



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

PARTICIPACIÓN DE LA ESTRUCTURA DE LA CROMATINA EN LA ACTIVIDAD
DEL *ENHANCER* α -GLOBINA DE POLLO

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

ESTELA GUADALUPE GARCÍA GONZÁLEZ

TUTOR PRINCIPAL

DR. FÉLIX RECILLAS TARGA

INSTITUTO DE FISIOLOGÍA CELULAR, UNAM

MIEMBROS DEL COMITÉ TUTOR

DRA. TZVETANKA DIMITROVA DINKOVA

FACULTAD DE QUÍMICA, UNAM

DRA. VIVIANA DEL CARMEN VALADEZ GRAHAM

INSTITUTO DE BIOTECNOLOGÍA, UNAM

Ciudad de México. Abril, 2016



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

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

Estela Guadalupe García González recibió una beca durante la realización de sus estudios de Doctorado por parte del Consejo Nacional de Ciencia y Tecnología (CONACyT, México) con número de becario 207428 y registro CVU 218481. Del 01/06/2011 al 30/11/2011 el CONACyT le otorgó la Beca Mixta para realizar una estancia doctoral en el extranjero en el laboratorio del Dr. Kami Ahmad en la Universidad de Harvard Medical School, Massachusetts, USA.

Este trabajo fue apoyado por la DGAPA, UNAM (IN209403, IN203811 y IN201114), el CONACyT (42653-Q, 128464 y 220503) y la Dirección General de Estudios de Posgrado-Universidad Nacional Autónoma de México (DGEP). El apoyo adicional fue proporcionado por el Programa de Posgrado de Doctorado, "Doctorado en Ciencias Biomédicas" para el Instituto de Fisiología Celular y la Universidad Nacional Autónoma de México.

Agradecimientos

Al Dr. Félix Recillas Targa por haberme permitido realizar esta tesis doctoral bajo su dirección. Su orientación, continua discusión y rigor científico durante la realización de este proyecto son base fundamental de mi formación académica.

A la Dra. Tzvetanka Dimitrova Dinkova (miembro del comité tutor) por haberme permitido discutir mis resultados y aportar ideas durante la realización de este proyecto.

A la Dra. Viviana del Carmen Valadez Graham (miembro del comité tutor) por sus valiosos comentarios y discusiones que permitieron un mejor abordaje de mi trabajo de investigación.

A la Biol. Georgina Guerrero Avendaño por su importante participación técnica en el desarrollo de esta tesis, en particular, su disponibilidad, generosidad y paciencia.

Al Dr. Fernando Suaste Olmos por su asistencia técnica.

A la Dra. Catherine Farrell, al Dr. Paul Delgado Olgúin, al Dr. Martín Escamilla-Del-Arenal, al Dr. José Manuel Hernández Hernández y al LCG Rodrigo Arzate Mejía por sus comentarios y sugerencias sobre los manuscritos que culminaron en su publicación.

Al Programa de Apoyo a los Estudios de Posgrado (PAEP) por las facilidades otorgadas para realizar la estancia doctoral en el extranjero en el laboratorio del Dr. Kami Ahmad en la Universidad de Harvard Medical School, Massachusetts, USA.

A la M. en C. Norma Trejo Medina, Secretaria Técnica del programa de Maestría y Doctorado en Ciencias Bioquímicas de la UNAM.

A Leticia García Gutiérrez, Asistente de Procesos del programa de Maestría y Doctorado en Ciencias Bioquímicas de la UNAM.

A la Dra. Laura Ongay Larios y la M.C. Minerva Mora Cabrera de la Unidad de Biología Molecular (UBM) del IFC.

A Juan Manuel Barbosa Castillo de la unidad de cómputo del IFC.

A todos los miembros del 122 N.

Papá José:

Tú eres mi mejor fuente de inspiración

Conchita, Adrián, Omar y Adris:

Gracias por su infinito amor

***“El mundo que hemos creado
es producto de nuestros pensamientos.***

***No se puede cambiar sin cambiar
nuestra forma de pensar”***

Albert Einstein

ÍNDICE

Índice de abreviaturas.....	8
Resumen.....	11
Summary.....	12
Introducción.....	13
I. Regulación transcripcional.....	13
<u>I.1 Elemento promotor.....</u>	<u>14</u>
<i>I.1.1 Iniciador o INR.....</i>	<i>16</i>
<i>I.1.2 Caja TATA.....</i>	<i>16</i>
<i>I.1.3 Elemento de reconocimiento del factor TFIIB o BRE.....</i>	<i>16</i>
<i>I.1.4 Elemento río abajo del core o DPE.....</i>	<i>16</i>
<i>I.1.5 Elemento motivo 10 o MTE.....</i>	<i>17</i>
<u>I.2 Elementos de regulación a distancia.....</u>	<u>17</u>
<i>I.2.1 Potenciador o enhancer.....</i>	<i>17</i>
<i>I.2.2 Región de control del locus o LCR.....</i>	<i>20</i>
<i>I.2.3 Silenciadores.....</i>	<i>21</i>
<i>I.2.4 Delimitadores o insulators.....</i>	<i>22</i>
<hr/>	
II. Regulación epigenética.....	23
<u>II.1 Estructura de la cromatina.....</u>	<u>24</u>
<u>II.2 Posicionamiento nucleosomal asociado a promotores y enhancers.....</u>	<u>25</u>
<i>II.2.1 Características de las secuencias de DNA asociadas a promotores y enhancers.....</i>	<i>28</i>
<u>II.3 Variantes de histonas.....</u>	<u>29</u>
<u>II.4 Complejos remodeladores de la cromatina dependientes de ATP.....</u>	<u>31</u>

II.4.1 Remodeladores de la familia SWI/SNF.....	32
II.4.2 Remodeladores de la familia ISWI.....	33
II.4.3 Remodeladores de la familia CHD.....	34
II.4.4 Remodeladores de la familia INO80.....	34
<u>II.5 Modificaciones post-traduccionales en las histonas.....</u>	35
II.5.1 La marca en la histona H3K4me1 se asocia con elementos tipo enhancers en todo el genoma.....	36
II.5.2 Las acetil-transferasas de histonas p300 y CBP.....	37
II.5.3 La modificación en la histona H3K27ac y el estado de actividad de los enhancers.....	37
<u>II.6 Complejos Polycomb y Trithorax.....</u>	39
<u>II.7 Metilación del DNA asociado a enhancers.....</u>	40
III. Organización espacial del genoma al interior del núcleo.....	41
IV. Modelos de acción de los enhancers.....	45
<u>IV.1 Modelos de comunicación entre promotores y enhancers.....</u>	45
IV.1.1 Modelos de asas o looping.....	46
IV.1.2 Modelo de vinculación o linking.....	47
IV.1.3 Combinación entre los modelos de asas y encarrilamiento.....	49
<u>IV.2 Modelos de procesamiento de la información de los enhancers.....</u>	49
IV.2.1 Modelo de enhanceosoma.....	50
IV.2.1 Modelo billboard.....	51
<u>IV.3 Modelos de acción de los enhancers.....</u>	52
IV.3.1 Modelo progresivo/gradual de acción de los enhancers.....	53
IV.3.2 Modelo binario/probabilidad de acción de los enhancers.....	54
IV.3.3 Implicaciones de los modelos progresivo y binario.....	55

<u>IV.4. Súper-enhancers</u>	56
<u>IV.5 RNA pol II y transcripción de enhancers</u>	57
<u>IV.6 RNAs que actúan como enhancers</u>	58
V. Modelo de estudio y antecedentes particulares	60
<u>V.1 Dominio multigénico α-globina de pollo</u>	60
<u>V.2 Factores transcripcionales relacionados con la expresión de los genes globina</u>	61
<u>V.3 Elementos reguladores del dominio α-globina de pollo</u>	63
<u>V.4 Organización de la cromatina de los elementos reguladores ERAE-enhancer del dominio α-globina de pollo</u>	64
VI. Planteamiento del problema	67
VII. Hipótesis	67
VIII. Objetivo general	68
IX. Objetivos particulares	68
X. Resultados	68
<u>X.1 Generación de líneas estables con los plásmidos que contienen al enhancer, ERAE-enhancer y ERAE</u>	68
<u>X.2 El ERAE disminuye la expresión del gen reportero GFP en el contexto α-globina de pollo</u>	70
<u>X.3 La actividad del ERAE es contexto y tejido dependiente</u>	73
<u>X.4 El factor de transcripción RUNX-1 interacciona <i>in vitro</i> con la secuencia ERAE</u>	77
<u>X.5 La integridad del ERAE es necesaria para la actividad del enhancer 3'α-globina de pollo</u>	82
<u>X.6 La perturbación de la marca epigenética que se caracteriza por dos sitios de corte a la enzima MNasa y el posicionamiento nucleosomal entre el ERAE y enhancer genera la pérdida de la actividad transcripcional del gen reportero</u>	84

<u>X.7 El incremento en la distancia entre las dos hipersensibilidades al corte por la enzima MNasa entre el ERAE-<i>enhancer</i> 3 'α-globina de pollo afecta el posicionamiento de los nucleosomas</u>	85
<u>X.8 El incremento en la distancia entre las dos hipersensibilidades al corte por la enzima MNasa entre el ERAE-<i>enhancer</i> 3 'α-globina de pollo no afecta la unión del factor transcripcional CTCF</u>	87
<u>X.9 Modelo de la regulación de la actividad potenciadora mediante el posicionamiento de nucleosomas</u>	88
XI. Discusión	90
XII. Conclusiones	94
XIII Perspectivas	95
XIV. Bibliografía	96
XV. Anexos	117
<u>XV.1 Materiales y métodos</u>	117
<i>XV.1.1 Plásmidos y construcciones</i>	117
<i>XV.1.2 Cultivo celular</i>	117
<i>XV.1.3 Transfecciones transitorias y estables</i>	118
<i>XV.1.4 Aislamiento de clonas</i>	118
<i>XV.1.5 Citometría de flujo</i>	118
<i>XV.1.6 Southern blot</i>	119
<i>XV.1.7 Microscopía de Fluorescencia</i>	119
<i>XV.1.8 Ensayo de movilidad electroforética (EMSA) y súper-retardo</i>	120
<i>XV.1.9 Digestión con la endonucleasa MNasa</i>	120
<i>XV.1.10 Inmunoprecipitación de cromatina</i>	121
<u>XV.2 Tabla 1. Factor Fano</u>	123

Índice de abreviaturas

H3K4me1 = Mono-metilación de la lisina 4 en la histona 3

H3K27ac = Acetilación de la lisina 27 en la histona 3

CBP/p300 = Acetiltransferasas

ERAE = Elemento río arriba del enhancer

EMSA = Ensayo de retardo de movilidad electroforética

eRNA = enhancer RNA

mRNA = RNA mensajero

RNA pol II = Polimerasa II dependiente de RNA

FT = Factor de transcripción

TSS = Sitio de inicio de la transcripción

LCR = Región controladora de un dominio

DNA = Ácido desoxirribonucleico

RNA = Ácido ribonucleico

PIC = Complejo de pre-inicio de la transcripción

INF- β = Interferón β

DNasa I = Endonucleasa

DHS = Sitio hipersensible al corte por la DNasa I

3C = Captura conformacional de cromosomas

PAH = *Fenilalanina hidroxilasa*

NRE = Elementos reguladores negativos

CAT = Acetil-transferasa de cloranfenicol

PAI-2 = Inhibidor del activador del plasminógeno tipo 2

PcG = Proteínas del grupo de Polycomb

Trx = Proteína del grupo Trithorax

HAT = Acetil-transferasa de histonas

RNA-seq = Secuenciación masiva de RNA

mESC = Células troncales embrionarias murinas

NPC = Células progenitoras neurales

H3K27me3 = Tri-metilación de la lisina 27 en la histona 3

NDR/NFR = Región libre de nucleosomas

H3K4me2 = Di-metilación de la lisina 4 de la histona 3

DHT = 5- α -dihidrotestosterona
ATP = Adenosina trifosfato
A = Adenina
T = Timina
G = Guanina
C = Citosina
CNS = Sistema Nervioso Central
PRC2 = Complejo 2 de proteínas del grupo de Polycomb
PRC1 = Complejo 1 de proteínas del grupo de Polycomb
FMR = Regiones de DNA completamente metiladas
UMR = Regiones de DNA no metiladas
LMR = Regiones de DNA de baja metilación
NP = Progenitor neural
5mC = 5-metilcitosina
5hmC = 5-hidroximetilcitosina
5fC = 5-formilcitosina
5caC = 5-carboxilcitosina
BER = Mecanismo de escisión de bases
HiC = Método para detectar interacciones de DNA de largo alcance en todo el genoma
TAD = Dominio topológico
Mb = Megabase
Kb = Kilobase
pb = Pares de bases
CT = Territorio cromosómico
RNAi = RNA de interferencia
ChIP = Inmunoprecipitación de cromatina
qPCR = Amplificación de DNA en tiempo real
PSA = Antígeno prostático específico
AR = Receptor de andrógenos
MTV = Tumor mamario de ratón
cREDS = Elementos reguladores en *cis* con rasgos dinámicos
lncRNA = RNAs intergénicos largos no codificantes

ORF = Marco abierto de lectura

H3K36me3 = Tri-metilación de la lisina 36 en la histona 3

10dRBC = Eritrocitos circulantes de 10 días

GFP = Proteína verde fluorescente

FACS = Citometría de flujo

FF = Factor fano

MNasa = Nucleasa micrococcal

Resumen

La presencia de secuencias regulatorias en el genoma como promotores y potenciadores o *enhancers* facilitan la expresión diferencial de genes. El panorama cromatínico asociado con los *enhancers* se conoce en detalle, por ejemplo, la presencia de marcas en las histonas, como la mono-metilación de la lisina 4 en la histona 3, su asociación con acetil-transferasas de histonas o proteínas estructurales como las cohesinas. Sin embargo, la contribución de la estructura de la cromatina local asociada con la actividad de los *enhancers* ha sido menos explorada. El análisis funcional de los *enhancers* sigue siendo validado mediante el estudio de la secuencia mínima de DNA que genera la mayor actividad transcripcional de los genes reporteros mientras que las regiones adyacentes son frecuentemente omitidas. En este trabajo se describe la función del elemento río arriba del *enhancer* (ERAE) del dominio 3' α -globina de pollo. Este elemento de regulación, se localiza a un costado del *enhancer* en la región 3' de los genes α -globina de pollo y modula su actividad en el contexto de la cromatina. La presencia del ERAE reduce principalmente las fluctuaciones en la expresión del gen reportero de acuerdo al modelo progresivo de acción de los *enhancers* e interacciona *in vitro* con el factor transcripcional RUNX-1. Usando construcciones mutantes que disminuyen drásticamente la actividad transcripcional del gen reportero, se observó una pérdida de la estructura de la cromatina que correlaciona con cambios en el posicionamiento de un nucleosoma localizado en la secuencia del ERAE. Basado en nuestros resultados, proponemos que el ERAE es un componente esencial en la actividad del *enhancer* 3' α -globina de pollo y genera una configuración local en la estructura de la cromatina que le permite modular la actividad transcripcional de su gen blanco.

Summary

The presence of regulatory sequences in the genome like promoters and enhancers facilitate differential patterns of gene expression. The chromatin landscape associated with enhancers is widely known as the presence of histone marks like the mono-methylation of lysine 4 on histone 3, its association with histone acetyl-transferases or structural proteins as cohesins. However, the contribution of local chromatin structure associated with enhancer activity remains less explored. The functional analysis of enhancers remains validated by studying the minimum DNA sequence that generates the highest transcriptional activity of the reporter gene whereas adjacent regions are often omitted. Here we describe the role of the upstream enhancer element (UEE) from the 3' α -globin chicken gene. This regulatory element is located next to the enhancer in the downstream region of the α -globin chicken genes and modulates the enhancer activity. Within chromatin, the presence of UEE mainly reduces fluctuations in the reporter gene expression according to the progressive model for enhancer activity and physically interacts with the RUNX-1 transcription factor. Using mutant constructs that dramatically decrease the transcriptional activity of the reporter gene, we observed a correlation with associated changes in chromatin structure affecting the nucleosome positioning associated with the UEE sequence. Based on our results, the UEE presence is essential for the enhancer activity of the 3' α -globin genes and creates a local chromatin configuration necessary to stimulate target gene transcription.

Introducción

La expresión de genes es un proceso central para la sobrevivencia de las células, debido a que sus propiedades biológicas dependen del tipo y cantidad de proteínas expresadas en un momento dado del desarrollo y en un tejido en particular (Recillas-Targa y Escamilla-Del-Arenal, 2004). A nivel molecular, la transcripción se refiere a la síntesis de moléculas de RNA mensajero (mRNA) a partir de segmentos del genoma correspondientes a genes donde las enzimas RNA polimerasas (RNA pol II) avanzan a lo largo de la hebra de DNA y catalizan la formación de enlaces fosfodiéster entre nucleótidos. Diversas señales externas e internas controlan a los factores de transcripción (FT) que operan sobre distintas secuencias reguladoras y por lo tanto determinan cuáles deben ser activadas y transcritas en el tiempo y espacio correctos (Greive y von Hippel, 2005).

