriquecimiento de ciertas marcas post-traduccionales de histonas sobre el transgén en la línea 005 que contiene el plásmido p $\alpha^D$ 3 a los 100 días de cultivo, sin *insulators*. **B.** Gráfica de

enriquecimiento de marcas de histonas en la línea 104 que contiene al plásmido pC2GC2 que incorpora al *insulator* 5' cHS4 a los 100 días de cultivo. **C.** Perfil de enriquecimiento de marcas de histonas sobre el transgén de la línea 506 que contiene al plásmido pG $\alpha^D$ 1.4+1.4. **D.** Gráfico de enriquecimiento de marcas post-traduccionales de histonas sobre el transgén en la línea 225 que tiene integrado al plásmido pG $\alpha^D$ 1.2+1.2. **E.** Gel representativo de amplificación de la región experimental (sobre el gen reportero EGFP) y la región control (región de heterocromatina del dominio  $\beta$ -globina de pollo). Cada reacción fue realizada por triplicado, en las gráficas se muestra el promedio +/- el error estándar.

Si observamos el perfil de marcas post-traduccionales de las histonas para la línea que contiene al *insulator* 5' cHS4 (línea 104) a los 100 días de cultivo, podemos notar como en este caso, el transgén incorpora marcas de cromatina abierta, en particular la H3K4me2 (Fig. 26B). Para el caso de las marcas de cromatina cerrada, podemos observar como no hay presencia significativa de H4K20me3, sin embargo existe un enriquecimiento en la H3K9me3.

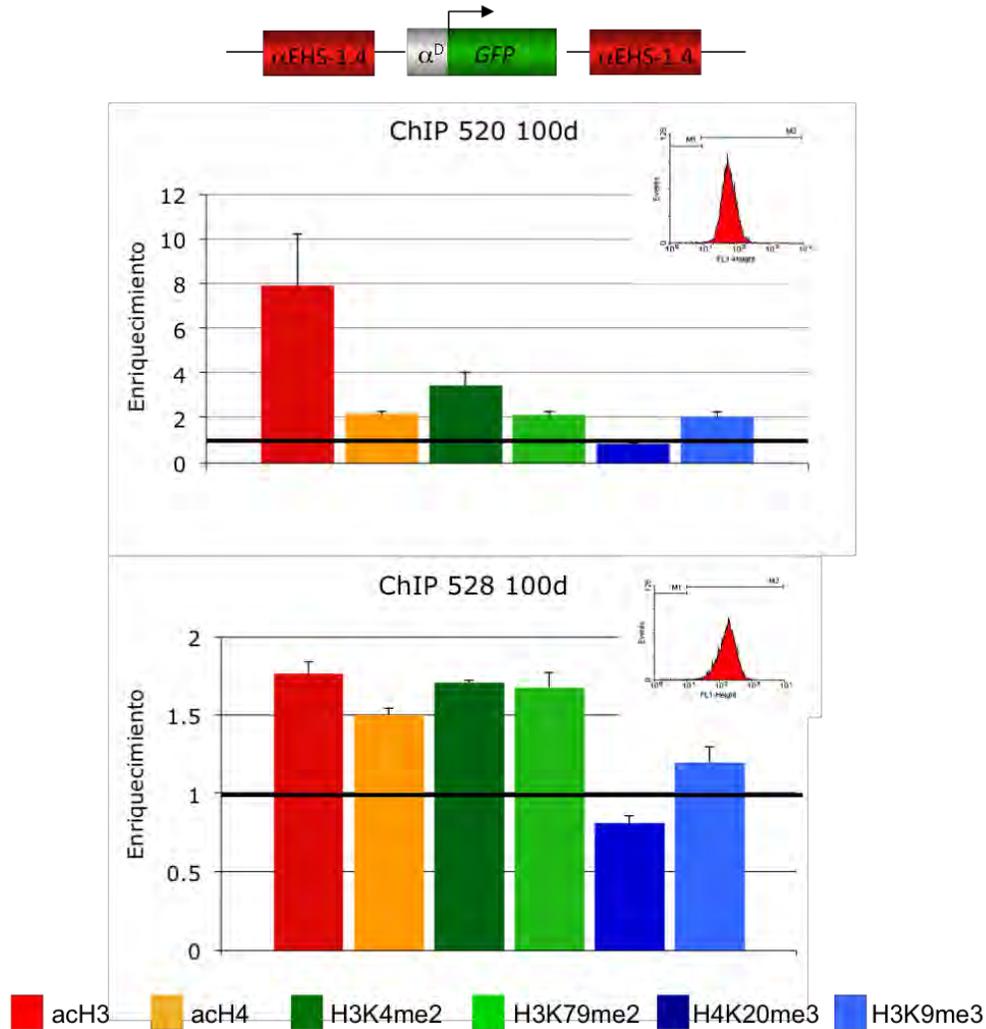
El resultado que se obtuvo para la línea 104 es el esperado ya que ha sido reportado que la protección brindada por el *insulator* 5' cHS4 del dominio  $\beta$ -globina de pollo, conlleva el enriquecimiento de marcas post-traduccionales de histonas correspondientes a cromatina permisiva. Es interesante recalcar la presencia de la H3K9me3 en esta línea, la cual no se observa enriquecida en las líneas con el *insulator*  $\alpha$ EHS-1.4 (Fig. 26C).

La línea 225 que contiene al elemento  $\alpha$ CHS-1.2 no presenta ningún enriquecimiento significativo de las marcas post-traduccionales de histonas analizadas (Fig. 26D). Esto coincide con la baja expresión de la EGFP en dicha línea desde el día 0 de cultivo. Cabe señalar que nos resulta difícil interpretar este resultado sin embargo, la falta de marcas puede significar la ausencia de nucleosomas en la región, o una disminución generalizada en las marcas de las histonas. En el laboratorio hemos observado en otras ocasiones la pérdida de expresión ligada a la disminución de marcas de apertura sin un enriquecimiento significativo en marcas represivas (De La Rosa *et al.*, 2007).

Por último, la línea 506 que contiene al transgén delimitado por el *insulator*  $\alpha$ EHS-1.4 muestra un enriquecimiento pronunciado en marcas de cromatina abierta mientras

que, a diferencia de lo observado para el *insulator* 5'5' cHS4, no encontramos un enriquecimiento de marcas correspondientes a cromatina más compacta. El resultado anterior nos muestra como el *insulator*  $\alpha$ EHS-1.4 del dominio  $\alpha$ -globina de pollo tiene la capacidad de mantener la expresión de un transgén a lo largo del tiempo (>100 días) y contener con los efectos de posición cromosomal conservando un perfil de apertura de la cromatina. Un rasgo adicional del *insulator*  $\alpha$ EHS-1.4 se desprende de experimentos de inmunoprecipitación de cromatina (ChIP) que se han llevado a cabo en el laboratorio (datos de Eria Rebollar, tesis de maestría) en donde se muestra que en esta región en el dominio endógeno, existe un enriquecimiento notable de marcas de cromatina abierta en la línea de eritroblastos HD3 de manera semejante a lo observado en las clonas estables. Este rasgo es consistente con la presencia de un elemento *insulator* capaz de contener con los efectos de posición cromosomal (Litt *et al.*, 2001; Recillas-Targa *et al.*, 2002). Por otro lado sobre el elemento  $\alpha$ CHS-1.2 endógeno en células HD3, se observa la ausencia de marcas de cromatina abierta y esto coincide con la actividad que observamos en las clonas estables.

A continuación se realizaron ChIPs en dos líneas adicionales protegidas con el *insulator*  $\alpha$ EHS-1.4 para así poder comparar los perfiles de marcas post-traduccionales de las histonas en tres contextos de integración genómica diferentes que presentan expresión sostenida de la *EGPF* (línea 506, Fig. 26C y líneas 528 y 520, Fig. 27).



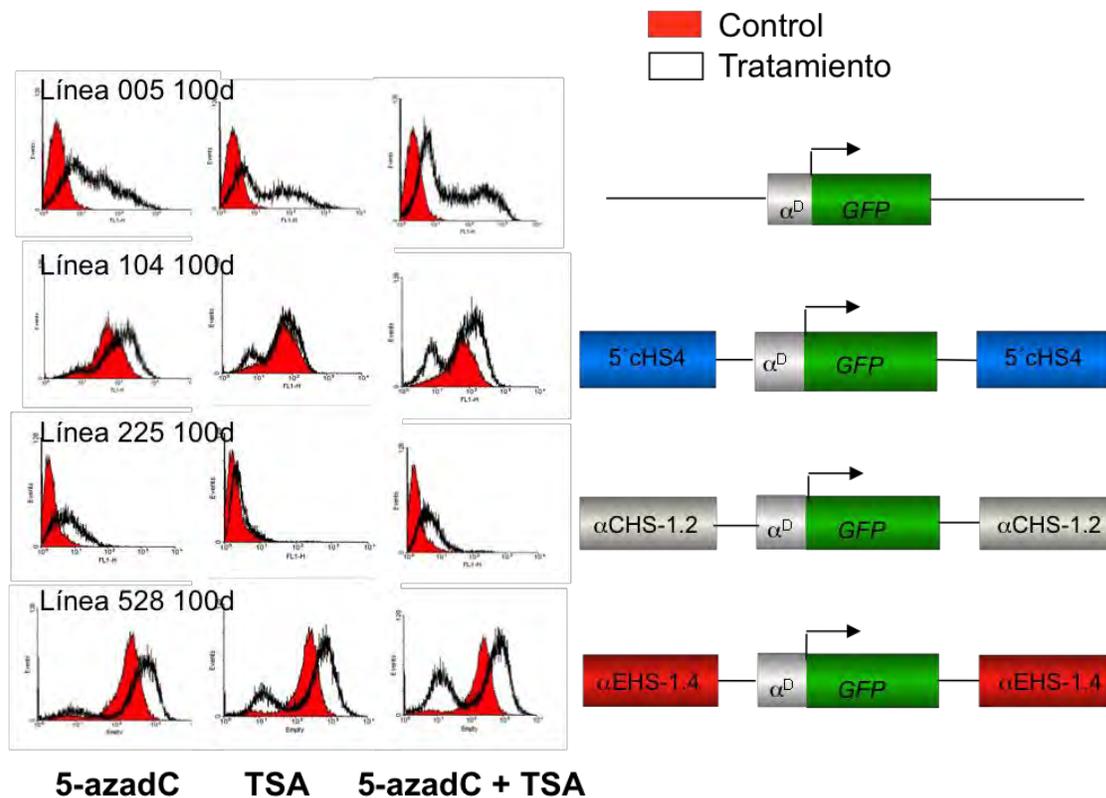
**Figura 27.** La cromatina de las clonas 520 y 528 presentan un perfil de marcas post-traduccionales de histonas correspondiente a cromatina permisiva. Puede observarse como ambas clonas están enriquecidas en marcas de cromatina abierta aun cuando el perfil y nivel de enriquecimiento es distinto entre ellas y en comparación con la línea 506 (ver Fig. 26C). Se presentan los promedios +/- el error estándar.

Puede observarse como las líneas HD3 528 y 520 que contienen al transgene  $pG\alpha D1.4+1.4$ , presentan un enriquecimiento en marcas de apertura de la cromatina al igual que la clona 506. Sin embargo el perfil observado difiere entre líneas, lo cual refleja que aun cuando el *insulator*  $\alpha\text{EHS-1.4}$  tiene la capacidad de protección contra los efectos de posición cromosomal, el contexto de integración influye en el perfil de marcas post-traduccionales de las histonas.

A partir de los datos anteriores podemos concluir que el *insulador*  $\alpha$ EHS-1.4 protege a transgenes de los EPC, favoreciendo el enriquecimiento de marcas permisivas de histonas sobre la región que enmarca.

### 7.3 Los EPC conllevan a la desacetilación de histonas y a la metilación del ADN de los transgenes

A continuación se abordó si acaso los EPC descritos están dados por la desacetilación de histonas y metilación del ADN ya que en estudios previos, se ha descrito que los EPC incluyen a dichos procesos (Pikaart *et al.*, 1998). Se realizaron tratamientos con TSA (Tricostatina-A) y 5-azadC (5-aza3-deoxicitidina) para inhibir a las desacetilasas de histonas (HDACs) y a la metilación de ADN, respectivamente (Fig. 28).



**Figura 28. Reactivación de las líneas 005, 104, 225 y 528 con TSA y 5-azadC.** Existe un efecto de posición cromosomal en todas las líneas que conlleva desacetilación de histonas y metilación del ADN ya que en todos los casos existe una reactivación de la fluorescencia con el tratamiento. El único caso diferente es en la línea 225 la cuál parece encontrarse metilada en el ADN pero no

presentar desacetilación de histonas. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células.

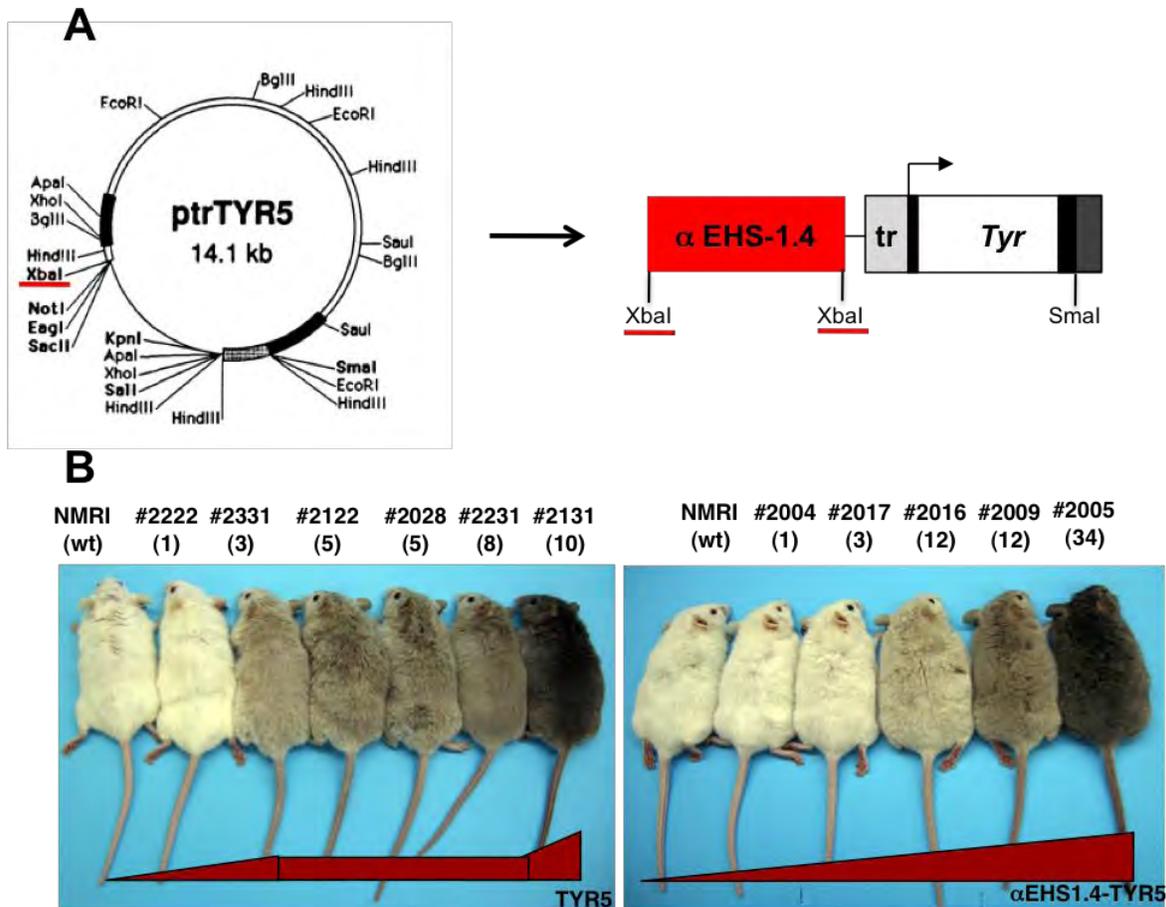
Al realizar los tratamientos con TSA y 5-azadC se pudo observar que en todos los casos, incluso los aislados con los *insulators* 5' cHS4 y  $\alpha$ EHS-1.4, presentan desacetilación de histonas y metilación del ADN de tal forma que, al incubar con los inhibidores de ambos procesos, existe un corrimiento hacia la derecha de la fluorescencia que indica ganancia de expresión de la *EGFP*. Esto significa que, en el caso de la línea con el elemento  $\alpha$ CHS-1.2, la desacetilación de histonas parece no tener un papel fundamental en el efecto de silenciamiento que causa este elemento sobre el gen reportero mientras que la metilación del ADN parece ser más importante (comparar reactivación con TSA y 5-azadCdC sobre la línea 225, Fig. 28). Esta observación deberá ser estudiada con mayor detalle para elucidar la función del elemento  $\alpha$ CHS-1.2 en el dominio  $\alpha$ -globina de pollo.

## **7.4 El *insulator* $\alpha$ EHS-1.4 protege contra EPC en ratones transgénicos**

Con base en el conjunto de datos anteriores, podemos concluir que el elemento  $\alpha$ EHS-1.4 es un *insulator* que posee la capacidad de proteger a un transgén contra los efectos de posición cromosomal. Para evaluar si el *insulator*  $\alpha$ EHS-1.4 puede proteger contra los EPC en un sistema diferente a las células HD3 de pollo, se decidió (a través de una colaboración con el grupo del Dr. Lluís Montoliu del Centro Nacional de Biotecnología CNB en Madrid, España) probar su capacidad de protección generando ratones transgénicos con el vector ptrTYR5 con y sin *insulators*  $\alpha$ EHS-1.4, en una línea de ratones albinos NMRI/Hsd. Dicho vector contiene la versión mínima del gen que codifica para la tirosinasa, enzima implicada en la biosíntesis de melanina. En los ratones que se logra la expresión de este transgén, el fenotipo albino se rescata presentando una coloración agouti del pelaje y ojos negros (Potts *et al.*, 2000). Se sabe que la eficiencia en la expresión de este transgén sin ningún elemento regulatorio adicional es muy baja, es decir que es susceptible a los efectos de posición cromosomales, dando como resultado una gran variabilidad en la pigmentación de los animales transgénicos generados. Así mismo se ha

demostrado que aislar a dicho transgén con el *insulator* 5' cHS4 del dominio  $\beta$ -globina de pollo homogeniza su expresión (Potts *et al.*, 2000). Para probar si el *insulator*  $\alpha$ EHS-1.4 tiene la capacidad de protección contra los EPC en este sistema, se construyó el plásmido ptrTYR5 $\alpha$ EHS-1.4 que contiene al *insulator*  $\alpha$ EHS-1.4 en el sitio XbaI del vector ptrTYR5. Se integró solo en un costado del gen reportero ya que se ha documentado que al generar ratones transgénicos las construcciones se integran en multicopias en tándem (Giraldo *et al.*, 2003). El vector ptrTYR5 $\alpha$ EHS-1.4 construido se utilizó para generar ratones transgénicos inyectándolo en pronúcleos de ovocitos fertilizados de ratón (Fig. 29).

Se llevaron a cabo las micro-inyecciones de los plásmidos control (ptrTYR5) y el plásmido con el *insulator*  $\alpha$ EHS-1.4 (ptrTYR5 $\alpha$ EHS-1.4). Se generaron cinco líneas independientes de ratones transgénicos ptrTYR5 $\alpha$ EHS-1.4 y seis con plásmido control ptrTYR5. Al comparar la coloración del pelaje de ambos grupos de ratones transgénicos, se derivaron varias observaciones: Por un lado, los ratones control (ptrTYR5) presentaron una pigmentación del pelaje bastante abundante de manera contrastante con lo previamente reportado en donde se presenta una alta variegación en los transgenes generados y una pigmentación muy escasa en la mayoría de los casos (Potts *et al.*, 2000). En nuestro caso, la pigmentación resultó bastante homogénea sin ninguna evidencia de mosaicismo observado en los reportes previos (Potts *et al.*, 2000). A pesar de lo anterior pudimos concluir que, en los ratones con el *insulator*  $\alpha$ EHS-1.4, existe una correlación entre el número de copias insertadas en el genoma y la intensidad de la coloración del pelaje a diferencia de lo que ocurre con los ratones control (Fig. 29B). La integridad y número de copias fueron analizados por southern blot (no mostrado). Para confirmar estos datos, a continuación se analizó la expresión del gen reportero *Tyr* en diversos tejidos de los ratones en presencia y ausencia de *insulators* por medio de PCR en tiempo real (qPCR).

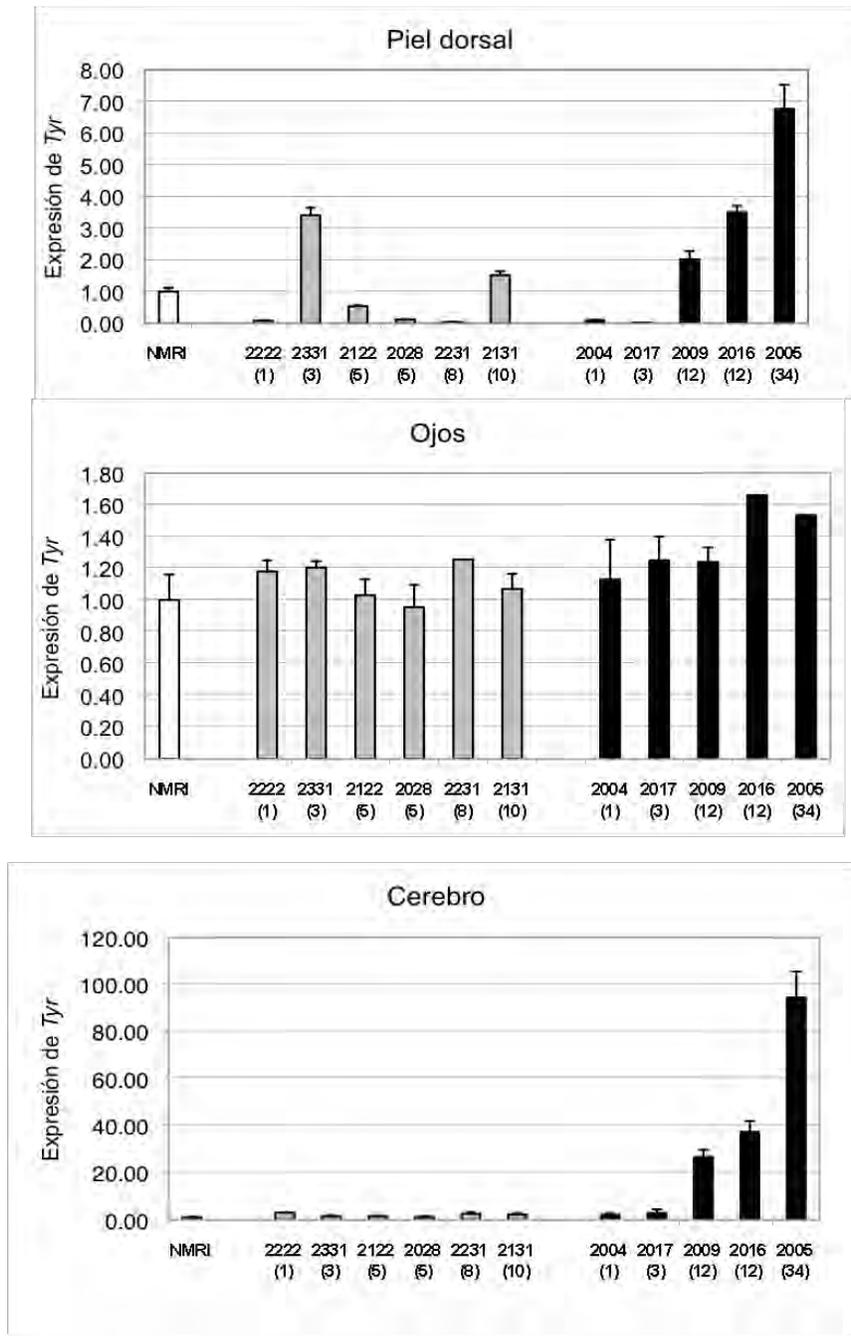


**Figura 29. Ratones transgénicos con los plásmidos ptrTYR5 y ptrTYR5 $\alpha$ EHS-1.4.** A. Esquema del vector ptrTYR5 el cual contiene una versión reducida del gen de la *tirosinasa* que comprende 250 pb del promotor (seguido del sitio XbaI) (tr), el exón 1 (caja negra) seguido del intrón 1 (caja blanca), los exones 2-5 fusionados (caja negra cerca del sitio SmaI) y las señales de procesamiento y poly(A) de SV40 (caja gris) (Potts *et al.*, 2000). El *insulator*  $\alpha$ EHS-1.4 se clonó en el sitio XbaI. B. Fotografía de los ratones transgénicos. Individuos representativos de las generaciones F1 y F2 de ratones transgénicos para el plásmido ptrTYR5 (izquierda) y ptrTYR5 $\alpha$ EHS-1.4 (derecha) junto con el control no transgénico NMRI (wt) utilizado como cepa receptora. Se muestra el número de la línea transgénica y el número de copias integradas entre paréntesis.

Como puede observarse, la expresión del gen de la *tirosinasa* en los diferentes tejidos analizados (piel dorsal, ojos y cerebro) de los ratones transgénicos ptrTYR $\alpha$ EHS-1.4 muestra una clara correlación con el número de copias insertadas en cada caso mientras que, en los ratones transgénicos ptrTYR5, la expresión es variable y no correlaciona con el número de copias insertadas del transgén utilizado (Fig. 30). Dicho fenómeno muestra como, al enmarcar al gen *Tyr* con el *insulator*  $\alpha$ EHS-1.4, la expresión que se obtiene se ve influenciada por el sitio de inserción en el genoma y responde proporcionalmente al

número de copias ya que, mientras más copias del transgén se integran, mayor es la expresión que se genera sin importar el tejido analizado. En cambio, en los ratones transgénicos que contienen la construcción sin *insulators*, la expresión de tirosinasa es muy variable y está influenciada por el sitio de inserción de tal forma que no existe una correlación con el número de copias integradas del transgén. Además, pudimos observar que la expresión de *Tyr* en los ratones sin *insulator* está sujeta al contexto celular utilizado de tal forma que, en un tejido como el ojo en donde el gen *Tyr* se expresa de manera robusta, observamos una expresión mayor mientras que en el cerebro en donde éste no se expresa (Gimenez *et al.*, 2003), la expresión es escasa (Fig. 30).

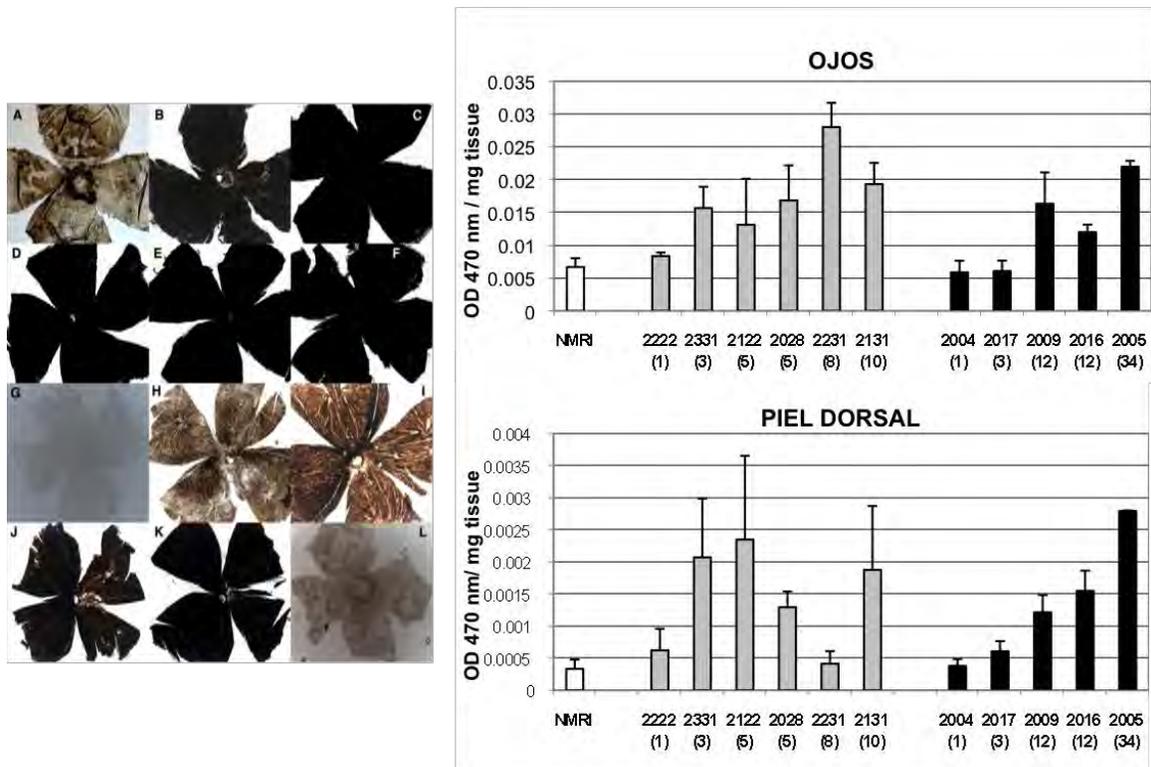
De manera interesante, al analizar los niveles de expresión de la *tirosinasa* en los diferentes tejidos, notamos que en el tejido neural (cerebro) en donde el gen *Tyr* normalmente no se expresa, el *insulator*  $\alpha$ EHS-1.4 confiere una expresión muy robusta al gen reportero (de manera dependiente al número de copias). Datos no publicados del laboratorio (Delgado *et al.*, sometido), muestran que el factor CTCF (analizado a través de hibridación *in situ* en el pez zebra) se expresa de manera muy abundante en el sistema nervioso central. Posiblemente la abundancia relativa de CTCF en este tejido influya en la actividad del *insulator*  $\alpha$ EHS-1.4 ya que dicho factor juega un papel esencial en sus funciones como hemos definido en el transcurso del presente trabajo de investigación (ver más adelante, sección 7.5.1).



**Figura 30. Expresión del gen reportero *Tyr* en diferentes tejidos de los ratones transgénicos ptrTYR5 y ptrTYR5αEHS-1.4.** Gráficos de la expresión del gen *Tyr* en piel dorsal (arriba), ojos (en medio) y cerebro (abajo). Los niveles de expresión fueron determinados mediante PCR en tiempo real (qPCR) normalizando contra la expresión del gen *TBP* (por sus siglas en inglés TATA Binding Protein) en cada caso. En el eje horizontal se muestra cada ratón transgénico (barras grises, ratones ptrTYR5, barras negras ratones ptrTYR5αEHS-1.4 y barra blanca ratón control, NMRI) y el número de copias entre paréntesis. El resultado se presenta como el promedio +/- la desviación estándar.

A continuación se realizaron dos experimentos adicionales para corroborar las observaciones previas con respecto a la capacidad del  $\alpha$ EHS-1.4 de proteger contra EPC en ratones transgénicos. Por un lado, se realizaron preparaciones de retinas de los ratones transgénicos para evaluar directamente la cantidad de células epiteliales pigmentadas y los melanocitos presentes en el corioide que son un reflejo de la expresión de la *Tirosinasa* (Fig. 31). En las retinas de los ratones transgénicos control (ptrTYR5), se observó una saturación de células pigmentadas sin importar el número de copias del transgén. Éste es el mismo comportamiento detectado en la cuantificación de la expresión de *Tyr* por qPCR (Fig. 30), en donde la expresión es elevada y constante entre los distintos ratones transgénicos. Al analizar las retinas de los ratones ptrTYR5 $\alpha$ EHS-1.4 observamos una clara correlación con el número de copias en donde la presencia del pigmento aumenta al aumentar el número de transgenes integrados (Fig. 31).

Por otro lado, se realizó la cuantificación del contenido total de melanina en los ojos y piel dorsal de todos los ratones transgénicos. La melanina es el producto final de la vía de síntesis en donde participa la tirosinasa y, por lo tanto, es un reflejo de la abundancia de la misma. El contenido de melanina de los ratones ptrTYR5 fue relativamente uniforme mientras que para los ratones ptrTYR5 $\alpha$ EHS-1.4 la cantidad de melanina correlaciona nuevamente con el número de copias insertadas del transgén (Fig. 31). En conclusión, el conjunto de datos obtenidos muestran que el *insulator*  $\alpha$ EHS-1.4 es capaz de proteger contra EPC en un sistema heterólogo como el ratón.



**Figura 31. Células pigmentadas en retina y contenido de melanina en los ratones transgénicos ptrTYR5 y ptrTYR5αEHS-1.4.** A la izquierda se muestran las preparaciones de retinas completas de los ratones ptrTYR5 (A-F) y ptrTYR5αEHS-1.4 (G-K) en orden creciente con respecto al número de copias insertadas del transgén. El control NMRI se ubica en el panel (L). A la derecha de muestran los gráficos del contenido de melanina en ojos (arriba) y piel dorsal (abajo) para los ratones transgénicos ptrTYR5 (barras grises) y ptrTYR5αEHS-1.4 (barras negras) y control (barras blancas). Los datos se muestran en promedio +/- la desviación estándar.

## 7.5 CTCF es necesario en la actividad del *insulator* αEHS-1.4

Una vez determinada la capacidad de protección contra EPC por parte del *insulator* αEHS-1.4, nos interesamos en investigar si CTCF juega un papel importante en dicha función. De manera clásica, la capacidad de protección contra los efectos de posición cromosomal por parte de los *insulators*, no se adjudica al factor CTCF ya que en el *insulator* mejor estudiado (el 5' cHS4), CTCF no contribuye en esta función (Recillas-Targa *et al.*, 2002 y West *et al.*, 2004). Sin embargo, algunas evidencias recientes presentadas en la introducción del trabajo (sección 1.3.1 y 1.3.2), sugieren que CTCF puede jugar un papel

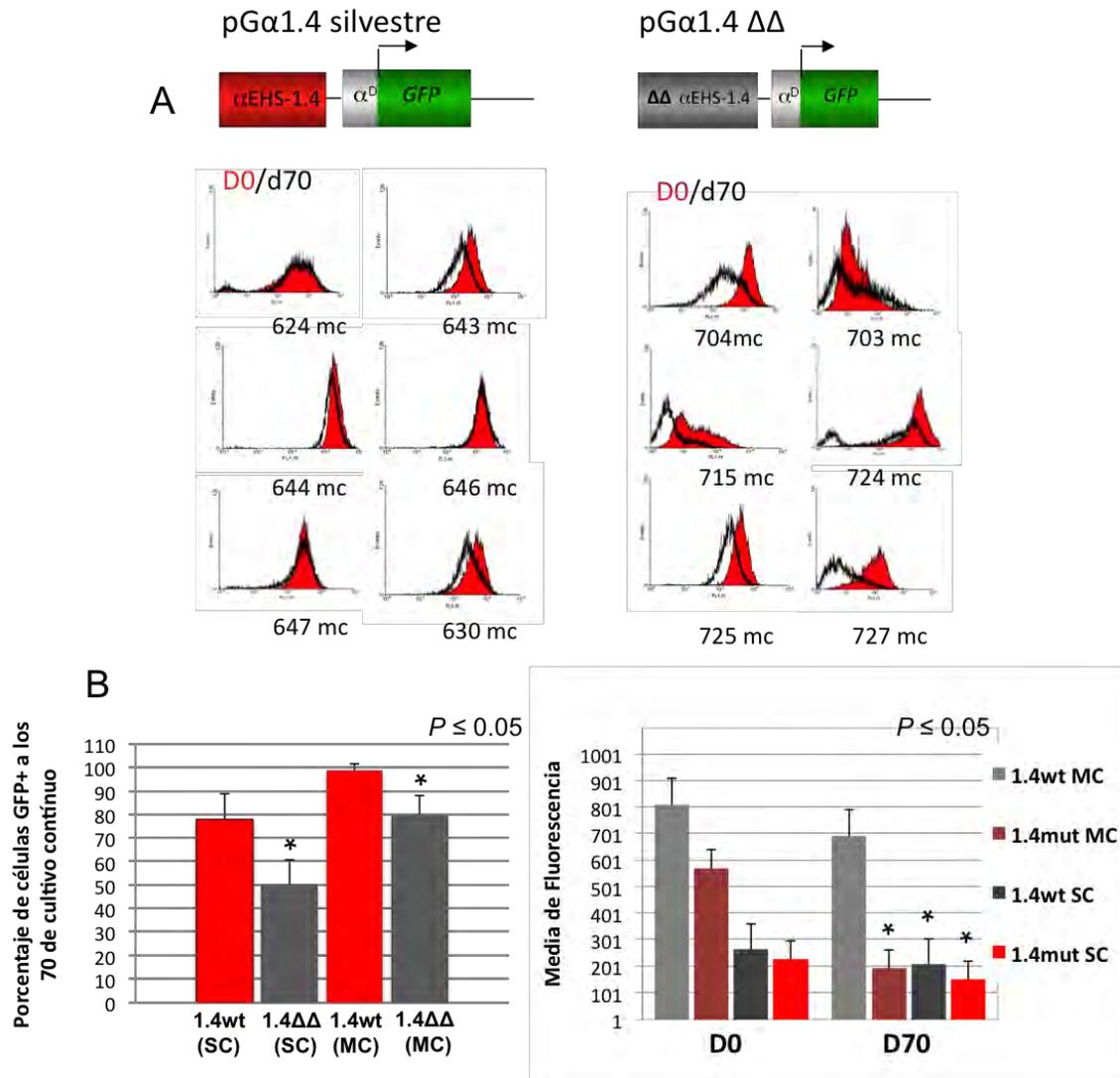
importante en esta propiedad de los *insulators* por lo cual decidimos evaluarlo. Para distinguir si CTCF es el factor responsable de la protección contra el efecto de posición cromosomal del *insulator*  $\alpha$ EHS-1.4 planteamos diferentes estrategias presentadas a continuación.

### **7.5.1 Mutaciones puntuales sobre la secuencia de unión de CTCF en el *insulator* $\alpha$ EHS-1.4 interfieren en su actividad**

Para evaluar el papel de CTCF en la actividad del insualtor generamos un *insulator*  $\alpha$ EHS-1.4 que contiene dos mutaciones en el sitio de unión de CTCF por medio de PCRs anidadas (ver métodos, sección 6.4). Dichas mutaciones ( $\Delta$ CAG y  $\Delta$ CTAG) fueron establecidas en ensayos de retardo confirmando la pérdida de interacción con CTCF (Valadez-Graham *et al*, 2004). El *insulator*  $\alpha$ EHS-1.4 $\Delta\Delta$  (que contiene ambas mutaciones), fue clonado en el plásmido pG $\alpha^D$  generando el vector pG $\alpha^D$ 1.4 $\Delta\Delta$ . Este fue linearizado y sin desfosforilar, se transfectó de manera estable en células HD3 para favorecer multicopias y así reproducir la condición en donde dos *insulators* enmarcan al transgén.

Al comparar los perfiles de fluorescencia de las clonas pG $\alpha^D$ 1.4 $\Delta\Delta$  al día 0 con el día 70 de cultivo continuo, pudimos observar una disminución en la expresión de la *EGFP* indicativa de la incapacidad de protección por parte del *insulator*  $\alpha$ EHS-1.4 $\Delta\Delta$  en el cual el sitio de unión al factor CTCF ha sido mutado. Estos perfiles asemejan a los que obtenidos para las clonas pG $\alpha^D$ 3 (línea 001) las cuales no incorporan a los *insulators* (Fig. 23A), lo cual apoya que CTCF tiene una función importante en la actividad de protección por parte del *insulator*  $\alpha$ EHS-1.4. Así mismo al observar los perfiles de las líneas que contienen al *insulator*  $\alpha$ EHS-1.4 silvestre pudimos notar la estabilidad de la fluorescencia a los 70 días de cultivo reiterando la capacidad del elemento para contender contra los efectos de posición cromosomales (Fig. 32A y B izquierda). Al cuantificar la intensidad de la fluorescencia al día 0 y día 70 del cultivo continuo para las líneas generadas, encontramos que la mayor intensidad promedio se consigue en las líneas que poseen al plásmido con el

*insulator*  $\alpha$ EHS-1.4 silvestre en multicopia mientras que en el resto de las condiciones (con el *insulator*  $\alpha$ EHS-1.4 silvestre en copia sencilla y con el *insulator*  $\alpha$ EHS-1.4 $\Delta\Delta$  en copia sencilla o multicopia), la intensidad de fluorescencia al día 70 es significativamente menor (Fig. 32B, derecha).



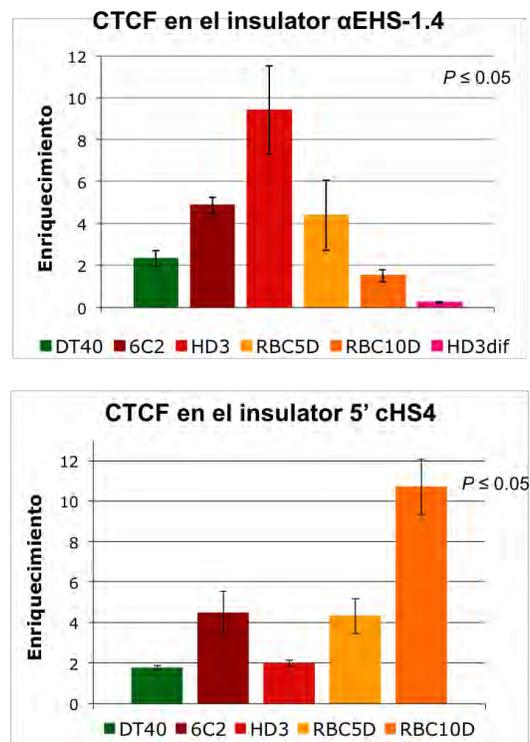
**Figura 32. Generación y análisis de clonas pG $\alpha$ <sup>D</sup>1.4 silvestre (líneas 600) y clonas pG $\alpha$ <sup>D</sup>1.4 $\Delta\Delta$  (líneas 700).** **A.** Histogramas de FACS que muestran la fluorescencia de diferentes líneas establecidas con el *insulator*  $\alpha$ EHS-1.4 silvestre (izquierda) y  $\alpha$ EHS-1.4 $\Delta\Delta$  (derecha) a los 0 y 70 días de cultivo continuo. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células. **B.** Cuantificación del porcentaje de células GFP<sup>+</sup> a los 70 días de cultivo (izquierda) y de la intensidad de fluorescencia (derecha) para las líneas con el *insulator*  $\alpha$ EHS-1.4 silvestre en copia sencilla (SC) y copia múltiple (MC) al igual que para el *insulator*  $\alpha$ EHS-1.4 $\Delta\Delta$ . Los datos se muestran en promedio +/- la desviación estándar.

$\alpha$ EHS-1.4 silvestre MC n=18,  $\alpha$ EHS-1.4 silvestre SC n=16,  $\alpha$ EHS-1.4  $\Delta\Delta$  MC n=22,  $\alpha$ EHS-1.4 $\Delta\Delta$  SC n=22.

Los datos anteriores indican que CTCF juega un papel importante en la actividad de protección contra EPC ejercida en el contexto del *insulator*  $\alpha$ EHS-1.4.

### 7.5.2 La asociación de CTCF al *insulator* $\alpha$ EHS-1.4 endógeno durante la diferenciación eritroide y el desarrollo del organismo es dinámica

A continuación, decidimos evaluar la abundancia relativa de CTCF sobre el *insulator*  $\alpha$ EHS-1.4 endógeno en diferentes estadios de la diferenciación eritroide y el desarrollo del pollo. Estos experimentos fueron realizados por Eria Rebollar en el laboratorio durante sus estudios de maestría.



**Figura 33. Ensayo de CHIP contra CTCF en los *insulator*  $\alpha$ EHS-1.4 y 5' cHS4.** El gráfico superior muestra el enriquecimiento relativo de CTCF sobre el insulator  $\alpha$ EHS-1.4 en los diferentes estadios de diferenciación eritroide y en eritrocitos primarios de 5 y 10 días de desarrollo. El gráfico inferior muestra el enriquecimiento relativo de CTCF sobre el *insulator* 5' cHS4. Todos los datos se presentan como promedios +/- la desviación estándar. DT40: Línea linfóide, 6C2: Pre-eritroblasto, HD3: Eritroblasto, HD3dif: Eritroblasto diferenciado, RBC5D y RBC10D: Eritrocitos primarios de 5 y 10 días de desarrollo del pollo. Gráfico superior: DT40 n= 3, 6C2 n=5, HD3 n=4, RBC5d n=3, RBC10d n=3, HD3dif n=4. Gráfico inferior: DT40 n= 3, 6C2 n=3, HD3 n=3, RBC5d n=3, RBC10d n=6.

Al realizar el ChIP contra CTCF observamos que la abundancia relativa de CTCF sobre el *insulator*  $\alpha$ EHS-1.4 cambia durante la diferenciación eritroide. Pudimos notar que existe un pico de enriquecimiento en las células HD3 (correspondientes al estadio de eritroblastos) y que conforme la diferenciación avanza, la presencia de CTCF disminuye en las condiciones de RBC 10d y HD3 dif en donde los genes adultos  $\alpha^D$  y  $\alpha^A$  son transcritos (Fig. 33). Al igual que para el *insulator*  $\alpha$ EHS-1.4 analizamos la presencia de CTCF sobre el *insulator* 5' cHS4 del dominio  $\beta$ -globina. La dinámica de enriquecimiento de CTCF sobre este *insulator* es distinta a la del  $\alpha$ EHS-1.4 y en este caso, se presenta un pico de enriquecimiento del factor en eritrocitos maduros de 10 días de desarrollo (RBC 10d). Esto indica que la abundancia de CTCF puede variar dependiendo el tejido analizado así como del *insulator* en estudio mostrando que el *insulator*  $\alpha$ EHS-1.4 y 5' 5' cHS4 pueden tener mecanismos diferentes que responden a las necesidades propias del entorno genómico en donde se localizan. Cabe señalar que estas diferencias parecen no estar relacionadas con variaciones en la abundancia general de CTCF en los diferentes estadios. Esto se confirmó por medio de western blot comparando células HD3 con células HD3 inducidas a diferenciar (Fig. 38A, comparar HD3 con HD3d).

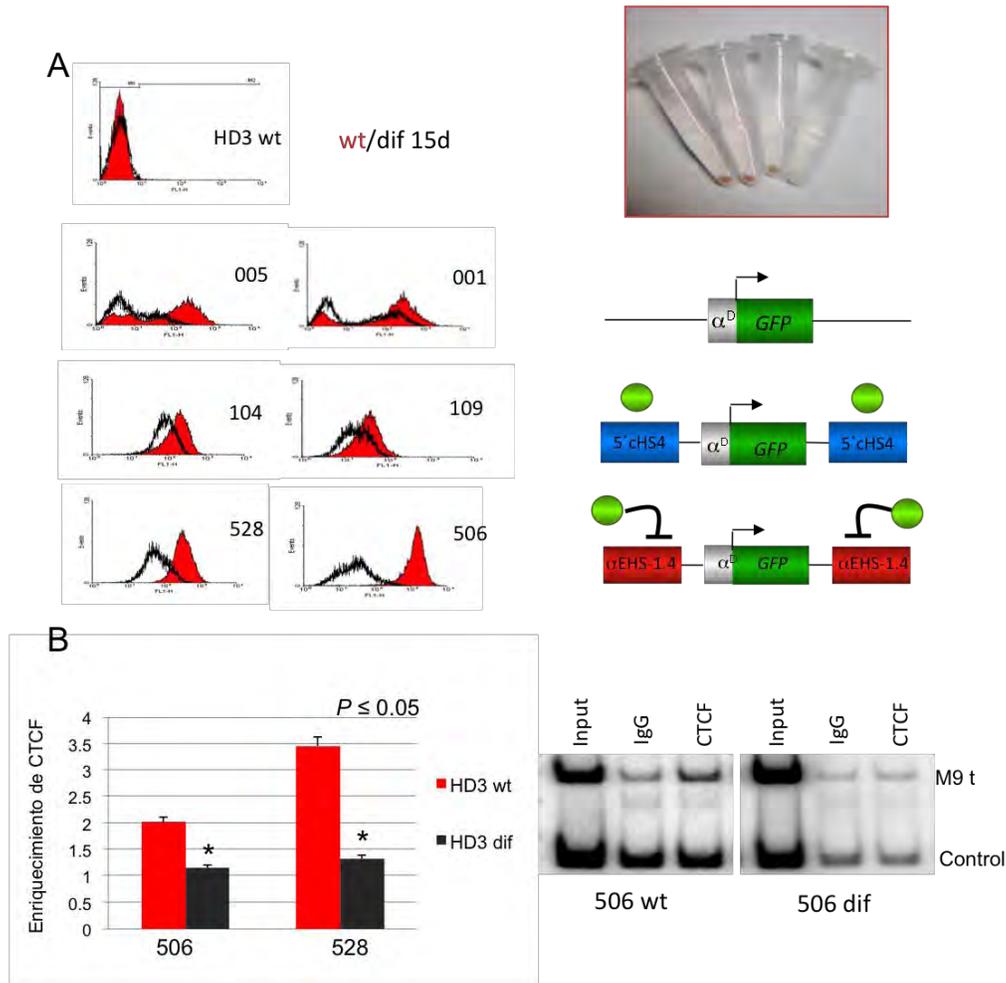
Dada la unión dinámica del factor CTCF en el *insulator*  $\alpha$ EHS-1.4 durante la diferenciación eritroide, a continuación decidimos evaluar la función de éste en condiciones diferenciadas en donde CTCF se desprende del *insulator*.

## **7.6 La actividad del *insulator* $\alpha$ EHS-1.4 correlaciona con la asociación de CTCF durante la diferenciación eritroide**

Para abordar la función del *insulator*  $\alpha$ EHS-1.4 en condiciones diferenciadas donde CTCF no está unido a dicho elemento, llevamos a cabo el protocolo de diferenciación

sobre nuestras clonas estables protegidas con este *insulator* y evaluamos la expresión de la *EGFP* por citometría de flujo (Fig. 34A).

Al realizar el protocolo de diferenciación notamos varias cosas interesantes; por un lado el silenciamiento de las líneas sin *insulators* ocurre con mayor rapidez (a un número menor de días de cultivo continuo) que en el contexto de HD3 sin diferenciar (Fig. 34A, clonas 001 y 005). Lo anterior nos habla del ambiente de la cromatina en los eritrocitos diferenciados en donde probablemente existe una vasta heterocromatinización del genoma con excepción de algunos dominios que se expresan en forma constitutiva y otros que son eritroide específicos, como los globina. Por otro lado las líneas aisladas con el *insulator* 5' cHS4 del dominio  $\beta$ -globina no muestran un silenciamiento muy pronunciado (Fig. 34A, clonas 104 y 109) mientras que las aisladas con el *insulator*  $\alpha$ EHS-1.4 presentan un silenciamiento prominente de la expresión de la *EGFP* (Fig. 34A, líneas 528 y 506). Lo anterior concuerda con la presencia y ausencia de CTCF en el *insulator*  $\alpha$ EHS-1.4 endógeno en eritroblastos y eritrocitos de 10 días de desarrollo, respectivamente. Estos datos coinciden con estudios previos que demuestran que, la protección contra EPC del *insulator* 5' cHS4 no depende de CTCF sino de USF1/USF2 entre otros factores (West *et al.*, 2004).



**Figura 34. La capacidad de protección contra CPE por parte del *insulator*  $\alpha$ EHS-1.4 se pierde al diferenciar las células HD3 a eritrocitos y CTCF se desprende del *insulator*.** **A.** Las líneas 001 y 005 ( $pG\alpha^D3$ ), 104 y 109 ( $pG\alpha^D$  5' cHS4 + 5' cHS4) y 528 y 506 ( $pG\alpha^D1.4+1.4$ ) (diagramas a la derecha, CTCF se representa con una esfera verde) se diferenciaron durante 10-15 días y se obtuvieron los perfiles de FACS (izquierda). En rojo se muestra el perfil sin diferenciar y en contorno negro el diferenciado. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células. a la derecha se muestran los pellets de células diferenciadas en comparación con células control. Las células diferenciadas se tornan de color rojizo. **B.** ChIP contra CTCF amplificando sobre el sitio de unión de éste en el *insulator*  $\alpha$ EHS-1.4 en el transgén (M9t). A la izquierda se presenta el gráfico de enriquecimiento relativo y a la derecha un gel representativo. Los datos se muestran en promedios  $\pm$  el error estándar. Línea 506silvestre  $n=3$ , 506dif  $n=3$ , 528silvestre  $n=3$  y 528dif  $n=3$ .

Como se ha mencionado, la pérdida de protección contra EPC en las líneas diferenciadas coincide con el desprendimiento de CTCF del *insulator*  $\alpha$ EHS-1.4 en el dominio  $\alpha$ -globina. Para corroborar que dicha pérdida sucede también a nivel del transgén, realizamos ChIPs contra CTCF en ambas condiciones, amplificando un fragmento del transgén que incluye el sitio de unión de CTCF en el *insulator*  $\alpha$ EHS-1.4 (ver Tabla 1) (Fig. 34B). Al igual que en el contexto endógeno, CTCF se desprende del *insulator* en el transgén al inducir la diferenciación de las líneas generadas en HD3 (eritroblastos) y esto corresponde con la pérdida parcial de protección contra EPC.

## 7.7 Alteración de la abundancia y la función de CTCF

La dinámica de CTCF sobre el *insulator*  $\alpha$ EHS-1.4 sugiere que el *insulator* posee un papel que depende del estadio de diferenciación eritroide. Para profundizar en cuanto a esta función en el contexto del dominio  $\alpha$ -globina de pollo, decidimos alterar la abundancia y función de CTCF para evaluar su efecto a nivel de la regulación del dominio  $\alpha$ -globina.

### 7.7.1 Generación de líneas celulares *knockdown* contra CTCF

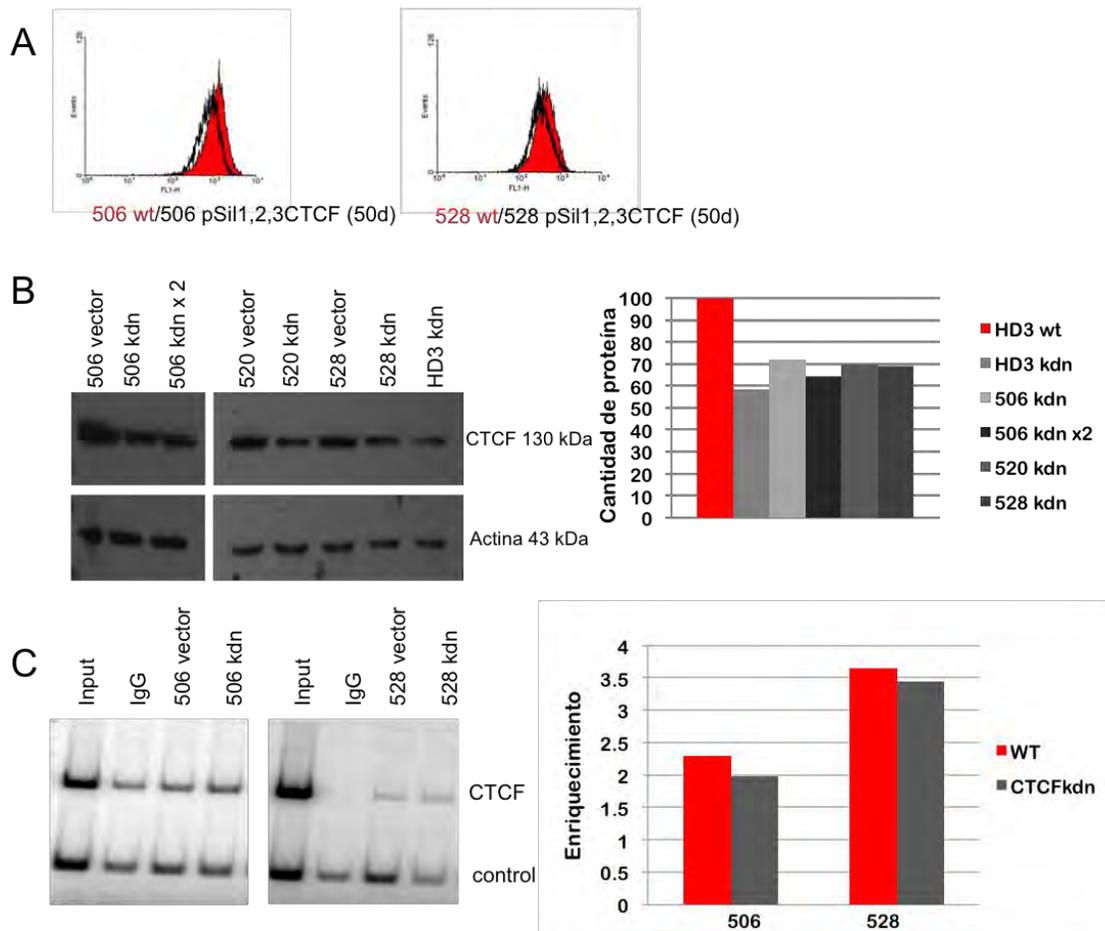
Para abatir la abundancia de la proteína de CTCF se desarrollaron varias estrategias experimentales. En primera instancia, llevamos a cabo el diseño de 3 shRNAs. Estas moléculas al ser transcritas, forman estructuras tallo-asa que son procesadas por la maquinaria de RNA de interferencia e interfieren con la traducción de la proteína en cuestión asociándose al mensajero e induciendo su degradación o impidiendo su traducción. Se diseñaron 3 oligonucleótidos de 21 pares de bases que hibridan con el mensajero de CTCF en las primeras 500 pb en relación al codón de inicio. Estos oligonucleótidos con estructura tallo-asa fueron clonados en los plásmidos pSilencer-3.1 Hygro (Ambion) generando así los plásmidos pSilCTCFkdn 1, 2 y 3. El diseño de los RNAs

interferentes fue realizado en colaboración con Fabián Flores del laboratorio del Dr. Luis Vaca (IFC-UNAM).

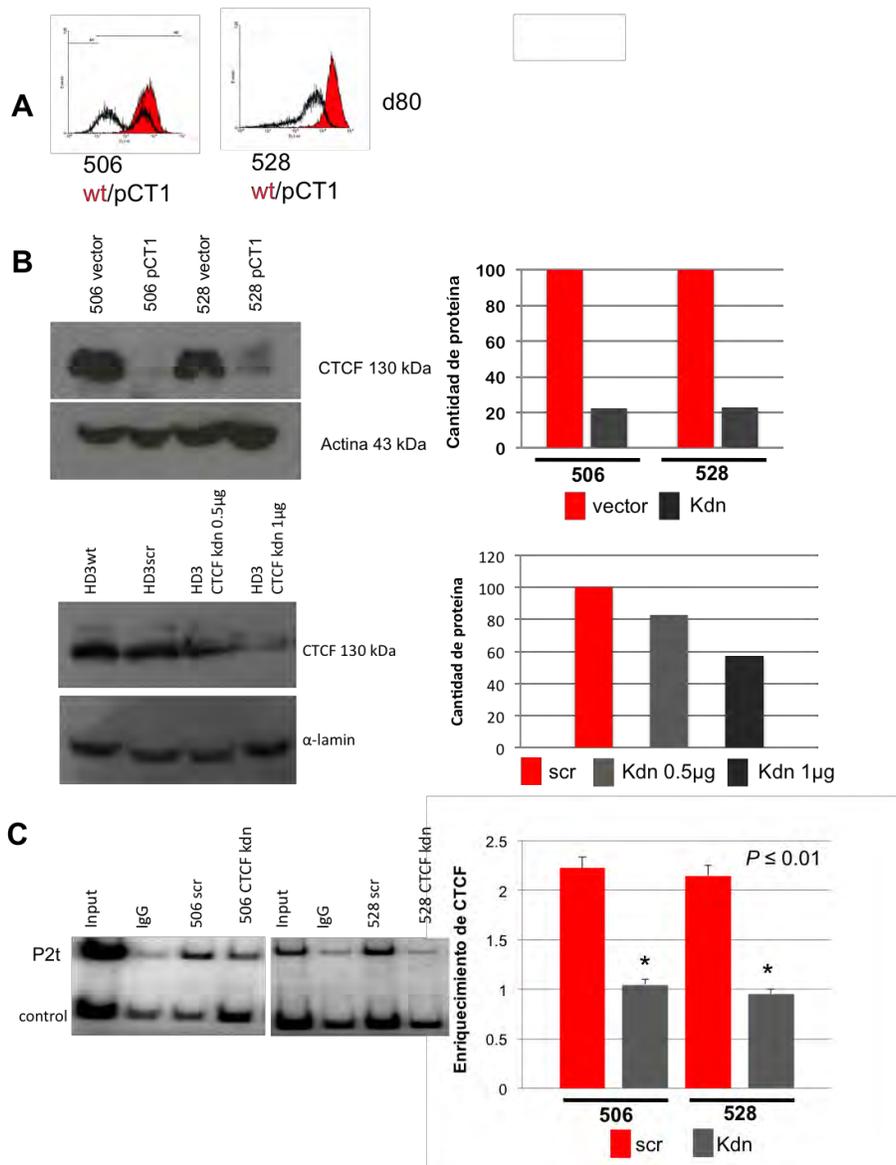
Una combinación de estos tres shRNAs fue transfectada establemente en las clonas 506, 520 y 528 (pGα<sup>D</sup>1.4+1.4) y la concentración relativa de la proteína CTCF se analizó por western blot y el nivel de fluorescencia se evaluó por citometría de flujo (Fig. 35B). La disminución conseguida con estos shRNAs no fue suficiente para notar un efecto a nivel de CHIP y de extinción de la fluorescencia (Fig. 35A y C).

A continuación se realizó otra serie de experimentos utilizando al plásmido pCT1 (Ishihara *et al.*, 2006). Con este vector se obtuvieron resultados variables, desde una disminución de alrededor del 80% en los niveles de CTCF hasta una de únicamente el 30% (Fig. 36B). En aquellas líneas en las que el porcentaje de disminución alcanzaba alrededor de un 50% y fue posible estimar de manera reproducible un abatimiento de CTCF en la región del *insulator* por medio de CHIP, se procedieron a realizar más experimentos incluyendo RT-PCRs, la inducción de la diferenciación de dichas células y el CHIP contra algunas marcas de histonas para analizar cambios en la región del *insulator* ante la pérdida de CTCF (ver adelante, sección 7.8).

Para intentar generar un sistema más robusto, decidimos diseñar nuevos shRNAs a partir de oligonucleótidos publicados que mostraron abatir al factor CTCF de pollo (Lefevre *et al.*, 2007). Una combinación de estos tres nuevos shRNAs en conjunto con el pCT1 fueron transfectados establemente en las células HD3 obteniendo nuevamente un porcentaje de disminución de alrededor de un 50% (Fig. 37). Con este material se repitieron los experimentos de CHIP para analizar la estructura de la cromatina endógena y las RT-PCRs confirmando los datos obtenidos anteriormente con el vector pCT1 (ver adelante, sección 7.8).