I. Regulación transcripcional en células eucariotas

La expresión de genes es un proceso complejo que ocurre en varios pasos, sin embargo, se ha propuesto que la activación de la transcripción es uno de los pasos limitantes en la regulación transcripcional en células eucariotas (Barberis y Petrascheck, 2003; Chakalova *et al*, 2005, Greive y von Hippel, 2005). En la secuencia del DNA, además de la información genética necesaria para la síntesis de proteínas es posible localizar secuencias o elementos reguladores que actúan en *cis* y que son reconocidos por factores de transcripción que actúan en *trans*. Las regiones reguladoras de un gen pueden hallarse cercanas al sitio de inicio de la transcripción (*TSS*) como el *core* del promotor y sus elementos proximales; o bien, alejadas del *TSS* conocidos como elementos de regulación a distancia. Dentro de estos últimos, se encuentran los potenciadores o *enhancers*, silenciadores, delimitadores o *insulators* y regiones controladoras de un dominio o *LCRs*; estas secuencias suelen hallarse en número variable y tienen como finalidad aumentar o disminuir los niveles de expresión de su(s) gen(es) blanco (Fig. 1).

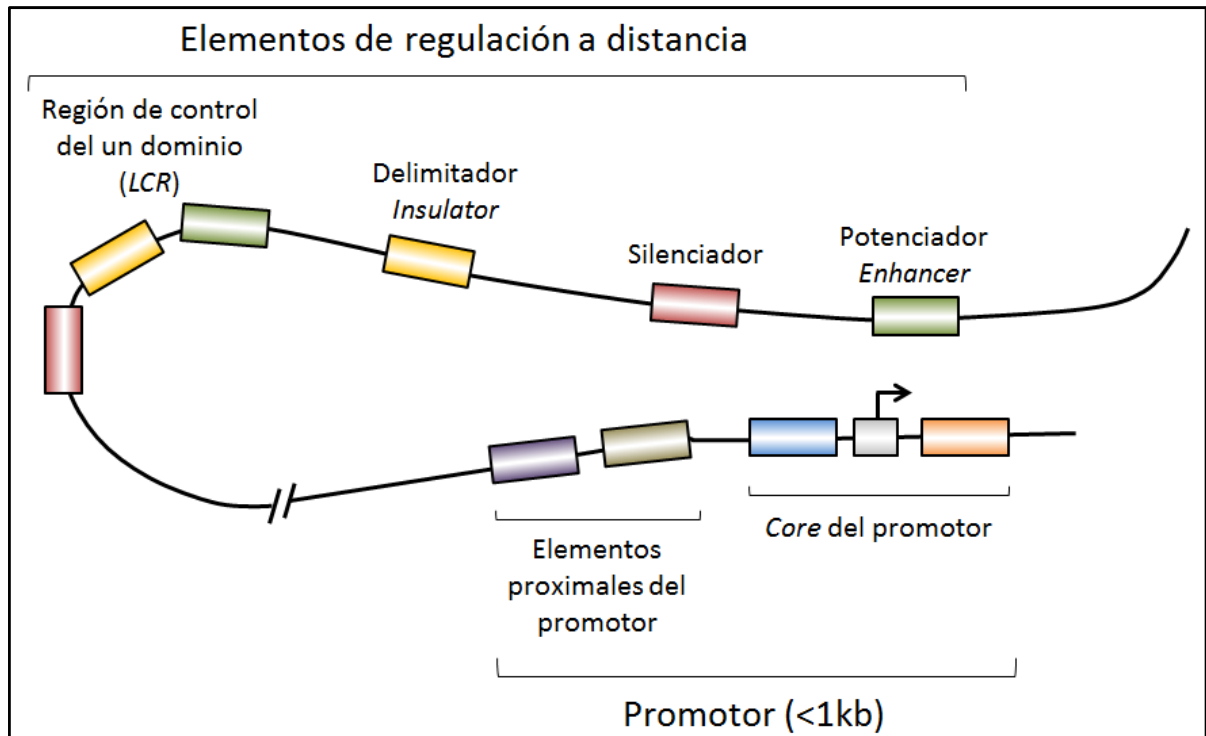


Figura 1. Regiones reguladoras de un gen. El promotor tiene sitios de unión a elementos proximales tejido específicos y factores generales de la transcripción que le permiten a la RNA pol II posicionarse correctamente en el TSS y comenzar la transcripción de un gen. Muchos genes necesitan además elementos de regulación a distancia que incluyen *enhancers*, silenciadores, delimitadores o *insulators* y regiones controladoras de un dominio o *LCRs*. Los elementos distales pueden contactar al promotor mediante la formación de asas de cromatina (Tomado y modificado de Maston *et al*, 2006).

A continuación, se describen las principales características asociadas con las secuencias reguladoras presentes en el DNA.

I.1 Elemento promotor

Un elemento regulador tipo promotor, se define como una secuencia de DNA continua y mínima necesaria para iniciar de manera regulada la transcripción de un gen con la intervención de factores generales y tejido específicos, así como de la enzima RNA pol II. Este elemento de regulación se encuentra cercano al TSS, está constituido por el *core* del promotor donde se unen los factores generales de la transcripción y por elementos proximales sobre los cuales se asocian factores específicos y activadores de la transcripción. Todos estos factores

favorecen el reclutamiento del complejo de pre-inicio de la transcripción y facilitan el posicionamiento de la RNA pol II sobre el promotor (Greive y von Hippel, 2005; Green, 2005; Maston *et al*, 2006). Mediante proteínas de fusión ha sido posible determinar experimentalmente que los activadores transcripcionales tienen al menos dos dominios, uno de unión al DNA y otro de activación (Green, 2005). El dominio de activación es necesario para la transcripción *in vivo* porque el DNA de las células eucariontes está almacenado en forma de cromatina al interior del núcleo y es menos accesible para los factores de transcripción (Green, 2005). Por lo tanto, el dominio de activación se encarga del reclutamiento de diversos co-activadores de la transcripción como acetil-transferasas de histonas (CBP/p300) y otras proteínas que forman el complejo de pre-inicio de la transcripción o PIC (*Pre-Initiation Complex*). Todos estos elementos en conjunto determinan el sitio de inicio de la transcripción o *TSS* (*Transcription start site*) (Green, 2005). Dentro del *core* de los promotores, se han descrito varias secuencias motivo que son reconocidas por diversos factores de transcripción (Juven-Gershon *et al*, 2008). Sus principales características se describen a continuación (Fig. 2).

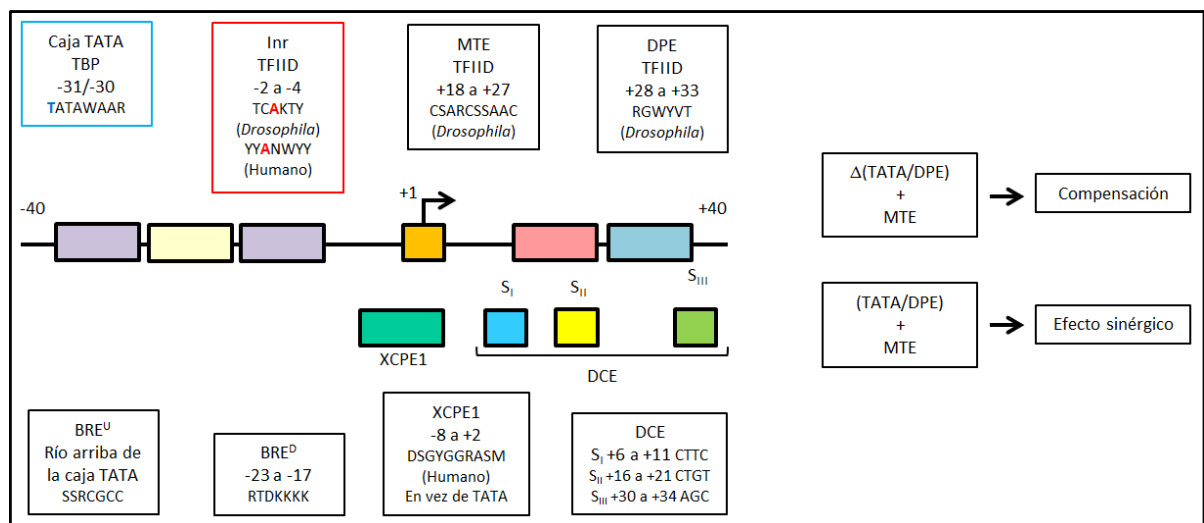


Figura 2. Secuencias motivo presentes en el *core* del promotor. Los factores de transcripción y los motivos de unión que reconocen en el DNA se muestran dentro de cajas. Se ha determinado experimentalmente que el motivo MTE puede compensar la ausencia de caja TATA y DPE, pero si los tres motivos están presentes tienen un efecto sinérgico sobre la transcripción (Tomado y modificado de Juven-Gershon *et al*, 2008).

1.1.1 Iniciador o Inr

Esta secuencia motivo coincide con el sitio de inicio de la transcripción. Mediante ensayos funcionales se determinó que la secuencia consenso es YYANWYY en humanos y TCAKTY en *Drosophila* (donde Y = Citosina o Timina, N = cualquier base, A = Adenina, W = Adenina o Timina, de acuerdo con la nomenclatura de la IUPAC para ácidos nucleicos). El nucleótido adenina dentro del consenso usualmente corresponde al TSS (+1) (letra A en color rojo en el esquema de la Fig. 2). Esta secuencia es reconocida por el complejo TFIID (Juven-Gershon *et al*, 2008).

1.1.2 Caja TATA

Su secuencia consenso es TATAWAAR (donde W = Adenina o Timina, R = Adenina o Guanina). La timina del extremo 5' comúnmente se localiza a -31 o -30 respecto del TSS (letra T de color azul en la Fig. 2). TFIID reconoce la caja TATA mediante la proteína TBP (TATA *binding protein* o proteína de unión a la caja TATA). En vertebrados no es común encontrar motivos TATA, sino islas CpG (Juven-Gershon *et al*, 2008).

1.1.3 Elemento de reconocimiento del factor TFIIB o BRE

Esta secuencia se localiza río arriba o río abajo de la caja TATA y los consensos son SSRSGCC Y RTDKKKK, respectivamente (donde S = Guanina o Citosina, R = Adenina o Guanina, D=Adenina, Guanina o Timina y K = Guanina o Timina). Estos motivos unen al factor general de la transcripción TFIIB (Juven-Gershon *et al*, 2008).

1.1.4 Elemento río abajo del core o DPE

A este motivo se une el factor general de la transcripción TFIID. Es un motivo conservado desde *Drosophila* a humanos, se ubica entre +28 y +33 relativo al TSS. El elemento DPE funciona de forma sinérgica con el Inr y la distancia entre ambos elementos es crítica para que se lleve a cabo la transcripción (Juven-Gershon *et al*, 2008).

1.1.5 Elemento motivo 10 o MTE

Su consenso corresponde a CSARCSSAAC (donde S = Guanina o Citosina y R = Adenina o Guanina) y abarca entre +18 a +27 respecto del Inr y coopera con este motivo. También se ha visto que la distancia entre ellos es crucial para su funcionamiento. Este elemento es capaz de compensar la falta de DPE o caja TATA (Juven-Gershon *et al*, 2008).

Es importante mencionar que los promotores no funcionan mediante un mecanismo universal para todos los genes, por el contrario, existe una gran diversidad relacionada con la estructura y función de cada uno de ellos. Esta combinatoria genera la especificidad necesaria para activar de forma regulada a cada promotor.

1.2 Elementos de regulación a distancia

Estos elementos reguladores en *cis* se localizan linealmente alejados del TSS y ayudan a regular en tiempo y espacio la actividad transcripcional de un gen (Thurman *et al*, 2012). Dentro de estos elementos se han descrito *enhancers*, regiones controladoras de un dominio (LCRs), silenciadores y delimitadores (*insulators*).

1.2.1 Potenciador o enhancer

Un *enhancer* representa un elemento de regulación a distancia cuya secuencia de DNA interacciona con combinatorias de activadores transcripcionales e incrementa los niveles de expresión basal de su gen blanco de manera específica, sin importar su distancia y orientación respecto al promotor (Bulger y Groudine, 2011). Estas secuencias de DNA fueron descritas por primera vez en tumores de mono causados por el virus del simio SV40 (Serfling *et al*, 1985). Este *enhancer* viral consiste en una secuencia repetida de 72 pares de bases (pb) de longitud que cuando se suprimió, abatió la expresión de proteínas expresadas en las primeras etapas de la infección, disminuyendo así la viabilidad del virus (Serfling *et al*, 1985). Pronto se demostró que el *enhancer* viral SV40 tenía además la capacidad de incrementar la actividad transcripcional del gen eucarionte que codifica para la cadena β -globina de conejo, incluso cuando se insertaba miles de pares de bases río arriba o río abajo del promotor de su gen

blanco (Banerji *et al*, 1981 y 1983; Khoury y Gruss, 1983; Patrusky, 1986). El descubrimiento de los *enhancers* celulares en genes humanos como la insulina y el interferón- β (IFN- β) condujeron al razonamiento de que los *enhancers* podrían definir la especificidad de tejido necesaria para generar patrones complejos de expresión de genes durante la diferenciación celular y el desarrollo de un organismo (Walker *et al*, 1983; Shöler *et al*, 1985; Patrusky, 1986). La capacidad de los *enhancers* de actuar independientemente de su posición u orientación con respecto al sitio de inicio de la transcripción representa una diferencia fundamental con los promotores (Banerji *et al*, 1981). Los promotores permiten a las RNA polimerasas posicionarse sobre el DNA y transcribir la información codificada en los genes generando mRNAs que posteriormente serán traducidos en proteínas. Sin embargo, recientemente se ha reportado que las secuencias *enhancer* también se transcriben y generan transcritos bidireccionales no codificantes conocidos como eRNAs cuya expresión coincide con la transcripción de su gen blanco y tiene implicaciones en su mecanismo de acción (Kim *et al*, 2010; Mousavi *et al* 2013; Lam *et al*, 2013; Li *et al*, 2013). En consecuencia, los *enhancers* ayudan a determinar cuándo, dónde y cómo será la magnitud de la respuesta para la correcta expresión de uno o más genes (Khoury y Gruss, 1983; Patrusky, 1986; Blackwood y Kadonaga 1998; Maston *et al*, 2006; Plank y Dean, 2014).

En *Drosophila* se ha reportado que cientos de *enhancers* muestran mayor especificidad por el *core* de promotores asociados con genes constitutivos (como genes ribosomales) que con genes que se expresan durante el desarrollo (como factores de transcripción) y viceversa (Zabidi *et al*, 2015). Los *enhancers* asociados con genes del desarrollo muestran mayor especificidad de tejido y los *TSS* muestran enriquecimiento de secuencias motivo como caja TATA, iniciador, DPE y MTE. Los *enhancers* asociados con genes constitutivos contienen el motivo de unión DRE, mientras que aquellos asociados con genes de desarrollo presentan motivos de unión al factor GAGA (Zabidi *et al*, 2015).

Los *enhancers* son sitios de mayor accesibilidad a la cromatina como consecuencia de su interacción con proteínas activadoras que se unen de forma específica, generando una configuración de cromatina abierta que permite su identificación experimental mediante una

alta frecuencia al corte con endonucleasas (Guarente, 1988; Felsenfeld *et al*, 1996). Estas observaciones han sido recientemente confirmadas por estudios a gran escala que muestran los patrones de corte por la enzima DNasa I en 125 tipos celulares y tejidos (Thurman *et al*, 2012). Todos estos sitios marcan elementos regulatorios presentes en el genoma como promotores, *enhancers*, *insulators*, *LCRs* y silenciadores (Sabo *et al*, 2006; Boyle *et al*, 2009; Thurman *et al*, 2012). Se identificaron aproximadamente 2,9 millones de sitios hipersensibles al corte por la DNasa I o DHSs. Alrededor de 970,100 sitios DHSs son específicos de un tipo celular, 1,920,642 de DHSs están presentes en dos o más tipos de células y la minoría (3,692) está presentes en todas la células analizadas (Thurman *et al*, 2012). El 95% de los DHSs específicos de un tejido se localizan alejados del *TSS*. Al comparar los patrones de accesibilidad de la cromatina con las señales de unión *in vivo* de 42 factores transcripcionales se observa que al menos el 94% los factores transcripcionales se localizan en regiones de cromatina accesible. Del examen detallado de los perfiles de DNasa I a través de muchos tipos celulares se observó que muchos *enhancers* se vuelven accesibles al mismo tiempo que aparece la hipersensibilidad en los promotores de los genes que se expresan en cada linaje celular (Thurman *et al*, 2012). Una forma de validar las interacciones físicas entre regiones genómicas es a través de la técnica de captura conformacional de cromosomas o 3C (*Chromosome Conformation Capture*). El 3C convierte las interacciones físicas de la cromatina en productos de ligación específicos, que se cuantifican individualmente por PCR. La frecuencia con la que dos fragmentos de restricción son ligados reflejan la frecuencia con la que interactúan en el núcleo (Dekker *et al*. 2002). El 3C resulta muy adecuado para estudios a pequeña escala centrados en el análisis de las interacciones entre un conjunto de elementos candidatos. Sin embargo, la detección de las interacciones de cromatina a gran escala requiere de la secuenciación masiva del DNA. Debido a la extrema complejidad de la biblioteca de DNA y la baja abundancia relativa de cada producto de ligación específica es necesario amplificarlo para facilitar su secuenciación. Este método es conocido como 3C-*Carbon Copy* o "5C" (Dostie *et al*, 2006). Usando esta metodología, se definió que la aparición de DHSs dentro del cuerpo del gen de la *fenilalanina hidroxilasa (PAH)* correlaciona con una mayor frecuencia de interacción entre el promotor y un *enhancer* río arriba previamente

definido como un elemento de regulación de este gen. Estas interacciones son además tejido específicas, pues sólo se observan en células hepáticas. Los DHSs también muestran conservación evolutiva funcional y predicen que la presencia de estos rasgos cromatínicos se asocian con regiones reguladoras en el genoma (Thurman *et al*, 2012). Aunque la asignación de DHSs es una forma precisa de validar la presencia de elementos de regulación en el genoma, no es posible distinguir promotores de *enhancers* o *insulators* usando solamente esta característica epigenética. Por esta razón, otros rasgos asociados con la cromatina como las modificaciones post-traduccionales en las histonas ayudan a identificar y definir secuencias tipo *enhancer*.

1.2.2 Región de control de un dominio o LCR

Son secuencias de DNA que contribuyen a la apertura de la cromatina e incrementan la expresión de su gen blanco a niveles fisiológicos. Son dependientes del número de copias pero independientes de su sitio de integración. Se han encontrado *LCRs* río arriba, río abajo o dentro de los genes que regulan (Qiliang *et al*, 2002). Los *LCRs* suelen ser accesibles al corte por la enzima DNasa I y sirven de plataforma para la unión de factores transcripcionales ubicuos y específicos de tejido (Thurman *et al*, 2012). Otra propiedad asociada con los *LCRs* es su habilidad de conferir expresión a un transgén independientemente del ambiente cromatínico donde se haya insertado. Esto supone que los *LCRs* mantienen una cromatina abierta que permite a los factores de transcripción reconocer su sitio en el genoma y se ha propuesto que la acetilación en las histonas asociadas a estas regiones favorece esta función (Ho *et al*, 2002). Un modelo propuesto para entender cómo los *LCRs* pueden activar genes que se encuentran a grandes distancias es el de formación de asas de cromatina entre el *LCR* y el promotor al que regulan, sin embargo, otros modelos como encarrilamiento y vinculación que se describen más adelante para los *enhancers*, también aplican para el caso de los *LCRs* (Bulger y Groudine, 1999 y 2011). Una de las principales características que distingue a los *enhancers* de los *LCRs*, es que estos últimos confieren expresión a un transgén dependiendo del número de copias (Grosveld *et al*, 1997; Qiliang *et al*, 2002; Recillas-Targa y Escamilla del Arenal, 2004; Gaszner y Felsenfeld, 2006).

1.2.3 Silenciadores

En un principio estos elementos reguladores del genoma se definieron como secuencias de DNA capaces de reprimir la actividad transcripcional de manera independiente de su orientación y posición; sin embargo la caracterización más detallada de estos elementos sugirió que su posición y/o orientación puede ser determinante en su actividad (Ogbourne y Antalis, 1998). Con base en lo anterior, se ha propuesto clasificar a los silenciadores en aquellos cuya actividad represiva se ejerce independientemente de su posición; silenciadores clásicos y aquellos que dependen de la posición; elementos reguladores negativos (NRE) (Fig. 3) (Ogbourne y Antalis, 1998).

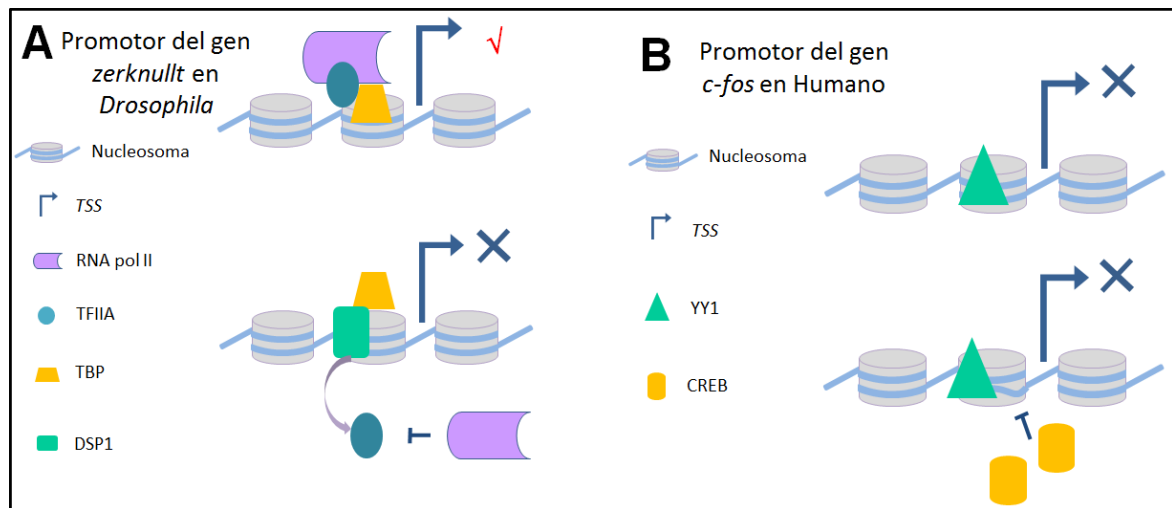


Figura 3. Clasificación de silenciadores con base en su mecanismo de acción. **A)** Silenciadores clásicos desplazan activamente factores generales de la transcripción. Como ejemplo se muestra el promotor del gen *zerknüllt* en *Drosophila* donde la unión directa de la proteína DSP1 (*Dorsal switch protein 1*) a la proteína TBP desplaza al factor general de la transcripción TFIIA y por lo tanto evita el reclutamiento de la RNA pol II al promotor. **B)** Elementos reguladores negativos (NRE) ejercen sus efectos de manera pasiva. El promotor del gen *c-fos* en humano tiene un NRE donde la unión del factor transcripcional YY1 induce una curvatura en el DNA que impide la unión de la proteína CREB a otras regiones del promotor que le permiten activarse (Antalis *et al*, 1996).

El mecanismo de acción de los silenciadores clásicos es activo ya que reclutan proteínas que interfieren directamente con los factores generales de la transcripción (GTF) que interactúan con el promotor. Un ejemplo de silenciador clásico se ha identificado en el promotor del gen *zerknüllt* en *Drosophila* donde la unión directa de la proteína DSP1 (*Dorsal switch protein 1*) a

la proteína TBP desplaza al factor general de la transcripción TFIIA y por lo tanto evita el reclutamiento de la RNA pol II al promotor evitando así su activación (Fig. 3) (Antalis *et al*, 1996). Los NRE ejercen sus efectos de manera pasiva ya que se unen a proteínas represoras que inhiben físicamente la interacción de los factores de transcripción (TF) con sus sitios de unión en el DNA localizados río arriba del promotor o bien, interfieren señales positivas de la transcripción como la poliadenilación (Ogbourne y Antalis, 1998). En el promotor del gen *c-fos* ha sido identificado un NRE que se une al factor transcripcional YY1 y reprime la actividad transcripcional del gen (Fig. 3) (Natesan y Gilman, 1993).

1.2.4 Delimitadores o insulators

La organización del genoma eucarionte origina que dominios cromatínicos con actividades transcripcionales opuestas se encuentren próximos (Gaszner y Felsenfeld, 2006). Los delimitadores o *insulators* son secuencias genómicas que previenen interacciones inadecuadas entre dominios cromatínicos adyacentes (Gaszner y Felsenfeld, 2006). La importancia de estos elementos reside en mantener aislados a los genes de señales externas de activación o represión para protegerlos de entornos adversos de la cromatina, en consecuencia ayudan a regular correctamente la expresión de genes. Dentro de la región circunscrita por *insulators* puede haber uno o más genes (West y Fraser, 2005).

Los *insulators* han sido caracterizados en función de dos propiedades funcionales separables. La primera es bloquear la acción de un *enhancer* siempre que el delimitador se encuentre entre el *enhancer* y su promotor, como en el caso del retrotransposón *gypsy* en *Drosophila*, que cuando se inserta dentro del locus *yellow* que contiene varios *enhancers*, bloquea su acción (Gaszner y Felsenfeld, 2006). La segunda propiedad es su capacidad de actuar como una barrera en contra de la expansión de la cromatina represiva o heterocromatina (Fig. 4) (Sun y Elgin, 1999). Esta propiedad se ha visto asociada con la unión específica de factores como USF1 y USF2 que a su vez atraen complejos que acetilan y metilan a las histonas, evitando la propagación de la heterocromatina (Fig. 4) (Gaszner y Felsenfeld, 2006).

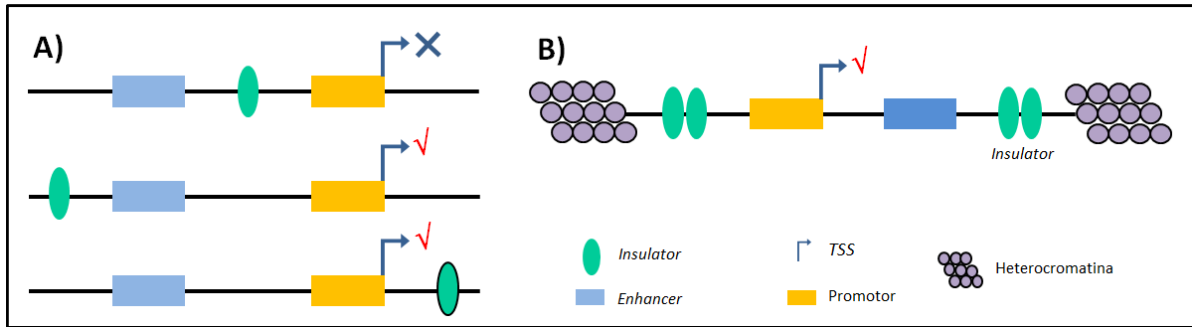


Figura 4. Propiedades funcionales de los delimitadores (*insulators*). **A)** Bloquear la comunicación entre un promotor y un *enhancer*, siempre que el delimitador se localice entre ambos elementos de regulación. **B)** Protección contra efectos de posición dictado por ambientes cromatínicos represivos (Tomado y modificado de Gaszner y Felsenfeld, 2006).

En resumen, la expresión regulada de los genes requiere de múltiples elementos de regulación presentes en el DNA. Entender los mecanismos y el comportamiento de estos elementos de regulación en el contexto de la cromatina resulta importante porque el DNA no se encuentra libre al interior del núcleo. El siguiente nivel de regulación de los genes se refiere a la regulación epigenética, representada por la estructura de la cromatina.

I.2. Regulación epigenética

La estructura de la cromatina está asociada íntimamente a la regulación epigenética que incluye los cambios en la regulación de la expresión génica sin que ocurran modificaciones en la secuencia del DNA (Bird, 2002; Jaenish y Bird, 2003; Vinci, 2011). Los principales procesos ligados a la regulación epigenética incluyen modificaciones en las histonas que afectan y regulan local y globalmente la estructura de la cromatina generando regiones de cromatina abierta y cerrada; la metilación del DNA por ejemplo tiene implicaciones importantes como la transmisión de patrones epigenéticos heredables en cada división celular y también contribuye a mantener la estabilidad del genoma (Robertson, 2005). Complejos como Polycomb (PcG) y Trithorax (Trx) que se relacionan con memoria epigenética y desarrollo también son parte de este tipo de regulación (Schuettengruber *et al*, 2007). Comprender cómo estos factores modulan la estructura de la cromatina para permitir o restringir el acceso a secuencias reguladoras en el genoma es clave para entender la regulación transcripcional.

II.1 Estructura de la cromatina

Los organismos eucariontes organizan su información genética dentro del núcleo en forma de cromatina. La cromatina es la estructura formada por DNA y proteínas básicas que ayudan a compactar la hebra de DNA (Fig. 5). La unidad fundamental de la cromatina es el nucleosoma, compuesto a su vez por el tetrámero (H3-H4)₂ y dos dímeros (H2A-H2B) alrededor del cual se asocian aproximadamente 147 pb de DNA. Cuando el DNA se asocia con el octámero de histonas se forma la fibra de 10 nanómetros (nm) o “collar de perlas” y con la incorporación de la histona de unión o *linker* también conocida como histona H1, la fibra de 10 nm se condensa hasta la fibra de 30 nm que a su vez se pliega en estructuras más complejas hasta formar cromosomas metafásicos necesarios para segregar el material genético durante la división celular (Felsenfeld y Groudine, 2003).

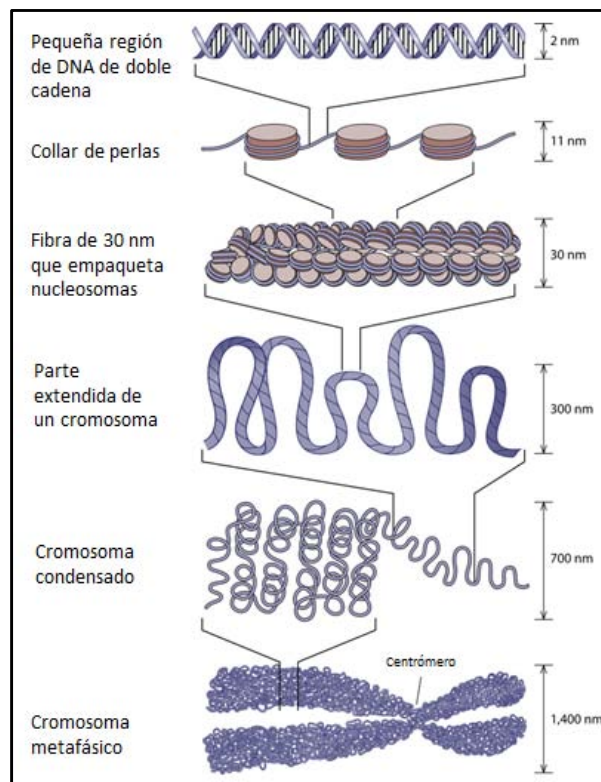


Figura 5. Empaquetamiento del DNA y su organización en cromatina. El primer nivel de organización es el nucleosoma, donde el DNA abarca dos vueltas alrededor del octámero de histonas (147 pb de DNA). Los nucleosomas se conectan uno al otro por medio de histonas de unión (H1 o *linker*). En el siguiente nivel de organización los nucleosomas se pliegan en una

fibra de 30 nm que a su vez se ordenan en estructuras más complejas hasta alcanzar la forma de un cromosoma metafásico (Tomado de Felsenfeld y Groudine, 2003).

II.2 Posicionamiento nucleosomal asociado a promotores y *enhancers*

La posición que ocupan los nucleosomas en el genoma afecta el reconocimiento de los sitios de unión a factores de transcripción (Beato y Eisefeld, 1997; Jiang y Pugh, 2009). La posición de los nucleosomas cambia según la actividad de los genes, y es un elemento importante de la regulación transcripcional (Shones *et al*, 2008). Aunque la accesibilidad de la cromatina depende del linaje celular es interesante encontrar patrones similares de posicionamiento nucleosomal *in vivo* asociado a elementos de regulación como promotores y *enhancers* en organismos tan diversos como la levadura y humano (Yuan *et al*, 2005; Mavrigh *et al*, 2008; Shones *et al*, 2008; He *et al*, 2010; Core *et al*, 2014).

Recientemente se han obtenido mapas a gran escala del posicionamiento nucleosomal *in vivo* de distintos organismos como humano, *Drosophila*, ratón, levadura, entre otros (Yuan *et al*, 2005; Mavrigh *et al*, 2008; Shones *et al*, 2008; Teif *et al*, 2012). La mayoría de estos análisis proviene de la secuenciación masiva de DNA aislado a partir de la digestión genómica de DNA con endonucleasas (MNasa o DNasa) que suelen dejar casi intacto al DNA unido al nucleosoma pero degrada los fragmentos entre nucleosomas (o DNA *linker*) (Jansen y Verstrepen, 2011; Hughes y Rando, 2014). Una conclusión general que ha derivado de estos estudios es que los nucleosomas no se distribuyen uniformemente en el genoma (Yuan *et al*, 2005; Mavrigh *et al*, 2008; Shones *et al*, 2008; Teif *et al*, 2012; Ozonov y Nimwegen, 2013). En diferentes organismos se favorece un mecanismo común donde el arreglo nucleosomal se comparte, es decir, se observa la presencia de una región carente o libre de nucleosomas (*NDR*, *nucleosome depleted region* o *NFR*, *nucleosome free region*) en promotores y *enhancers* (Fig. 6) (Yuan *et al*, 2005; Mavrigh *et al*, 2008; Shones *et al*, 2008; He *et al*, 2010; Teif *et al*, 2012; Core *et al*, 2014).

En el análisis de las secuencias promotoras de genes, se observa la formación de un patrón de nucleosomas característico en ambos extremos del *NDR*; siendo el nucleosoma +1 (o sea, el primer nucleosoma después del *TSS*) el que se encuentra mejor posicionado (Fig. 6). Lo

anterior significa que a nivel poblacional, la mayoría de los promotores tienen en esa posición genómica un nucleosoma asociado (Fig. 6) (Jansen y Verstrepén, 2011; Hughes y Rando, 2014). La región correspondiente al *NDR* de un promotor abarca aproximadamente 150 pb, se asocia con factores generales y específicos de la transcripción, además de la RNA pol II (Yuan *et al*, 2005; Shones *et al*, 2008).