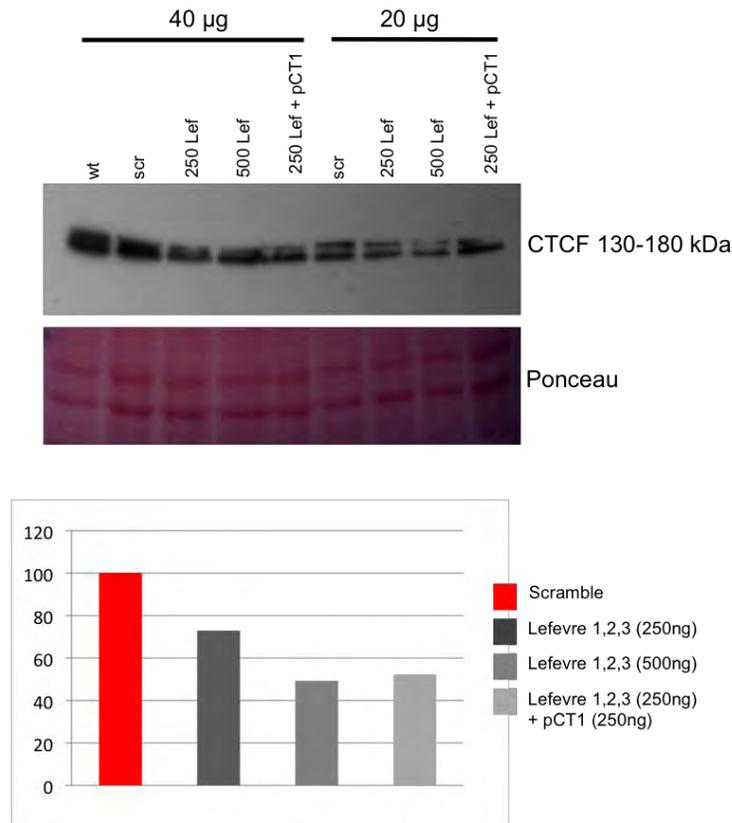


**Figura 35. Análisis de las líneas *knockdown* contra CTCF usando diferentes shRNAs.** **A.** Perfiles de citometría de las clonas pSilCTCFkdn 1,2,3. Los tres diferentes shRNAs contra CTCF (1, 2 y 3) se transfectaron en dos diferentes líneas pG $\alpha^{D1.4+1.4}$  (506 y 528). Como puede observarse, los perfiles de citometría son prácticamente idénticos entre la clonas silvestres (rojo) y las clonas transfectadas con los plásmidos pSilCTCFkdn 1,2 y 3 (negro). En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células. Este resultado sugiere que la disminución de CTCF que consiguen estos shRNAs no es suficiente para generar un efecto en la expresión del transgén. **B.** Western blot contra CTCF en las líneas con los pSilCTCFkdn 1,2,3. Se transfectaron diferentes líneas (506, 520 y 528) y en la línea 506 se realizó un experimento transfectando dos rondas de los shRNAs (506 kdn x 2). Como puede observarse la disminución de CTCF es de aproximadamente 30%. **C.** ChIP contra CTCF en las clonas 506 y 528 pSil 1,2,3 CTCF. No se observa una disminución en la asociación de CTCF del *insulator*  $\alpha$ EHS-1.4.



**Figura 36. Análisis de las líneas *knockdown* contra CTCF generadas con el shRNAi pCT1. A.** Perfiles de citometría de flujo de las clonas 506 y 528 silvestre (rojo) y pCT1 (negro) a los 80 días de cultivo continuo. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células. El efecto del shRNAi pCT1 en reducir en un 80% la abundancia de CTCF afecta la capacidad de protección del *insulator*  $\alpha$ EHS-1.4 y esto se ve reflejado en la pérdida de expresión del transgén. **B.** Western blot de las diferentes líneas *knockdown* para CTCF utilizando el shRNAi pCT1. (arriba) Western blot contra CTCF y gráfico sobre las líneas 506 y 528 transfectadas con el vector vacío (vector) *versus* el vector con el shRNAi contra CTCF (pCT1). La disminución de CTCF en éste caso fue de aproximadamente del 80% en ambas líneas (normalizado con la señal de el control de actina). (abajo) Western blot y gráfico representativo de los experimentos subsecuentes. La disminución máxima de CTCF al utilizar el plásmido pCT1 en diferentes concentraciones es de aproximadamente el 50%, (normalizado con el

control de lamina). **C.** ChIP contra CTCF en las líneas 506 y 528 *knockdown*. Se puede observar la pérdida de CTCF sobre el *insulator*  $\alpha$ EHS-1.4 en el transgén (P2t representa el par de oligos que amplifican a la región de unión de CTCF sobre el insulator  $\alpha$ EHS-1.4 en el transgén y , ver Tabla 1) (n=3), se muestran los promedios +/- el error estándar.



**Figura 37. Evaluación de las líneas *knockdown* contra CTCF utilizando los shRNAs Lefevre 1, 2 y 3.** Western blot contra CTCF en las líneas *knockdown* utilizando los plásmidos pSiLefevre 1, 2 y 3 (250 ng), pSiLefevre 1, 2 y 3 (500 ng) y la combinación pSiLefevre 1,2 y 3 + pCT1 (250 ng). El plásmido pScramble fue utilizado como control. Se muestra un duplicado cargando 40 y 20 µg de proteína total. La disminución obtenida de CTCF fue de alrededor del 50%. Con este material se repitieron los experimentos de ChIP y RT-PCRs para confirmar los datos obtenidos con el *knockdown* generado con el plásmido pCT1.

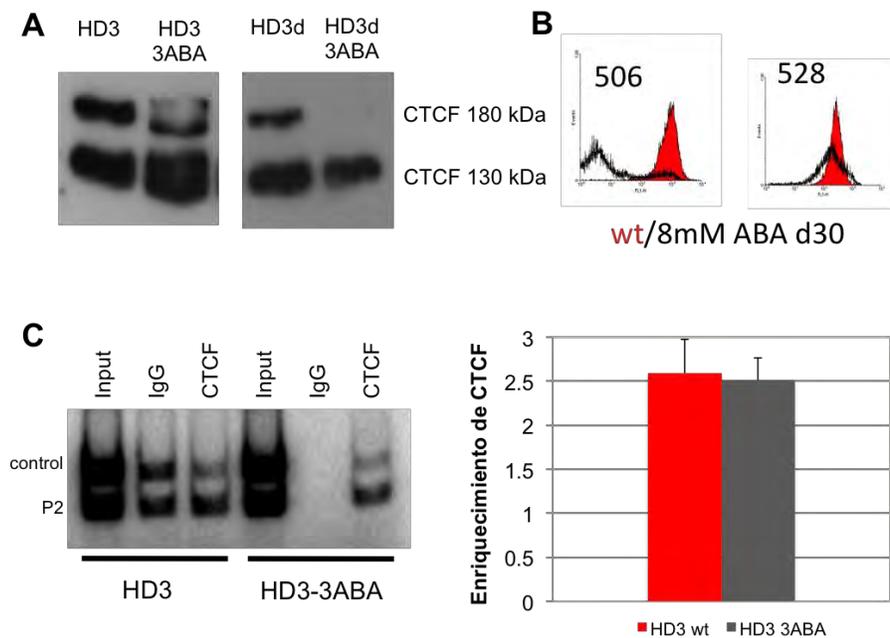
Por último se realizaron transfecciones utilizando todos los shRNAs disponibles (los 7 diferentes plásmidos; pSilCTCFkdn 1, 2 y 3, pSiLefevre 1, 2 y 3 y pCT1). Sin embargo, los resultados obtenidos con el pCT1 sólo ó con la combinación de pSiLefevre 1, 2 y 3 no fueron superados (datos no mostrados).

La variabilidad en la eficiencia del abatimiento de CTCF al utilizar los diferentes shRNAs puede deberse a diversas razones. Por un lado, toda la colección de secuencias utilizadas son complementarias al extremo 5' del mensajero maduro de CTCF (región que codifica para el extremo amino terminal). Probablemente el acceso a dicha región en el mensajero no sea la óptima debido a la formación de estructuras secundarias y convenga probar secuencias que se aparezcan en regiones del mensajero que codifican para zonas más centrales o hacia el carboxilo de la proteína CTCF. Por otro lado, sería aconsejable generar clonas estables para seleccionar el mejor escenario de disminución de CTCF ya que en todos los casos presentados en este trabajo, se utilizaron el conjunto de células transfectadas con los plásmidos de interés y esto suma la eficiencia de transfección y los efectos de posición a las variables para generar el *knockdown*. Otra posibilidad sería diseñar oligonucleótidos pequeños (siRNAs) que no tengan que pasar por el procesamiento de la maquinaria de RNA de interferencia y se dirijan directamente al mensajero blanco de CTCF.

### **7.7.2 La inhibición de la poli(ADPribosil)asa polimerasa (PARP-1) interfiere con la función del *insulator* $\alpha$ EHS-1.4**

Para interferir con la función de CTCF sin alterar los niveles de proteína presente en la célula, utilizamos 3-aminobenzamida (3-ABA), un inhibidor de la enzima poli(ADPribosil)asa-polimerasa (PARP-1). PARP-1, como se mencionó en la introducción de este trabajo, es una enzima que modifica post-traduccionamente a CTCF agregando residuos de ADP a su extremo amino terminal y dicha modificación es necesaria para que CTCF lleve a cabo su función (Yu *et al.*, 2004). Dependiendo de la cantidad de residuos de ADP presentes en CTCF se generan dos migrantes electroforéticos de dicho factor; una a 130 kDa con menos de 20 residuos y una de 180 kDa que contiene 20 o más residuos de ADP (Farrar *et al.*, 2009).

Al llevar a cabo el tratamiento con 8 mM de 3-ABA por 3-7 días de cultivo, pudimos observar una disminución importante en la cantidad de la forma de 180 kDa de CTCF (Fig. 38A). Al incubar con 3-ABA a las líneas celulares 506 y 528 que contienen al transgen *EGFP* delimitado por el *insulator*  $\alpha$ EHS-1.4 notamos una pérdida de protección contra EPC por parte del  $\alpha$ EHS-1.4 de manera similar a lo que ocurre en el *knockdown* (Fig. 38B) lo cual indica que el *insulator*  $\alpha$ EHS-1.4 requiere esta poli(ADPribosil)ado para ser funcional. Sin embargo, al analizar la presencia de CTCF por medio de CHIP sobre el *insulator*  $\alpha$ EHS-1.4 endógeno, no observamos una disminución significativa de éste en relación al control (Fig. 38C). Esto indica que la pérdida de la poli(ADPribosil)ación de CTCF no impide que éste se una al ADN.



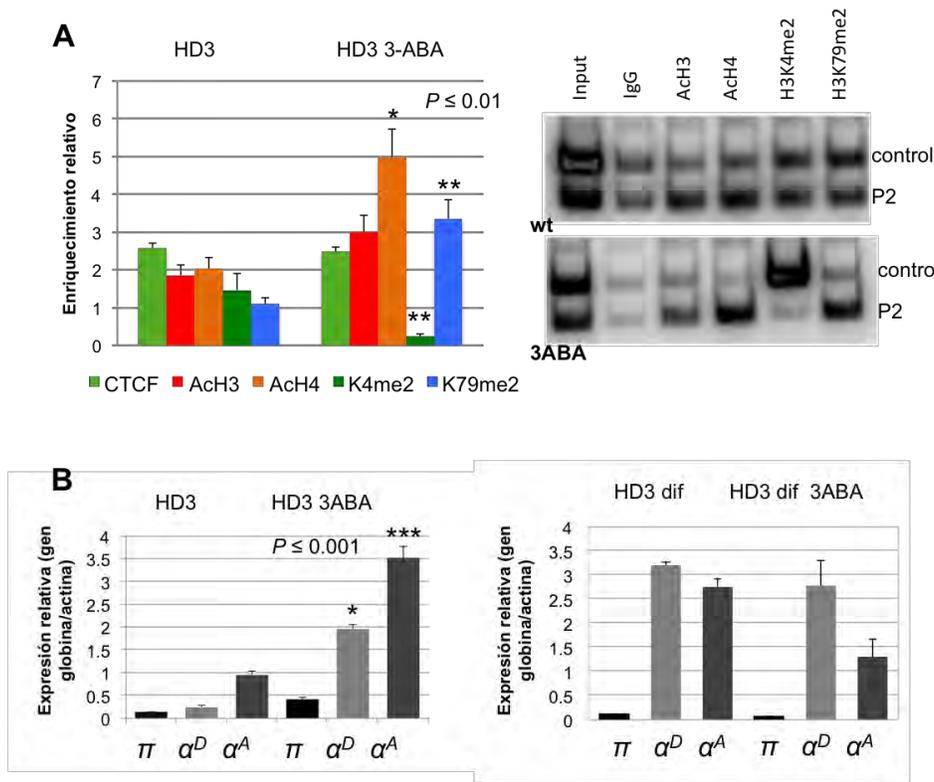
**Figura 38. Evaluación del tratamiento con 3-ABA.** **A.** Western blot contra CTCF en extractos proteicos de células HD3 y HD3 tratadas con 3-ABA así como células HD3 diferenciadas (HD3d) y células HD3 diferenciadas tratadas con 3-ABA (HD3d 3ABA). Existe una disminución en la forma de 180 kDa de CTCF **B.** Perfiles de FACS de las líneas celulares con el transgen *GFP* protegido por el *insulator*  $\alpha$ EHS-1.4 (506 y 528) tratadas con 3-ABA (gráfica abierta) y controles (gráficas rojas). En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células. **C.** Ensayo de CHIP contra CTCF en las células HD3 y células HD3

tratadas con 3-ABA. CTCF no disminuye sobre el *insulator*  $\alpha$ EHS-1.4 (gel representativo a la izquierda, cuantificación en el gráfico derecho, n=3) (P2 representa el par de primers que amplifican la región de unión de CTCF en el *insulator*  $\alpha$ EHS-1.4 y control representa el par de oligos que amplifican la región control, ver Tabla 1 y 2). Se muestran los promedios +/- el error estándar.

## **7.8 La alteración de la función del *insulator* $\alpha$ EHS-1.4 genera una configuración activa de la cromatina y una desregulación de la expresión de los genes $\alpha$ -globina adultos**

A continuación, evaluamos el efecto al inhibir la función del *insulator*  $\alpha$ EHS-1.4 por medio del tratamiento con 3-ABA a nivel de la estructura de la cromatina del dominio y de la expresión de los genes  $\alpha$ -globina (los ensayos de RT-PCRs se realizaron en colaboración con Georgina Guerrero).

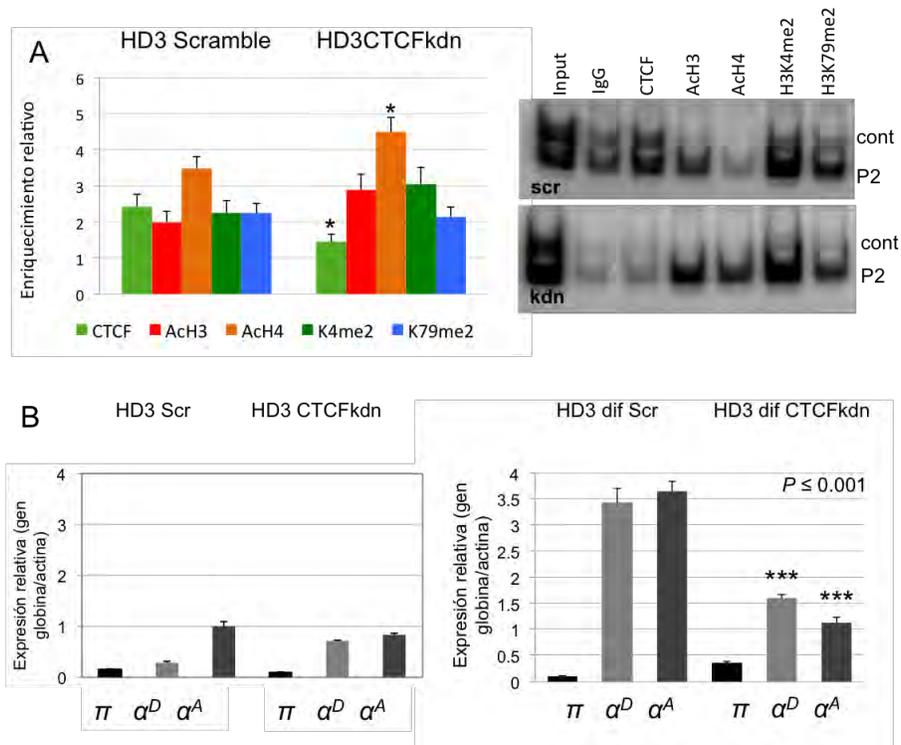
Al evaluar el estado de la cromatina en la región del *insulator*  $\alpha$ EHS-1.4, observamos que de manera general, las marcas post-traduccionales de las histonas que corresponden a marcas de cromatina permisiva, presentan un enriquecimiento en las células tratadas con 3-ABA en relación a las células HD3 sin tratamiento (Fig. 39A). Posteriormente analizamos la expresión de los genes  $\alpha$ -globina y notamos que existe un incremento en la expresión de los genes  $\alpha$ -globina adultos en las células tratadas con 3-ABA (HD3 3-ABA) en relación con las células HD3 sin tratamiento en donde existe una expresión basal de los genes (Fig. 39B, izquierda). Cuando analizamos la expresión de los genes en condiciones diferenciadas con y sin tratamiento (HD3d 3-ABA y HD3d, respectivamente), no observamos diferencias significativas en la activación de la expresión de los genes adultos (Fig. 39B, derecha).



**Figura 39. Estructura de la cromatina y la expresión de los genes  $\alpha$ -globina en células tratadas con 3-ABA. A.** ChIP contra marcas post-traduccionales de histonas que corresponden a cromatina permisiva en células HD3 y HD3 tratadas con 3-ABA (HD3 3-ABA) sobre el *insulator*  $\alpha$ EHS-1.4 endógeno (P2 y control representan los pares de oligonucleótidos utilizados para amplificar la región de unión de CTCF en el *insulator*  $\alpha$ EHS-1.4 y la región control, respectivamente. Ver Tabla 1 y 2) (n=3, se muestra el promedio +/- la desviación estándar). **B.** Expresión de los genes  $\alpha$ -globina en células HD3 y HD3 diferenciadas (HD3 dif) con y sin tratamiento con 3-ABA (n=9, se muestra el promedio +/- la desviación estándar).

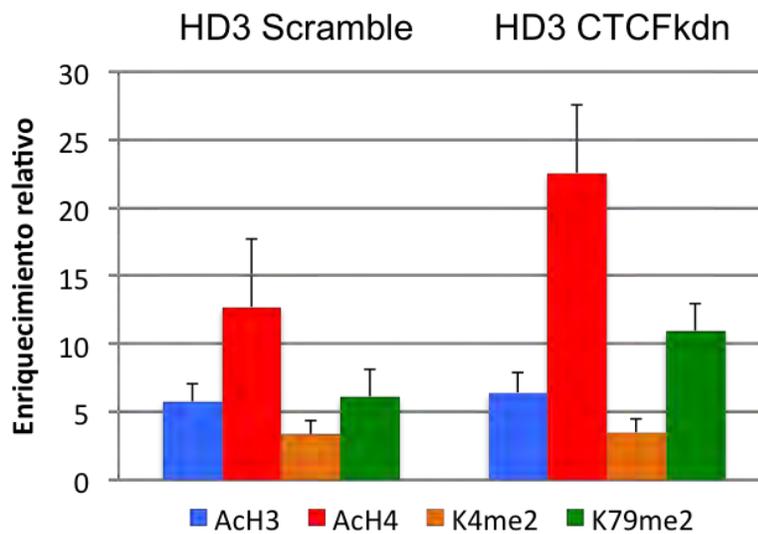
Posteriormente realizamos el mismo conjunto de experimentos para analizar la estructura de la cromatina sobre la región del *insulator*  $\alpha$ EHS-1.4 y la expresión de los genes  $\alpha$ -globina en las células *knockdown* para CTCF. Al estudiar diferentes marcas post-traduccionales de las histonas correspondientes a cromatina permisiva, observamos como cuando los niveles de CTCF disminuyen, existe un incremento general de dichas marcas en relación a las células HD3 control (Fig. 40A), de la misma forma que al tratar a las células con 3-ABA. De manera interesante, cuando analizamos la expresión de los genes  $\alpha$ -globina notamos que en este caso, la reducción de CTCF genera una disminución significativa en la

expresión de los genes adultos en condiciones diferenciadas (HD3 dif 3-ABA) en relación a las condiciones control (HD3 dif) (Fig. 40). Sin embargo en los eritroblastos *knockdown* para CTCF (HD3 vs HD3 CTCFkdn) no hay cambios significativos en la expresión de los genes con respecto al control a diferencia de lo observado en el tratamiento con 3-ABA.



**Figura 40. Estructura de la cromatina y la expresión de los genes  $\alpha$ -globina en células *knockdown* contra CTCF. A.** Ensayo de ChIP contra marcas post-traduccionales de histonas que corresponden a cromatina permisiva sobre la región del *insulator*  $\alpha$ EHS-1.4 (P2) (n=6, se muestran los promedios +/- el error estándar) en células HD3 transfectadas con el plásmido control (pScramble) y las células *knockdown* contra CTCF (HD3 CTCFkdn). Gel representativo de lado derecho (P2 y control representan los pares de oligonucleótidos utilizados para amplificar la región de unión de CTCF en el *insulator*  $\alpha$ EHS-1.4 y la región control, respectivamente. Ver Tabla1 y 2). **B.** Expresión de los genes  $\alpha$ -globina en eritroblastos (HD3 Scr) y eritroblastos diferenciados (HD3-dif Scr) transfectados con el plásmido control (pSil-scramble) y células *knockdown* contra CTCF (HD3 CTCFkdn y HD3 dif CTCFkdn) (n=9, se muestran los promedios +/- el error estándar).

Además, analizamos el perfil de marcas post-traduccionales de histonas sobre una región localizada 3' del *insulator*  $\alpha$ EHS-1.4 que corresponde al elemento  $\alpha$ CHS-1.2 en las células HD3 vs células HD3 CTCFkdn (Fig. 41) y al igual que en la región del *insulator*  $\alpha$ EHS-1.4, notamos un enriquecimiento de las marcas de apertura sobre esta región.



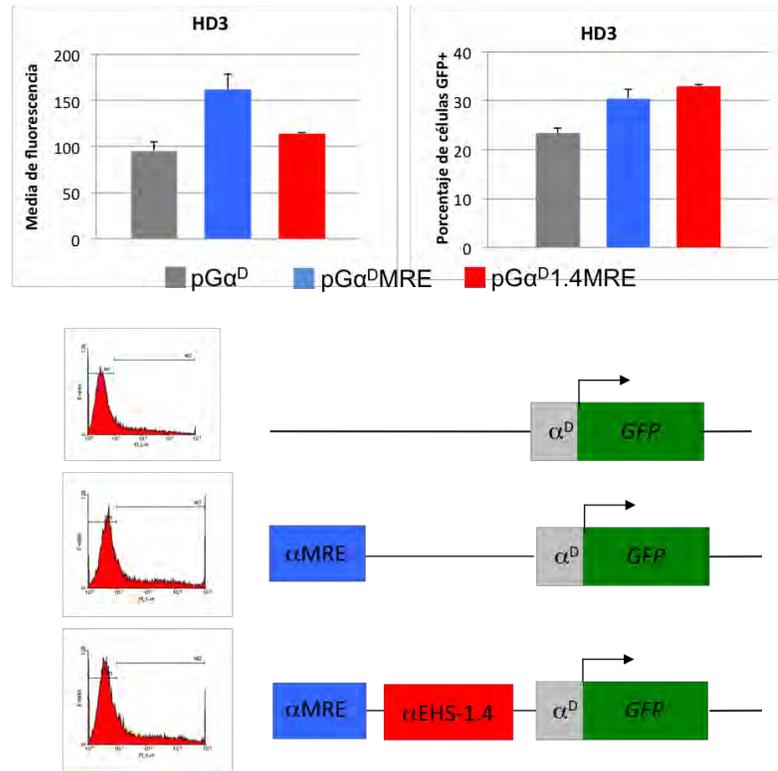
**Figura 41. Análisis del perfil de marcas post-traduccionales de histonas sobre la región 3' del *insulator*  $\alpha$ EHS-1.4 correspondiente al  $\alpha$ CHS-1.2.** Existe una relajación en la cromatina de la zona al reducir los niveles de CTCF (n=6, se muestran los promedios +/- el error estándar) al igual que lo observado en el *insulator*  $\alpha$ EHS-1.4.

Cuando analizamos los resultados anteriores en el contexto del dominio  $\alpha$ -globina nos sugieren que el *insulator*  $\alpha$ EHS-1.4 juega un papel en el estadio de eritroblastos adulto previniendo la trans-activación de los genes por parte de el elemento  $\alpha$ MRE localizado río arriba de éste (ver Antecedentes, sección 2.2.1). Cuando se altera la función de CTCF mediante la inhibición de PARP-1 o el abatimiento de sus niveles de expresión, existe una apertura prematura de la cromatina en la región del  $\alpha$ EHS-1.4 y regiones aledañas. Al analizar los niveles de expresión de los genes  $\alpha$ -globina de pollo en células HD3 tratadas con 3-ABA notamos una trans-activación incorrecta de los genes adultos lo cual correlaciona con la relajación de la cromatina sobre el *insulator*. Sin embargo, los resultados para la condición *knockdown* de CTCF fueron diferentes. En este caso, en las células HD3 no hubieron diferencias significativas en la expresión de los genes pero, al inducir su diferenciación a eritrocitos maduros, la trans-activación de los genes se pierde dramáticamente (Fig. 40B, derecha). Estas diferencias nos sugieren que, al reducir los niveles de CTCF, la estructura del dominio debe alterarse drásticamente ya que el *insulator*  $\alpha$ EHS-1.4 no es el único sitio de unión de CTCF en el dominio (Martin *et al.*,

sometido), de tal forma que aun induciendo la diferenciación de los eritroblastos (HD3) los genes adultos no consiguen trans-activarse. En cambio, es probable que no todos los CTCFs unidos al dominio  $\alpha$ -globina de pollo sean modificados post-traduccionalmente por PARP-1, por lo cual, al inhibir su actividad, generamos una estructura del dominio similar a la de eritrocitos maduros, favoreciendo así una trans-activación prematura de los genes adultos (ver Fig. 44).

## **7.9 La transactivación génica por parte del $\alpha$ MRE se potencia ante la disociación de CTCF del insulador $\alpha$ EHS-1.4**

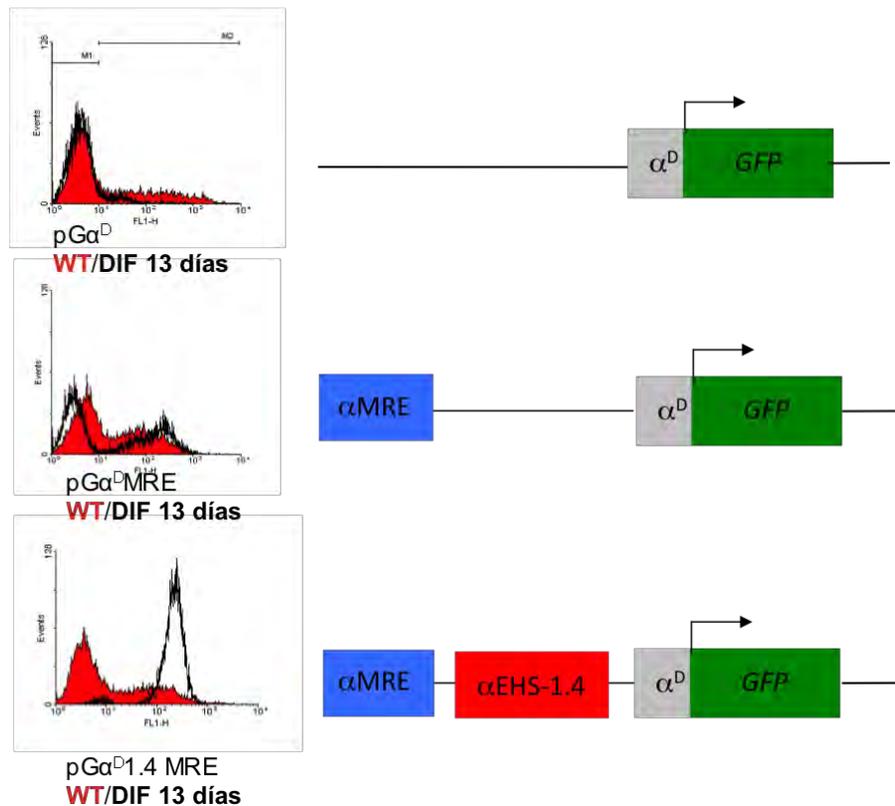
Para profundizar en la hipótesis generada a partir de los resultados mostrados previamente, generamos construcciones con el elemento  $\alpha$ MRE en presencia y ausencia del *insulador*  $\alpha$ EHS-1.4 (éstas fueron realizadas en colaboración con Georgina Guerrero y Edgar González Buendía en el laboratorio).



**Figura 42. Transfecciones transitorias evaluando la acción del elemento  $\alpha$ MRE en presencia y ausencia del *insulator*  $\alpha$ EHS-1.4.** En la parte superior se muestran gráficos de la media de fluorescencia (izquierda) y el porcentaje de células GFP<sup>+</sup> (derecha) presentada por la población transfectada de manera transitoria con cada una de las construcciones indicadas. En la parte inferior se muestran los histogramas de FACS de las poblaciones transfectadas con cada construcción. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células.

Al realizar transfecciones transitorias en células HD3 de las tres construcciones generadas (Fig. 42) , pudimos observar que el elemento  $\alpha$ MRE incrementa la media de fluorescencia de la población así como el porcentaje de células fluorescentes en relación con el plásmido pGα<sup>D</sup> que únicamente posee al promotor  $\alpha$ <sup>D</sup>. Al incluir al *insulator*  $\alpha$ EHS-1.4 entre el elemento  $\alpha$ MRE y el gen reportero, observamos una disminución en la media de la fluorescencia de la población que refleja el bloqueo de *enhancer* (el  $\alpha$ MRE) ejercido por el *insulator* (Fig. 42). Sin embargo el porcentaje de células fluorescentes no se modifica por la presencia del *insulator*. Finalmente, decidimos realizar las mismas transfecciones pero de manera estable para así poder someter a las células HD3

(eritroblastos) transfectantes al protocolo de diferenciación hacia eritrocitos maduros ya que previamente describimos que CTCF se desprende del *insulator*  $\alpha$ EHS-1.4 en esta condición así como en eritrocitos de 10d en donde la cromatina de la región está abierta y suponemos el elemento  $\alpha$ MRE es más activo (Fig 42).



**Figura 43. Transfecciones estables evaluando la acción del elemento  $\alpha$ MRE en presencia y ausencia del *insulator*  $\alpha$ EHS-1.4.** A la izquierda se muestran los histogramas mostrando el nivel de fluorescencia de las poblaciones transfectantes con cada una de las construcciones (a la izquierda). En rojo se muestra el histograma de las células HD3 silvestre y en negro el histograma de las células sometidas a diferenciación por 13 días. En los histogramas de FACS el eje de las abscisas corresponde al nivel de fluorescencia y el de las ordenadas al número de células.