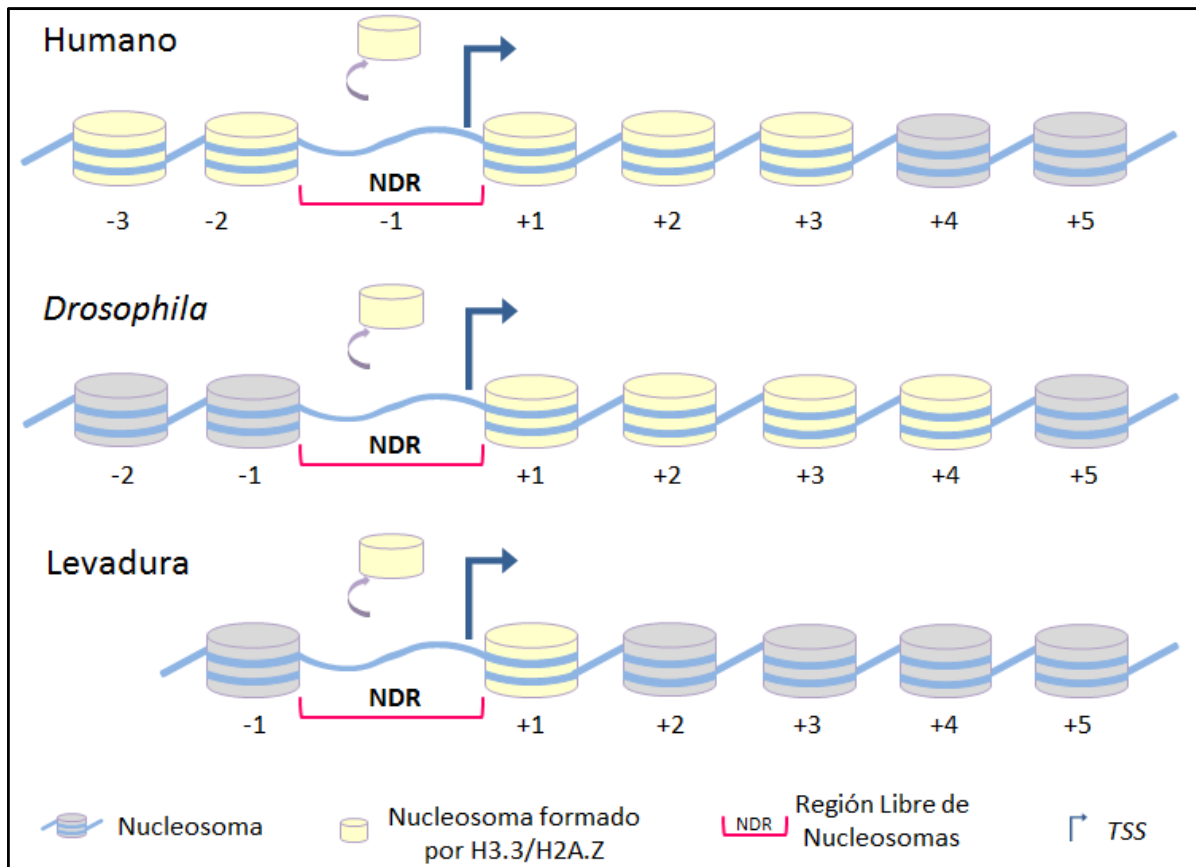


Figura 6. Posicionamiento nucleosomal sobre las regiones correspondientes a promotores en distintos organismos. **A)** Se muestra la posición que ocupan los nucleosomas en la región del promotor en humanos, *Drosophila* y levaduras. **B)** Se muestra la simbología usada para representar el posicionamiento de los nucleosomas en la región del promotor. La numeración que se ha dado a los nucleosomas en diferentes organismos es distinta, en humano la región correspondiente al *NDR* coincide con el nucleosoma -1 mientras que en *Drosophila* y levadura el nucleosoma -1 corresponde al primer nucleosoma río arriba del *NDR* (Tomado y modificado de Jiang y Pugh, 2009).

El posicionamiento nucleosomal asociado con los promotores muestra sutiles diferencias entre organismos, por ejemplo la posición del TSS en *S. cerevisiae* se localiza ~10 pb dentro del nucleosoma +1; en *Drosophila* el borde 5' del nucleosoma +1 se encuentra ~60 pb después del

TSS y en humanos el borde 5' del nucleosoma se localiza a ~40 pb después del TSS para genes transcripcionalmente activos y ~10 pb para genes inactivos. En levadura y *Drosophila*, por ejemplo, la formación del NDR no se limita a la región 5' de los genes sino que también está presente en la región 3' (Mavrigh *et al*, 2008; Jansen y Verstrepen, 2011). El tamaño del DNA linker también varía según el organismo, en levadura es de ~18 pb, en *Drosophila* ~28 pb y en humano de ~38 pb (Jiang y Pugh, 2009). El 95% de genes en levadura presentan NDRs, sin embargo, en humano la formación de estas estructuras correlaciona con genes transcripcionalmente activos o que están próximos a transcribirse y están unidos a la RNA pol II no fosforilada presente en genes pausados cuyo promotor es transcripcionalmente inactivo (Jansen y Verstrepen, 2011; Shones *et al*, 2008). Un aspecto relevante para nuestro estudio es que las zonas libres de nucleosomas también están presentes en elementos de regulación como los *enhancers* (Escamilla-del-Arenal y Recillas Targa, 2008; Shones *et al*, 2008; Gaffney *et al*, 2012; Jansen y Verstrepen, 2011).

Usando una línea celular de cáncer de próstata humano que responde a la estimulación por andrógenos se generó un mapa de alta resolución del posicionamiento nucleosomal en regiones regulatorias asociadas a *enhancers* en todo el genoma (Fig. 7) (He *et al*, 2010). En las células de cáncer de próstata, donde el receptor de andrógenos se une principalmente a *enhancers*, encontraron que el tratamiento con andrógenos permite la salida de un nucleosoma central presente en los sitios del genoma donde se une este receptor. La región genómica que ocupa el receptor de andrógenos está delimitada por un par de nucleosomas post-traduccionalmente marcados con la H3K4me2. Después de la unión del receptor de andrógenos, el nucleosoma central que contiene la variante de histona H2A.Z (en vez de la histona canónica H2A) además de la marca de histona H3K4me2 se desestabiliza y se genera una región libre de nucleosomas; sin embargo, los dos nucleosomas que limitan el NDR se posicionan de manera más estable (Fig. 7) (He *et al*, 2010). Esta estructura de cromatina característica de los *enhancers* ha sido recientemente confirmada mediante el análisis de RNAs nacientes donde los *enhancers* se caracterizan por la formación de un NDR y la simetría de dos nucleosomas que delimitan pares de transcripción divergentes (Core *et al*, 2014).

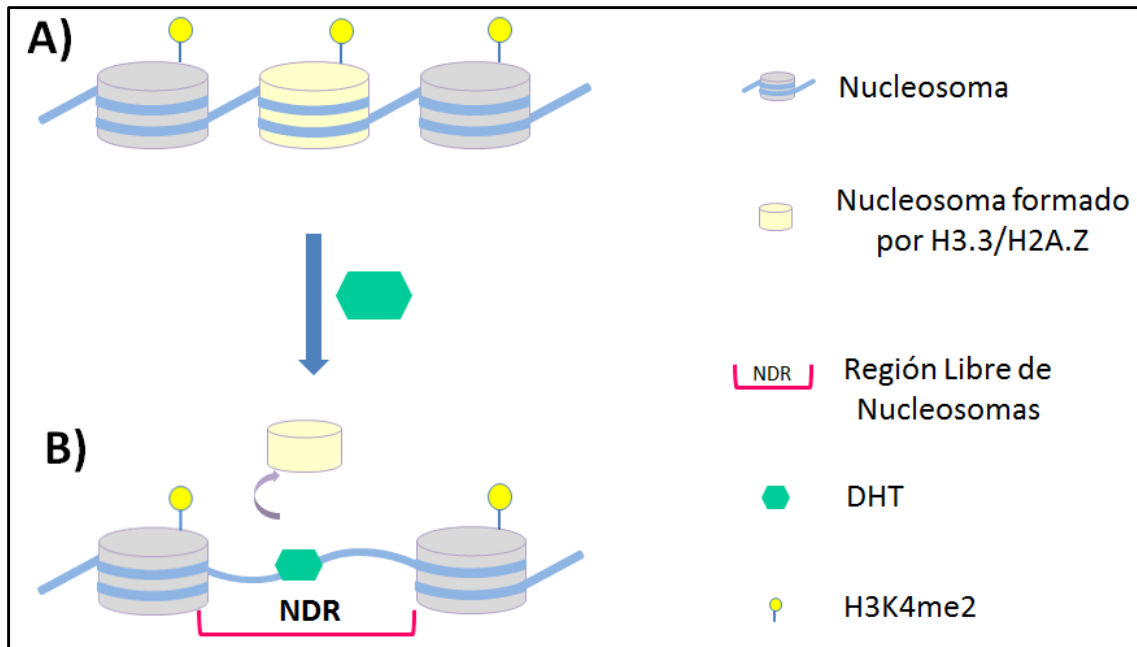


Figura 7. Dinámica nucleosomal de *enhancers* presentes en la línea celular LNCaP en respuesta al tratamiento con andrógenos. **A)** Antes del tratamiento con 5- α -dihidrotestosterona (DHT) muchos *enhancers* en el genoma de las células están marcados por la H3K4me2. **B)** Después de 4 horas de tratamiento con DHT, el nucleosoma central que ocluye el sitio de unión al receptor de andrógenos se elimina generando un NDR mientras que los nucleosomas que limitan el NDR están mejor posicionados (He *et al*, 2010).

Varios factores afectan la posición que ocupan los nucleosomas en el genoma como los factores de transcripción, la secuencia de DNA, las variantes de histonas y los complejos de remodelación de la cromatina dependientes de ATP (Jiang y Pugh, 2009; Struhl y Segal, 2013). Estos factores pueden favorecer el posicionamiento de los nucleosomas en lugares específicos como promotores y *enhancers* o bien, en sentido opuesto, facilitar la formación de regiones libres de nucleosomas.

II.2.1 Características de las secuencias de DNA presentes en promotores y *enhancers*

En levaduras, la principal señal de posicionamiento nucleosomal en *cis* presente en los *NDRs* de promotores se origina en secuencias ricas en A/T, en particular, en fragmentos de DNA conocidos como poli (dA:dT) que actúan previniendo fuertemente la formación de nucleosomas debido a la rigidez de la secuencia (Yuan *et al*, 2005). En contraste, en *Drosophila* se observa una periodicidad dictada por secuencias ricas en CC/GG, este patrón reproduce la

posición exacta del nucleosoma +1, por lo tanto, favorece la formación de nucleosomas (Mavrich *et al*, 2008). En humanos, las secuencias subyacentes de DNA para posicionar nucleosomas no demuestran ser suficientes para predecir el lugar que ocupan los nucleosomas *in vivo* (Lee *et al*, 2008; Zhang *et al*, 2009). En relación con esto, muchas regiones reguladoras como promotores y *enhancers* que se unen a factores de transcripción tienden a estar ocupadas por nucleosomas *in vivo* (Tillo *et al*, 2010). Por esta razón se propone que los organismos metazoarios son más dependientes de la actividad asociada con los complejos ATP-dependientes que remodelan la estructura de la cromatina (Mavrich *et al*, 2008).

El análisis de la composición de nucleótidos presentes en la secuencia de DNA proveniente de *enhancers* que responden a la estimulación con andrógenos en la línea celular de cáncer de próstata LNCaP es congruente con lo reportado en los promotores de genes. Es decir, las combinaciones de dinucleótidos de adenina y timina (AA / TT / TA / AT) por lo general se asocian con regiones libres de nucleosomas o NDRs presentes en los *enhancers* mientras que los nucleótidos guanina y citosina muestran la tendencia opuesta (He *et al*, 2010).

11.3 Variantes de histonas

Con excepción de la histona H4, el resto de las histonas canónicas H3, H2A y H2B presentan diversas isoformas que se conocen como variantes de histonas (Fig. 8) (Skene y Henikoff, 2013). Las histonas canónicas se sintetizan y depositan al mismo tiempo que ocurre la replicación del DNA durante la fase S del ciclo celular; en cambio, las variantes de histonas pueden integrarse a la cromatina en forma independiente a la replicación (Ahmad y Henikoff, 2002; Skene y Henikoff, 2013; Biterge y Schneider, 2014).

Estas variantes se diferencian de las histonas canónicas en sus extremos amino y carboxilo terminal, siendo el *core* altamente conservado. En algunos casos, la diferencia de aminoácidos entre histonas canónicas y sus variantes es mínima, por ejemplo, la variante H3.3 se diferencia de la H3 canónica por tan solo 4/5 aminoácidos, en cambio la variante de histona macroH2A tiene un dominio globular en el extremo carboxilo que la hace casi tres veces mayor en tamaño que la histona canónica (Fig. 8) (Hake *et al*, 2005; Biterge y Schneider, 2014).

El papel funcional de las variantes de histonas en distintas regiones de cromatina es remarcable, por ejemplo, el DNA puede sufrir roturas de doble cadena (*DBS*) y cuando esto ocurre, la histona H2A es reemplazada por la variante H2A.X y sirve como blanco para la fosforilación en la serina 139, esto permite el reclutamiento de la maquinaria de reparación del DNA y de complejos remodeladores de cromatina como INO80 y SRW1 (Ver más adelante en sección II.4) (Biterge y Schneider, 2014). Aunque los cambios en la secuencia de aminoácidos entre H3 y H3.3 pueden parecer sutiles, la variante H3.3 puede reemplazar a la H3 en promotores de genes transcripcionalmente activos durante el ciclo celular por un complejo proteico llamado HIRA, favoreciendo una cromatina activa sin marcas represivas, en cambio la H3 solamente se deposita acoplada a la replicación del DNA por el complejo CAF-1 (Ahmad y Henikoff, 2002; Tagami *et al*, 2004; Henikoff, 2008). En varios organismos, la histona H3.3 además co-localiza con otra variante de histonas, la H2A.Z en promotores de genes activos y *enhancers* (Mavrigh *et al*, 2008; Lee *et al*, 2008; Jin *et al*, 2009; Tolstorukov *et al*, 2009; He *et al*, 2010). En *Drosophila*, el 85% de los genes codificantes contienen al menos un nucleosoma con H2A.Z cercano al *TSS* (Fig. 6) (Mavrigh *et al*, 2008).

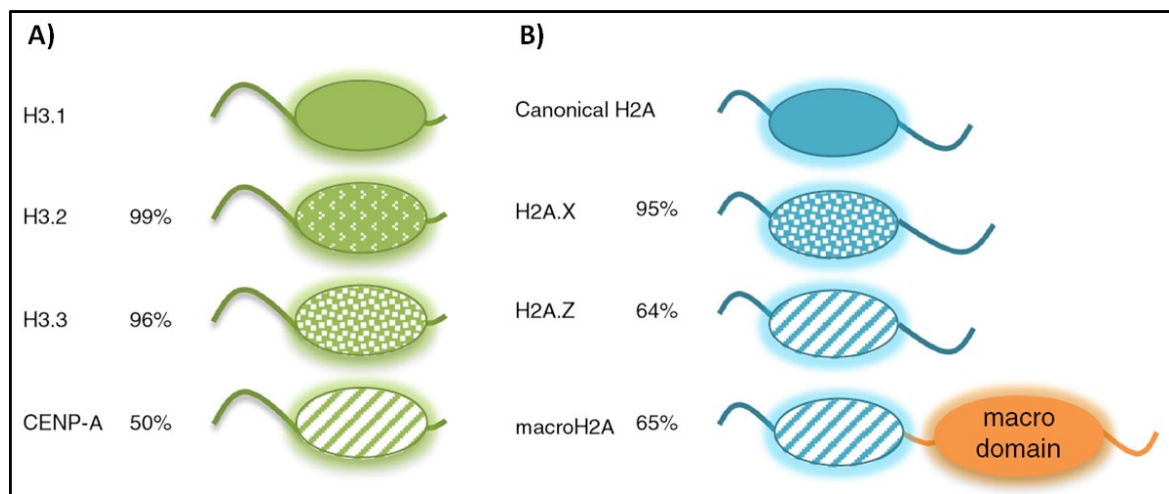


Figura 8. Representaciones esquemáticas de las variantes de la histona H3 y H2A. **A)** Se muestran las variantes de la histona H3.1. **B)** Se muestran las variantes de la histona H2A.Z. Los dominios globulares se muestran como óvalos y los extremos amino y carboxilo terminales como líneas, excepto para la variante macroH2A que tienen un dominio globular grande. El porcentaje de similitud con las histonas canónicas se muestra del lado izquierdo de los óvalos (Tomado de Biterge y Schneider, 2014).