Al realizar las transfecciones de manera estable en las células HD3 y someterlas a diferenciación, pudimos observar que en la construcción que contiene al *insulator*  $\alpha$ EHS-1.4 entre el  $\alpha$ MRE y el gen reportero, existe una trans-activación pronunciada de éste en relación a la misma construcción en células HD3 sin diferenciar (Fig. 43). Al analizar la construcción que sólo contiene al elemento  $\alpha$ MRE con el gen reportero, notamos que

existe una trans-activación al diferenciar a las células HD3. Sin embargo, esta es menor a la observada en presencia del *insulator*  $\alpha$ EHS-1.4. El control pG $\alpha^D$  se comportó como previamente descrito mostrando una disminución de la fluorescencia de la población al diferenciar a las células (Fig. 43). Los resultados previos sugieren que, al diferenciar a las células la conformación  $\alpha$ MRE- $\alpha$ EHS-1.4 (que asemeja a la configuración endógena de los elementos) favorece la trans-activación del gen reportero y esto coincide con la pérdida de bloqueo (pérdida de CTCF) del *insulator*  $\alpha$ EHS-1.4 en este estadio. Por otro lado, los resultados anteriores muestran como el elemento  $\alpha$ MRE por si sólo (por lo menos el fragmento de 500 pb utilizado en estos experimentos) genera una trans-activación moderada de la expresión de la *GFP* en condiciones estables y que el *insulator*  $\alpha$ EHS-1.4 favorece la trans-activación por parte del  $\alpha$ MRE en el estadio terminalmente diferenciado.

## 8. CONCLUSIONES

La expresión génica se encuentra regulada a múltiples niveles. Diferentes elementos de regulación contenidos dentro de un dominio cromatínico actúan de manera coordinada para conseguir la correcta expresión de genes en el tiempo y espacio adecuados. En el presente trabajo de investigación abordamos la contribución del *insulator*  $\alpha$ EHS-1.4 en la regulación de la estructura de la cromatina y de la expresión génica al interior del dominio  $\alpha$ -globina de pollo.

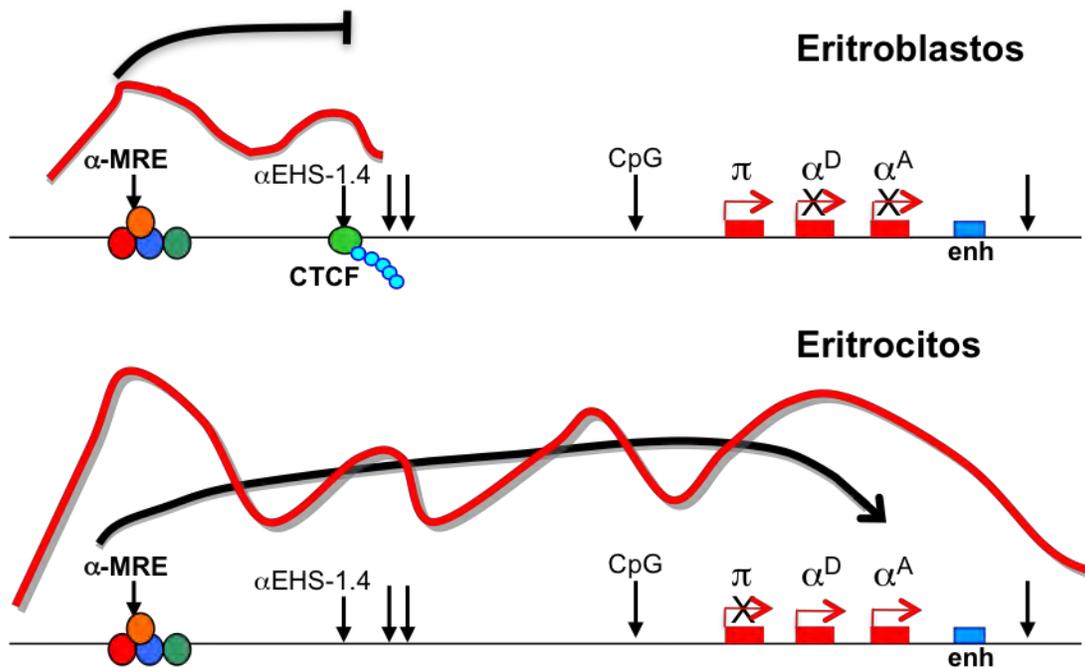
Trabajo previo del laboratorio demostró que el *insulator*  $\alpha$ EHS-1.4 y el elemento  $\alpha$ CHS-1.2 tienen la capacidad de bloqueo de *enhancer* en ensayos transitorios (Valadés-Graham *et al.*, 2004). Los resultados del presente trabajo muestran que además, el *insulator*  $\alpha$ EHS-1.4 tiene la capacidad de contender contra efectos de posición cromosomal tanto en líneas celulares de pollo como en ratones transgénicos, mientras que el elemento  $\alpha$ CHS-1.2 actúa como un represor del gen reportero en transfecciones estables. A partir de estos resultados, decidimos enfocarnos en el elemento  $\alpha$ EHS-1.4. Sin embargo, será interesante estudiar la contribución del elemento  $\alpha$ CHS-1.2 por si mismo y en conjunto con el  $\alpha$ EHS-1.4 y el otro sitio de hipersensibilidad de la región a la regulación del dominio, ya que diversos datos del laboratorio han sugerido que puede jugar un papel relevante en la regulación del dominio. Por un lado, al analizar la estructura de la cromatina a lo largo de la región identificamos un pico de enriquecimiento en marcas de histonas correspondientes a cromatina permisiva sobre este elemento en el estadio de eritrocitos de 10 días de desarrollo. Esto sugiere que en el estadio de eritrocitos terminalmente diferenciados en esta región pueden existir factores activos que jueguen un papel relevante en la regulación del dominio. Por otro lado, datos de Cristian Arriaga (estudiante de doctorado adscrito al laboratorio), han mostrado la presencia de transcritos intergénicos no-codificantes a lo largo del dominio  $\alpha$  globina. De manera interesante, el transcrito en sentido de los genes globina parece comenzar en la región del elemento  $\alpha$ CHS-1.2. Será importante estudiar la regulación de dicho transcrito en relación a la estructura de la cromatina y a la actividad observada para el elemento  $\alpha$ CHS-1.2.

Nuestros datos muestran que el *insulator*  $\alpha$ EHS-1.4 posee tanto la capacidad de bloqueo de *enhancer* como la de contender contra efectos de posición cromosomal. Inicialmente se asumía que estas dos propiedades eran independientes ya que éste es el caso del *insulator* 5' 5' cHS4 del dominio  $\beta$ -globina sobre el cual se realizaron los estudios pioneros acerca de *insulators* en vertebrados (Recillas-Targa *et al.*, 1999). En el *insulator* 5' cHS4 el factor asociado a la protección contra CPE es USF (West *et al.*, 2004). Sin embargo, el *insulator*  $\alpha$ EHS-1.4 presenta ambas propiedades y éstas parecen atribuirse, al menos en parte, al factor CTCF. Lo anterior sugiere que, para el caso del *insulator*  $\alpha$ EHS-1.4 y probablemente para varios otros *insulators*, ambas capacidades pueden derivarse del mismo mecanismo en donde el bloqueo de señales, tanto de apertura (cómo en el caso del contexto endógeno del *insulator*  $\alpha$ EHS-1.4), como de cierre de la cromatina (como en el caso de los transgenes integrados de manera estable y aleatoria en el genoma), explican la protección observada. El bloqueo de señales que generan una cromatina permisiva se ha reportado para otros *insulators* (Zhao *et al.*, 2004; Gomes y Espinosa, 2010).

Por otro lado, se ha sugerido también que la capacidad de CTCF para asociarse a sí mismo u otros factores y formar asas de cromatina, puede ser la función que subyace a todas las propiedades que se han atribuido a dicho factor y a los *insulators* a los que se asocia (revisado en Phillips y Corces, 2009). Sin embargo, al parecer este no es el caso del *insulator*  $\alpha$ EHS-1.4. Trabajo reciente de Razin y colaboradores en donde se estudió la arquitectura en asas del dominio  $\alpha$ -globina tanto en células precursoras eritroides como en células terminalmente diferenciadas, muestra que el *insulator*  $\alpha$ EHS-1.4 no forma parte de los puntos de anclaje de las asas que comprenden al dominio (Gavrilov y Razín, 2008). Sin embargo, en este trabajo no se abordó la función de CTCF en dichas conformaciones. Es posible que otros sitios de unión a CTCF en el dominio detectados mediante CHIP-Seq (Martin *et al.*, sometido), participen en la estructuración estadio específica del locus.

Al estudiar la unión de CTCF en el *insulator*  $\alpha$ EHS-1.4 durante la diferenciación eritroide, observamos que la abundancia relativa de CTCF varía dependiendo del estadio analizado. Esto sugiere que el *insulator*  $\alpha$ EHS-1.4 actúa de manera estadio específica

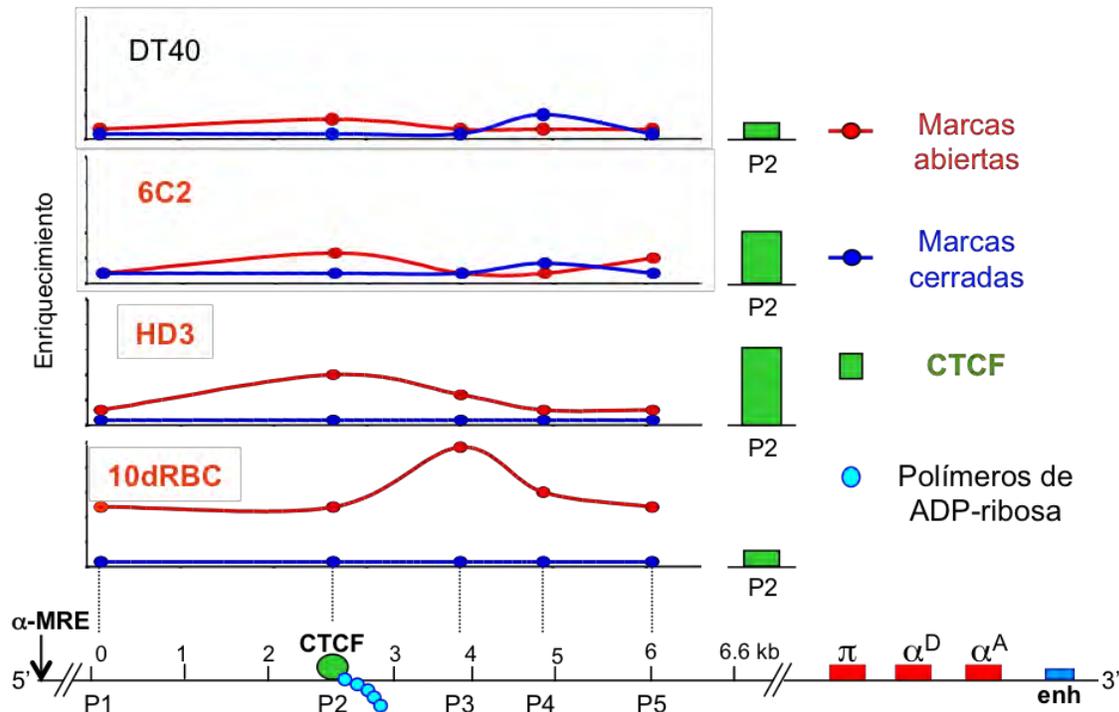
bloqueando las señales provenientes del  $\alpha$ MRE en la etapa de eritroblastos adultos. En eritrocitos maduros en cambio, CTCF ya no se encuentra unido al *insulator* favoreciendo la trans-activación de los genes adultos por parte del  $\alpha$ MRE (ver modelo Fig. 44).



**Figura 44. Modelo de acción del *insulator*  $\alpha$ EHS-1.4 dentro del dominio  $\alpha$ -globina de pollo.** En el panel superior se muestra la conformación del dominio en eritroblastos adultos (HD3) en donde ninguno de los genes globina (cuadros rojos) tienen expresión. CTCF se encuentra unido al *insulator* por lo cual las señales de apertura provenientes del LCR se bloquean y no transactivan a los genes adultos. En el panel inferior se muestra la situación de los eritrocitos adultos terminalmente diferenciados en donde existe expresión de los genes adultos. En ausencia de CTCF el LCR ahora es capaz de transactivar la expresión de estos genes. La línea roja representa las marcas de apertura de la cromatina y su comportamiento en los dos estadios. Se muestran a los elementos enhancer 3' (caja azul) y a la isla CpG (que corresponde al promotor del gen antisentido *C16orf35*, ver antecedentes, sección 2.2.1).

El modelo anterior se sustenta con base en diversas evidencias. Una de ellas es la estructura de la cromatina en la región del *insulator* en los diferentes estadios de diferenciación (datos de Eria Rebollar, en la figura 45 se muestra un resumen de estos datos). Las marcas de cromatina permisiva muestran un pico de enriquecimiento sobre el *insulator*  $\alpha$ EHS-1.4 en la etapa de eritroblastos que correlaciona con la presencia de CTCF. Cuando la diferenciación procede y CTCF se disocia del *insulator*, la estructura de la

cromatina se relaja de manera generalizada en la región, mostrando una zona de enriquecimiento que se recorre en dirección de los genes globina.



**Figura 45. Resumen de los perfiles de modificaciones post-traduccionales de las histonas y CTCF en la región de 6.6 kb que contiene al *insulator*  $\alpha$ EHS-1.4.** A medida que la diferenciación eritroide avanza existe un aumento en las marcas post-traduccionales de histonas que corresponden a cromatina permisiva sobre el *insulator*  $\alpha$ EHS-1.4 y el sitio de unión de CTCF (P2) que se expanden e incrementan aun más en el estudio de eritrocitos de 10 días (10dRBC) en donde los genes globina adultos se expresan. La presencia de CTCF sobre el *insulator* aumenta hasta llegar al punto más alto en eritroblastos para después desaparecer en el estadio terminalmente diferenciado.

Otra evidencia proviene de las consecuencias de perturbar o abatir la función de CTCF. Al alterar la función de CTCF, las marcas de cromatina permisiva en la región del *insulator*  $\alpha$ EHS-1.4 aumentan y los genes globina adultos se trans-activan de manera prematura semejando al estadio de eritrocitos maduros. Al realizar el *knockdown* de CTCF, también se observa una preapertura de la región del *insulator* similar a la observada en el tratamiento con 3-ABA. Sin embargo, en este caso no existe una trans-activación prematura de los genes adultos, sino una dramática disminución en su expresión en el estadio diferenciado, lo cual sugiere que, al reducir la abundancia de CTCF el efecto es

más drástico ya que probablemente, estamos afectando la conformación general del dominio y no solo a los CTCFs sujetos a poli(ADPribosil)ación como el  $\alpha$ EHS-1.4.

Los datos previos sugieren que los *insulators* dependientes de CTCF pueden regularse. Como se ha descrito en la sección acerca de CTCF en la introducción, éste se encuentra sujeto a distintas modificaciones post-traduccionales como la poli(ADPribosil)ación y posee distintas parejas de interacción que son necesarias para llevar a cabo su función como son las cohesinas y la proteína CHD8. Asimismo, se ha observado que la unión de CTCF a su secuencia blanco en el ADN también está sujeta a regulación, en particular desestabilizándose ante el paso de la RNA Polimerasa II. La presencia de transcritos intergénicos que recorren la región del *insulator*  $\alpha$ EHS-1.4 (datos no publicados del laboratorio), hacen de éste un mecanismo posible de regulación de la unión de CTCF en el *insulator*  $\alpha$ EHS-1.4. Es importante recalcar que, todos los procesos descritos regulan la función de CTCF sólo en un subconjunto de los sitios a donde éste se une. Por lo tanto, la combinatoria de mecanismos de regulación posiblemente genere subtipos de *insulators* que pueden tener características particulares. Esta noción es muy novedosa y existe muy poca información con respecto a sus consecuencias funcionales. Las diferencias observadas en nuestros datos con respecto a la inhibición de la poli(ADPribosil)ación en comparación con el abatimiento general de CTCF mediante un *knockdown*, favorecen la visión de que, al menos dentro del dominio  $\alpha$ -globina del pollo, existen distintos *insulators* (probablemente formados por complejos diferentes) y no todos ellos sufren modificaciones por parte de PARP-1. Dicha hipótesis se encuentra en estudio actualmente en el laboratorio (ver perspectivas). La modulación de los diferentes CTCFs al interior del dominio pueden jugar un papel central en estructurar y dirigir a los diferentes elementos de regulación a su gen o genes blanco en el espacio y tiempo adecuados. Esta hipótesis propone a los *insulators* como participantes activos en la regulación de la expresión génica al interior de un dominio en lugar de representar elementos pasivos que sólo delimitan y aíslan a loci completos de manera constitutiva.

Finalmente, otra evidencia que apoya fuertemente el modelo presentado proviene de las transfecciones estables con las construcciones que incluyen al  $\alpha$ MRE, el *insulator*  $\alpha$ EHS-1.4 y al gen reportero *GFP*. La trans-activación de la expresión de *GFP* en las células diferenciadas con dicha construcción en comparación con la trans-activación observada en presencia del  $\alpha$ MRE sólo (Fig. 43), sugieren que el elemento  $\alpha$ EHS-1.4 una vez que CTCF se desprende, actúa favoreciendo la estructuración necesaria de la región para promover la trans-activación por parte del  $\alpha$ MRE. Dicha propiedad es dependiente de la presencia del  $\alpha$ MRE ya que, en las líneas que contienen sólo al *insulator*  $\alpha$ EHS-1.4 sometidas a diferenciación, la expresión del *GFP* se pierde (Fig. 43). Los datos anteriores proponen que el *insulator*  $\alpha$ EHS-1.4 contribuye, en eritrocitos terminalmente diferenciados, en promover la estructuración correcta del dominio para que la trans-activación robusta de los genes adultos por parte del  $\alpha$ MRE se lleve a cabo. Lo anterior propone al *insulator*  $\alpha$ EHS-1.4 como un elemento de regulación dual que actúa dependiendo del estadio de diferenciación.

## 9. PERSPECTIVAS

A partir de los resultados obtenidos en este trabajo, surgen diversos aspectos interesantes que deben ser abordados para profundizar y complementar la visión actual de la función del *insulator*  $\alpha$ EHS-1.4 y otros elementos de regulación en el dominio  $\alpha$ -globina de pollo.

Por un lado, una de las perspectivas principales que surge de los resultados presentados será analizar la participación de los tres sitios de hipersensibilidad contenidos en la región de 6.6 kb (al  $\alpha$ EHS-1.4, el  $\alpha$ CHS-1.2 y el  $\alpha$ EHS-1.0) en conjunto sobre la regulación del dominio  $\alpha$ -globina, tanto a nivel de la estructura de la cromatina como de la expresión de los genes globina y de los transcritos intergénicos que recorren al dominio. En el contexto endógeno, los sitios de hipersensibilidad son contiguos y, por los datos mostrados previamente, el *insulator*  $\alpha$ EHS-1.4 y el elemento  $\alpha$ CHS-1.2 poseen propiedades contrarias (al menos en transfecciones estables). Utilizando el sistema de transferencia de cromosomas (establecido en el laboratorio por Georgina Guerrero), ó cromosomas bacterianos (BACS) que contienen al dominio completo (trabajo de Gabriel Moreno y Christian Valdés en el laboratorio), será posible eliminarlos en conjunto y de manera independiente estos elementos de regulación y analizar el efecto de su ausencia en la función del dominio durante la diferenciación eritroide.

Con respecto a la función específica del *insulator*  $\alpha$ EHS-1.4, a continuación se deberá abordar cual es el elemento mínimo del *insulator* y que otros factores (en particular factores eritroides), además de CTCF, participan en su función. Además será interesante abordar si la composición de el complejo presente en el  $\alpha$ EHS-1.4 varía dependiendo del estadio de diferenciación. Posiblemente este sea el caso ya que por un lado, la unión de CTCF es dinámica y por otro, nuestros datos muestran que la función del  $\alpha$ EHS-1.4 es diferente dependiendo del estadio de diferenciación.

Hemos observado que la actividad del CTCF asociado al *insulator*  $\alpha$ EHS-1.4 se encuentra modulada por PARilación. Resultará interesante estudiar que *insulators* en el

dominio asociados a CTCF (datos no publicados) presentan este tipo de modificación y asociar la combinatoria de factores a la presencia y ausencia de ésta. Esta propuesta surge de la noción de que los polímeros de ADP pueden funcionar como una plataforma de reclutamiento de factores específicos y probablemente, la identidad de los complejos presentes varíe en cada caso. Para abordar lo anterior pueden realizarse ChIPs contra PARP-1 (la enzima encargada de realizar la poli(ADPribosil)ación de diversos factores incluido CTCF) ya que se ha visto que dicha enzima normalmente se recluta a los sitios en donde existen factores poli(ADPribosil)ados como el caso del *insulator*  $\alpha$ EHS-1.4. Esta estrategia se está llevando a cabo en el laboratorio.

Adicionalmente, será importante analizar las consecuencias de abatir los niveles de CTCF a nivel de la estructura tridimensional del dominio (por medio de 3C) así como comparar dicho efecto con el generado por la inhibición de la PARP-1 ya que, dados nuestros resultados, es posible que las conformaciones que se obtengan sean diferentes.

Por otro lado y como se ha comentado previamente, la región del  $\alpha$ EHS-1.4 se transcribe. Será interesante estudiar si la transcripción de la zona tiene una participación activa en la función del *insulator*, en particular en la regulación de la unión de CTCF.

Finalmente, será muy relevante estudiar la función del  $\alpha$ MRE de manera independiente y en conjunto con el  $\alpha$ EHS-1.4 a lo largo de la diferenciación eritroide y el desarrollo del pollo ya que nuestros datos sugieren que existe una interrelación en la función de ambos elementos. Diversos datos provenientes de nuestro grupo y otros (Razin y colaboradores, Higgs y colaboradores) muestran que la regulación de los genes  $\alpha$ -globina de pollo es compleja y nuestra hipótesis actual sugiere que cada gen se encuentra modulado por un elemento de regulación particular a diferencia de lo que ocurre en el dominio  $\alpha$ -globina en humano. Además, se ha reportado recientemente en pollo, que el dominio eritroide se extiende en su costado 3' no-codificante (más de 12 kb) incluyendo al gen TMEM-8 adyacente (Philonenko *et al.*, 2009) dentro del conjunto de asas de cromatina que lo abarcan. Resultará muy interesante entender como es que los elementos de regulación como el  $\alpha$ EHS-1.4 (y el resto de los sitios de hipersensibilidad

descritos en el dominio), contribuyen en regular la acción de los *enhancers* ( $\alpha$ MRE y *enhancer 3'* ) sobre los genes y organizar al dominio durante la diferenciación eritroide y el desarrollo del pollo.

## 10. REFERENCIAS

1. Ahmad y Henikoff. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* (2002) vol. 9 (6) pp. 1191-200
2. Allis, Jenuwein, Reinberg y Caparros. *Epigenetics*. Cold Spring Harbor Laboratory Press (2006) pp. 15-16, NY.
3. Anguita *et al.* Identification of a conserved erythroid specific domain of histone acetylation across the alpha-globin gene cluster. *Proc Natl Acad Sci USA* (2001) vol. 98 (21) pp. 12114-9
4. Brown *et al.* Association between active genes occurs at nuclear speckles y is modulated by chromatin environment. *J Cell Biol* (2008) vol. 182 (6) pp. 1083-97
5. Brown *et al.* Coregulated human globin genes are frequently in spatial proximity when active. *J Cell Biol* (2006) vol. 172 (2) pp. 177-87
6. Bulger y Groudine. *Enhancers: the abundance y function of regulatory sequences beyond promoters*. *Developmental Biology* (2010) vol. 339 (2) pp. 250-7
7. Campos y Reinberg. Histones: annotating chromatin. *Annu Rev Genet* (2009) vol. 43 pp. 559-99
8. Capelson y Corces. Boundary elements and nuclear organization. *Biol Cell* (2004) vol. 96 (8) pp. 617-29
9. Chao *et al.* CTCF, a cyidate trans-acting factor for X-inactivation choice. *Science* (2002) vol. 295 (5553) pp. 345-7
10. Chen *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* (2008) vol. 133 (6) pp. 1106-17

11. Chernukhin *et al.* CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Molecular and Cellular Biology* (2007) vol. 27 (5) pp. 1631-48
12. Clapier y Cairns. The biology of chromatin remodeling complexes. *Annu Rev Biochem* (2009) vol. 78 pp. 273-304
13. Cook. A Model for all Genomes: The Role of Transcription Factories. *J Mol Biol* (2010) vol. 395 (1) pp. 1-10
14. Cosgrove *et al.* Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* (2004) vol. 11 (11) pp. 1037-43
15. Cremer y Cremer. Chromosome territories. *Cold Spring Harb Perspect Biol* (2010) vol. 2 (3) pp. a003889
16. Cuddapah *et al.* Global analysis of the *insulator* binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Research* (2009) vol. 19 (1) pp. 24-32
17. De Gobbi *et al.* Tissue-specific histone modification and transcription factor binding in alpha globin gene expression. *Blood* (2007) vol. 110 (13) pp. 4503-10
18. De La Rosa-Velázquez *et al.* Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res* (2007) vol. 67 (6) pp. 2577-85
19. Dekker *et al.* Capturing chromosome conformation. *Science* (2002) vol. 295 (5558) pp. 1306-11
20. Farrar *et al.* Mutational Analysis of the Poly(ADP-ribosyl)ation Sites of the Transcription Factor CTCF Provides an Insight into the Mechanism of its Regulation by Poly(ADP-ribosyl)ation. *Molecular and Cellular Biology* (2009) pp.

21. Fathallah *et al.* Epigenetic analysis of the human alpha- y beta-globin gene clusters. *Blood Cells Mol Dis* (2008) vol. 40 (2) pp. 166-73
22. Felsenfeld y Groudine. Controlling the double helix. *Nature* (2003) vol. 421 (6921) pp. 448-53
23. Filippova *et al.* CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive *insulator* at the DM1 locus. *Nat Genet* (2001) vol. 28 (4) pp. 335-43
24. Flint *et al.* Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. *Hum Mol Genet* (2001) vol. 10 (4) pp. 371-82
25. Garrick *et al.* The role of the polycomb complex in silencing alpha-globin gene expression in nonerythroid cells. *Blood* (2008) vol. 112 (9) pp. 3889-99
26. Gaszner y Felsenfeld. *Insulators*: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* (2006) vol. 7 (9) pp. 703-13
27. Gavrilov *et al.* Mapping of the nuclear matrix-bound chromatin hubs by a new M3C experimental procedure. *Nucleic Acids Research* (2010) pp.
28. Gavrilov y Razin. Spatial configuration of the chicken  $\alpha$ -globin gene domain: immature and active chromatin hubs. *Nucleic Acids Research* (2008) pp. 12
29. Gomes y Espinosa. Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes & Development* (2010) vol. 24 (10) pp. 1022-34
30. Henikoff *et al.* Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet* (2004) vol. 20 (7) pp. 320-6
31. Hughes *et al.* Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. *Proc Natl Acad Sci USA* (2005) vol. 102 (28) pp. 9830-5

32. Ishihara *et al.* CTCF-dependent chromatin *insulator* is linked to epigenetic remodeling. *Mol Cell* (2006) vol. 23 (5) pp. 733-42
33. Janicki *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* (2004) vol. 116 (5) pp. 683-98
34. Kiefer *et al.* Epigenetics of beta-globin gene regulation. *Mutat Res* (2008) vol. 647 (1-2) pp. 68-76
35. Klochkov *et al.* A CTCF-dependent silencer located in the differentially methylated area may regulate expression of a housekeeping gene overlapping a tissue-specific gene domain. *Molecular y Cellular Biology* (2006) vol. 26 (5) pp. 1589-97
36. Lefevre *et al.* The LPS-Induced Transcriptional Upregulation of the Chicken Lysozyme Locus Involves CTCF Eviction and Noncoding RNA Transcription. *Molecular Cell* (2008) vol. 32 (1) pp. 129-139
37. Li *et al.* Highly compacted chromatin formed in vitro reflects the dynamics of transcription activation in vivo. *Mol Cell* (2010) vol. 38 (1) pp. 41-53
38. Lieberman-Aiden *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* (2009) vol. 326 (5950) pp. 289-293
39. Litt *et al.* Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* (2001) vol. 293 (5539) pp. 2453-5
40. Margueron y Reinberg. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* (2010) vol. 11 (4) pp. 285-96
41. McKittrick *et al.* Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci USA* (2004) vol. 101 (6) pp. 1525-30
42. Mito *et al.* Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* (2005) vol. 37 (10) pp. 1090-7

43. Ohlsson *et al.* Does CTCF mediate between nuclear organization and gene expression?. *Bioessays* (2010) vol. 32 (1) pp. 37-50
44. Osborne *et al.* Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* (2004) vol. 36 (10) pp. 1065-71
45. Phillips y Corces. CTCF: Master Weaver of the Genome. *Cell* (2009) vol. 137 (7) pp. 1194-1211
46. Philonenko *et al.* TMEM8 - a non-globin gene entrapped in the globin web. *Nucleic Acids Research* (2009) pp. 13
47. Pikaart *et al.* Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by *insulators*. *Genes & Development* (1998) vol. 12 (18) pp. 2852-62
48. Potts et al. Chicken beta-globin 5'HS4 insulators function to reduce variability in transgenic founder mice. *Biochem Biophys Res Commun* (2000) vol. 273 (3) pp. 1015-8
49. Ragoczy y Groudine. Getting connected in the globin interactome. *Nat Genet* (2010) vol. 42 (1) pp. 16-17
50. Razin *et al.* Genomic domains and regulatory elements operating at the domain level. *Int Rev Cytol* (2003) vol. 226 pp. 63-125
51. Razin *et al.* Functional analysis of DNA sequences located within a cluster of DNase I hypersensitive sites colocalizing with a MAR element at the upstream border of the chicken alpha-globin gene domain. *J Cell Biochem* (1999) vol. 74 (1) pp. 38-49
52. Recillas-Targa *et al.* Prospects and implications of using chromatin *insulators* in gene therapy y transgenesis. *Bioessays* (2004) vol. 26 (7) pp. 796-807

53. Recillas-Targa *et al.* Position-effect protection and *enhancer* blocking by the chicken beta-globin *insulator* are separable activities. Proc Natl Acad Sci USA (2002) vol. 99 (10) pp. 6883-8
54. Recillas-Targa y Razin. Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. Crit Rev Eukaryot Gene Expr (2001) vol. 11 (1-3) pp. 227-42
55. Rister y Desplan. Deciphering the genome's regulatory code: the many languages of DNA. Bioessays (2010) vol. 32 (5) pp. 381-4
56. Ruthenburg *et al.* Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol (2007) vol. 8 (12) pp. 983-94
57. Schoenfelder *et al.* Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet (2010) vol. 42 (1) pp. 53-61
58. Sutherly y Bickmore. Transcription factories: gene expression in unions?. Nat Rev Genet (2009) vol. 10 (7) pp. 457-66
59. Talbert y Henikoff. Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol (2010) vol. 11 (4) pp. 264-75
60. Tang *et al.* Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. Progress in Biophysics and Molecular Biology (2010) vol. 102 (2-3) pp. 122-128
61. Theo Sijtse Palstra. Close encounters of the 3C kind: long-range chromatin interactions y transcriptional regulation. Briefings in Functional Genomics and Proteomics (2009) vol. 8 (4) pp. 297-309
62. Tsiftoglou *et al.* Erythropoiesis: model systems, molecular regulators, and developmental programs. IUBMB Life (2009) vol. 61 (8) pp. 800-30

63. Tufarelli *et al.* Comparative analysis of the alpha-like globin clusters in mouse, rat, and human chromosomes indicates a mechanism underlying breaks in conserved synteny. *Genome Research* (2004) vol. 14 (4) pp. 623-30
64. Valadez-Graham *et al.* CTCF-dependent *enhancer* blockers at the upstream region of the chicken alpha-globin gene domain. *Nucleic Acids Research* (2004) vol. 32 (4) pp. 1354-62
65. Vernimmen *et al.* Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS -40). *Blood* (2009) vol. 114 (19) pp. 4253-60
66. West y Fraser. Remote control of gene transcription. *Hum Mol Genet* (2005) vol. 14 Spec No 1 pp. R101-11
67. West *et al.* Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell* (2004) vol. 16 (3) pp. 453-63
68. Yu *et al.* Poly(ADP-ribosylation) regulates CTCF-dependent chromatin insulation. *Nat Genet* (2004) vol. 36 (10) pp. 1105-10
69. Zhao *et al.* Nuclear neighborhoods and gene expression. *Curr Opin Genet Dev* (2009) vol. 19 (2) pp. 172-9
70. Zhao y Dean. An *insulator* blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an *enhancer* and gene. *Nucleic Acids Research* (2004) vol. 32 (16) pp. 4903-19

## **11. ANEXO 1. Identificación de nuevos elementos de regulación en el dominio $\alpha$ -globina de pollo.**

De manera adicional al trabajo principal que se desarrolló durante este proyecto de investigación (presentado en la parte central de la tesis), se llevaron a cabo dos series de experimentos para estudiar la estructura global de la cromatina en los costados 5' y 3' no-codificantes del dominio  $\alpha$ -globina de pollo y tratar de identificar elementos de regulación adicionales a los previamente descritos en el dominio  $\alpha$ -globina de pollo.

Para llevar a cabo lo anterior, se realizaron ensayos de digestión con la enzima DNasa I (ver métodos). Dicho experimento permite analizar regiones grises (de varias kilobases) del genoma en cuanto a su conformación cromatínica global (que tan accesibles son al corte por la DNasa I) así como determinar si existen elementos de regulación reconocidos por ser hipersensibles al corte por esta enzima. Los experimentos de esta sección fueron realizados en colaboración con Eria Rebollar y Martín Escamilla en el laboratorio.

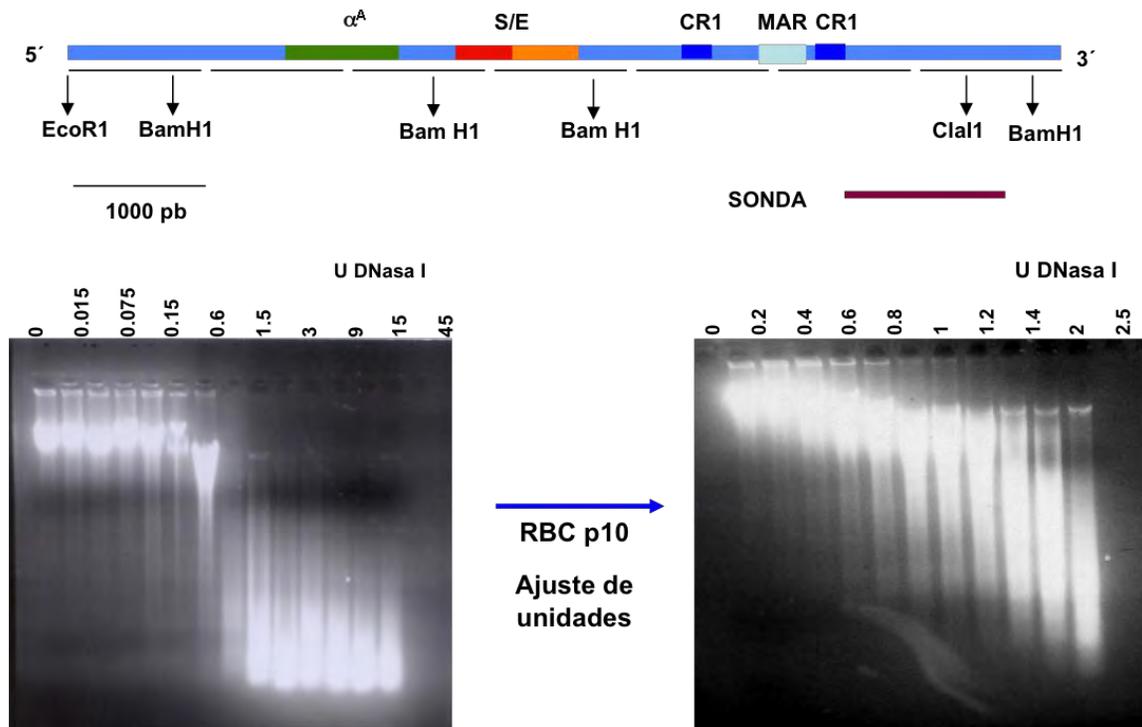
### **11.1 Caracterización de la región 3' no-codificante del dominio $\alpha$ -globina de pollo.**

La región 3' no-codificante del dominio  $\alpha$ -globina de pollo (zona río abajo del *enhancer* 3' del dominio) no ha sido estudiada y se desconoce si pudiera contener elementos de regulación adicionales que contribuyan a la regulación del dominio  $\alpha$ -globina o bien elementos tipo *insulator* que demarquen el límite 3' del mismo.

Experimentos realizados por Martín Escamilla en el laboratorio, mostraron la presencia de algunos sitios de hipersensibilidad alrededor de 1 kb río abajo del *enhancer* 3' (datos no mostrados). Con base en las observaciones anteriores, se decidió repetir el experimento para tratar de localizar de manera precisa a dichos sitios de hipersensibilidad y evaluar su posible función sobre la regulación del dominio  $\alpha$ -globina. Inicialmente se

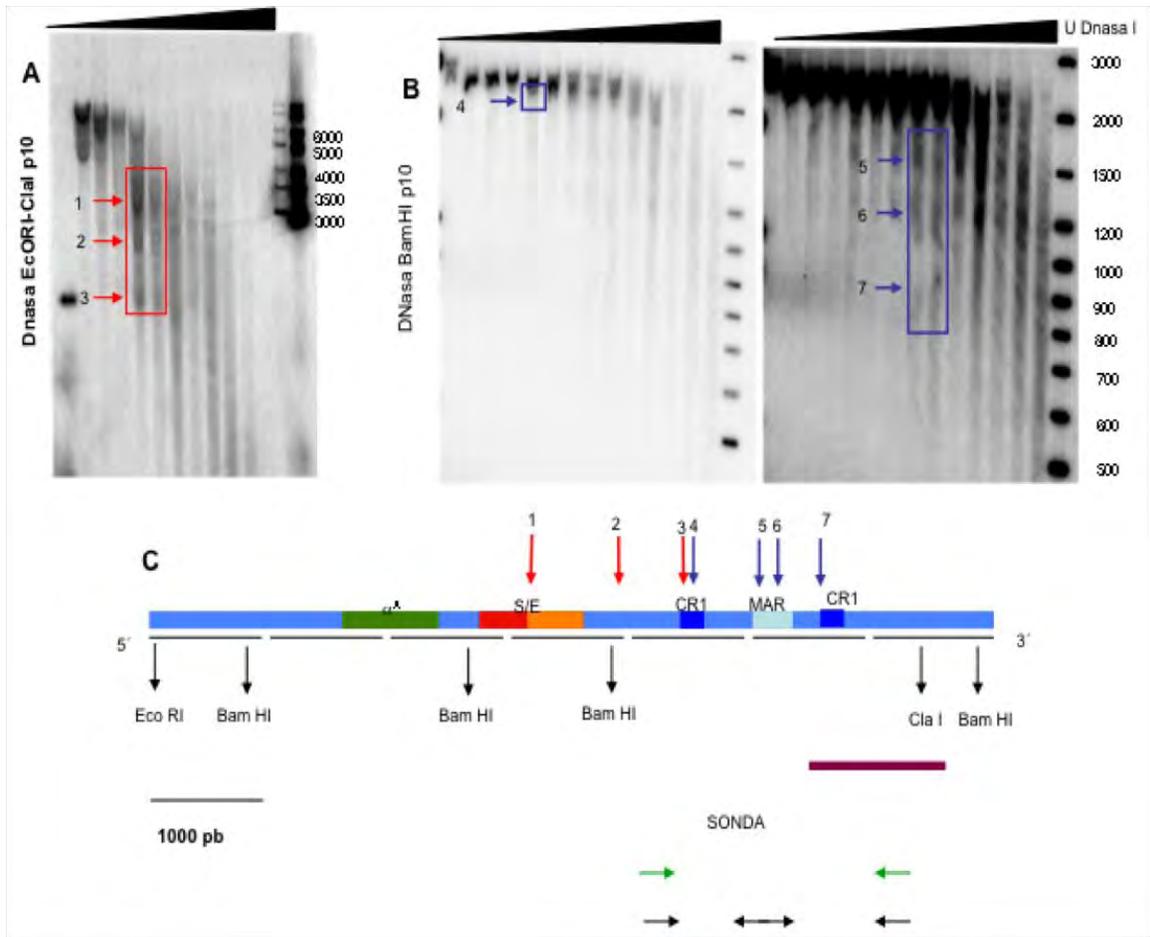
ajustaron las concentraciones de DNasa I para observar una escalera de digestión adecuada al utilizar los núcleos de eritrocitos de 10 días (RBC 10d) ya que la prueba inicial resultó en una digestión excesiva después de las 3 unidades (Fig. 46).

### 11.1.1 Ensayo de hipersensibilidad al corte por la DNasa I en la región 3' no-codificante del dominio $\alpha$ -globina de pollo.



**Figura 46. Esquema de la región de estudio en el costado 3' del dominio  $\alpha$ -globina de pollo y ajuste de unidades de DNasa I.** Arriba. Esquema de la región de estudio. Se indican los elementos que se encuentran en el extremo 3' como son el gen  $\alpha^A$ , el "enhancer" (S/E) (rojo y naranja), las repetidas CR1 y las secuencias de unión a matriz nuclear (MAR). Se muestran también los sitios de restricción (flechas) y la sonda utilizada (barra morada). Abajo. Ajuste de las unidades de DNasa I para generar una escalera de digestión adecuada para realizar el ensayo.

Posteriormente se escogió un sonda ubicada en el extremo 3' de la región de análisis (Fig. 46) y las enzimas de restricción para delimitar la región (EcoRI-ClaI en un caso y BamHI en otro). A continuación se realizó el ensayo de protección a la DNasa I hibridó con la sonda indicada (Fig. 47).



**Figura 47. Ensayo de accesibilidad a la DNasa I de la región 3' no-codificante del dominio  $\alpha$ -globina.** **A.** Digestión con DNasa I y EcoRI-ClaI. Las flechas rojas indican los sitios de hipersensibilidad (SH) identificados. **B.** Digestión con DNasaI y con BamHI. Las flechas azules indican los (SH) identificados. El panel izquierdo y derecho de B son el mismo blot con diferentes intensidades de señal. **C.** Esquema de la región en estudio. Se ubican el gen  $\alpha^A$ , el *enhancer* 3', las repetidas CRI y la región MAR. En rojo se localizan los sitios de hipersensibilidad de la DNasa I EcoRI- Cla I (A) y en azul los sitios de hipersensibilidad de la DNasaI BamHI (B). En morado se marca la sonda utilizada. Las flechas verdes y negras horizontales indican los sitios en donde se localizan los oligonucleótidos diseñados para amplificar la región (ver adelante).

Al realizar los ensayos de accesibilidad a la DNasa I pudimos distinguir tres nuevos sitios de hipersensibilidad (aparentemente eritroides por datos de Martín Escamilla) alrededor de 1 kb río abajo del *enhancer* 3' del dominio  $\alpha$ -globina. Estos sitios sugieren la presencia de nuevos elementos de función desconocida que podrían jugar un papel importante en la regulación de la expresión de los genes globina. Así mismo, los sitios de

hipersensibilidad podrían representar algún elemento estructural ó límite 3' del dominio, así como el sitio de inicio de algún transcrito no-codificante.

Los datos anteriores fueron confirmados a través de la re-hibridación de las membranas con una sonda diferente ubicada en el límite 3' de la región y los resultados fueron muy similares (datos no mostrados), por lo cual, la presencia de los sitios de hipersensibilidad pudo ser confirmada y su localización mapeada con base a dos ensayos independientes.

Posteriormente se realizó una búsqueda bioinformática en el programa TESS (Transcription Element Search System) (<http://www.cbil.upenn.edu/cgi-bin/tess/>) para identificar cydatos de unión a la región que abarca a los sitios de hipersensibilidad (Fig. 48).

FACTOR	LINAJE	FACTOR	LINAJE
GATA-1	eritroide	GATA-3	linfoide
NF-like	eritroide	NP-TCII	linfoide
CAC-binding protein	eritroide	cETS-2	linfoide
PPUR	eritroide	AML1	mieloide
PU-1	eritroide	C/EBP	FTU
NF-ATp	linfoide	YY1	FTU
IK-1	linfoide	Sp1	FTU
C-myb	linfoide	TBP	FTU

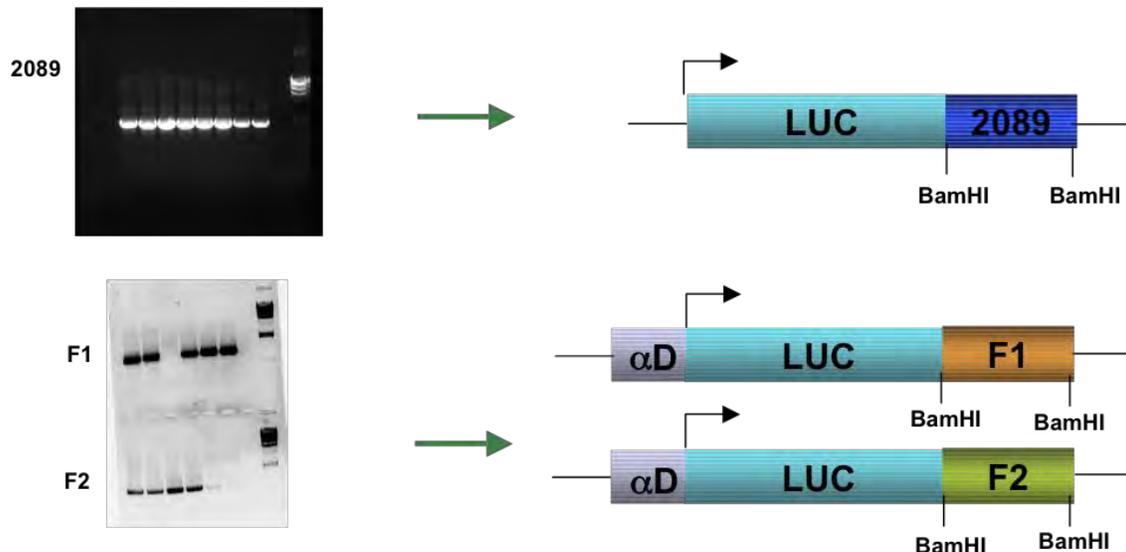
**Figura 48.** Lista de factores hematopoyéticos identificados en la región que contiene a los sitios de hipersensibilidad de la región 3' no-codificante del dominio  $\alpha$ -globina. Se marcan los factores eritroides, linfoides y los factores transcripcionales ubicuos (FTU).

Como se muestra en la figura 48, existen en la zona varias secuencias de unión a factores vinculados con la hematopoyesis en general y con la diferenciación eritroide en particular. La información que a continuación se obtenga con respecto a la actividad de

dichos sitios de hipersensibilidad encontrados podrá evaluarse teniendo en mente la serie de factores cyidatos que podrían estar presentes en la zona.

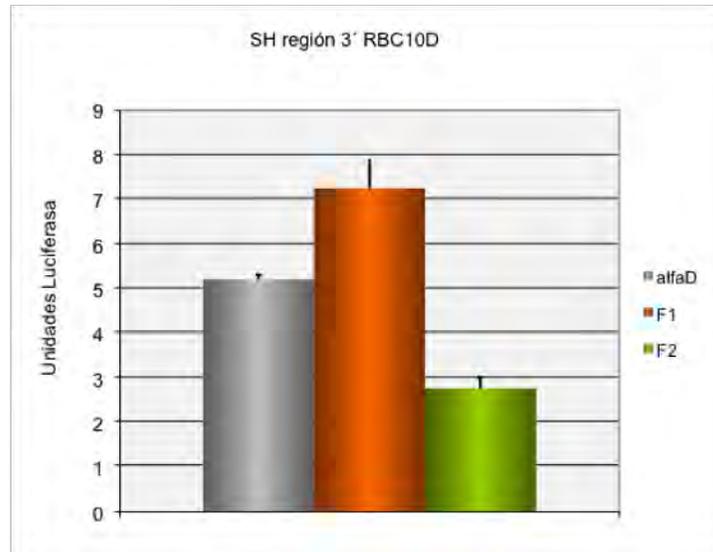
### 11.1.2 Evaluación de la función de los elementos identificados en la región 3' no-codificante del dominio $\alpha$ -globina de pollo.

Para evaluar la función de los elementos identificados se utilizaron los oligonucleótidos mostrados (Fig. 47C) para amplificar la región completa de 2089 pb (flechas verdes) y dos fragmentos de 1045 pb denominados F1 y F2 (flechas negras) por PCR para clonarlos a los plásmidos pL3B y pL $\alpha^D$  con el fin de realizar transfecciones transitorias (ver abajo). Estos plásmidos contienen al gen de la luciferasa bajo el control de un promotor a escoger (para el pL3B) y bajo el gen  $\alpha^D$  (para el pL $\alpha^D$ ). Los fragmentos fueron clonados en sitios BamHI río abajo del gen reportero para evaluar su posible actividad sobre del promotor  $\alpha^D$  (Fig. 49).



**Figura 49. Amplificación de los fragmentos que contienen a los sitios de hipersensibilidad de la región 3' del dominio  $\alpha$ -globina.** Geles representativos de la amplificación de los diferentes fragmentos (izquierda). Esquemas de los plásmidos pL2089 (contiene a los tres sitios de hipersensibilidad) , pL $\alpha^D$ F1 (contiene al primer sitio de hipersensibilidad en dirección 5' - 3') y pL $\alpha^D$ F2 (contiene a segundo y tercer sitio de hipersensibilidad en dirección 5' - 3') (derecha).

Utilizando los plásmidos  $pL\alpha^D F1$  y  $pL\alpha^D F2$  se realizaron transfecciones transitorias en RBC de 10 días para evaluar el efecto de dichas secuencias sobre el promotor  $\alpha^D$  en esta etapa (Fig. 50).



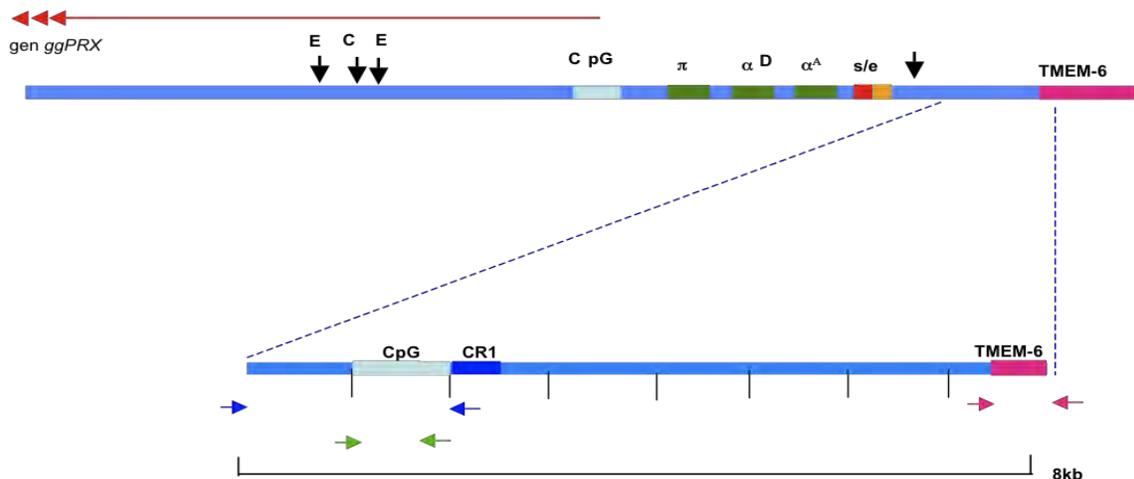
**Figura 50. Transfección transitoria de los plásmidos  $pG\alpha^D F1$  y  $pG\alpha^D F2$  en RBC de 10 días.** Puede observarse una ligera trans-activación por parte del fragmento F1 y una ligera disminución de la actividad del reportero por parte de F2 con respecto al  $pL\alpha^D$ .

Como puede observarse, el efecto del fragmento correspondiente al primer sitio de hipersensibilidad (F1) posee una actividad trans-activadora modesta con respecto al control (plásmido  $pL\alpha^D$ ) mientras que los otros dos sitios de hipersensibilidad (F2) muestran un comportamiento ligeramente represor. Sin embargo, las veces de incremento y disminución con respecto a la actividad reportada para el control no son muy pronunciadas, por lo cual, se requerirán de otras aproximaciones para tratar de entender la función de dichos elementos en el contexto genómico en donde se encuentran.

Por un lado, será importante evaluar la orientación de los fragmentos y el efecto de la transfección en la orientación contraria. Por otro lado, es posible que al separar a los elementos el efecto neto de la región se pierda, por lo cual, la evaluación del fragmento completo dará información adicional si este fuera el caso. Además, será relevante evaluar la actividad de los fragmentos utilizando al promotor del gen embrionario  $\pi$  así como en

otros estadios del desarrollo (RBC de 5-7 días) y diferenciación eritroide (HD3). Por último, resultará interesante realizar la transfección transitoria en una línea linfoide (DT40), para así verificar si su acción se encuentra restringida al linaje eritroide. La continuación de este proyecto la llevará a cabo Andrea Casasola, estudiante de doctorado en el laboratorio.

En la región subsiguiente a la analizada en los ensayos anteriores en dirección 3' del dominio, existen 8 kb hasta el inicio del siguiente gen denominado *TMEM-6 like* ó *TMEM-8* (Flint *et al.*, 2001, Philonenko *et al.*, 2009) en donde se localiza una isla CpG (rica en dinucleótidos CG) y varios sitios putativos de unión a CTCF identificados por el Dr. Recillas (Fig. 51). La función del gen *TMEM-8* no ha sido descrita, sin embargo, trabajos con respecto a una proteína transmembranal semejante, reportan que su expresión es preferentemente linfoide-eritroide (Motohashi *et al.*, 2000). Recientemente, Philonenko y colaboradores han reportado que *TMEM-8* forma parte del "hub" comatínico del dominio  $\alpha$ -globina y que éste contacta al *enhancer* 3' del dominio  $\alpha$ -globina. El significado de esta observación está siendo evaluado por Georgina Guerrero y Christian Valdés en el laboratorio.



**Figura 51. Representación del dominio  $\alpha$ -globina de pollo con una amplificación de la región 3' no codificante que contiene a la isla CpG.** En la amplificación se puede observar la localización de la isla CpG que abarca 1 kb seguida de una repetida CR-1. Las flechas indican pares de oligonucleótidos que fueron diseñados para la amplificación por PCR de los diferentes fragmentos.

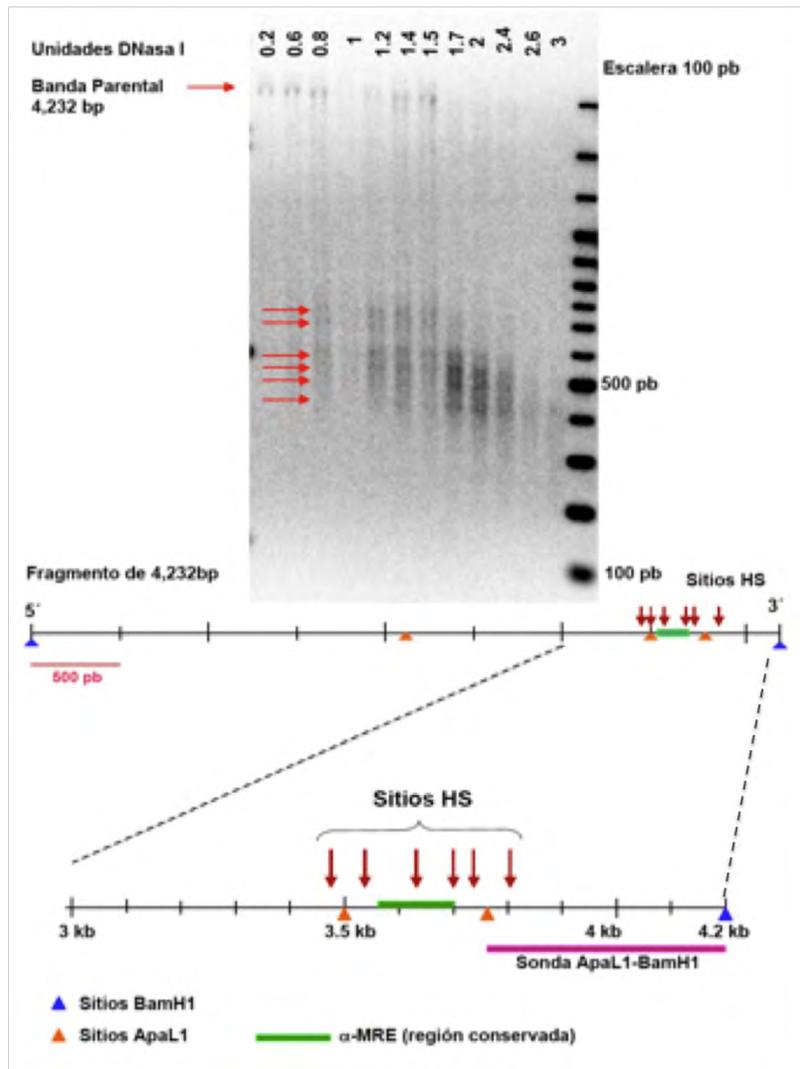
La presencia de una isla CpG en esta región así como la identificación de sitios putativos de unión a CTCF, nos llevó a diseñar oligonucleótidos para amplificar la región y evaluar su función en ensayos transitorios como los presentados anteriormente. Por otro lado, los oligonucleótidos serán utilizados para realizar ensayos de ChIP y así comparar el estado de la cromatina en la isla CpG con la región promotora del gen *TMEM-8* que, como se ha mencionado, parece estar co-regulado con el dominio  $\alpha$ -globina en células eritroides. Dichos experimentos serán realizados por Yrea Casasola en el laboratorio como parte de su proyecto de doctorado.

En resumen, hemos identificado que el costado 3' no codificante del dominio  $\alpha$ -globina de pollo presenta una serie de elementos de regulación de función desconocida representados por sitios de hipersensibilidad eritroides así como una isla CpG con sitios de unión a CTCF que pudieran contribuir en regular ó separar a los dominios  $\alpha$ -globina y *TMEM-8*.

## **11.2 Ensayo de accesibilidad al corte por la DNasa I en la región 5' no-codificante del dominio $\alpha$ -globina de pollo.**

A diferencia del costado 3' no-codificante del dominio  $\alpha$ -globina de pollo, en donde no había evidencias reportadas de la presencia de elementos de regulación eritroides, en el costado 5' no-codificante del dominio, como se ha mencionado en la sección de antecedentes y discusión de esta tesis, existen diversas evidencias que sugieren la presencia de un elemento muy conservado denominado  $\alpha$ MRE. Estudios de alineamientos múltiples han determinado una región de sintenia conservada en la cuál posiblemente estén contenidos todos los elementos regulatorios en *cis* del dominio  $\alpha$ -globina. Dentro de ésta, el  $\alpha$ MRE fue identificado mediante la búsqueda de sitios de unión a factores transcripcionales relacionados con la eritropoyesis, en particular proteínas Mafs (Flint *et al.*, 2001). Para determinar la función de este elemento en el dominio  $\alpha$ -globina de pollo, estudios previos en los cuales se llevaron a cabo transfecciones transitorias en





**Figura 53. Ensayo de hipersensibilidad a la DNasa I en eritrocitos de 10 días de desarrollo.** Se utilizó la sonda ApaL1-BamH1 que se muestra (morado). Los seis sitios de hipersensibilidad encontrados (HS) se ubican en el esquema de abajo en relación a la secuencia del dominio  $\alpha$ -globina de pollo y al elemento  $\alpha$ MRE en verde.

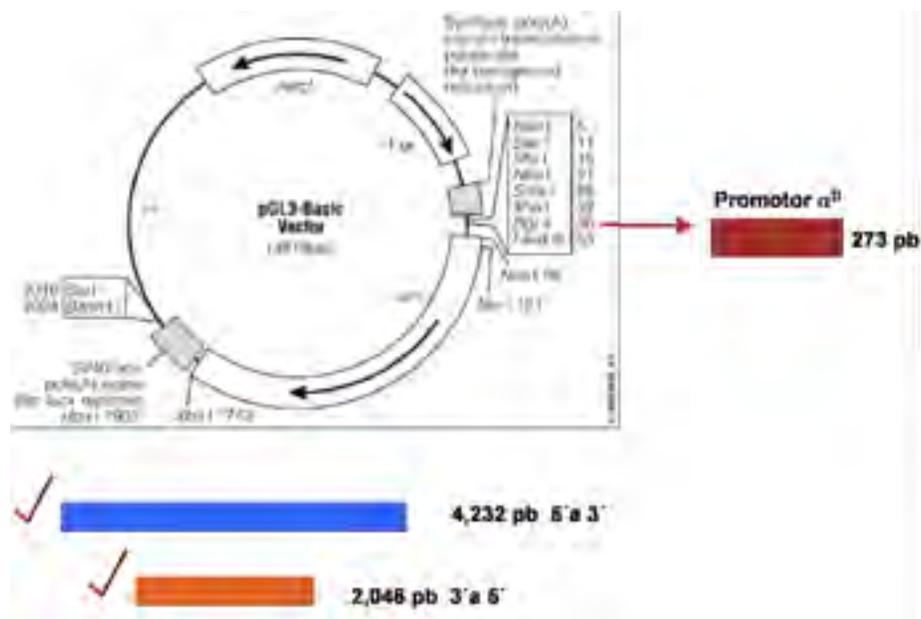
Mediante el ensayo de hipersensibilidad al corte por la DNasa I en eritrocitos de 10 días de desarrollo, fue posible determinar la presencia de al menos seis sitios de hipersensibilidad localizados en la secuencia reconocida como el  $\alpha$ MRE (Fig. 53). En el estudio realizado previamente en células HD3 correspondientes al estadio de eritroblastos, se identificó la presencia de sólo un sitio de hipersensibilidad (Flint *et al.*, 2001). Esto sugiere que el estado de la cromatina en eritrocitos terminalmente

diferenciados es más relajada en esta zona y que existen un conjunto de sitios hipersensibles que pueden representar factores de unión a la cromatina que le confieran actividad al  $\alpha$ MRE en esta etapa de la diferenciación eritroide en donde los genes  $\alpha$ -globina adultos están activos.

### 11.2.2 Evaluación de la función del elemento $\alpha$ MRE identificado en la región 5' no-codificante del dominio $\alpha$ -globina.