La sustitución de las histonas canónicas por sus variantes es otra forma de alterar la estructura y propiedades termodinámicas de los nucleosomas, por ende, sus funciones regulatorias en el genoma (Jin y Felsenfeld, 2007; Jin *et al*, 2009). Un posible efecto de la incorporación de las variantes de histonas H3.3/H2A.Z en los nucleosomas de promotores y *enhancers* es la desestabilización de las interacciones entre el DNA y los nucleosomas. En experimentos *in vitro*, los nucleosomas que contienen H3.3/H2A.Z tienden a disociarse más rápidamente cuando el solvente tiene una fuerza iónica elevada, por ejemplo, en presencia de cantidades altas de sales (Jin *et al*, 2007). Por lo anterior, la sustitución de H2A por H2A.Z en promotores y *enhancers* podría favorecer o proporcionar señales para la formación de regiones libres de nucleosomas (Jin *et al*, 2007; Goldberg *et al*, 2010).

II.4 Complejos remodeladores de la cromatina dependientes de ATP

Los complejos remodeladores de la cromatina son enzimas que utilizan la energía proveniente de la hidrólisis del ATP para alterar la asociación física entre las histonas y el DNA, por lo tanto, son responsables de la modificación y alteración de la estructura de la cromatina en forma local o global (Vignali *et al*, 2000; Smith y Peterson, 2005).

Se ha propuesto que estos complejos pueden mover, desestabilizar, desalojar o re-estructurar la composición de los nucleosomas. A pesar de esto, los mecanismos de remodelación aún están bajo discusión pero el posicionamiento *in vivo* de los nucleosomas se ha atribuido ampliamente a la presencia de estos complejos (de-la-Serna *et al*, 2006; Clapier y Cairns, 2009; Brown *et al*, 2011; Lorch *et al*, 2011). Para alterar la estructura de la cromatina, estas enzimas deben ser reclutadas a través de interacciones proteína-proteína pues a diferencia de los factores de transcripción, estos complejos de remodelación de la cromatina no se unen directamente al DNA sino que contienen dominios que reconocen otras proteínas o modificaciones post-traduccionales en las histonas permitiendo exponer en mayor o menor medida secuencias reguladoras en el DNA (de-la-Serna *et al*, 2006, Clapier y Cairns, 2009). Todas estas enzimas comparten un dominio ATP-dependiente tipo helicasa, similar al de la familia SWI2/SNF2 en levaduras (Smith y Peterson, 2005). Se han subdividido en cuatro

familias principales de acuerdo a los dominios característicos que poseen: SWI/SNF, ISWI/SWI, CHD e INO80 (Fig. 9) (Clapier y Cairns, 2009).

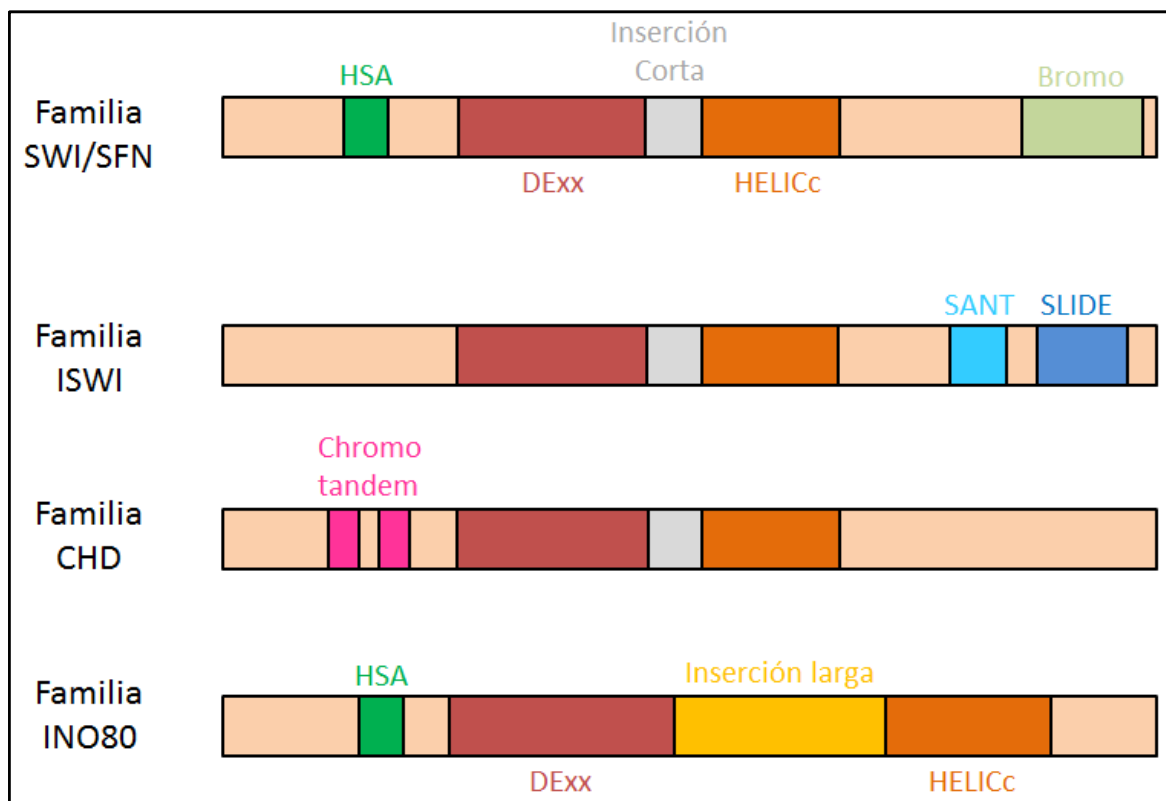


Figura 9. Dominios proteicos asociados con las familias de remodeladores de la cromatina dependientes de ATP. Todas las familias tienen en común un dominio ATPasa y se distinguen por los otros dominios que poseen. Las familias SWI/SNF, ISWI y CHD tienen una inserción pequeña (gris) dentro del dominio ATPasa, a diferencia de la familia INO80 que contiene una inserción larga (amarillo). En verde claro se muestra el bromodominio de la familia SWI/SNF. El dominio HSA se muestra en verde oscuro. El módulo SANT-SLIDE de la familia ISWI en azul y el cromodominio en magenta (Tomado de Clapier y Cairns, 2009)

II.4.1 Remodeladores de la familia SWI/SNF.

Los miembros de la familia SWI/SNF poseen al menos un bromodominio (dominio proteico que reconoce residuos de lisina metilados) y se asocian con otras proteínas como *ARPs* (*actin-related proteins*) (Yun *et al*, 2011; Becker y Hörz, 2002). SWI/SNF y RSC se describieron por primera vez en levadura mediante análisis genéticos, pero sus homólogos están presentes en otros organismos como *Drosophila* (Brahma) y humano (complejos PBAF y BRG1) (Becker y Hörz, 2002). Esta familia puede deslizar o expulsar nucleosomas en muchos locus y por

distintas razones, pero no pueden ensamblar cromatina (Clapier y Cairns, 2009). En humano, la formación de oligodendrocitos, un tipo de célula especializada para la mielinización en el sistema nervioso central (CNS), juega un papel importante en el desarrollo del cerebro y la función neuronal. En este contexto, se ha reportado que el factor transcripcional Olig2 dirige la unión del complejo Smarca4/BRG1 (SWI/SNF) a *enhancers* que exhiben un enriquecimiento en la marca de histonas H3K27ac (Rada -Iglesias *et al*, 2011; Yu *et al*, 2013). Lo anterior promueve la progresión y maduración del linaje de los oligodendrocitos para permitir la mielinización de sistema nervioso central (CNS) en los seres humanos (Yu *et al*, 2013). En este modelo, la mutante Brg1 ATPasa-*dead* (K798R), inhibe la expresión de genes de mielina lo que sugiere que la actividad ATPasa es esencial en la función de Brg1 para inducir el programa de mielinización (Yu *et al*, 2013).

II.4.2 Remodeladores de la familia ISWI

Esta familia contiene la ATPasa ISWI, fue descubierta en *D. melanogaster* debido a sus similitudes con el complejo Brahma. Los integrantes de esta familia son NURF, ACF y CHRAC. En humano se han encontrado los homólogos hWICH y hNoRC. En el extremo carboxilo terminal de la familia ISWI residen dos dominios característicos denominados SANT y *SLIDE*, que en conjunto reconocen histonas no modificadas y DNA. Los complejos ACF y CHRAC pueden ensamblar cromatina y reprimir genes mientras que NURF tiene el efecto opuesto (Clapier y Cairns, 2009).

En humano, hay evidencias experimentales que muestran que los *enhancers* funcionales se asocian con cambios en el posicionamiento nucleosomal (Yasui *et al*, 2002; Shones *et al*, 2008). SATB1 ("*special AT-rich sequence binding 1*") una proteína que se encuentra predominantemente en los timocitos (linfocitos T en desarrollo) dirige la unión *in vivo* de las subunidades ACF1 e ISWI al *enhancer* μ localizado dentro del gen del receptor de la interleucina-2 (IL-2Ra) y regula el posicionamiento nucleosomal dentro de este locus (Yasui *et al*, 2002). Una clara diferencia en el patrón de corte por la nucleasa micrococcal en la región del *enhancer* entre el control y una mutante nula para la proteína SATB1 muestra una clara diferencia en la distribución de los nucleosomas sobre el *enhancer*. La ausencia del complejo

remodelador dependiente de ATP tiene como consecuencia la represión del gen de IL-2Ra y coincide con cambios en la accesibilidad de la cromatina del dominio, en particular, una menor accesibilidad del elemento regulador *enhancer* (Yasui *et al*, 2002).

II.4.3 Remodeladores de la familia CHD

Esta familia es muy interesante porque además de poseer la ATPasa Mi-2, contiene cromodominios (dominios proteicos que reconocen preferencialmente la acetilación de las lisinas) que reconocen los extremos amino terminales en las histonas (Fig. 9) (Clapier y Cairns, 2009; Zentner y Henikoff, 2013). Ciertos remodeladores CHD deslizan o expulsan nucleosomas para promover la transcripción (Clapier y Cairns, 2009). CHD7, es uno de los miembros de esta familia que se requiere para el desarrollo normal de los mamíferos (Schnetz *et al*, 2010). Para hacerse una idea de la función de CHD7, se determinó su ubicación genómica en células troncales de ratón. La unión de CHD7 correlaciona con *enhancers* distales al TSS, con altos niveles de H3K4me1 dentro de regiones DHS, p300 y la presencia de reguladores maestros en células troncales como OCT4, SOX2 y NANOG (Schnetz *et al*, 2010). Aunque el mecanismo de acción no fue explorado, la regulación negativa parece ser el efecto más directo de la unión de CHD7 a los *enhancers*. Una implicación importante que deriva de este estudio es que en el Síndrome de CHARGE causado por mutaciones en el gen *CHD7* existe una expresión inapropiada de genes que deben ser silenciados durante el desarrollo (Schnetz *et al*, 2010). El efecto represivo de esta familia ha sido ampliamente documentado en otros complejos como NuRD, que desacetila a las histonas presentes en regiones de DNA metilado debido a su acoplamiento con la proteína MBD2 que reconoce el DNA metilado en forma preferencial (Feng y Zhang, 2001; Becker y Hörz, 2002; Clapier y Cairns, 2009).

II.4.4 Remodeladores de la familia INO80

La característica de esta familia es que posee una larga inserción en medio del dominio ATPasa donde se unen las proteínas Rvb1/2 y ARP (Fig. 9). Entre sus funciones se encuentran la activación de la transcripción y reparación del DNA (Clapier y Cairns, 2009).

Una función alternativa y descrita recientemente para los complejos remodeladores de cromatina dependientes de ATP, es su capacidad para intercambiar histonas, en particular variantes de histonas, en nucleosomas localizados en zonas específicas del genoma, con consecuencias funcionales (Clapier y Cairns, 2009). Recientemente se han identificado a Swr1 e INO80, en el intercambio de dímeros H2A.Z-H2B por H2A-H2B en forma independiente de la replicación (Krogan *et al*, 2003). A través de experimentos *in vitro* se determinó que Swr1 cataliza la sustitución unidireccional del dímero H2A-H2B por H2A.Z-H2B (Luk *et al*, 2010), mientras que INO80 cataliza la reacción opuesta (Papamichos-Chronakis *et al*, 2011).

En células madre de ratón (mESCs), el complejo INO80 se asocia con los promotores de genes pluripotentes. La unión de INO80 depende de la presencia de Oct4 y WDR5, un componente del complejo metiltransferasa H3K4. La presencia del complejo INO80 favorece una estructura de cromatina accesible en la región del promotor de los genes de pluripotencia y tiene la capacidad de reclutar a componentes del Mediador y la RNA pol II [Wang *et al*, 2014]. Sin embargo, aún falta determinar si los *enhancers* de estos genes son regulados de la misma manera.

En resumen, los estudios a gran escala del genoma muestran que los promotores y *enhancers* de genes activos suelen asociarse con variantes de histonas (H2A.Z/H3.3) y generar regiones libres de nucleosomas o NDRs (Jiang y Pugh, 2009; He *et al*, 2010; Calo y Wysocka, 2013; Biterge y Schneider, 2014; Core *et al*, 2014). Aunque los complejos remodeladores de cromatina dependientes de ATP facilitan el ensamblaje de la maquinaria de transcripción, los mecanismos que permiten un adecuado posicionamiento nucleosomal en regiones reguladoras como los *enhancers* no se conocen en detalle. En esta tesis, los resultados sugieren que la región denominada ERAE contribuye de forma importante con el posicionamiento nucleosomal del enhancer 3' α -globina de pollo.

II.5 Modificaciones post-traduccionales en las histonas

En forma general, las modificaciones post-traduccionales en los extremos amino terminales de las histonas pueden afectar directamente la estructura de la cromatina o bien, establecer

una superficie específica o “código de histonas” que les permite interactuar con otros complejos proteicos (Strahl y Allis, 2000; Berger, 2002, Kouzarides, 2007). La acetilación por ejemplo, reduce la afinidad de las histonas por el DNA cargado negativamente y ayuda en la relajación de la fibra de cromatina (Berger, 2007). Regiones asociadas al proceso de transcripción como promotores y *enhancers* suelen encontrarse altamente acetiladas, mientras que las regiones silenciadas están desprovistas de esta modificación en las histonas (Li *et al*, 2007; Berger, 2007; Wang *et al*, 2008). La distribución de diferentes modificaciones en las histonas o las enzimas responsables de depositarlas a lo largo de todo el genoma ha demostrado ser capaz de identificar *enhancers* (Roh *et al*, 2005, 2007; Heintzman *et al*, 2007), en diferentes líneas celulares (Heintzman *et al*, 2007, 2009; Wang *et al*, 2008; Creyghton *et al*, 2010), *in vivo* (Visel *et al*, 2009) y durante el desarrollo (Creyghton *et al*, 2010; Zhou *et al*, 2011).

II.5.1 La marca en la histona H3K4me1 se asocia con elementos enhancer en todo el genoma

La mono-, di- y tri metilación en las histonas son marcas de cromatina que han sido analizadas a nivel global (Heintzman *et al*, 2007 y 2009). La primera modificación en las histonas a escala genómica en ser relacionada con elementos reguladores a distancia fue la mono-metilación en la lisina 4 de la histona 3 (H3K4me1) pues muestra un enriquecimiento variable y tejido específico que coincide con la presencia de elementos tipo *enhancer* (Heintzman *et al*, 2009). En contraste, la tri-metilación en la lisina 4 de la histona H3 (H3K4me3) muestra un enriquecimiento en regiones promotoras de genes activos pero niveles muy bajos en regiones correspondientes a *enhancers* (Heintzman *et al*, 2009, Pekowska *et al*, 2011). Las diferencias a nivel de modificaciones post-traduccionales en las histonas en el genoma sugieren fuertemente que los *enhancers* participan en la regulación de patrones complejos de expresión de genes. En levadura se ha descrito que la proteína Set1 es la única metil-transferasa que mono-, di- y tri-metila a las histonas; sin embargo, en mamíferos se han reportado al menos seis complejos que contienen la subunidad catalítica Set1/COMPASS (Herz *et al*, 2012; Lee *et al*, 2013). Trx y Trr en *Drosophila* y MLL3/MLL4 en mamíferos son las principales mono-metil-transferasas asociadas con regiones tipo *enhancer* (Herz *et al*, 2012;

Hu *et al*, 2013; Lee *et al*, 2013; Tie *et al*, 2014); mientras que las proteínas dSet1 en *Drosophila* y SET1A/SET1B en mamíferos, depositan la modificación en la histona H3K4me3 en los promotores de genes (Lee *et al*, 2013).

II.5.2 Las acetil-transferasas de histonas p300 y CBP

Anteriormente se mencionó que la acetilación de la cromatina resulta importante para la activación génica, por ello resulta útil conocer si sobre las regiones *enhancer* se asocian proteínas que llevan a cabo dicha actividad enzimática para usarlas como un predictor de dichas secuencias. Las proteínas p300 y CBP son acetil-transferasas de histonas (HATs) que se expresan de forma ubicua y se relacionan con el inicio competente de la RNA pol II sobre algunas regiones promotoras (Cho *et al*, 1998) pero sobre todo en *enhancers* (Merika *et al*, 1998; Heintzman *et al*, 2007; Visel *et al*, 2009; Kim *et al*, 2010). En concordancia, un análisis global del genoma durante el desarrollo en humanos muestra que el 95 % de los sitios de unión a p300 no corresponden a promotores, sino a *enhancers* (Rada-Iglesias *et al*, 2011). Datos experimentales han confirmado que el enriquecimiento *in vivo* de p300 en el genoma es un indicador confiable de secuencias tipo *enhancer*. Además, la mayoría de las regiones asociadas con p300 están conservadas y correlacionan con los patrones de expresión de genes propios del desarrollo de embriones de ratón (Visel *et al*, 2009). A pesar de que la acetil-transferasa de histonas p300 es un buen predictor de secuencias tipo *enhancer*, no todos los *enhancers* se unen a esta acetil-transferasa pero sí a otras proteínas como TRAP220 que forma parte del complejo Mediador (ENCODE Project Consortium *et al*, 2007). Este hallazgo evidencia la importancia de explorar otras modificaciones en la cromatina que ayuden a definir con precisión a los *enhancers*.