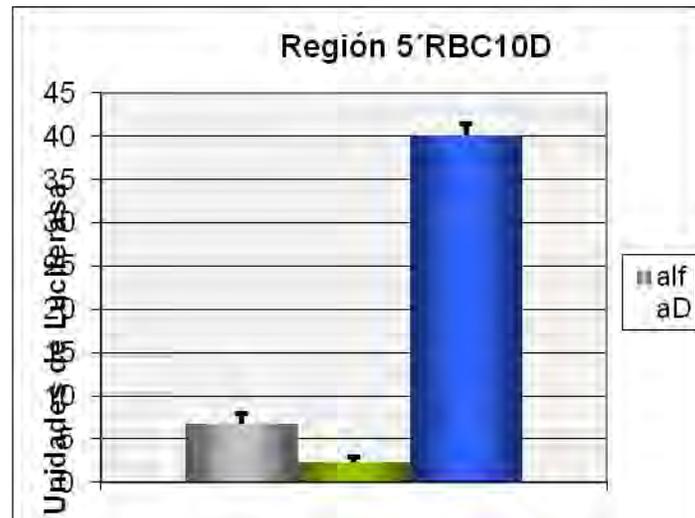
A continuación decidimos evaluar la función del elemento  $\alpha$ MRE identificado a partir de los ensayos de DNasa I en el estadio de eritrocitos maduros de 10 días de desarrollo ya que en estudios previos, su función se evaluó en eritroblastos en donde los genes  $\alpha$ -globina no se expresan.

Para realizar el análisis de la actividad del  $\alpha$ MRE se generaron plásmidos similares a los utilizados para evaluar los elementos de la región 3' no codificante, es decir, que contienen al gen reportero que codifica para la luciferasa bajo el promotor  $\alpha^D$ , de tal forma que, al clonar diferentes regiones de interés río arriba del promotor, se puede evaluar su función realizo transfecciones transitorias y midiendo la actividad de la luciferasa (Fig. 54).



**Figura 54. Esquema de las construcciones realizadas.** Ambas contienen al promotor  $\alpha^D$ , una contiene al fragmento de 4,232 pb que abarca al elemento  $\alpha$ MRE y la otra contiene al fragmento de 2,046 pb localizado entre el  $\alpha$ MRE y *insulador*  $\alpha$ EHS-1.4.

Utilizando ambas construcciones generadas, se realizaron transfecciones transitorias en eritrocitos primarios de embriones de pollo 10 días de desarrollo (Fig. 55).



**Figura 55. Transfección transitoria en eritrocitos de 10 días de desarrollo.** En gris se muestra el plásmido control con la expresión basal generada por el promotor  $\alpha^D$ . En verde se muestra el resultado para el vector con la secuencia de 2,046 pb y en azul el resultado para el vector con el fragmento de 4,232 pb que contiene al  $\alpha$ MRE.

En la figura 55 se muestra como la secuencia de 2,046 pb que se ubica entre el elemento  $\alpha$ MRE contenido en la secuencia de 4,232 pb y el *insulador*  $\alpha$ EHS-1.4 no presenta ninguna actividad sobre la expresión basal de la luciferasa bajo el promotor  $\alpha^D$ . Sin embargo, la secuencia de 4,232 pb que contiene al  $\alpha$ MRE y por lo tanto a el conjunto de sitios de hipersensibilidad que hemos detectado en este estadio, presenta una clara trans-activación de la actividad del reportero luciferasa bajo el promotor  $\alpha^D$ . Este resultado apoya la noción de que el  $\alpha$ MRE es un elemento fuerte e importante que promueve la trans-activación de los genes  $\alpha$ -globina adultos en los estadios que éstos se expresan. A continuación será relevante evaluar el efecto de dicho elemento en diferentes estadios de la diferenciación eritroide y el desarrollo del pollo (en particular en eritrocitos de 5 días de desarrollo en donde se expresa el gen embrionario  $\pi$ ), así como estudiar su función utilizando los diferentes promotores del dominio  $\alpha$ -globina. También, como se ha

planteado en las perspectivas de este trabajo, será interesante evaluar su relación con el resto de los elementos de regulación del dominio incluyendo al *insulator*  $\alpha$ EHS-1.4.

## 12. ANEXO 2. Publicaciones.

### Publicaciones relacionadas con la tesis

- **Furlan-Magaril**, Rebollar, Guerrero, Fernández, Moltó, González-Buendía, Cantero, Montoliu y Recillas-Targa . An insulator embedded in the chicken  $\alpha$ -globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic Acids Research* (2010).
- **Furlan-Magaril** y Recillas-Targa. Chromatin remodeling and epigenetic regulation during development. *Topics in Animal and Plant Development: From Cell Differentiation to Morphogenesis* (2010).
- **Furlan-Magaril**, Rincón-Arano y Recillas-Targa. Sequential chromatin immunoprecipitation protocol: ChIP-reChIP. *Methods Mol Biol* (2009) vol. 543 pp. 253-66
- Rincón-Arano, **Furlan-Magaril** y Recillas-Targa. Protection against telomeric position effects by the chicken cHS4 beta-globin insulator. *Proc Natl Acad Sci USA* (2007) vol. 104 (35) pp. 14044-9
- Guerrero, Delgado-Olguín, Escamilla-del-Arenal, **Furlan-Magaril**, Rebollar, De La Rosa-Velázquez, Soto-Reyes, Rincón-Arano, Valdés-Quezada, Valadez-Graham y Recillas-Targa. Globin genes transcriptional switching, chromatin structure and linked lessons to epigenetics in cancer: a comparative overview. *Comp Biochem Physiol, Part A Mol Integr Physiol* (2007) vol. 147 (3) pp. 750-60
- Martín, Pantoja, Fernández Miñán, Valdés-Quezada, Moltó, Matesanz, Calle-Mustienes, Bogdanovi, Domínguez, Taher, **Furlan-Magaril**, Alcina, Cañón, Catala-Rabasa, Blasco, Pereira, Ovcharenko, Recillas-Targa, Montoliu, Manzanares, Guigó, Serrano, Casares, Gómez-Skarmeta. Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid in the identification of disease associated genes in humans. Sometido

### Publicaciones no relacionadas con la tesis

- Dávalos, **Furlan-Magaril**, González-Buendía, Valdés-Quezada y Recillas-Targa. Gain of DNA methylation is enhanced in the absence of CTCF at the human retinoblastoma gene promoter. Sometido
- Hernández-Hernández et al. Sox9 represses alpha-sarcoglycan gene expression in early myogenic differentiation. *J Mol Biol* (2009) vol. 394 (1) pp. 1-14



# An insulator embedded in the chicken $\alpha$ -globin locus regulates chromatin domain configuration and differential gene expression

Mayra Furlan-Magaril<sup>1</sup>, Eria Rebollar<sup>1</sup>, Georgina Guerrero<sup>1</sup>, Almudena Fernández<sup>2,3</sup>, Eduardo Moltó<sup>2,3</sup>, Edgar González-Buendía<sup>1</sup>, Marta Cantero<sup>2,3</sup>, Lluís Montoliu<sup>2,3</sup> and Félix Recillas-Targa<sup>1,\*</sup>

<sup>1</sup>Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, México D.F., México, <sup>2</sup>Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Departamento de Biología Molecular y Celular, Campus de Cantoblanco and <sup>3</sup>Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, Madrid, Spain

Received July 2, 2010; Revised July 29, 2010; Accepted August 4, 2010

## ABSTRACT

Genome organization into transcriptionally active domains denotes one of the first levels of gene expression regulation. Although the chromatin domain concept is generally accepted, only little is known on how domain organization impacts the regulation of differential gene expression. Insulators might hold answers to address this issue as they delimit and organize chromatin domains. We have previously identified a CTCF-dependent insulator with enhancer-blocking activity embedded in the 5' non-coding region of the chicken  $\alpha$ -globin domain. Here, we demonstrate that this element, called the  $\alpha$ EHS-1.4 insulator, protects a transgene against chromosomal position effects in stably transfected cell lines and transgenic mice. We found that this insulator can create a regulated chromatin environment that coincides with the onset of adult  $\alpha$ -globin gene expression. Furthermore, such activity is in part dependent on the *in vivo* regulated occupancy of CTCF at the  $\alpha$ EHS-1.4 element. Insulator function is also regulated by CTCF poly(ADP-ribosylation). Our results suggest that the  $\alpha$ EHS-1.4 insulator contributes in organizing the chromatin structure of the  $\alpha$ -globin gene domain and prevents activation of adult  $\alpha$ -globin gene expression at the erythroblast stage via CTCF.

## INTRODUCTION

Differential gene expression in genomic domains is regulated at multiple levels. Not only does it require

chromatin structure remodeling at local regulatory elements, but also the establishment of a chromatin architecture defining the domain itself. Consequently, more than proximal and distal regulatory elements are required for a gene's coordinated transcription in order to first demarcate the genomic region in which it is located. Once the chromatin domain is delimited and organized, local regulatory actions take place to achieve expression of specific genes (1).

Insulators have been considered as regulatory elements that act as boundaries to the action of external enhancers and repressive chromatin and participate in the formation and maintenance of chromatin domains (1,2). Accumulated evidence suggests that insulators, particularly CTCF-linked insulators, do not act exclusively as boundary elements (3). For example, CTCF in coordination with cohesins, mediate the formation of chromatin loops favoring optimal spatial organization of gene domains (4,5). This vision is further supported by the presence of insulator bodies that are important for domain formation at the nuclear periphery of *Drosophila* cells. Insulator bodies are formed of polymorphic protein complexes including dCTCF, Su(Hw) and CP190, besides other proteins (6). Interestingly, such insulator-mediated topological organization of the genome is dynamic. For instance, CTCF and cohesin mediate chromosomal *cis* contacts in a regulated manner in several loci. Such loci include the *Igf2/H19* imprinted locus (7,8), the cytokine locus including *IL4*, *IL5*, *IL13* and *IFNG* (9), the mouse  $\beta$ -globin locus, in which the spatial organization of regulatory components to form an active chromatin hub is facilitated by CTCF during tissue-specific activation of globin genes (10), the MHC class II genes (11) and more recently, the  $\beta$ -globin locus, which interacts with flanking olfactory receptor genes on the human chromosome 11

\*To whom correspondence should be addressed. Tel: +52 55 56 22 56 74; Fax: +52 55 56 22 56 30; Email: frecilla@ifc.unam.mx

(5). Notoriously, such interactions are not exclusively intrachromosomal, CTCF also mediates interchromosomal contacts between one allele of the *Igf2/H19* imprinting control region on chromosome 7 and one allele of the *Wsb1/Nf1* imprinted locus on chromosome 11 (12). Insulator action through CTCF enhancer blocking activity can also be regulated by other mechanisms as by thyroid hormone (13) and post-translational modifications as poly(ADP-ribosylation) (14) among others. Thus, CTCF is implicated in diverse regulatory and structural functions (3).

We previously described a regulatory element located in the 5' non-coding region of the chicken  $\alpha$ -globin domain, ~14-kb upstream of the embryonic  $\pi$  gene (Figure 1A) (15). This element, named  $\alpha$ EHS-1.4, which corresponds to a DNA fragment of 1.4kb is an insulator with CTCF-dependent enhancer-blocking activity (15). This insulator is located within an intron of the *C16orf35* gene antisense transcript (16,17), between the distal locus control region known as the  $\alpha$ -major regulatory element ( $\alpha$ -MRE), and the  $\alpha$ -globin genes (Figure 1A). Since the  $\alpha$ EHS-1.4 is located 3' from the  $\alpha$ -MRE it seemed unlikely that the  $\alpha$ EHS-1.4 insulator element acts as a boundary for the 5'-side of the  $\alpha$ -globin domain, shielding it from silent chromatin expansion or from the action of enhancers outside the domain. Consequently, we addressed the contribution of this element to the chromatin configuration of the  $\alpha$ -globin domain and for regulated gene expression during erythropoiesis and development.

Here, we demonstrate that the  $\alpha$ EHS-1.4 element shields a transgene from chromosomal position effects (CPE) and gene silencing in stably transfected cell lines and transgenic mice. In contrast to the chicken 5' cHS4  $\beta$ -globin insulator (18), the protective function of the  $\alpha$ EHS-1.4 insulator is CTCF-dependent. Moreover, this element mediates the chromatin conformation of the 5' non-coding region of the domain during developmentally regulated gene expression. Subsequently, we demonstrate that prior to erythroid differentiation, the  $\alpha$ EHS-1.4 insulator is bound by CTCF and prevents adult  $\alpha$ -globin gene expression, and, as differentiation proceeds, CTCF is deployed and the blockage is released allowing the activation of the adult  $\alpha$ -globin genes in fully developed erythroid cells. Our results provide insight into the function of insulators as organizers of chromatin domains and regulators of differential gene expression during cell differentiation and development.

## MATERIALS AND METHODS

### Culture of stable cell lines

Avian arrested cell lines were maintained at 37°C, 5% CO<sub>2</sub> with the following corresponding media: 6C2 (pre-erythroblasts) cells with  $\alpha$ -MEM (GIBCO) supplemented with 10% (v/v) fetal bovine serum (FBS) (Multicell), 2% (v/v) chicken serum (CS), 1 mM HEPES (pH 7.2) (SIGMA) and 50 mM  $\beta$ -mercaptoethanol. The avian erythroblastosis virus-transformed chicken erythroblast cell line HD3 and the chicken lymphoid cell-line DT40 were cultured in DMEM (GIBCO) supplemented

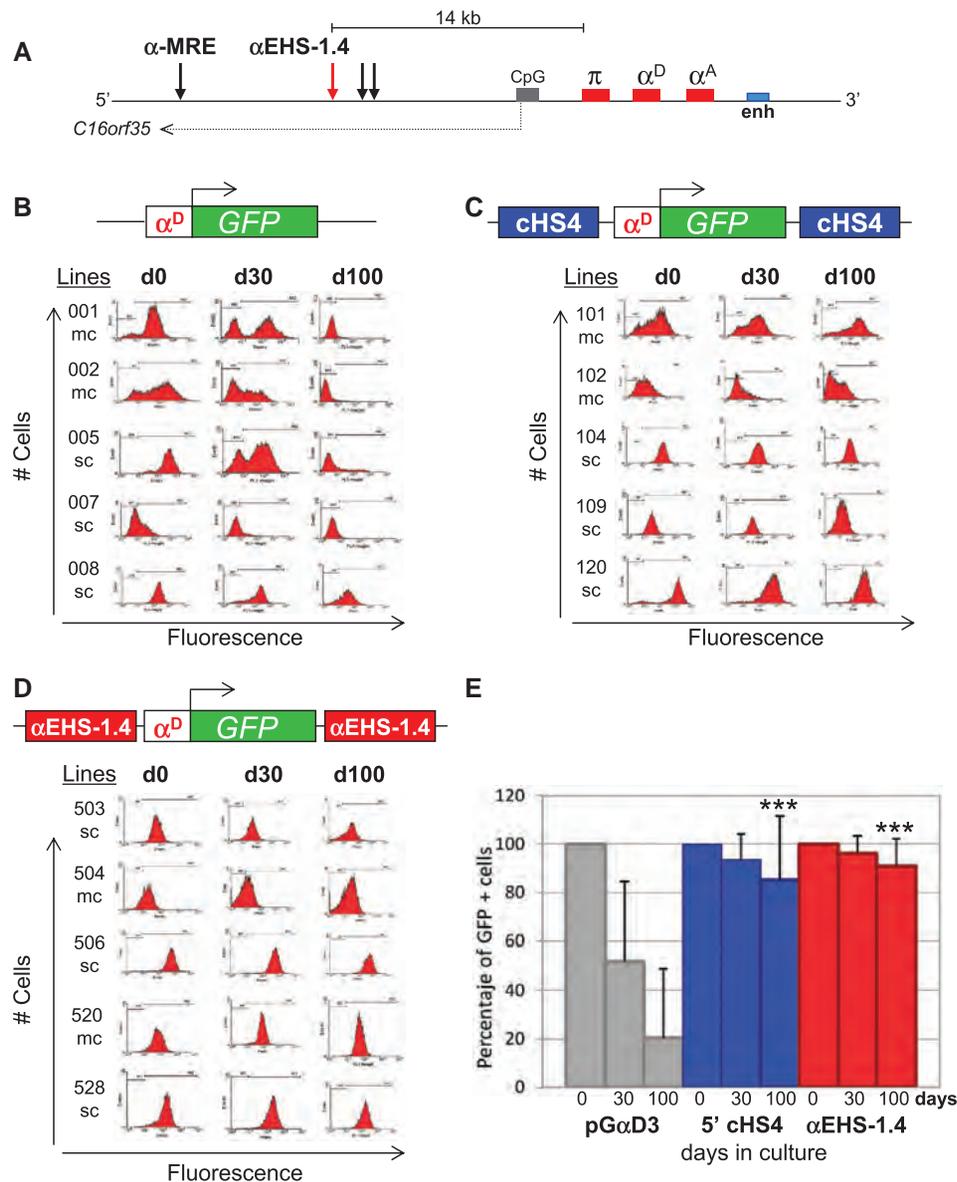
with 8% FBS and 2% CS (19,20). All media contain penicillin/streptomycin (500 U/ml) (GIBCO). To differentiate HD3 cells they were cultured with the same media with 2% anemic CS at 42°C for 7–15 days, as described (21,22).

### Protection against chromosomal position effects assay

The protection against CPE assay was performed as previously described (18,23). Stable transfection of the different plasmids was performed with Lipofectamine (Invitrogen) following manufacturer instructions. The transfected plasmids were pG $\alpha^D$ , containing the *EGFP* under the chicken  $\alpha^D$  gene promoter, pG $\alpha^D$ HS4 (containing the  $\beta$ -globin 5'cHS4 insulator on each side of the  $\alpha^D$ -*EGFP*) and pG $\alpha^D$ 1.4 (containing the  $\alpha$ EHS-1.4 insulator on each side of the  $\alpha^D$ -*EGFP*). Individual integrants were obtained by using a semisolid media (Methocel Fluka, Buchs, Switzerland) with neomycin at 0.9 mg/ml (18). After 10–14 days of selection, individual clones were isolated and their expression was determined by FACS (at Day 0). Copy number and transgene integrity was evaluated by Southern blot analysis using a DNA fragment of the *EGFP* as a probe. Transgene expression was evaluated twice a month until completion of at least 100 days of continuous cell culture without neomycin. The mutant  $\alpha$ EHS1.4 $\Delta\Delta$  construct was obtained by nested PCR using primers containing the  $\Delta$ CAG and  $\Delta$ CTAG base changes at the CTCF-binding site defined previously (15) and then using primers to amplify the complete 1.4-kb DNA fragment (all primers are available upon request).

### Generation of transgenic mice

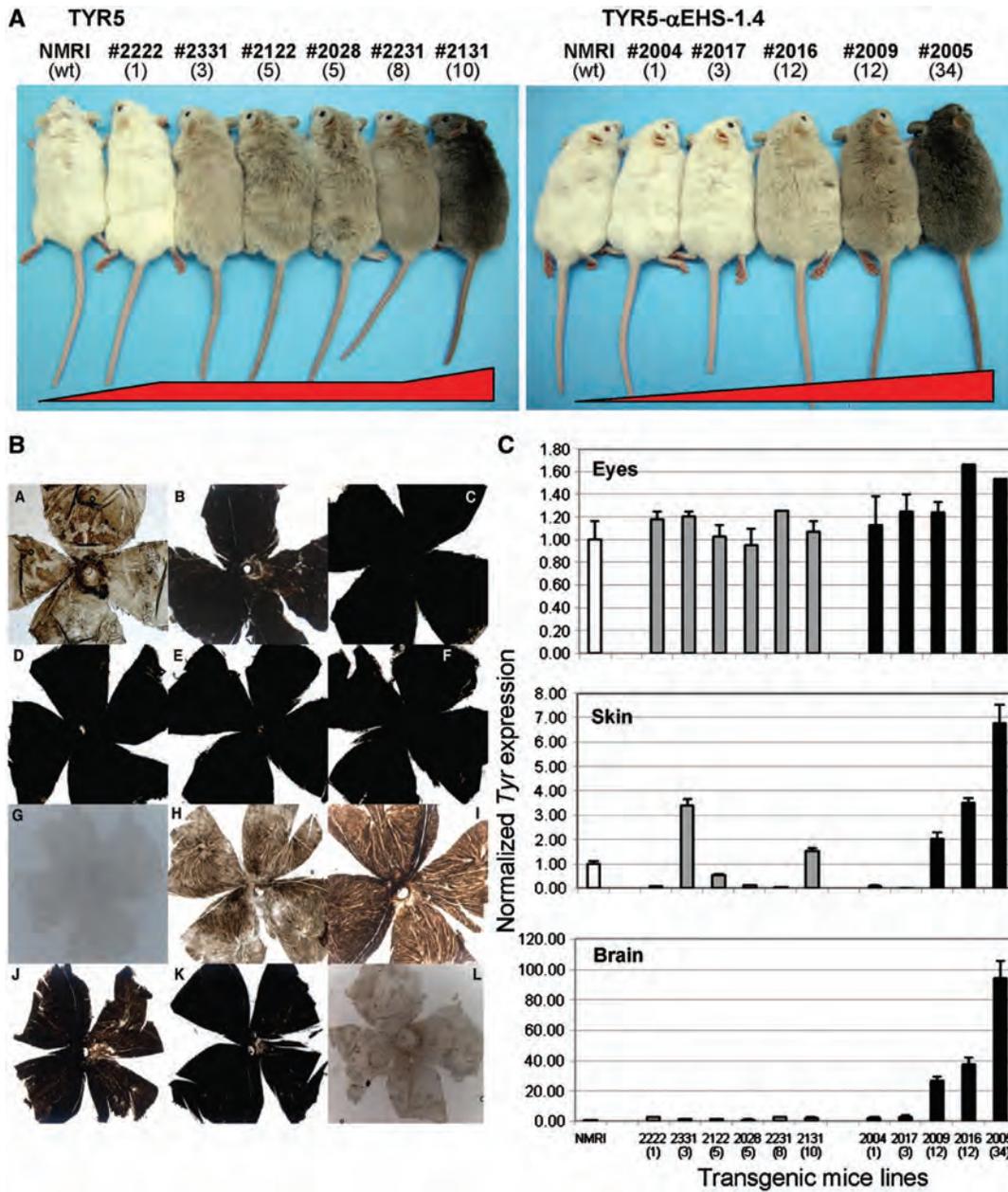
The plasmid pTYR5-1.4 containing the  $\alpha$ EHS-1.4 element was cloned in a standard 5'–3' orientation, at the 5' end in the unique *Xba* I site of the *Tyr* (tyrosinase)-reporter construct ptrTYR5 (24); it was digested with *Ecl* XI and *Sal* I enzymes, thereby releasing the 13-kb DNA band with the insert that was purified following described standard procedures (25). In parallel, the corresponding *Ecl* XI-*Sal* I 11.6-kb DNA band from ptrTYR5, without the insulator element, was also purified and used as the reference control for transgenesis. The resulting DNA fragments and transgenes were named as: TYR5 and TYR5- $\alpha$ EHS-1.4, respectively. For the generation of transgenic mice, suitable DNA solutions from each of the two transgenes were microinjected into the pronucleus of albino outbreed NMRI/Hsd fertilized mouse oocytes (Harlan Interfauna Iberica SL, Barcelona, Spain) using standard procedures (25). Transgenic founder mice were detected at birth due to obvious pigmentation (24) and confirmed by PCR analysis, using tyrosinase minigene-specific primers, as described before (24) and also by Southern blot analysis, using the pmTyrE5 tyrosinase probe that reveals polymorphic *Hind* III DNA bands between the transgene (3.4 kb) and the endogenous tyrosinase locus (2.2 kb), as reported (26). Transgenic mouse lines were established and maintained, as hemizygous, in the albino outbreed NMRI/Hsd genetic background. Transgene copy-number was determined from the DNA of F<sub>1</sub> individuals by Southern blot analysis, as reported before (26). For the TYR5- $\alpha$ EHS-1.4 construct, 289 fertilized oocytes



**Figure 1.** Protection against position effects by the  $\alpha$ EHS-1.4 insulator. (A) Scheme of the  $\alpha$ -globin gene domain. The  $\alpha$ -MRE constitutes the putative locus control region of the domain. The  $\alpha$ EHS-1.4 insulator is located  $\sim$ 14-kb upstream of the embryonic  $\pi$  globin gene. CpG, represents a CpG-island located  $\sim$ 4-kb upstream of the embryonic  $\pi$  globin gene which represents the first CTCF-binding site identified *in vivo*, that corresponds to the promoter of the antisense transcript *C16orf35* (47). Vertical arrows indicate the position of DNase I hypersensitive sites and the 3'-side enhancer (21). (B) Flow cytometry histograms (FACS) showing the GFP fluorescence level of five representative and independent stably transfected cell lines containing the control vector without insulators at days 0, 30 and 100 of continuous cell culture. (C) Fluorescence of independent cell lines incorporating the insulated plasmid with the chicken 5'cHS4 insulator of the  $\beta$ -globin gene locus. (D) Fluorescence of independent cell lines containing the plasmid flanked on each side of the transgene by the  $\alpha$ EHS-1.4 insulator. (E) Chart showing the mean percentage of GFP expressing cells ( $\pm$ SD) at different time points during culture for the three different transgenes (pG $\alpha$ D3  $n = 15$ , pG $\alpha$ D3 5'cHS4  $n = 10$  and pG $\alpha$ D3 EHS-1.4  $n = 10$ ). sc = single-copy and mc = multi-copy integrants. \*\*\* $P \leq 0.001$  with respect to the pG $\alpha$ D3 (100 days) control (ANOVA).

were microinjected, 192 embryos were transferred to pseudopregnant females and resulted in the birth of 17 pups, out of which five independent transgenic mouse lines were established. For the TYR5 transgene, 581 fertilized oocytes were microinjected, 284 embryos were transferred and gave rise to 17 newborns, out of which three were transgenic. Subsequently, two of these mouse lines were found to carry multiple transgene insertions that could be segregated in the F1 generation, thus finally resulting in a total of six independent transgenic

mouse lines with the TYR5 construct. Representative F1 individuals from each transgenic mouse line were briefly anesthetized with 1 ml/100 g of a mixture of Ketamine (10 mg/ml) and Xylazine (2 mg/ml) for obtaining the arranged pictures (Figure 2A). All the procedures that required the use of animals complied with Spanish and European legislation concerning vivisection and the use of genetically modified organisms, and the protocols were approved by the local ethics committees on animal experimentation.



**Figure 2.** Analysis of TYR5 versus TYR5-αEHS-1.4 transgenic mice. (A) Coat color of transgenic mice. Representative F1 individuals from TYR5 (left) and TYR5-αEHS-1.4 (right) transgenic mice are shown alongside non-transgenic albino NMRI littermates (wt, for wild-type), used as the recipient mouse strain for transgenesis. The number of the independent transgenic line and number of integrated transgene copies (in brackets) are shown on the top. (B) Retinas of TYR5 and TYR5-αEHS-1.4 transgenic mice. Whole-mount retinas showing pigmentation levels due to the retinal pigment epithelium and the overlying choroid with melanocytes, from adult TYR5 (panels A to F: A, line #2222; B, line #2332; C, line #2122; D, line #2028; E, line #2231; F, line #2131) and TYR5-αEHS1.4 (panels G to K: G, line #2004; H, line #2017; I, line #2016; J, line #2009; K, line #2005) transgenic mice. A retina from a non-transgenic albino NMRI littermate is included for reference (panel L). (C) Analysis of tyrosinase gene expression. Tyrosinase (*Tyr*) gene expression was measured by quantitative RT-PCR using specific TaqMan and reference (Tbp) probes in all transgenic mouse lines carrying the TYR5 (grey bars) and the TYR5-αEHS-1.4 (solid black bars) construct in total RNA isolated from three different tissues: eyes (top chart), dorsal skin (middle), and whole brain (bottom). Two F1 or F2 adult representative animals from each transgenic mouse line were used for measurements. Results are shown as averages ± SD. The number of each of the transgenic mouse lines and their corresponding transgene copy-number (in brackets) is indicated below each bar. A non-transgenic albino NMRI littermate (white bar) was used as a control for all tissues.

### Analysis of transgenic mice

Total RNA was prepared from whole brain, dorsal skin and eyes from at least three adult F<sub>1</sub> or F<sub>2</sub> individuals

from each of the transgenic mouse lines carrying the TYR5 and TYR5-αEHS-1.4 constructs, using the RNeasy Kit (Qiagen). Tyrosinase (*Tyr*) expression was assessed by quantitative RT-PCR using specific mouse

TaqMan (Applied Biosystems) *Tyr* (Mm00495817\_m1) and *Tbp* (Mm00446973\_m1, included as a reference) probes, on an ABI-Prism 7000 (Applied Biosystems), using TaqMan Universal PCR Master Mix (Applied Biosystems), as described before (27). Whole-mount retinas from adult F<sub>1</sub> or F<sub>2</sub> individuals from each of the transgenic mouse lines carrying the TYR5 and TYR5- $\alpha$ EHS-1.4 constructs were prepared according to previously described methods (27,28). Quantification of the total melanin contents from eye and skin extracts was carried out from at least two adult F<sub>1</sub> or F<sub>2</sub> individuals from each of the transgenic mouse lines carrying the TYR5 and TYR5- $\alpha$ EHS-1.4 constructs, according to reported protocols (28).

### Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) analysis were performed and analyzed as described (29). Duplex PCR was used for evaluating histone marks and CTCF enrichment at different points in the  $\alpha$ EHS-1.4 insulator region covering the CTCF-binding site (see scheme in Figure 4), the  $\alpha$ EHS-1.4 CTCF-binding site in the transgene and a DNA fragment of *EGFP*. As controls we used regions in which the analyzed histone covalent modification or nuclear factor is absent; for permissive chromatin histone marks and CTCF we used two heterochromatin-associated sequences located 5' of the chicken  $\beta$ -globin gene domain (30), for the repressive histone marks we used the  $\beta$ -globin 5' cHS4 insulator sequence that is enriched with permissive marks in all cell types used (30). Enrichment was calculated by using the following equation: Ab test (test region/control region)/IgG (test region/control region). The antibodies employed were the following: antibodies against acetylated histone H3 and acetylated histone H4 (06-599 and 06-866) were obtained from Upstate Biotechnology (Lake Placid, NY). The antibody against H3K4me2 (ab7766) was purchased from Abcam (Cambridge, MA). The antibody against H3K79me2 was kindly provided by Fred van Leeuwen (Nederland Kanker Instituut–Antoni van Leeuwenhoek Ziekenhuis, Amsterdam, The Netherlands). Antibodies against the trimethylated versions of histone 3 on K9 and histone 4 on K20 were kindly provided by Thomas Jenuwein (Max Planck Institute of Immunology, Freiburg, Germany). The antibody against chicken CTCF [anti-cCTCF(86–233)] was previously obtained as described (15).

### CTCF knock down and 3-aminobenzamide treatment

For the CTCF knock down series of experiments, three different sets of small hairpin RNA interference (shRNAi) sequences against CTCF were tested with overall similar results. To generate the CTCF knock down cells we used a collection of four plasmids containing different sequences to generate shRNAi against the 5' region of chicken CTCF mRNA. Three of them were designed from the reported sequences in (31) (pSiLefevre1, 2 and 3) and one was kindly provided by Ishihara (pCTI). Even though the pCTI shRNAi was designed against human CTCF (32), it reduced CTCF levels in the chicken

cell lines. Plasmids containing the hairpin DNAs were stably transfected as described previously, and integrants were selected with 1  $\mu$ g/ml puromycin (SIGMA) for 7–15 days. CTCF reduction was assessed by western blot analysis (Supplementary Figure S4A). For 3-aminobenzamide (3-ABA) treatment, 8 mM 3-ABA was added to the culture media for 3–6 days prior to evaluation by western blotting and ChIP analysis, as described (14). The CTCF antibody used to perform the western blot (15) recognizes both the 130- and 180-kDa forms of CTCF as shown in Supplementary Figure S4A.

### Statistics

Two-tailed analysis of variance (ANOVA) or multivariate analysis of variance (MANOVA) statistical analyses were performed using SPSS program (version 18) considering as significant: \* $P \leq 0.05$ , \*\* $P \leq 0.01$  and \*\*\* $P \leq 0.001$ . Scheffe or Games-Howell *post hoc* tests were used when appropriate to distinguish which components analyzed are causing the significant variance between conditions.