II.5.3 La modificación en la histona H3K27ac y el estado de actividad de los enhancers

Se ha observado que la presencia de la marca en la histona H3K4me1 coincide con la acetil-transferasa de histonas p300 en forma tejido específica en el genoma, pero esto no es suficiente para predecir si los *enhancers* son activos o no (Heintzman *et al*, 2009; Creighton *et al*, 2010). En un intento por evaluar la posible actividad de los *enhancers*, se analizó la

expresión de genes cercanos mediante el uso de microarreglos o secuenciación masiva de RNA (*RNA-seq*). El análisis reportó una alta correlación entre la actividad transcripcional de los genes cercanos y los sitios enriquecidos con p300 (Heintzman *et al*, 2009; Visel *et al*, 2009). Sin embargo, otros grupos reportaron que los genes asociados con la acetil-transferasa de histonas p300 no se transcriben de manera más activa que el promedio de genes no asociados con p300 (Creyghton *et al*, 2010; Bonn *et al*, 2011; Rada-Iglesias *et al*, 2011). En contraste, se reportó que sí existe una correlación positiva entre la acetilación de la lisina 27 en la histona 3 (H3K27ac) asociada con *enhancers* cercanos a genes transcripcionalmente activos (Creyghton *et al*, 2010; Zentner *et al*, 2011; Bonn *et al*, 2012). Durante la diferenciación de células troncales embrionarias murinas (mESC) a células diferenciadas (*Neural Progenitor Cells* o *NPC*) se determinó que la marca H3K4me1 se encuentra presente en ambos tipos celulares, mientras que la modificación en la H3K27ac está presente únicamente en células progenitoras neurales (*NPC*). Lo anterior sugiere que cuando las ESC comienzan a diferenciarse, la adquisición de la marca en la histona H3K27ac en regiones *enhancer* correlaciona con el nuevo estado de diferenciación celular (Creyghton *et al*, 2010; Rada-Iglesias *et al*, 2011).

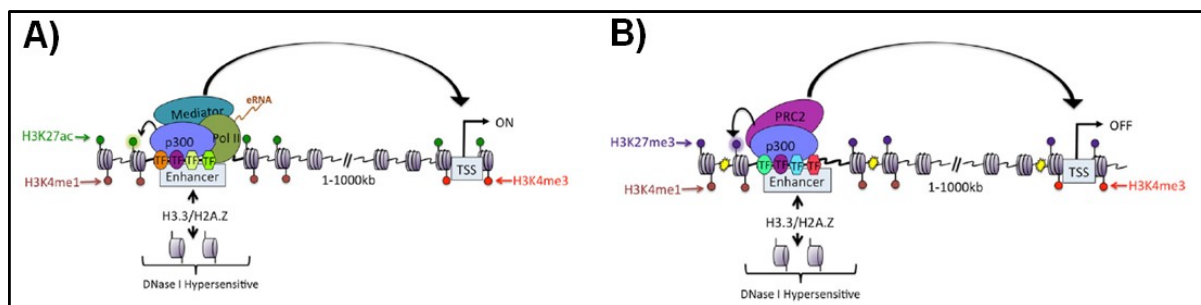


Figura 10. Características epigenéticas relacionadas con distintos estados de actividad de los *enhancers*. **A)** *Enhancers* activos. Los *enhancers* están asociados con las variantes de histonas H3.3 y H2A.Z, que compiten con los factores de transcripción por el DNA. Los factores de transcripción pueden atraer otras proteínas co-activadoras que modifican la estructura de la cromatina. Los nucleosomas que delimitan el *enhancer* tienen las marcas en las histonas H3K4me1 y H3K27ac. **B)** *Enhancers* predeterminados en células troncales. Los nucleosomas que delimitan el *NDR* tienen las marcas en las histonas H3K4me1 y H3K27me3. Se propone que parte del mecanismo de represión de los genes en células troncales está mediado por el complejo PRC2 (Tomado de Calo y Wysocka, 2013).

En resumen, las modificaciones post-traduccionales específicas en las histonas están enriquecidas en elementos reguladores importantes para la transcripción, como promotores y

enhancers (Barski *et al*, 2007; Heintzman *et al*, 2007; Roh *et al*, 2005 y 2007). El uso de combinatorias de marcas en los extremos amino terminales de las histonas como la H3K27ac, la H3K4me1 y la presencia de p300 se han usado para predecir nuevos *enhancers* (Heintzman *et al*, 2007) e inferir si son activos o no (Fig. 10). Sin embargo, ya se ha discutido que otros procesos epigenéticos como el posicionamiento nucleosomal y el recambio de histonas canónicas por sus variantes a lo largo del genoma constituyen aspectos relevantes que también afectan la expresión de los genes.

II.6 Complejos Polycomb y Trithorax

Se ha visto que complejos multi-proteicos pertenecientes al grupo de Polycomb y Trithorax son reguladores clave de la expresión de genes implicados en el desarrollo de diversos organismos, memoria epigenética y control del ciclo celular (Schuettengruber *et al*, 2007; Steffen y Ringrose, 2014). Ambos complejos regulan a los genes homeóticos, un conjunto de factores de transcripción que especifican la identidad celular a lo largo del eje anteroposterior en animales segmentados y también se han identificado en mamíferos (Steffen y Ringrose, 2014). El silenciamiento mediado por Polycomb es quizás el mejor ejemplo caracterizado de silenciamiento epigenético e implica la tri-metilación en la lisina 27 de la histona H3 (H3K27me3) por *Enhancer* de Zeste E(Z), la metil-transferasa presente en el complejo PRC2 (Polycomb-group Repressive Complex 2). La proteína Pc que forma parte del complejo PRC1 (Polycomb Repressive Complex 1), reconoce la marca H3K27me3. El silenciamiento por Polycomb es antagonizado por las proteínas Trithorax, un grupo diverso de reguladores de la cromatina que se requieren para el mantenimiento del estado transcripcionalmente activo y a largo plazo de los genes que regulan (Pasini *et al*, 2010; Tie *et al*, 2014).

Los patrones distintos de modificaciones en las histonas asociado con los *enhancers* está directamente relacionado con su actividad (Zentner *et al*, 2010). Esto ha permitido dividirlos en *enhancers* predeterminados o pausados que contienen H3K4me1 más la H3K27me3 y *enhancers* activos asociados con H3K4me1 y H3K27ac (Fig. 10) (Creyghton *et al*, 2010; Rada-Iglesias *et al*, 2011). En las células troncales embrionarias humanas y de ratón, los *enhancers* de genes transcritos activamente están marcados por la presencia de H3K4me1 y H3K27ac

(Creyghton *et al*, 2010; Rada-Iglesias *et al* 2011; Zentner *et al*, 2011). Sin embargo, los *enhancers* de genes de desarrollo, que se activan en etapas posteriores están marcados por la H3K4me1 y la marca represiva H3K27me3 impuesta por PcG en lugar de la marca H3K27ac de activación depositada por p300/CBP (Rada-Iglesias *et al*, 2011). Lo anterior es consistente con el hecho de que las modificaciones H3K27me3 y H3K27ac son marcas excluyentes en las histonas (Tie *et al*, 2012). En *Drosophila* se ha demostrado que los niveles de acetilación en la histona H3K27 dependen de la acción de otras proteínas trithorax como UTX que se encarga de desmetilar de forma específica a la histona H3K27 además de asociarse físicamente con CBP y el complejo remodelador ATP dependiente BRM (Tie *et al*, 2012). La asociación de estas proteínas proporciona un mecanismo de desmetilación de la H3K27 acoplado a la acción de una acetil-transferasa de histonas (CBP) necesaria para activar a los *enhancers*.

II.7 Metilación de DNA en secuencias *enhancer*

La metilación de las citosinas es una modificación epigenética esencial para los genomas de mamíferos (Stadler *et al*, 2011). Tomando en cuenta la cantidad de metilación en el DNA se han identificado regiones completamente metiladas (*FMR*), no-metiladas (*UMR*) y de baja metilación (*LMR*) (Stadler *et al*, 2011). Estas últimas corresponden a sitios alejados del TSS, son hipersensibles al corte por la DNasa I, se unen a factores de transcripción, tienen presencia de marcas asociadas con *enhancers* (como la presencia de H3K4me1 y la unión de p300) y estimulan la actividad de genes reporteros, es decir, correlacionan con elementos *enhancer*. Una comparación entre los metilomas de células troncales de ratón y progenitores neuronales (*Neuronal Progenitors/NP*) confirma que la aparición de *LMRs* es específica de tejido, es decir, en los *NP* se forman nuevos *LMRs* adyacentes a genes importantes para el desarrollo neuronal. En concordancia con lo anterior, las secuencias de los nuevos *LMRs* contienen motivos de unión a factores neurales como Sox17 y Pax6 en vez de factores de pluripotencia propios de células troncales como Nanog, Oct4 y Klf4 (Stadler *et al*, 2011).

En general, los *enhancers* suelen estar desprovistos de metilación en la posición 5 de la citosina (5mC) (Hon *et al*, 2014). En los genomas de mamíferos, las enzimas TET oxidan sucesivamente la 5-metilcitosina (5mC) a 5-hidroximetilcitosina (5hmC), 5-formilcitosina (5fC),

y 5-carboxilcitosina (5caC). Los sustratos que pueden ser removidos por la TDG (*thymine-DNA glycosilase*) y reparados por el mecanismo de escisión de bases (*BER*) son la 5fC y la 5caC, con lo que se regeneran citosinas no modificadas. A este mecanismo se le conoce como desmetilación activa (Yu *et al* , 2012, Shen *et al* , 2013, Song *et al* , 2013). Mapas de distribución de los productos de oxidación de la 5mC en todo el genoma han sido generados en células troncales de ratón y humano. Se ha podido determinar que la 5hmC correlaciona con *enhancers* predeterminados (H3K4me1) y activos (H3K4me1 y H3K27ac) (Stroud *et al*, 2011; Yu *et al*, 2012) mientras que la 5fC exhibe mayor preferencia por *enhancers* predeterminados (Shen *et al*, 2013; Song *et al*, 2013). El papel de 5hmC sobre la función de los *enhancers* aún queda por determinarse, sin embargo, algunas proteínas reconocen preferencialmente la marca 5hmC. En las células neuronales progenitoras (*NP*) por ejemplo, la expresión del gen *Uhrf2* aumenta con la diferenciación y el producto peptídico se une específicamente a la 5hmC (Frauer *et al*, 2011; Mellén *et al*, 2012; Spruijt *et al*, 2013). Se necesitan más estudios para demostrar la función de alguno de los productos de oxidación de la 5mC pues su ausencia no es exclusiva de elementos *enhancer* sino que más bien parece ser una característica general de elementos regulatorios que además cambia durante la diferenciación (Rönnerblad *et al*, 2014).

En resumen, estudios del genoma completo muestran que existe una correlación negativa entre la metilación del DNA y elementos de regulación en el genoma identificados por su sensibilidad al corte con la enzima DNasa I (Stadler *et al*, 2012; Thurman *et al*, 2012). Los *enhancers* están sujetos a la acción de muchas proteínas que regulan local y globalmente su accesibilidad, sin embargo también son parte de la organización tridimensional del genoma al interior del núcleo. El tercer nivel de regulación de los genes está relacionado con la organización espacial del genoma al interior del núcleo.

III. Organización espacial del genoma al interior del núcleo

El núcleo no solo es una estructura que contiene al DNA de las células eucariontes, sino un sub-organelo celular dinámico (Cremer y Cremer, 2001). Con ayuda de la microscopía electrónica y de fluorescencia ha sido posible determinar experimentalmente que al interior

del núcleo existe una fracción de la cromatina que se pliega en una fibra más compacta denominada heterocromatina y otra menos compacta conocida como eucromatina (Misteli, 2007). Durante el período de interfase, el DNA se organiza en sub-dominios que se caracterizan por un enriquecimiento de factores específicos como la RNA pol II, la maquinaria de *splicing* o proteínas presentes en la heterocromatina que hacen que porciones del genoma sean diferencialmente permisivas a la expresión génica (Fig.10) (Kress *et al*, 2010; Qureshi y Mehler, 2010).

A lo largo de la diferenciación celular, ocurren cambios globales en la arquitectura nuclear, por ejemplo, el cromosoma humano 19 que contiene una alta cantidad de genes, reside preferentemente en el interior del núcleo mientras que genes con poca cantidad de genes como el cromosoma humano 18, tiene más probabilidades de localizarse en la región periférica asociado con la lámina nuclear (Croft *et al*, 1999). Este tipo de observaciones permitieron establecer que cada cromosoma ocupa una porción relativamente bien definida dentro del núcleo, llamado territorio cromosómico donde las posiciones dentro de estos territorios parecen estar relacionados con el control de la expresión génica y otras actividades nucleares (Fig. 11) (Cremer y Cremer, 2001).

Una visión muy reciente acerca de la estructuración del DNA al interior de núcleo deriva de la implementación de un método para identificar interacciones de la cromatina en todo el genoma conocido como Hi-C (van Berkum *et al*, 2010). El Hi-C es una variante del 3C, por lo tanto, las células se fijan con formaldehído, dando lugar a enlaces covalentes entre segmentos de cromatina espacialmente adyacentes. La cromatina se digiere con una enzima de restricción y los extremos cohesivos resultantes se rellenan con nucleótidos, uno de los cuales está biotinilado (van Berkum *et al*, 2010). La ligación de los extremos resultantes se lleva a cabo bajo condiciones extremadamente diluidas favoreciendo mayoritariamente aquellos eventos de ligación intramolecular. El DNA se purifica usando perlas de estreptavidina y los fragmentos que interactúan son identificados por secuenciación masiva de DNA resultando en un catálogo de fragmentos de todo el genoma que interactúa entre sí (van Berkum *et al*, 2010).

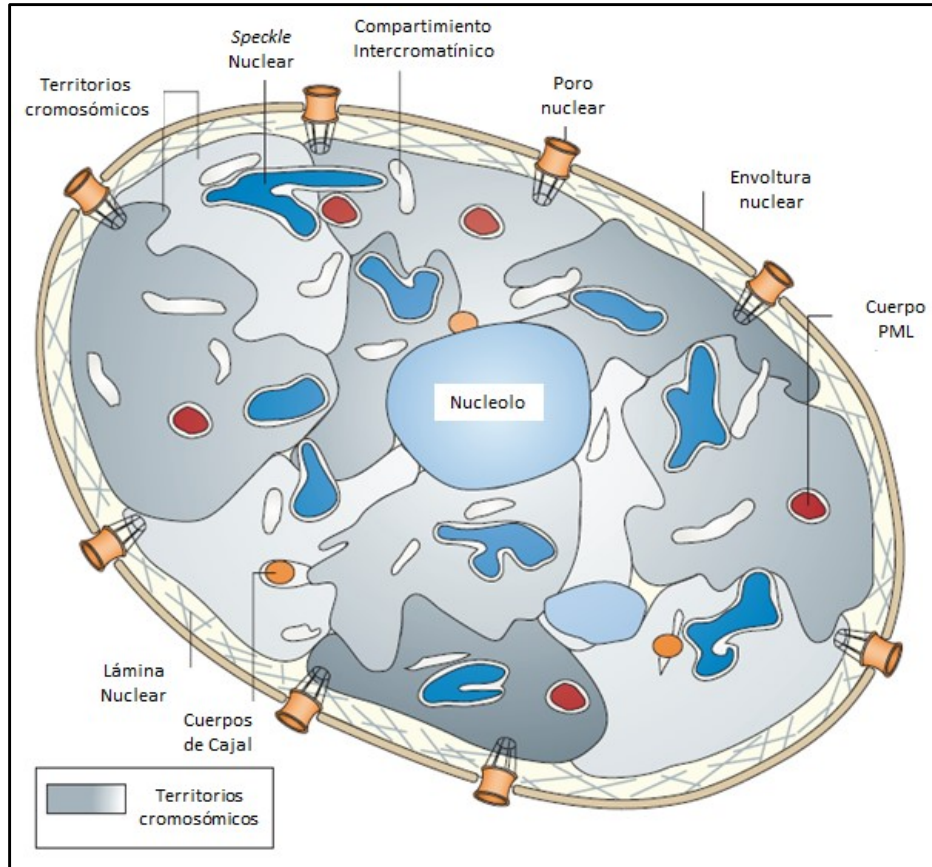


Figura 11. Representación esquemática del interior del núcleo eucarionte en interfase. Se muestran distintos sub-dominios al interior del núcleo que corresponden a territorios cromosómicos, *nuclear speckles*, cuerpos PML, cuerpos de Cajal, nucleolo y poros nucleares (Tomado de Lanctot *et al*, 2007).

Aplicando la metodología de Hi-C en células madre embrionarias y células terminalmente diferenciadas se identificaron grandes dominios de interacción de cromatina conocidos como dominios topológicos o *TADs* (*Topological Associated Domain*), cuyo tamaño es de 1 megabase (Mb) en promedio y limitan la propagación de la heterocromatina (Fig. 12) (Dixon *et al*, 2012; Nora *et al*, 2012). Los límites de los dominios topológicos están enriquecidos en la proteína de unión CTCF, genes constitutivos, RNAs de transferencia y retrotransposones tipo SINE (Dixon *et al* 2012). La mayoría de los *TADs* ya están formados en las células madre embrionarias y son estables en diferentes tipos celulares, incluso entre especies, sin embargo, parece existir una extensa reorganización estructural al interior de cada *TAD* durante la

diferenciación celular (Dixon *et al*, 2012; Philips-Cremens *et al*, 2013; Nora *et al*, 2013; Rao *et al*, 2014).

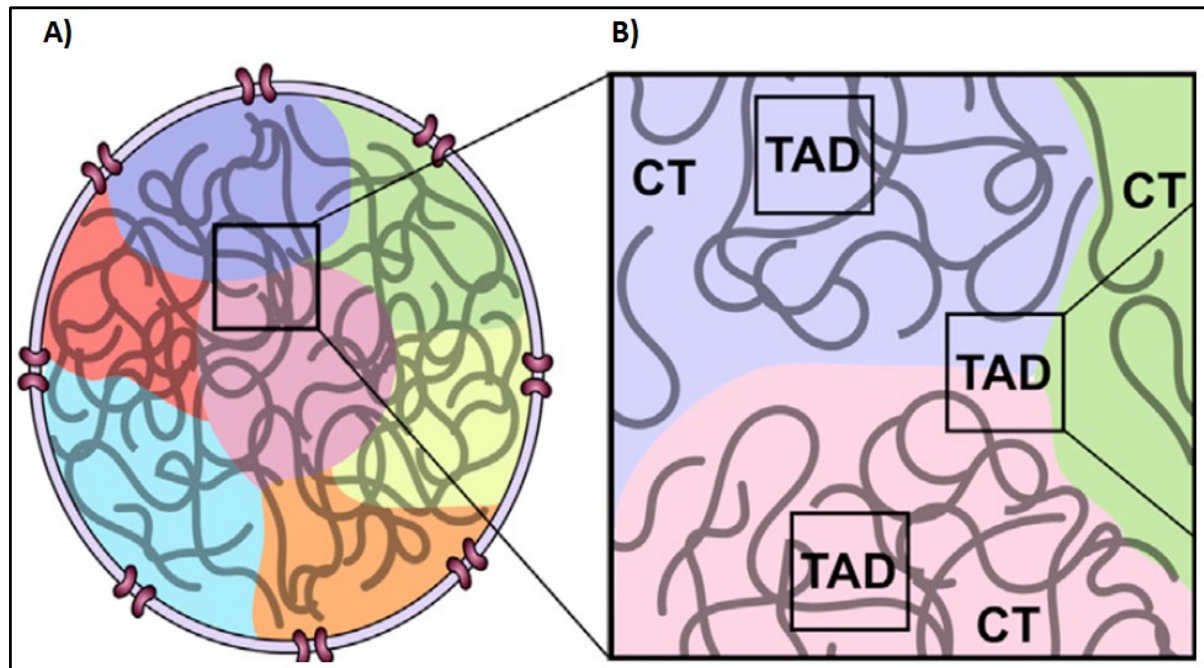


Figura 12. Representación esquemática de los dominios topológicos al interior del núcleo. **A)** En colores se muestra la región que ocupan los cromosomas dentro del núcleo en interfase. **B)** En el lado derecho de la imagen se muestra un acercamiento de la zona que muestra en mayor detalle a los Dominios Topológicos Asociados (*TADs*) y los territorios cromosómicos (*CT*) (Tomado y modificado de Fanucchi *et al*, 2013).

Los dominios topológicos pueden reflejar una propiedad inherente del genoma de los mamíferos y ayudar a limitar la distancia sobre la que operan los *enhancers* (de Laat y Duboule, 2013). En relación con esta idea, uno de los ejemplos extremos de regulación a larga distancia corresponde al gen *Shh* que se localiza linealmente a 1 megabase de distancia del *enhancer* que lo regula (Visel *et al*, 2009). De forma interesante y con la reciente descripción de los *TADs*, se ha reportado que ambos elementos de regulación están contenidos dentro del mismo *TAD* y es probable que esto garantice los contactos entre promotor-*enhancer* a pesar de la distancia lineal entre ellos. Esto es consistente con el hecho de que las frecuencias de interacción dentro de un *TAD* son mayores que entre distintos dominios topológicos (de Laat y Duboule, 2013).

IV. Modelos de acción de los *enhancers*

IV. 1 Modelos de comunicación entre promotores y *enhancers*

Las interacciones entre elementos reguladores que se localizan distantes dentro del genoma determinan en gran medida la especificidad de la expresión de genes durante el desarrollo de los organismos (Ong y Corces, 2011). La capacidad de los *enhancers* de contactar físicamente con su gen blanco plantea la necesidad de generar mecanismos de comunicación que se ajusten a los datos experimentales (Bulger y Groudine, 2011). Al respecto, se han propuesto varios mecanismos de comunicación entre promotores y *enhancers* como encarrilamiento o *tracking*, vinculación o *linking* y asas o *looping*. En la siguiente sección se describen sus principales características (Fig. 13).

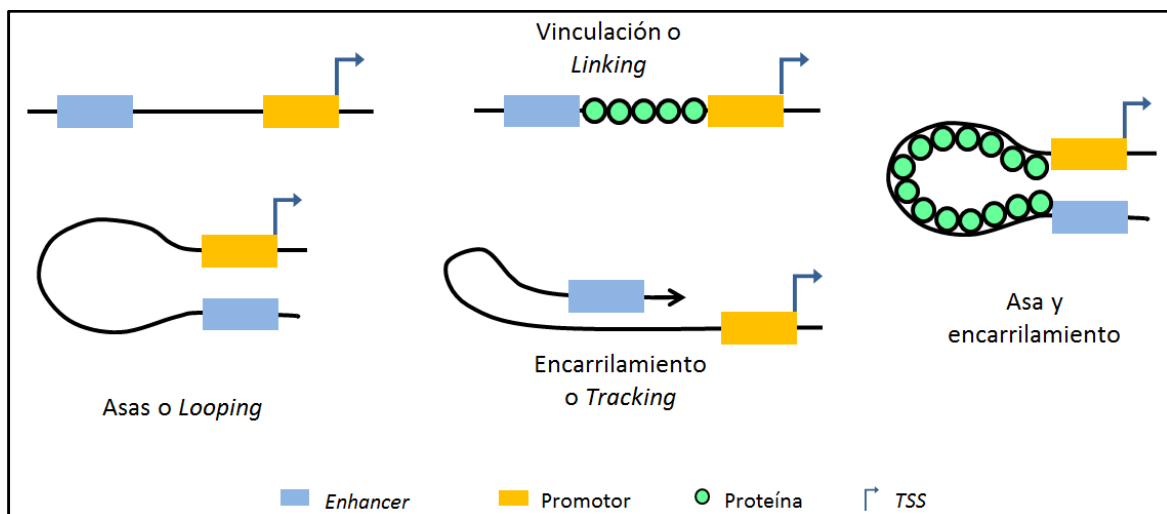


Figura 13. Modelos de comunicación entre un promotor y un *enhancer*. En la forma más simple de asa, *enhancer* y promotor difunden libremente hasta contactarse gracias a su unión con proteínas que regulan los contactos de larga distancia. En el modelo de encarrilamiento, el *enhancer* explora al DNA en busca del promotor. Los modelos de asas y encarrilamiento predicen que la secuencia del *enhancer* está próxima al gen que se va a transcribir. En contraste, en el modelo de vinculación no es necesaria la proximidad entre los elementos reguladores pues el complejo proteico que permite la comunicación entre regiones distales se propaga a lo largo de la fibra de cromatina desde el *enhancer* hacia el promotor. Estos modelos no son excluyentes (Bulger y Groudine, 2002).

IV.1.1 Modelo de asas o looping

El modelo de asas implica el acercamiento de dos secuencias de DNA linealmente distales a través de interacciones proteína-proteína con la flexión de la cromatina situada entre el promotor y el *enhancer* (Bulger y Groudine, 2002). El modelo de asas, por ejemplo, ha sido favorecido con el desarrollo de nuevas metodologías que permiten detectar la frecuencia de asociación entre las distintas regiones del genoma como el 3C o Hi-C (Dekker *et al*, 2002; Tolhuis *et al*, 2002). En los mamíferos, los genes globina residen en un conjunto lineal que refleja su expresión secuencial durante el desarrollo y están bajo la influencia de una serie de sitios hipersensibles al corte por la DNasa I conocido como *LCR* (Grosveld *et al*, 1993). En humanos el *LCR* se localiza aproximadamente 50 kb río arriba del gen β -globina adulto, abarca de 20 a 25 kb y tiene sitios de unión para factores generales y específicos de la transcripción (Grosveld *et al*, 1993). A pesar de la distancia lineal entre el gen β -globina adulto y el *LCR*, este actúa como un elemento eritroide específico que influencia positivamente la transcripción (Grosveld *et al*, 1993; Bulger *et al*, 1999), por lo tanto, es un buen modelo para estudiar la formación de asas de cromatina (Tolhuis *et al*, 2002).

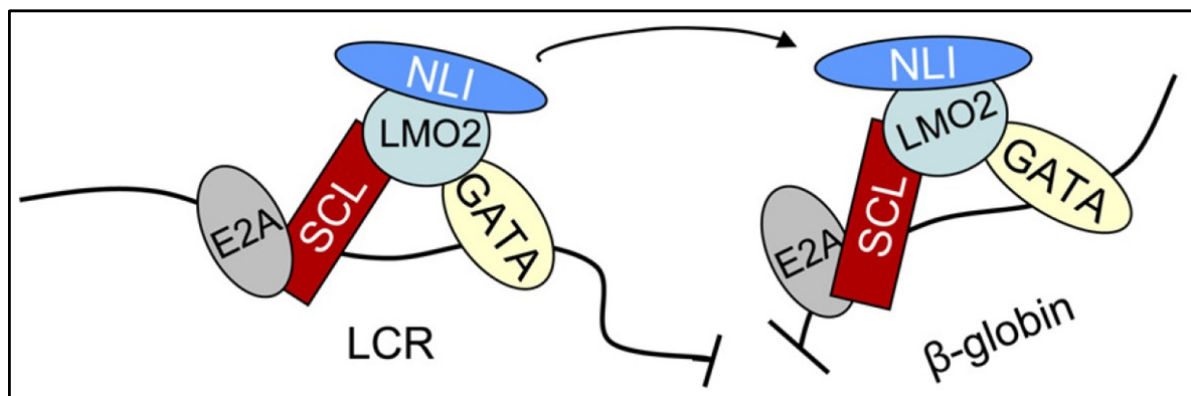


Figura 14. Representación gráfica de la comunicación entre el *LCR* β -globina de mamíferos y su gen blanco. Los factores transcripcionales están presentes en ambos elementos de regulación y forman complejos eritroide-específicos donde la proteína NLI regula la asociación de larga distancia entre ambos elementos de regulación (Tomado de Song *et al*, 2007).

En este contexto se demostró que la proteína nuclear NLI/Ldb1 es un regulador de interacciones de largo alcance (Fig. 14) (Song *et al*, 2007). En mamíferos, antes de la expresión de los genes adultos, el factor NLI/LDB1 además de los factores eritroide específicos GATA-

1/SCL/LMO2 se unen *in vivo* solamente en la región correspondiente al *LCR* del locus β -globina. Una vez que se induce la diferenciación de las células y comienza la expresión de los genes adultos, el complejo se detecta adicionalmente en la región promotora (Song *et al*, 2007). Las interacciones físicas entre ambos elementos de regulación son tejido específicas pues sólo se observan en los linajes hematopoyéticos. Para demostrar que la proteína NLI es responsable de las interacciones de larga distancia se redujo su expresión usando RNAs de interferencia (RNAi). En ausencia de la proteína NLI se pierde la expresión del gen β -globina adulto y la frecuencia de interacción entre el *LCR* y el promotor es menor (Song *et al*, 2007). Estos resultados son consistentes con interacciones tejido específicas entre elementos distales proporcionando evidencia de que las interacciones proteína-proteína facilitan la activación de genes a larga distancia dentro de un contexto tridimensional (Song *et al*, 2007).

IV.1.2 Modelo de vinculación o linking

La comunicación entre dos elementos de regulación a distancia también puede producirse a través de contactos entre proteínas desde el *enhancer* hasta llegar a la región promotora (Bulger y Groudine, 2002). Los dominios β -globina de ratón y humanos están altamente conservados tanto a nivel de la secuencia, como a nivel estructural (Bulger *et al*, 1999). Usando nuevamente como modelo de estudio el dominio β -globina de ratón se reportó el mecanismo por el cual el complejo MLL2 es reclutado a elementos de regulación durante la diferenciación terminal de las células eritroides (Demers *et al*, 2007). Las proteínas MLL2 y ASH2L tienen actividad enzimática como metil-transferasas de histonas y forman parte del complejo MLL2. Mediante inmunoprecipitación de la cromatina (ChIP) acoplada a amplificación de DNA en tiempo real (qPCR) se determinó que las proteínas ASH2L y MLL2 son reclutadas en forma específica al *LCR* β -globina a través del factor NF-N2. Cuando ocurre la diferenciación eritroide, la proteína MLL2 además de estar unida *in vivo* al *LCR* se halla también presente a lo largo de la hebra de DNA que se localiza entre el *LCR* y el gen *β maj* (aproximadamente unas 38 kb) (Fig. 15). De forma interesante, la marca de metilación asociada con la proteína MLL2 se limita a una región de aproximadamente 400 pb después del TSS (Demers *et al*, 2007). Estos datos sugieren una comunicación a larga distancia entre *LCR* y

el gen de β -globina activo mediado por la proteína MLL2 a lo largo del locus, o sea, por el mecanismo de vinculación (Demers *et al*, 2007). Este mecanismo parecería contradecir el modelo de asas propuesto anteriormente para la regulación del dominio β globina de mamíferos. Sin embargo, es posible que el mecanismo mediado por el complejo MLL2 sea marcar diferencialmente la zona del promotor β^{maj} y no del resto de los genes. De esta forma garantizaría los contactos entre el LCR y el promotor únicamente en la etapa donde se expresa el gen adulto. En consecuencia, ambos modelos podrían no ser excluyentes sino necesarios para mantener la correcta expresión diferencial de los genes que codifican para las cadenas globina.

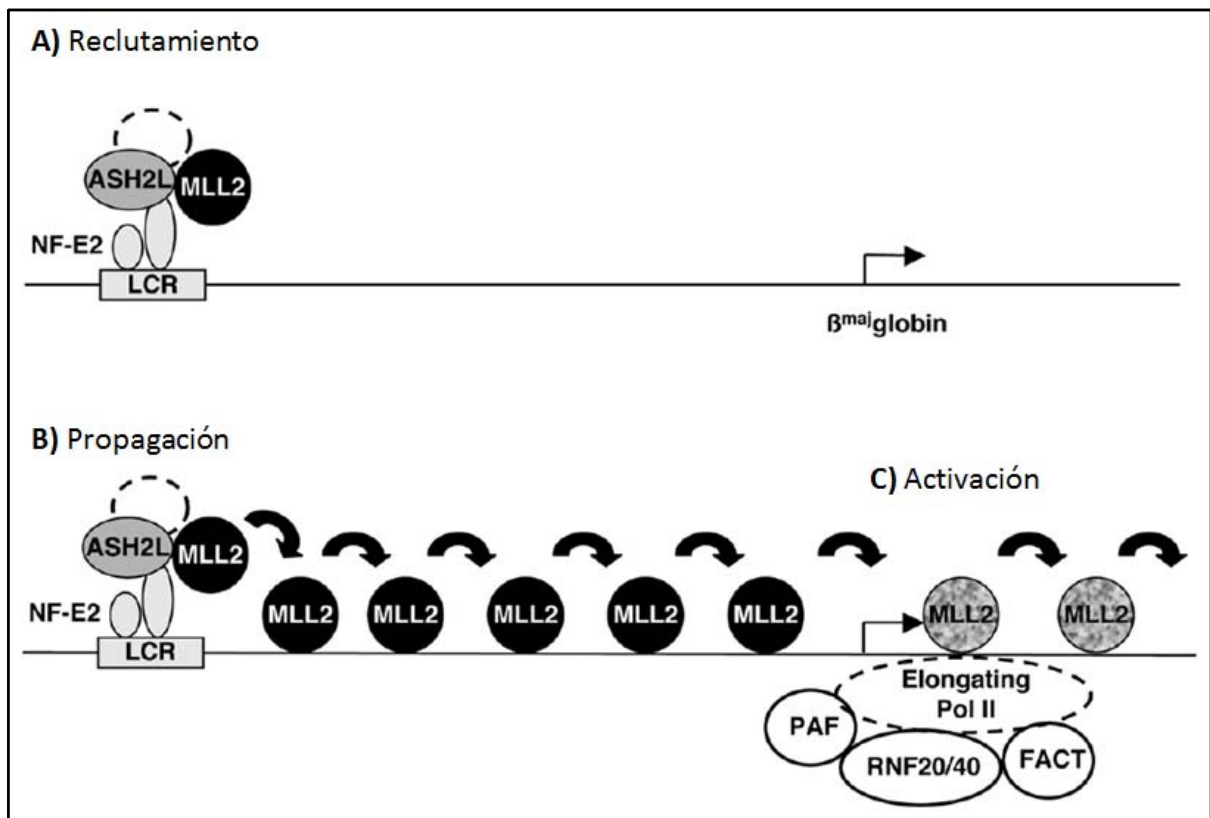


Figura 15. Representación gráfica de la comunicación entre el LCR β -globina murino y su gen blanco a través del complejo MLL2. **A)** Los factores transcripcionales son reclutados al LCR. **B)** Cuando se expresa el gen adulto, la proteína MLL2 avanza desde el LCR hasta alcanzar el TSS. **C)** En la región del promotor el complejo MLL2 es activo y se asocia con la marca H3K4me3 en las histonas (Tomado de Demers *et al*, 2007).