## RESULTS

### The $\alpha$ EHS-1.4 insulator protects against chromosomal position effects in stably transfected cell lines

To test if the  $\alpha$ EHS-1.4 insulator located 14-kb upstream of the embryonic  $\pi$  gene of the  $\alpha$ -globin locus protects a transgene against chromosomal position effects, we generated stably transfected HD3 erythroblasts expressing a *GFP* reporter driven by the chicken  $\alpha^D$  gene promoter ( $\alpha^D$ -*GFP*) (Figure 1B) (29), or  $\alpha^D$ -*GFP* flanked on each side by the  $\alpha$ EHS-1.4 insulator (Figure 1D). The  $\alpha^D$ -*GFP* flanked on each side by the 5'cHS4 core insulator of the  $\beta$ -globin locus was used as control (18). Transgene integrity and copy number (single- or multi-copy) was determined by Southern blot for a selected set of stable lines (data not shown). We measured GFP fluorescence by flow cytometry at different time points to assess transgene expression (18,23). As expected, the  $\alpha^D$ -*GFP* plasmid (without insulators) showed sensitivity to CPE, as transgene expression was extinct even before day 100 of continuous cell culture (Figure 1B). In contrast, the  $\alpha$ EHS-1.4 insulator robustly protected against CPE even after 100 days of continuous cell culture (Figure 1D). This protective activity was comparable to that exerted by the chicken  $\beta$ -globin 5'cHS4 insulator (Figure 1C–E). Next, we examined the profile of histone post-translational modifications present over the *GFP* transgene in distinct cell lines by ChIP. Modifications corresponding to a permissive chromatin environment were present only in insulated transgenes (Supplementary Figure S1). Thus, the  $\alpha$ EHS-1.4 insulator efficiently protects against CPE in chicken HD3 erythroblasts cells thereby favoring a transcriptionally permissive chromatin environment.

### The $\alpha$ EHS-1.4 insulator protects against chromosomal position effects in transgenic mice

To test the functionality of the  $\alpha$ EHS-1.4 element *in vivo* we addressed its role in preventing chromosomal position effects in transgenic mice. We used a reporter system in which the albino phenotype is rescued in transgenic mice expressing a functional tyrosinase gene (*Tyr*) (24), and which has been useful to validate the effect of the chicken 5' cHS4  $\beta$ -globin insulator in protecting transgenes from CPE (33). The  $\alpha$ EHS-1.4 insulator was cloned as a single copy at the 5' end of *Tyr* in the ptrTYR5 construct (TYR5- $\alpha$ EHS-1.4). We cloned the  $\alpha$ EHS-1.4 sequence only at the 5' end of the vector since constructs usually integrate as multi-copy tandem arrays; therefore, most transgene copies would end up shielded by  $\alpha$ EHS-1.4 elements (25). The parental plasmid, ptrTYR5, here called TYR5, was used as a reference.

Five independent transgenic mouse lines carrying the insulated TYR5- $\alpha$ EHS-1.4 construct (lines #2004, #2017, #2016, #2009 and #2005), and six carrying the control TYR5 construct (lines #2222, #2331, #2122, #2028, #2231 and #2131) were generated in the albino outbred strain NMRI/Hsd (Figure 2A). Transgene integrity and copy number was assessed by Southern blot analysis (data not shown). As anticipated, insertion of either construct rendered pigmented transgenic mice (Figure 2A). In all cases, pigmentation was uniform, without obvious signs of mosaicism [that were obvious in similar transgenic mice reported by Potts *et al.* (33)]. Mouse lines with higher number of TYR5 copies [i.e. #2131 (10 copies)] showed increased pigmentation, as compared to lines with fewer transgene copies [i.e. #2222 (one copy)]. However, there was little coat color variation in the remaining TYR5 transgenic lines with intermediate transgene copy-number (Figure 2A, lines #2331, #2122, #2028 and #2231). In contrast, the five transgenic mouse lines generated with the TYR5- $\alpha$ EHS-1.4 transgene, produced individuals with progressively darker pigmentation, mostly correlating with the corresponding transgene copy-numbers. Only one line, #2004 (one copy), showed no detectable pigmentation in skin and/or eyes (Figure 2A).

To better assess *Tyr* expression in transgenic mice we carried out three additional experiments. First, we prepared whole-mount retinas from representative F1 individuals from TYR5 and TYR5- $\alpha$ EHS-1.4 transgenic lines to directly visualize the retinal pigment epithelium cells and melanocytes present in the choroid. Pigment saturation was readily observed in retinas from TYR5 transgenic lines, without any obvious correlation with transgene copy number (Figure 2B and A–F). In contrast, progressively higher pigmentation correlated with transgene copy number in retinas from TYR5- $\alpha$ EHS-1.4 transgenic lines similar to coat color pigmentation (Figure 2B and G–K).

Second, we measured the *Tyr* gene expression by quantitative RT-PCR. *Tyr* expression was assessed in eyes and skin, where *Tyr* is normally expressed; and brain, where *Tyr* is not expressed (27). *Tyr* expression in TYR5 transgenic mouse lines was relatively uniform in the eye, erratic

in the skin, and absent in the brain (Figure 2C), and did not correlate with transgene copy number. In contrast, *Tyr* expression in TYR5- $\alpha$ EHS-1.4 transgenic lines was relatively proportional to transgene copy number in the eye, while it fully correlated with transgene copy number in the skin and, surprisingly, in the brain.

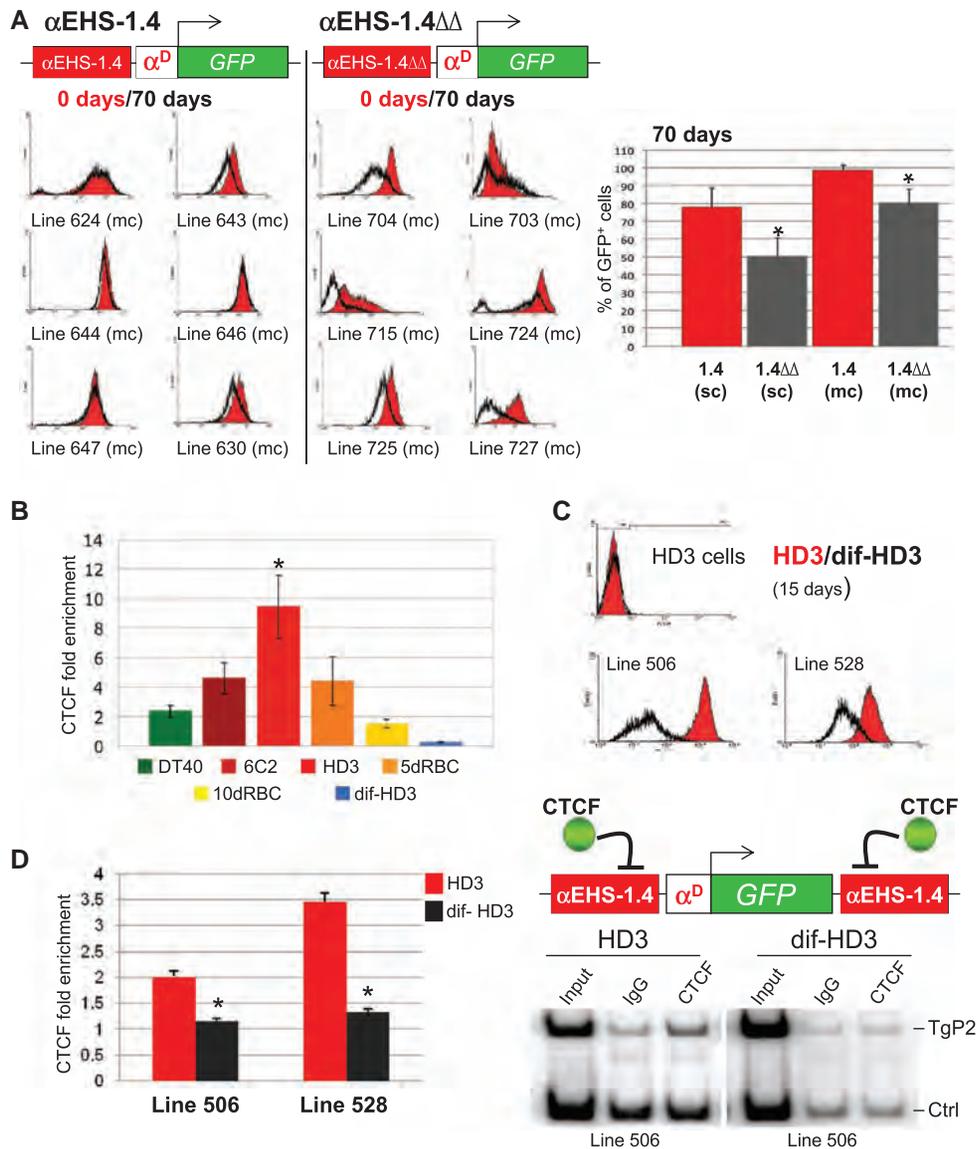
Finally, we quantified the total melanin contents in the eyes and skin of all transgenic lines (Supplementary Figure S2). Eye melanin contents of TYR5 transgenic mice were relatively uniform whereas in TYR5- $\alpha$ EHS-1.4 mice melanin contents correlated with transgene copy number. Likewise, melanin contents in the skin of TYR5 transgenic individuals were highly variable. In contrast, melanin contents in the dorsal skin of TYR5- $\alpha$ EHS-1.4 transgenic mice clearly correlated with transgene copy number. In conclusion, the  $\alpha$ EHS-1.4 insulator protects a transgene against chromosomal position effects and progressive silencing not only in stable cell lines but also in transgenic mice.

### CTCF-binding site is indispensable for $\alpha$ EHS-1.4 insulator activity

In order to address whether CTCF contributes to the protection against CPE by the  $\alpha$ EHS-1.4 insulator, we introduced two point mutations that abolish CTCF binding to this element (15) to generate the  $\alpha$ EHS-1.4 $\Delta\Delta$  insulator and assessed its protective function against CPE in stably transfected HD3 cells by flow cytometry. A very high proportion of cells from lines in which *GFP* was flanked on each side by the  $\alpha$ EHS-1.4 insulator showed high levels of fluorescence even after 70 days of continuous culture. In contrast, fluorescence levels decreased in an important proportion of cells carrying the mutant insulator  $\alpha$ EHS-1.4 $\Delta\Delta$  after 70 days (Figure 3A). These results are reflected in the percentages of *GFP*-expressing cells in selected stably transfected lines (Figure 3A). Interestingly, the protection against CPE by a single copy of the intact, or the mutant  $\alpha$ EHS-1.4, was less efficient than by multiple copies, suggesting that optimal insulation is reached only when the  $\alpha$ EHS-1.4 sequence flanks the transgene on each side (Figure 3A, compare single- and multi-copy cell lines). Thus, the mutant  $\alpha$ EHS-1.4 $\Delta\Delta$  insulator protects against CPE less efficiently than the  $\alpha$ EHS-1.4 insulator, suggesting that, unlike the chicken 5' cHS4  $\beta$ -globin insulator (18), CTCF binding is required for the insulator function of  $\alpha$ EHS-1.4.

### The activity of the $\alpha$ EHS-1.4 insulator correlates with CTCF occupancy during erythroid differentiation

Since the  $\alpha$ -MRE is located further upstream of the  $\alpha$ EHS-1.4 in the  $\alpha$ -globin locus, it seems unlikely that the  $\alpha$ EHS-1.4 acts as the 5' boundary limiting the  $\alpha$ -globin chromatin domain (Figure 1A). Therefore, we asked if the  $\alpha$ EHS-1.4 element regulates  $\alpha$ -globin gene expression and if it does so via CTCF. For these purpose we first addressed the CTCF binding to the endogenous  $\alpha$ EHS-1.4 during erythroid differentiation and chicken development by semi-quantitative ChIP. Interestingly, the occupancy of the  $\alpha$ EHS-1.4 insulator by CTCF varied



**Figure 3.** CTCF is necessary for  $\alpha$ EHS-1.4 insulator function and its *in vivo* occupancy varies during erythroid differentiation. **(A)** Transgene expression analysis by FACS at day 0 (red) and 70 (grey) in cultured cell lines with the respective transgene integrated in multi-copy ( $\alpha$ EHS-1.4 and mutant  $\alpha$ EHS-1.4 $\Delta\Delta$  insulator were placed only at the 5'-side of the transgene). Quantification of the average ( $\pm$  standard error) of GFP<sup>+</sup> cells at day 0 and 70 of culture of the set of single- and multi-copy clones.  $\alpha$ EHS-1.4 (sc)  $n = 16$ ;  $\alpha$ EHS-1.4 $\Delta\Delta$  (sc)  $n = 22$ ;  $\alpha$ EHS-1.4 (mc)  $n = 18$  and  $\alpha$ EHS-1.4 $\Delta\Delta$  (mc)  $n = 22$ . \* $P \leq 0.05$  with respect to the  $\alpha$ EHS-1.4 (ANOVA). **(B)** CTCF *in vivo* binding at the  $\alpha$ EHS-1.4 insulator sequence was estimated by semi-quantitative ChIP assays (29). We observed an enrichment peak of CTCF in HD3 cells (erythroblasts), and as differentiation proceeds (dif-HD3), CTCF dissociates from the insulator. \* $P \leq 0.05$ . The number of independent ChIP assays per cell line is: DT40  $n = 3$ , 6C2  $n = 5$ , HD3  $n = 4$ , 5dRBC  $n = 3$ , 10dRBC  $n = 3$  and dif-HD3  $n = 4$ . **(C)** Protection against CPE in two representative differentiated HD3 cell lines containing the plasmid insulated with the  $\alpha$ EHS-1.4 element (the same experiment was performed using the  $\beta$ -globin 5' *chs4* insulator, see [Supplementary Figure S3B](#)). **(D)** ChIP using an antibody against CTCF and primers recognizing the transgene. CTCF binds to the transgene in HD3 cells and is dissociated in differentiated HD3 cells mimicking what happens at the endogenous context (see **B**), consistent with a loss of insulator function in the differentiated clones. A representative gel of the radioactive duplex-PCR is shown. Results are shown as averages ( $\pm$  standard error). \* $P \leq 0.05$  with respect to the lines 506 or 528 (ANOVA).

depending on the differentiation stage analyzed. CTCF abundance peaked at the pre-erythroblasts stage, represented by HD3 cells, and, it drastically diminished as differentiation proceeds into a mature erythrocyte, represented by red blood cells from 10-day-old chicken embryos (10dRBCs), or differentiated HD3 cells (Figure 3B). Western blots showed similar CTCF total protein levels in undifferentiated and differentiated HD3

cells ([Supplementary Figure S4A](#), right panel), suggesting regulated binding of CTCF to the  $\alpha$ EHS-1.4 element. In contrast to the  $\alpha$ EHS-1.4 insulator, the CTCF's highest enrichment on the chicken 5' *chs4*  $\beta$ -globin insulator was detected in 10dRBCs ([Supplementary Figure S3](#)), suggesting a different mode of action of CTCF for each of these elements. Next, we analyzed the protective function of the  $\alpha$ EHS-1.4 insulator against CPE in undifferentiated

cells (HD3) and terminally differentiated erythroblasts (dif-HD3). Two HD3 cell lines carrying the  $\alpha^D$ -GFP transgene shielded by the  $\alpha$ EHS-1.4 (lines 506 and 528) were differentiated for 15 days (21), and GFP fluorescence was measured by flow cytometry. GFP fluorescence of the differentiated cells reproducibly decreased as compared to their undifferentiated counterpart (Figure 3C). This correlates with decreased occupancy of the endogenous  $\alpha$ EHS-1.4 insulator by CTCF in 10dRBCs (Figure 3B), and of the transgene incorporated in differentiated HD3 cells (Figure 3D). In contrast, differentiated HD3 lines carrying the 5'cHS4  $\beta$ -globin insulator did not show a significant fluorescence decrease (Supplementary Figure S3B). Therefore, the decreased protective effect of the  $\alpha$ EHS-1.4 against CPE might be due to decreased CTCF binding.

### Reduction of CTCF protein levels and inhibition of PARP-1 affects insulator function

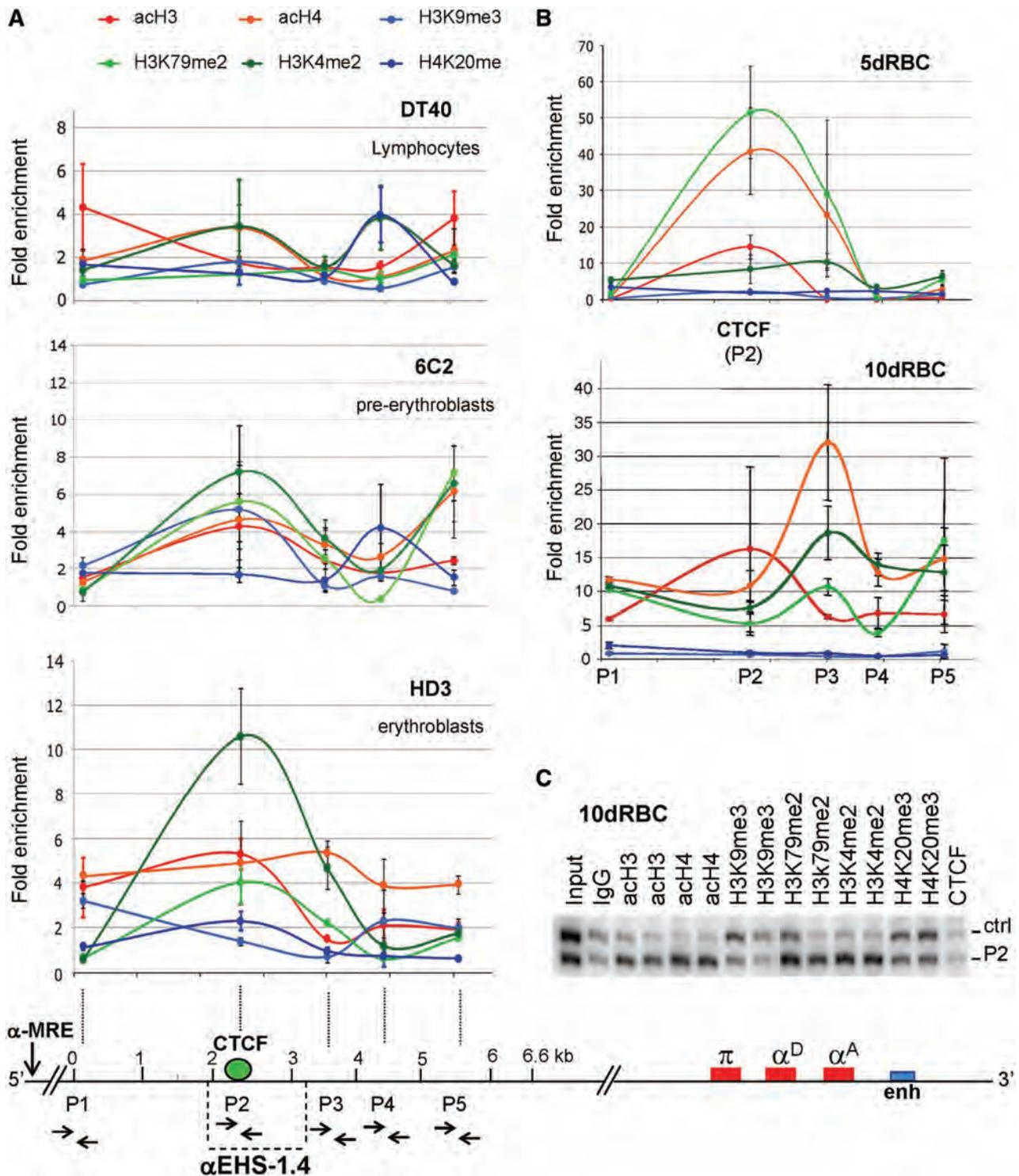
Based on the data supporting dynamic CTCF binding at the  $\alpha$ EHS-1.4 insulator during differentiation, we decided to knock down CTCF expression and inhibit its potential poly(ADP-ribosylation) (Supplementary Figure S4). First, we generated a series of cell lines in which CTCF was stably knocked down by transfection of three different sets of small hairpin RNAs. Decreased CTCF protein was confirmed by western blot (Supplementary Figure S4A, left panel). The reduction of CTCF decreased the protective effect of the  $\alpha$ EHS-1.4 insulator against CPE (Supplementary Figure S4B, left panel). ChIP assays confirmed decreased CTCF at the  $\alpha$ EHS-1.4 insulator upon CTCF knock down (Supplementary Figure S4C). Poly(ADP-ribosylation) has been shown to be needed for CTCF insulator function (14,34–37), therefore we addressed whether inactivation of CTCF by inhibition of the poly(ADP) ribose polymerase-1 (PARP-1) impacts on the protection against CPE conferred by the  $\alpha$ EHS-1.4 insulator. PARP-1 post-translationally modifies CTCF by adding ADP polymers to its N-terminal domain, resulting in two electrophoretic forms, one migrating at 130 kDa, which contains few ADP-ribose residues, and the other one migrating at 180 kDa, which contains more than 20 ADP-ribose residues (36). Treatment of the HD3 transgenic cell lines 506 and 528, which carry GFP flanked on each side by the  $\alpha$ EHS-1.4 insulator, with 3-ABA, an inhibitor of PARP-1 (36), resulted in a diminishment of transgene protection against CPE along with a decrease of the 180 kDa form of CTCF (Supplementary Figure S4A, right panel). Interestingly, inhibition of poly(ADP-ribosylation) did not affect CTCF binding to the  $\alpha$ EHS-1.4 insulator, as shown by ChIP (Figure 6A), implying that, while affecting its function on the  $\alpha$ EHS-1.4 insulator, inhibition of CTCF PARylation does not affect its DNA-binding properties.

Taken together, these results further support the notion that activity of the  $\alpha$ EHS-1.4 insulator is CTCF-dependent and that such activity can be, at least in part, modulated by post-translational modification.

### The genomic region encompassing the $\alpha$ EHS-1.4 insulator progressively structures as transcriptionally permissive chromatin during erythroid differentiation

As a first approach to understand the function of  $\alpha$ EHS-1.4 element on the chromatin organization of the  $\alpha$ -globin domain, we characterized the chromatin structure of a 6.6-kb genomic region encompassing the  $\alpha$ EHS-1.4 element during erythroid differentiation and development (Figure 4A). We performed ChIP assays for several histone post-translational modifications corresponding to transcriptionally permissive (acH3, acH4, H3K4me2 and H3K79me2) and repressive (H3K9me3 and H4K20me3) chromatin conformations. We found that the enrichment in histone modifications along the targeted genomic region varies during adult erythroid differentiation (Figure 4A). In 6C2 cells, which represent a pre-erythroblast differentiation stage, a modest enrichment in acetylated histones, as well as H3K4me2 and H3K79me2 was observed over the  $\alpha$ EHS-1.4 insulator (Figure 4A). As differentiation proceeds into an erythroblast stage represented by HD3 cells, there was a notorious peak of permissive histone marks overlapping with the CTCF-binding sequence at the  $\alpha$ EHS-1.4 element (Figure 4A). Importantly,  $\alpha$ -globin genes are not transcribed in 6C2 or HD3 cells (21). Thus, these results suggest that the  $\alpha$ EHS-1.4 element may contribute to the establishment of an open configuration of the domain prior to adult  $\alpha$ -globin gene expression. Such hypothesis is further supported by notorious enrichment of transcriptionally permissive histone modifications on the same genomic region in 5dRBC [Figure 4B, CTCF(P2)], in which the embryonic  $\pi$  gene is transcribed (21). Then, the 6.6-kb genomic region encompassing the  $\alpha$ EHS-1.4 insulator structures as transcriptionally permissive chromatin after globin gene expression switching, i.e. in 10dRBCs, when the adult  $\alpha$ -globin genes ( $\alpha^D$  and  $\alpha^A$ ) are robustly expressed, while the embryonic  $\pi$  gene is silenced (21). The DT40 lymphoid cell line, in which the  $\alpha$ -globin gene domain is inactive, was used as a control (21,38). Thus, the progressive arrangement of the 6.6-kb genomic region containing the  $\alpha$ EHS-1.4 insulator into transcriptionally permissive chromatin correlates with  $\alpha$ -globin locus activation and the onset of adult  $\alpha$ -globin gene expression. These results are in contrast to the timing of  $\alpha$ EHS-1.4 insulator function and CTCF occupancy, which rise in erythroblasts, and decrease in mature erythrocytes (Figure 3B).

Taken together, our results support a model in which the  $\alpha$ EHS-1.4 insulator, located between the domain's putative  $\alpha$ -MRE and the  $\alpha$ -globin genes might act by blocking the transcriptional influence of the  $\alpha$ -MRE over the  $\alpha$ -globin adult genes at the erythroblast stage. Then, as differentiation proceeds towards mature erythrocytes, CTCF dissociates from the  $\alpha$ EHS-1.4 insulator, releasing its blocking activity over the  $\alpha$ -MRE and allowing the establishment of transcriptionally active chromatin and *trans*-activation of the adult genes.



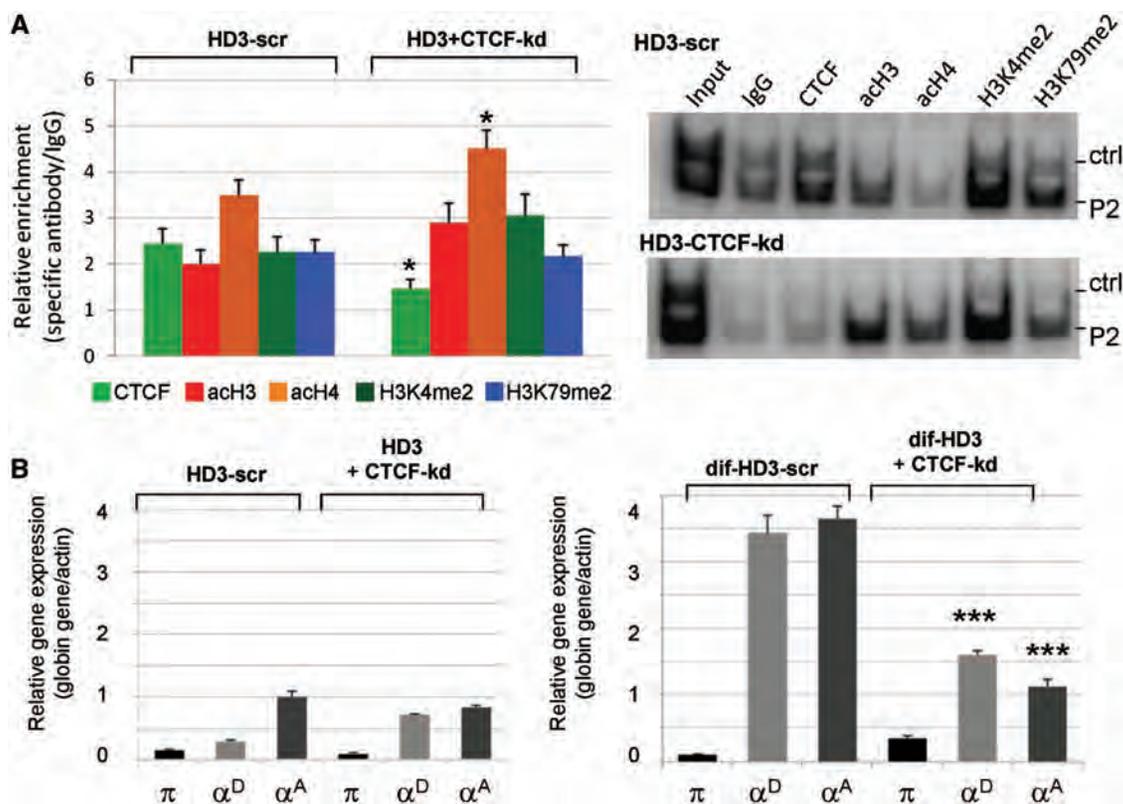
**Figure 4.** Distribution of histone post-translational modifications over the 6.6-kb genomic region containing the  $\alpha$ EHS-1.4 insulator. (A) Histone covalent modifications were evaluated semi-quantitatively using duplex-PCR with primers P1–P5 (primers locations are shown in the bottom scheme) in DT40 (lymphoid), 6C2 (pre-erythroblasts) and HD3 (erythroblasts) cells. (B) ChIP assays as in (A) using 5d and 10dRBCs. (C) Representative gel of the duplex-PCR amplification using primers P2 located over the CTCF-binding site with its corresponding control (Ctrl). The ChIP was performed on 10-day-old embryonic erythrocytes (10dRBC). All the histone modification enrichments are statistically significant (ANOVA with  $P \leq 0.001$ ) with respect to the control IgG except for the acH3 enrichment. Results are shown as averages ( $\pm$  standard error).

### Loss of $\alpha$ EHS-1.4 insulator function results in active chromatin configuration and precocious adult $\alpha$ -globin gene expression

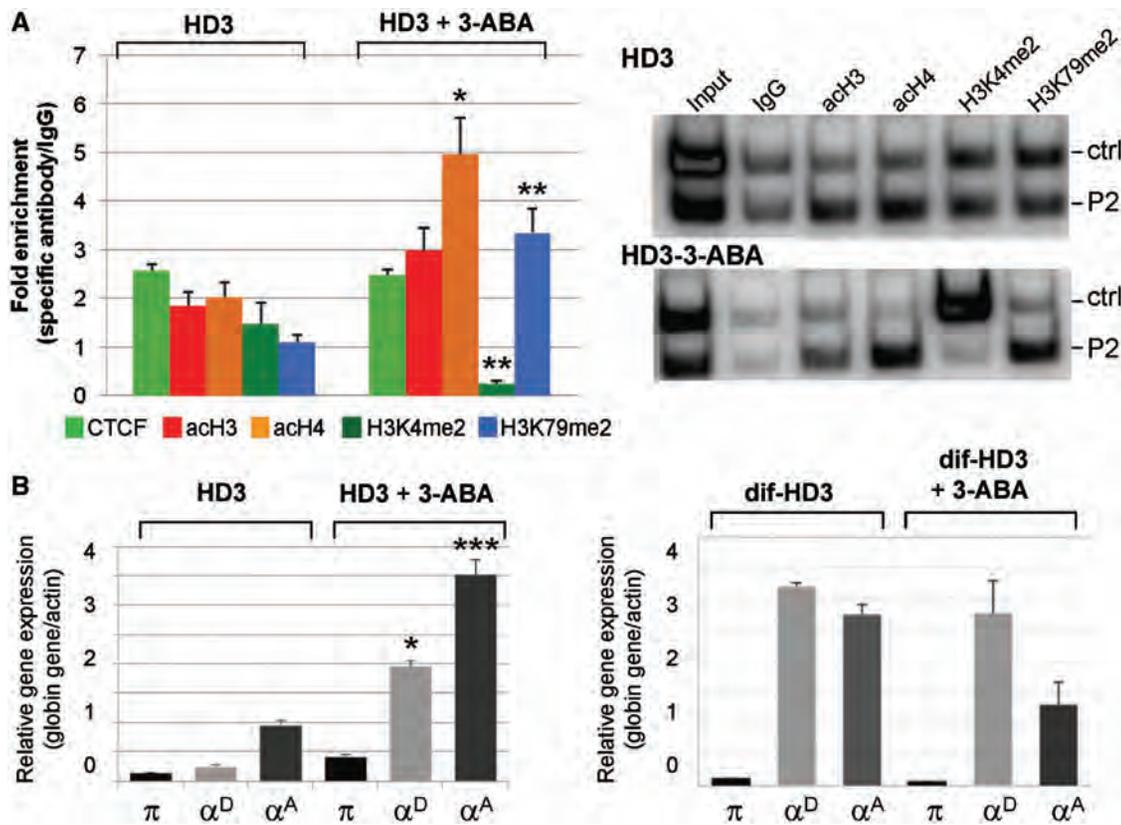
To start testing the proposed model in which the  $\alpha$ EHS-1.4 insulator blocks the  $\alpha$ -MRE over  $\alpha$ -globin gene expression in erythroblasts, we addressed the chromatin configuration as well as  $\alpha$ -globin gene expression in cells with decreased or inactive CTCF. Open chromatin histone modifications at the  $\alpha$ EHS-1.4 insulator region were increased in undifferentiated HD3 cells with knocked-down CTCF (Figure 5A). As expected,  $\alpha$ -globin gene expression was not detected in untreated HD3 cells (Figure 5B) (21). No changes were observed in undifferentiated HD3 cells with decreased CTCF (Figure 5B, left panel). In contrast, in differentiated HD3 cells with decreased CTCF the adult  $\alpha^D$  and  $\alpha^A$  globin genes did not get fully activated (Figure 5B, right panel).

A similar set of experiments was performed in cells treated with 3-ABA inhibitor of PARP-1 enzyme. Open chromatin histone modifications at the  $\alpha$ EHS-1.4 insulator region were generally increased in undifferentiated

HD3 cells in which CTCF function was inactivated by 3-ABA, as compared to untreated cells (Figure 6A). Accordingly, HD3-treated cells showed premature activation of the adult  $\alpha^D$  and  $\alpha^A$  globin genes in contrast with CTCF knock-down cells (Figure 6B). No significant differences in histone modifications or  $\alpha$ -globin gene expression were found in treated differentiated HD3 cells. Two main conclusions emerge from this set of experiments. First, inhibition of CTCF poly(ADP-ribosylation) in HD3 erythroblasts mimics, at least in part, the chromatin structure and the expression of the adult  $\alpha^D$  and  $\alpha^A$  globin genes in 10dRBCs (Figures 4B and 6). Second, knock down of CTCF perturbs more dramatically the chromatin structure of the  $\alpha$ -globin domain, as compared to inhibition of CTCF poly(ADP-ribosylation), while impeding full globin gene activation (Figure 5B, right panel). Thus, it is possible that poly(ADP-ribosylation) regulates the activity of CTCF on the  $\alpha$ EHS-1.4 insulator at its natural context. Together, these results support a model in which the  $\alpha$ EHS-1.4 promotes optimal chromatin configuration and prevents  $\alpha$ -MRE-mediated adult  $\alpha$ -globin genes *trans*-activation at the erythroblast stage.



**Figure 5.** Reduction of CTCF alters the profile of histone covalent modifications at the  $\alpha$ EHS-1.4 insulator region and  $\alpha$ -globin genes expression. (A) ChIP analysis of histone modifications in HD3 cells (HD3) and HD3 cells knocked down for CTCF (HD3+CTCF-kd). Each ChIP assay was performed in triplicate. Representative radioactive duplex PCR gels of the HD3-scramble (HD3-scr) and HD3 cells knocked down for CTCF (HD3+CTCF-kd) are shown. (B) Evaluation of  $\alpha$ -globin gene expression in HD3 cells (HD3) and HD3 cells knocked down for CTCF in normal conditions (undifferentiated, HD3+CTCF-kd) and differentiated HD3 cells (dif-HD3). The results are representative of nine independent PCR reactions from two independent total RNA purifications and reverse transcriptase reactions. There is a premature chromatin opening of the insulator region in the CTCF knock down cells, and the adult genes are unable to be expressed in the HD3 differentiated cells. \* $P \leq 0.05$  (MANOVA) and \*\*\* $P \leq 0.001$  (MANOVA) with respect to the corresponding histone covalent modification or expressed gene in HD3 or dif-HD3 scramble controls, respectively.



**Figure 6.** Poly(ADP-ribose) polymerase-1 inhibitor affects histone modifications associated with the  $\alpha$ EHS-1.4 insulator and  $\alpha$ -globin genes expression. (A) ChIP analysis of histone covalent modifications in HD3 cells (HD3) and cells incubated with the PARP-1 inhibitor (HD3+3-ABA) using P2 primers (Figure 4). Each ChIP was performed in triplicate. Representative radioactive duplex PCR gels for HD3 cells and HD3 cells incubated with the 3-ABA inhibitor are shown. (B) Semi-quantitative duplex RT-PCR of  $\alpha$ -globin genes expression in HD3 cells and HD3 cells treated with 8 mM of 3-ABA (left graph). The same experiment was performed in differentiated HD3 cells (dif-HD3) and differentiated HD3 cells incubated with 3-ABA (dif-HD3+3-ABA, right panel). The results are representative of nine independent PCR reactions from two independent total RNA purification and reverse transcriptase reactions. \* $P \leq 0.05$  (MANOVA), \*\* $P \leq 0.01$  (MANOVA) and \*\*\* $P \leq 0.001$  (MANOVA), with respect to the corresponding histone covalent modification or expressed gene in HD3 cells. Results are shown as averages ( $\pm$  standard error).

#### $\alpha$ -MRE *trans*-activation potentiation by CTCF release from the $\alpha$ EHS-1.4 insulator during differentiation

In an attempt to validate the proposed model we generated pools of stable HD3 cell lines comparing the activity of the  $\alpha^D$  gene promoter alone, the same promoter with the  $\alpha$ -MRE element and a construct with the  $\alpha$ MRE, the  $\alpha$ -EHS-1.4 and the  $\alpha^D$  gene promoter mimicking their endogenous organization (Figure 7).

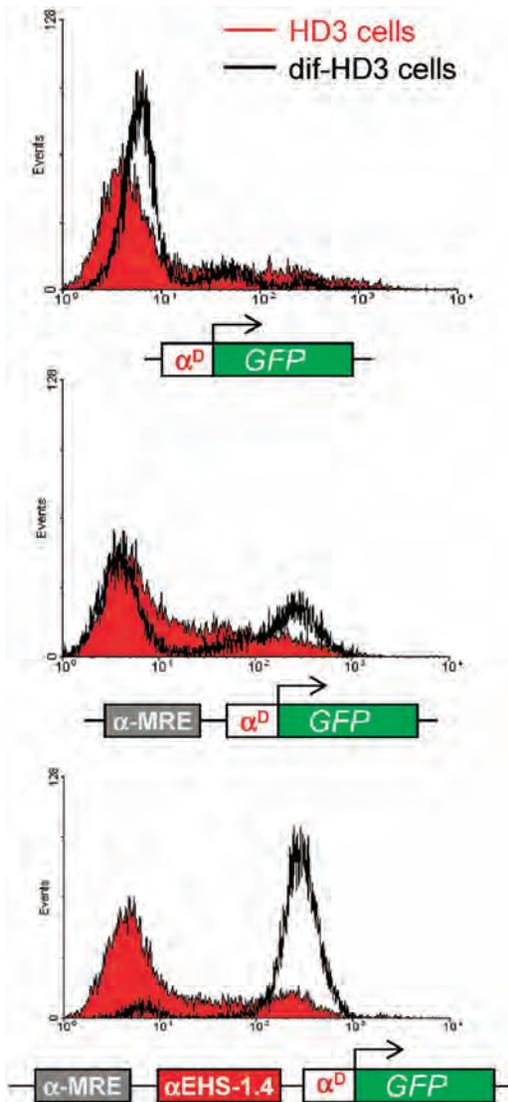
Based on our previous observations we propose that CTCF dissociation from the  $\alpha$ EHS-1.4 insulator in mature erythrocytes (differentiated HD3 cells) increase  $\alpha$ -MRE *trans*-activation potential. Indeed, we found that under non-differentiated conditions the majority of the reporter gene activity is not *trans*-activated, instead when HD3 cells are differentiated into mature erythrocytes we detected a robust *trans*-activation of the reporter gene (Figure 7) consistent with the dissociation of CTCF over the insulator in the transgene (Figure 3B and C).

These results further support the model in which the  $\alpha$ EHS-1.4 insulator regulates the  $\alpha$ -MRE *trans*-activating function over the adult  $\alpha$ -globin genes during adult erythroid differentiation.

#### DISCUSSION

Genomic organization into chromatin domains represents an important level of gene expression regulation achieved by regulatory elements located mainly at intergenic regions (2,3). Here we studied the function of the  $\alpha$ EHS-1.4 insulator in the chromatin organization and gene expression regulation of the  $\alpha$ -globin gene domain. We previously demonstrated that the  $\alpha$ EHS-1.4 insulator, embedded in the 5' non-coding region of the  $\alpha$ -globin domain, possesses enhancer blocking activity (15). With the aim to better understand the insulator properties of the  $\alpha$ EHS-1.4 element and its role in the chicken  $\alpha$ -globin domain, we first demonstrate that the  $\alpha$ EHS-1.4 element is able to shield a transgene from CPE in stably transfected cell lines and in mice in a CTCF-dependent manner. This differs from the 5'cHS4 insulator in which only enhancer blocking function is CTCF-dependent, whereas protection against CPE needs binding of USF1 and USF2 factors (18,39). Since no USF1/USF2 factors are present at the  $\alpha$ EHS-1.4 (data not shown), a different protein complex most likely regulates  $\alpha$ EHS-1.4 insulator activity.

Importantly, CTCF occupancy of the  $\alpha$ EHS-1.4 insulator varies during erythroid differentiation. This raises the



**Figure 7.**  $\alpha$ MRE element is able to *trans*-activate the  $\alpha^D$  promoter upon differentiation to the mature erythrocyte stage. Flow cytometry histograms (FACS) showing the GFP fluorescence level of pools of stably transfected HD3 cells with three different constructs,  $\alpha^D$ ,  $\alpha$ MRE- $\alpha^D$  and  $\alpha$ MRE- $\alpha$ EHS-1.4- $\alpha^D$  (showed underneath each graph). The full red graph corresponds to the erythroblast stage (HD3 cells) and the black empty graph corresponds to the same pool subjected to differentiation (HD3-dif).

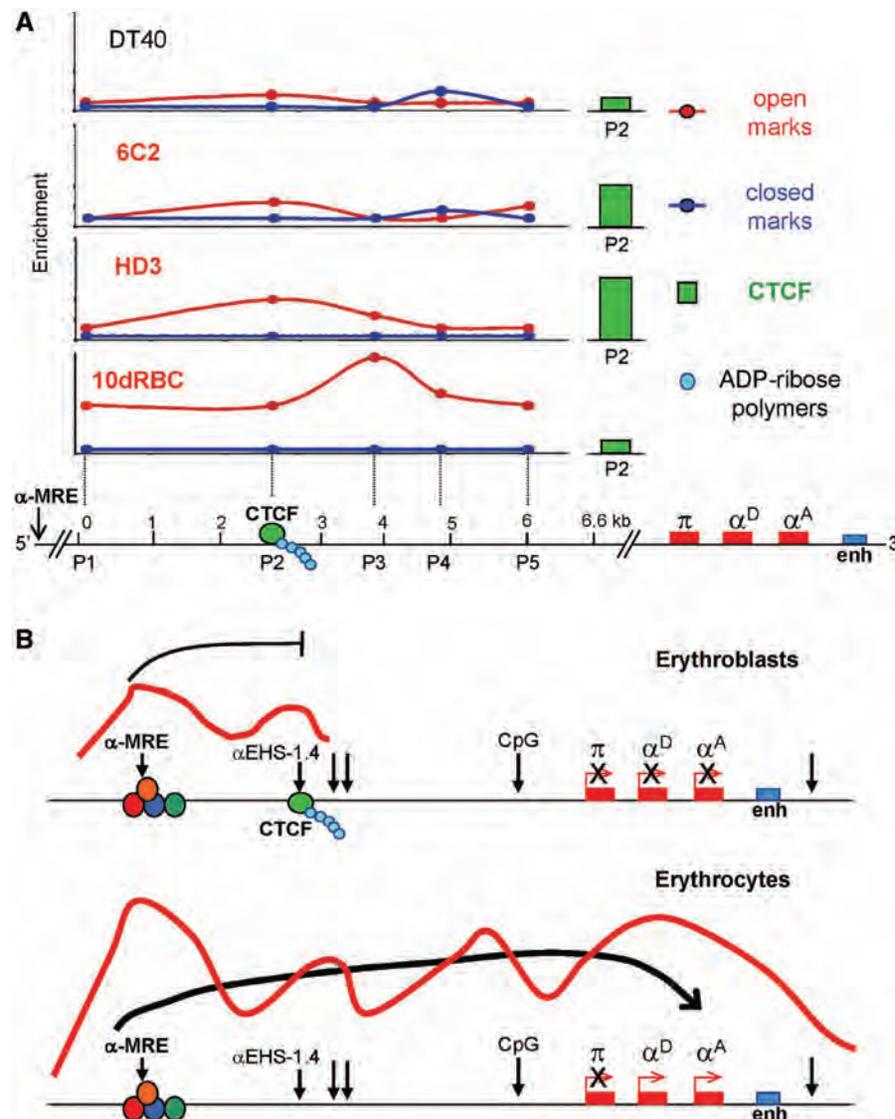
possibility that the  $\alpha$ EHS-1.4 acts as a stage-specific insulator probably influencing the activity of the  $\alpha$ -MRE on the expression of adult  $\alpha^D$  and  $\alpha^A$  genes. Indeed, the activity of  $\alpha$ EHS-1.4 is regulated, and peaks during the switching from early to late adult erythroid differentiation, when an erythroid-specific signal might favor CTCF dissociation. We propose that the function of the  $\alpha$ EHS-1.4 insulator is to block the activity of the  $\alpha$ -MRE distal element during early adult erythroid differentiation. Once differentiation takes place, CTCF dissociates and open chromatin propagates 3' towards the gene cluster (Figure 8A). Possibly the  $\alpha$ EHS-1.4 element regulates the poised state of the RNA polymerase II and pre-initiation complex components found at the  $\alpha$ -MRE

(40). The RNA polymerase II blockage by CTCF has been described in other loci (41,42). In such a hypothetical scenario, CTCF release from the  $\alpha$ EHS-1.4 insulator may favor the chromosomal conformation changes and the signal spreading needed for the interaction between the  $\alpha$ -MRE and the adult  $\alpha^D$  and  $\alpha^A$  gene promoter regions and their transcriptional activation (Figure 8B). This model represents a novel regulatory role for a CTCF-dependent chromatin insulator involved in a differentiation pathway. In summary, at the endogenous context the  $\alpha$ EHS-1.4 insulator may have enhancer blocking activity (presumably over the  $\alpha$ -MRE) and barrier function to contribute to chromatin structure changes at the domain scale. Such dual activity is thus in part CTCF-dependent, and probably can also be driven by different stage-specific nuclear factors.

A possible mechanism for CTCF release regulation could come from the presence of intergenic transcripts at the  $\alpha$ -globin domain (43 and unpublished results). It has been recently demonstrated that for lysozyme gene activation by pro-inflammatory stimuli there is a CTCF-cohesin complex eviction by the transient stimulation of a non-coding RNA (31).

Additionally, we demonstrated that poly(ADP-ribosylation) is modulating CTCF activity at the  $\alpha$ EHS-1.4 insulator (Supplementary Figure S4 and Figure 6). When PARP-1 is inhibited at the erythroblast stage we are able to mimic the effect of CTCF release during differentiation, i.e. an increase in histone open marks and the activation of the adult  $\alpha^D$  and  $\alpha^A$  globin genes (Figure 6). In contrast, CTCF knock down do not cause the activation of any  $\alpha$ -globin gene in erythroblasts (Figure 5B) and impedes  $\alpha$ -globin gene *trans*-activation at the differentiated stage. From these results several interpretations emerged. One predicts that CTCF poly(ADP-ribosylation) does not occur in all the CTCFs bound *in vivo* along the domain and that the CTCF located at the  $\alpha$ EHS-1.4 insulator is modified post-translationally by PARP-1 (36). Furthermore, it is possible that not all CTCF associated complexes located along the domain possess the same peptidic composition and activity. This is an aspect that is currently under investigation. We predict that some CTCF-associated complexes are more structural and others, like the one linked to the  $\alpha$ EHS-1.4, should have a regulatory function. Based on such predictions, it is now attractive to look for CTCF partners at different locations in the domain, including PARP-1. For example, CHD8, cohesins, lamin or CTCF itself among others (3,32) which would further support the proposal that poly(ADP-ribosylation) could form a docking site for multiple protein interactions with regulatory and chromatin functions (44).

A different interpretation, that does not exclude the previous one, has to do with the spatial organization of the chicken  $\alpha$ -globin domain. We propose that CTCF is possibly contributing to this 3D organization of the domain, and therefore when it is knocked down, there is a generalized reduction of the protein levels affecting the active chromatin hub formation and consequently the activation of the adult  $\alpha$ -globin genes (Figure 5). This is not surprising since it has previously been demonstrated that



**Figure 8.** Summary and model for the  $\alpha$ EHS-1.4 insulator role at the chicken  $\alpha$ -globin locus. (A) Summary of histone post-translational modifications and CTCF *in vivo* enrichment over the 6.6-kb genomic region containing the  $\alpha$ EHS-1.4 insulator. As adult erythroid differentiation proceeds there is an increment in histone modifications corresponding to a permissive chromatin environment over the  $\alpha$ EHS-1.4 insulator and the CTCF site (see P2 primer location) that expands and further increases over the entire region in 10dRBC when the adult genes  $\alpha^D$  and  $\alpha^A$  are switched on. CTCF is present during differentiation with a peak at the erythroblast stage (HD3 cells) and then released at the mature 10dRBC when the adult  $\alpha^D$  and  $\alpha^A$  genes are being expressed (see right side green graph bar). P1–P5 represent the primers covering the 6.6-kb genomic region used for ChIP assays and P2 are the primers located over the CTCF-binding site. (B) Model suggesting a mechanism of action for the  $\alpha$ EHS-1.4 insulator at the level of the chromatin structure at the domain scale and adult  $\alpha$ -globin gene expression. Vertical arrows represent DNase I hypersensitive sites. We propose that at the erythroblast stage the  $\alpha$ EHS-1.4 insulator blocks the effects coming from the  $\alpha$ -MRE concomitant with the fact that the adult  $\alpha^D$  and  $\alpha^A$  globin genes remain silenced. At the late erythrocyte stage CTCF dissociation from the insulator facilitates  $\alpha$ -MRE activity allowing *trans*-activation of  $\alpha$ -globin adult genes.

CTCF can mediate long-distance chromatin loop formation in the mouse  $\beta$ -globin locus (10). We predict that CTCF is not the only factor participating as it has been shown that LCR long-distance contacts can also be dependent on EKLf and GATA-1/FOG-1 (45,46), factors that also bind different regions of the  $\alpha$ -globin gene domain (21).

In this regard, the 3D structuring of the  $\alpha$ -globin domain was recently suggested by a chromatin conformation capture approach, which showed the formation of an active chromatin hub by the interaction between the

$\alpha$ -MRE, a DNase I hypersensitive site of unknown function located  $-9$ -kb upstream of the  $\pi$  gene ( $-9$ -kb DHS), the CpG-island corresponding to the promoter of the *C16orf35* gene and the promoter region of the adult  $\alpha^D$  gene (Figure 1) (38). Interestingly, the  $\alpha$ EHS-1.4 element, located around  $-14$ -kb upstream the embryonic  $\pi$  gene, is contained in the loop formed by contacts between the  $\alpha$ -MRE and the  $-9$ -kb DHS, thus the  $\alpha$ EHS-1.4 element appears not to contribute directly to hub formation in this context. Whether alternative CTCF sites along the  $\alpha$ -globin domain, which we recently identified

(unpublished observations), contribute to domain organization via interaction of CTCF and its co-factors, would require further investigation.

In summary, we propose that the  $\alpha$ EHS-1.4 insulator contributes to regulate the chromatin structure and differential gene expression of the chicken  $\alpha$ -globin gene domain throughout erythroid differentiation. We suggest that the regulated interaction of CTCF with its binding site could represent a way to control the function of the  $\alpha$ EHS-1.4 element over  $\alpha$ -globin genes expression, together with post-translational modifications like the poly(ADP-ribosylation) that modulate CTCF activity. In terms of the chromatin configuration of the locus, multiple CTCF-binding sites and their associated factors may facilitate the formation of a chromatin hub in a regulated manner. Our results suggest that the intergenic regulatory and structural components of a chromatin domain are critical to achieve regulated gene expression.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors thank Paul Delgado-Olguín, Catherine Farrell and Gary Felsenfeld for invaluable suggestions on the manuscript. The authors also thank Martín Escamilla-del-Arenal, Héctor Rincón-Arango and all members of Félix Recillas-Targa laboratory for constant scientific discussions. The authors thank Fabian Flores Jasso for his help generating the shRNA used against CTCF. The authors thank L. Ongay, G. Codiz and M. Mora from the Unidad de Biología Molecular, IFC-UNAM, for the DNA sequencing and FACS facility. The CNB Transgenic Facility is acknowledged for producing the transgenic mice reported in this article. The authors are most grateful to Almudena Tello, Diego Muñoz, Zahira Corrales, and Esther Zurita for technical assistance.

## FUNDING

Dirección General de Asuntos del Personal Académico-Universidad Nacional Autónoma de México (IN209403 and IN214407); Consejo Nacional de Ciencia y Tecnología (CONACyT: 42653-Q and 58767); PhD fellowship from CONACyT (170087 and 207989) and Dirección General de estudios de Posgrado-Universidad Nacional Autónoma de México (DGEP) (to M.F.-M. and E.G.-B.); PhD Graduate Program of Doctorado en Ciencias Biomédicas, UNAM; Spanish Ministry of Education and Science (BFU2006-12185); Spanish Ministry of Science and Innovation (BIO2009-12697) (to L.M.). Funding for open access charge: Universidad Nacional Autónoma de México (UNAM), Instituto de Fisiología Celular.

*Conflict of interest statement.* None declared.

## REFERENCES

- West, A.G. and Fraser, P. (2005) Remote control of gene transcription. *Hum. Mol. Genet.*, **14**, R101–R111.
- Wallace, J.A. and Felsenfeld, G. (2007) We gather together: insulators and genome organization. *Curr. Opin. Genet. Develop.*, **17**, 400–407.
- Phillips, J.E. and Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.
- Dorsett, D. (2007) Roles of the sister chromatin cohesion apparatus in gene expression, development and human syndromes. *Chromosoma*, **116**, 1–13.
- Hou, C., Dale, R. and Dean, A. (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc. Natl Acad. Sci. USA*, **107**, 3651–3656.
- Bushey, A.M., Ramos, E. and Corces, V.G. (2009) Three subclasses of a *Drosophila* insulator show distinct and cell-type-specific genomic distributions. *Genes Dev.*, **23**, 1338–1350.
- Murrell, A., Heeson, S. and Reik, W. (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. Genet.*, **36**, 889–893.
- Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenko, V., Reik, W. and Ohlsson, R. (2006) CTCF binding at the *H19* imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. *Proc. Natl Acad. Sci. USA*, **103**, 10684–10689.
- Ribeiro de Almeida, C., Heath, H., Krpic, S., Dingjan, G.M., van Hamburg, J.P., Bergen, I., van de Nobelen, S., Grosveld, F., Galjart, N. and Hendriks, R.W. (2009) Critical role of the transcription regulator CCCTC-binding factor in the control of the Th2 cytokine expression. *J. Immunol.*, **182**, 999–1010.
- Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the  $\beta$ -globin locus. *Genes Dev.*, **20**, 2349–2354.
- Majumder, P., Gomez, J.A., Chadwick, B.P. and Boss, J.M. (2008) The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J. Exp. Med.*, **205**, 785–798.
- Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qui, X.W., Cherry, A.M. and Hoffman, A.R. (2006) CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science*, **312**, 269–272.
- Lutz, M., Burke, L.J., LeFevre, P., Myers, F.A., Thorne, A.W., Crane-Robinson, C., Bonifer, C., Filippova, G.N., Lobanenko, V. and Renkawitz, R. (2003) Thyroid-hormone regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. *EMBO J.*, **22**, 1579–1587.
- Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavoosidana, G., Mukhopadhyay, R., Kanduri, C. *et al.* (2004) Poly(ADP-ribosylation) regulates CTCF-dependent chromatin insulation. *Nat. Genet.*, **36**, 1105–1110.
- Valadez-Graham, V., Razin, S.V. and Recillas-Targa, F. (2004) CTCF-dependent enhancer blockers at the upstream region of the chicken  $\alpha$ -globin gene domain. *Nucleic Acids Res.*, **32**, 1354–1362.
- Vyas, P., Vickers, M.A., Picketts, D.J. and Higgs, D.R. (1995) Conservation of position and sequence of a novel, widely expressed gene containing the major human  $\alpha$ -globin regulatory element. *Genomics*, **29**, 679–689.
- Flint, J., Tufarelli, C., Peden, J., Clark, K., Daniels, R.J., Hardison, R., Miller, W., Philipson, S., Tan-Un, K.C., McMorrow, T. *et al.* (2001) Comparative genome analysis delimits a chromosomal domain and identified key regulatory elements in the  $\alpha$  globin cluster. *Hum. Mol. Genet.*, **10**, 371–382.
- Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M. and Felsenfeld, G. (2002) Position-effect protection and enhancer blocking by the chicken  $\beta$ -globin insulator are separable activities. *Proc. Natl Acad. Sci. USA*, **99**, 6883–6888.
- Beug, H., von Kirchbach, A., Doderlin, G., Conscience, J.F. and Graf, T. (1979) Chicken haematopoietic cells transformed by seven

- strains of defective avian leukaemia viruses display three distinct phenotypes of differentiation. *Cell*, **18**, 375–390.
20. Dieken, E.S., Epner, E.M., Fiering, S., Fournier, R.E. and Groudine, M. (1996) Efficient modification of human chromosomal alleles using recombination-proficient chicken/human microcell hybrids. *Nat. Genet.*, **12**, 174–182.
  21. Rincón-Arano, H., Guerrero, G., Valdes-Quezada, C. and Recillas-Targa, F. (2009) Chicken  $\alpha$ -globin switching depends on autonomous silencing of the embryonic  $\pi$  globin gene by epigenetics mechanisms. *J. Cell. Biochem.*, **108**, 675–687.
  22. Iarovaia, O., Razin, S.V., Linarez-Cruz, G., Sjakste, N. and Scherrer, K. (2001) In chicken leukemia cells globin genes are fully transcribed but their RNAs are retained in the perinucleolar area. *Exp. Cell. Res.*, **270**, 159–165.
  23. Pikaart, M.J., Recillas-Targa, F. and Felsenfeld, G. (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.*, **12**, 2852–2862.
  24. Beermann, F., Ruppert, S., Hummler, E. and Schütz, G. (1991) Tyrosinase as a marker for transgenic mice. *Nucleic Acids Res.*, **19**, 958.
  25. Giraldo, P., Martínez, A., Regales, L., Lavado, A., García-Díaz, A., Alonso, A., Busturia, A. and Montoliu, L. (2003) Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. *Nucleic Acids Res.*, **31**, 6290–6305.
  26. Ganss, R., Montoliu, L., Monaghan, A.P. and Schütz, G. (1994) A cell-specific enhancer far upstream of the mouse tyrosinase gene confers high level and copy number-related expression in transgenic mice. *EMBO J.*, **13**, 3083–3093.
  27. Giménez, E., Lavado, A., Giraldo, P. and Montoliu, L. (2003) Tyrosinase gene expression is not detected in mouse brain outside the retinal pigment epithelium cells. *Eur. J. Neurosci.*, **18**, 2673–2676.
  28. Lavado, A., Olivares, C., García-Borrón, J.C. and Montoliu, L. (2005) Molecular basis of the extreme dilution mottled mouse mutation: a combination of coding and noncoding genomic alterations. *J. Biol. Chem.*, **280**, 4817–4824.
  29. Rincón-Arano, H., Furlan-Magaril, M. and Recillas-Targa, F. (2007) Protection against telomeric position effects by the chicken  $\chi$ HS4  $\beta$ -globin insulator. *Proc. Natl Acad. Sci. USA*, **104**, 14044–14049.
  30. Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N. and Felsenfeld, G. (2001) Transitions in histone acetylation reveal boundaries of three separately regulated neighbouring loci. *EMBO J.*, **20**, 2224–2235.
  31. Lefevre, P., Witham, J., Lacroix, C.E., Cockerill, P.N. and Bonifer, C. (2008) The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol. Cell*, **32**, 129–139.
  32. Ishihara, K., Oshimura, M. and Nakao, M. (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol. Cell*, **23**, 733–742.
  33. Potts, W., Tucker, D., Wood, H. and Martin, C. (2000) Chicken  $\beta$ -globin 5'HS4 insulators function to reduce variability in transgenic founder mice. *Biochem. Biophys. Res. Commun.*, **273**, 1015–1018.
  34. Torrano, V., Navascues, J., Docquier, F., Zhanf, R., Burke, L.J., Chernukhin, I., Farrar, D., León, J., Berciano, M.T., Renkawitz, R. et al. (2006) Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosyl)ation-dependent mechanism. *J. Cell. Sci.*, **119**, 1746–1759.
  35. Witcher, M. and Emerson, B. (2009) Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol. Cell*, **34**, 271–284.
  36. Farrar, D., Rain, S., Chernukhin, I., Jagodic, M., Ito, Y., Yammine, S., Ohlsson, R., Murrel, A. and Klenova, E. (2010) Mutational analysis of the poly(ADP-ribosyl)ation sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-ribosyl)ation. *Mol. Cell. Biol.*, **30**, 1199–1216.
  37. Soto-Reyes, E. and Recillas-Targa, F. (2010) Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene*, **15**, 2217–2227.
  38. Gavrilov, A.A. and Razin, S.V. (2008) Spatial configuration of the chicken  $\alpha$ -globin gene domain: immature and active chromatin hubs. *Nucleic Acids Res.*, **36**, 4629–4640.
  39. West, A.G., Huang, S., Gaszner, M., Litt, M.D. and Felsenfeld, G. (2004) Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol. Cell*, **16**, 453–463.
  40. Higgs, D.R., Vernimmen, D., Hughes, J. and Gibbons, R. (2007) Using genomics to study how chromatin influences gene expression. *Annu. Rev. Genomics Hum. Genet.*, **8**, 299–325.
  41. Zhao, H. and Dean, A. (2004) An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res.*, **32**, 4903–4914.
  42. Gomes, N.P. and Espinosa, J.M. (2010) Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev.*, **24**, 1022–1034.
  43. Borunova, V., Iarovaia, O.V., Vassetzky, Y.S. and Razin, S.V. (2005) The upstream area of the chicken  $\alpha$ -globin gene domain is transcribed in both directions in the same cells. *FEBS Lett.*, **579**, 4746–4750.
  44. Kleine, H. and Lüscher, B. (2009) Learning how to read ADP-ribosylation. *Cell*, **139**, 17–19.
  45. Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S. and de Laat, W. (2004) The active spatial organization of the  $\beta$ -globin locus requires the transcription factor EKLF. *Genes Dev.*, **18**, 2485–2490.
  46. Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J. and Blobel, G.A. (2005) Proximity among distant regulatory elements at the  $\beta$ -globin locus requires GATA-1 and FOG-1. *Mol. Cell*, **17**, 453–462.
  47. Klochkov, D., Rincón-Arano, H., Ioudinkova, E.S., Valadez-Graham, V., Gavrilov, A., Recillas-Targa, F. and Razin, S.V. (2006) A CTCF-dependent silencer located in the differentially methylated area may regulate expression of a housekeeping gene overlapping a tissue-specific gene domain. *Mol. Cell. Biol.*, **26**, 1589–1597.