IV.1.3 Combinación entre los modelos de asas y encarrilamiento.

La combinación de distintas formas de comunicación a distancia entre elementos reguladores ha sido descrita en la regulación del gen que codifica para el antígeno prostático específico (PSA). La activación de este gen está mediada por el reclutamiento del receptor de andrógenos (AR) y co-activadores que se unen a sus regiones reguladoras (Wang *et al*, 2005). Para determinar la dinámica del reclutamiento del complejo co-activador AP a las regiones reguladoras del gen PSA se hicieron ChIPs acoplados a la amplificación del DNA en tiempo real (qPCR) usando anticuerpos que reconocen específicamente a las proteínas AR, SRC1, GRIP1, AIB1, p300, TRAP220, CARM1, TBP y RNA pol II. Los niveles de enriquecimiento detectados para todos estos factores es mayor en la región del *enhancer* que sobre el promotor pero la RNA pol II además es detectable ente ambos elementos de regulación. La RNA pol II es reclutada primeramente al *enhancer*, luego a la región intermedia y finalmente al promotor, lo que sugiere que la RNA pol II avanza desde el *enhancer* hasta llegar al promotor. La fosforilación de la RNA pol II es importante para que avance desde un elemento de regulación al otro. Para demostrar el modelo de encarrilamiento, los autores generaron una construcción en presencia del *insulator*/delimitador del dominio β -globina de pollo (FII) entre el promotor y *enhancer* del gen PSA y observaron una disminución del gen reportero luciferasa. Adicionalmente midieron la frecuencia de interacción entre ambos elementos de regulación para confirmar su interacción y favorecer el modelo de asas. En conjunto, los datos soportan la combinación de los modelos de asas y encarrilamiento, donde la formación de una asa entre el promotor y el *enhancer* situado 4 kb río arriba del TSS del gen PSA le permite a la RNA pol II avanzar desde el *enhancer* hacia el promotor (Wang *et al*, 2005).

IV.2 Modelos de procesamiento de la información de los *enhancers*

Los modelos descritos anteriormente explican la manera en cómo los *enhancers* se comunican con su promotor blanco para aumentar su transcripción, pero hay otros modelos que analizan de qué forma los *enhancers* procesan la información que reciben de estímulos internos o externos para conducir la expresión de genes específicos. En relación con el procesamiento de la información que llevan a cabo los *enhancers*, se propone un modelo altamente cooperativo

y coordinado conocido como *enhanceosoma* y otro más flexible denominado *billboard* (Arnosti y Kulkarni, 2005).

IV.2.1 Modelo del *enhanceosoma*

El modelo del *enhanceosoma* propone que los contactos entre proteínas son indispensables para la acción del *enhancer* y cualquier cambio en la estereo-especificidad del complejo formado suprime su función (Arnosti y Kulkarni, 2005).

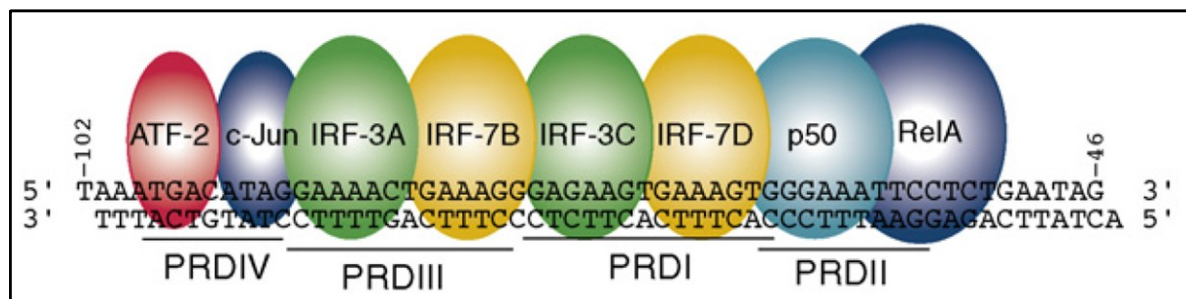


Figura 16. Representación gráfica de las proteínas que se unen al *enhanceosoma* del gen de *interferón-β*. La secuencia de DNA está subdividida en cuatro dominios reguladores positivos (PRD). Las proteínas IRF se unen a PRDI y III, NFκB a PRDII, y ATF-2/c-Jun a PRDIV (Panne, 2008).

El ejemplo mejor estudiado de *enhanceosoma* en mamíferos es el *enhancer* del gen de interferón-β y se activa en respuesta a las infecciones virales (Yie *et al*, 1999). Este *enhancer* corresponde a una secuencia de 65 pb que se une a las proteínas de la familia NF-κB, el heterodímero c-Jun/ATF-2, y proteínas de la familia IRF (Fig. 16) (Merika y Thanos, 2001). La unión de la proteína HMGA1a provee una superficie que facilita la unión de los demás factores (Yie *et al*, 1999). El *enhancer* está delimitado por dos nucleosomas, uno de los cuales restringe el acceso a la caja TATA y el TSS del gen (Agalioti *et al*, 2000). Tras la formación del complejo nucleoproteico, hay un reclutamiento secuencial de las proteínas GCN5 y CBP/p300, lo que conduce a la acetilación del nucleosoma y su remodelación por el complejo SWI/SNF que se encarga de reposicionar el nucleosoma que oculta la caja TATA. Esto permite el acceso de la proteína TBP y la maquinaria de la RNA pol II sobre el promotor (Merika *et al*, 1998; Panne, 2008). Mutaciones puntuales que alteran o eliminan el sitio de unión de las proteínas individuales afectan por completo la actividad del *enhancer*, lo que sugiere que la estructura

formada por proteínas y DNA es fundamental en su actividad (Thanos y Maniatis, 1995). Este modo de acción es considerado como el paradigma de sinergismo transcripcional porque el complejo de proteínas formado opera como una sola unidad de regulación (Panne, 2008).

IV.2.2 Modelo billboard

El segundo modelo relacionado con el procesamiento de la información de los *enhancers* es conocido como pizarra o *billboard*. Este modelo propone una forma más flexible de unión de los factores de transcripción donde los *enhancers* no actúan como una sola unidad funcional sino como elementos separados capaces de alterar la expresión en forma independiente, en este caso, la cooperación no es una necesidad como en el caso de los *enhanceosomas* (Arnosti y Kulkarni, 2005). Por el contrario, la información que contiene el *enhancer* a través de su unión con factores de transcripción es procesada por la maquinaria de transcripción donde cada elemento afecta de forma independiente la expresión de los genes. En este modelo, no es necesaria la formación de una arquitectura estrictamente definida sobre el *enhancer* (Arnosti y Kulkarni, 2005; Panne, 2008). Un ejemplo de este tipo de organización modular es el gen *endo 16* en *Strongylocentrotus purpuratus* (erizo de mar) el cual es regulado por un elemento en *cis* de 2.3 kb que se puede subdividir en seis elementos distintos donde la actividad generada por sus combinaciones genera diferentes respuestas, tal como el modelo *billboard* lo propone (Yuh *et al*, 1998; Arnosti y Kulkarni, 2005).

La información que integran los *enhancers* podría representar diferentes mecanismos biológicos que las células han desarrollado para controlar adecuadamente la expresión de sus genes. En el caso del gen de *IFN- β* que se activa en respuesta a una infección viral, resulta muy conveniente generar una estructura nucleoproteica que actúa como una unidad porque la naturaleza de la respuesta tiene que ser inmediata. Sin embargo, aquellos *enhancers* que regulan genes relacionados con otros procesos como el desarrollo pueden requerir respuestas más flexibles, que de igual forma aseguran una respuesta transcripcional adecuada.

IV.3 Modelos de acción de los *enhancers*

Como se ha mencionado con anterioridad, los *enhancers* pueden ser definidos e identificados con base en la correlación que existe entre las marcas de cromatina y los sitios de hipersensibilidad al corte por la DNasa I a nivel genómico. Históricamente se les ha caracterizado funcionalmente mediante ensayos experimentales que miden su actividad transcripcional a través de genes reporteros. A pesar de todas estas correlaciones, aún falta entender de qué manera estos elementos de regulación pueden incrementar los niveles de transcripción de su gen blanco. Los *enhancers* fueron inicialmente caracterizados gracias a su capacidad de aumentar los niveles de transcripción de los genes que regulaban. Sin embargo, este incremento puede corresponder a un aumento en el nivel de expresión de cada célula o que una mayor proporción de células se vuelven activas a la transcripción. Para distinguir experimentalmente entre ambos modelos es necesario realizar el análisis en células individuales. En la literatura existen ejemplos experimentales que apoyan ambos tipos de fenómenos y se han propuesto los modelos binario y progresivo que intentan explicar estas diferencias (Fig. 17) (Blackwood y Kadonaga, 1998; Fiering *et al*, 2000).

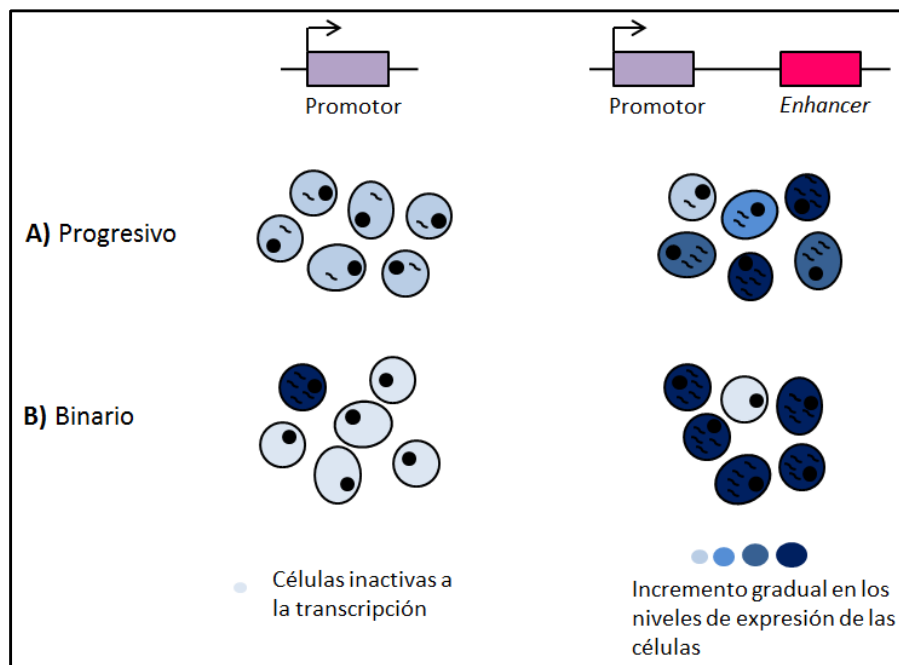


Figura 17. Modelo binario y progresivo para la activación transcripcional mediante *enhancers*.

A) En el modelo progresivo o gradual los genes son activados por los *enhancers* de manera

uniforme y la transcripción es proporcional a la fuerza del *enhancer*. **B)** En el modelo binario o de probabilidad los genes están en cualquier estado “apagado” o “encendido”. En este caso, los *enhancers* incrementan la probabilidad de que sus genes sean transcritos, pero no afectan los niveles de transcripción. La fracción de células cuyos genes están activados reflejan la fuerza del *enhancer*, que está en función del tipo y número asociado de factores de transcripción (Obtenido y modificado de Fiering *et al*, 2000).

IV.3.1 Modelo progresivo/gradual de acción de los *enhancers*

El modelo progresivo/gradual sugiere que los *enhancers* incrementan la tasa de inicio de transcripción del gen que regulan. Este modelo propone que la maquinaria asociada con la RNA pol II se posiciona con mayor frecuencia sobre el promotor, en consecuencia hay un incremento en el número de moléculas de RNA sin aumentar el número de células activas a la transcripción (Sandaltzopoulos y Becker, 1998). Para abordar si el incremento en la actividad transcripcional se ajusta al modelo binario o progresivo, el grupo de Guy *et al*, obtuvo ratones transgénicos que tienen *lacZ* como gen reportero ligado al promotor y *LCR* del gen β -globina de ratón. El análisis se realizó célula por célula mediante tinciones *in situ* en células hepáticas fetales usando el reactivo X-gal que genera células con una coloración azul (Guy *et al*, 1997). A simple vista se observa una gran variación en la intensidad de la tinción de las células individuales pero los resultados de densitometría confirman que las células provenientes del mismo ratón transgénico no contienen las mismas cantidades de proteína β -galactosidasa en un mismo momento (por ejemplo, a las 2 horas ver flechas en la Fig. 18). Al aumentar el tiempo de incubación se observa un aumento en la intensidad de la tinción que varía entre células (Fig. 18). Estos datos muestran un gradiente en la actividad de la enzima β -galactosidasa en cada célula y sugieren que en este caso el *LCR* funciona de acuerdo al modelo gradual (Guy *et al*, 1997). Otro ejemplo que se ajusta al modelo progresivo se reportó mediante el uso del promotor y *enhancer* del virus de tumor mamario en ratón (MTV) que se induce por hormonas glucocorticoides en células provenientes de tejido conectivo de ratón (Ltk-) (Ko *et al*, 1990). La expresión del gen reportero β -galactosidasa en las células transfectadas bajo la influencia de glucocorticoides muestra que la actividad enzimática en las células individuales es muy heterogénea entre células (Ko *et al*, 1990).

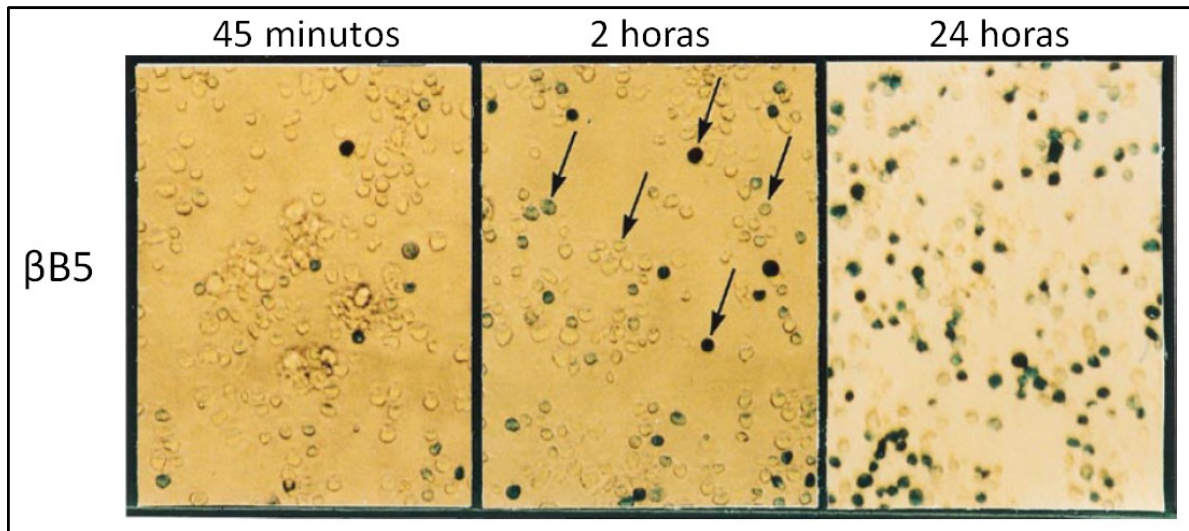


Figura 18. Cinética de aparición de células teñidas *in situ* con X-gal. Se muestran células de hígado fetal de la línea transgénica $\beta B5$. Las células se incubaron con X-gal y se fotografiaron en diferentes momentos (45 minutos, 2 horas y 24 horas). Las células individuales muestran niveles variables en la intensidad de la tinción (tomado de Guy *et al*, 1997).

IV.3.2 Modelo binario/probabilidad de los enhancers

El modelo binario o de probabilidad propone que los *enhancers* no afectan la tasa transcripcional de su gen blanco, sino que incrementan la proporción de células activas a la transcripción. Este modelo supone que el gen (es) en estudio puede (n) estar en un estado "activo" o "inactivo", pero la presencia de un *enhancer* favorece su actividad transcripcional (Fig. 17) (Walters *et al*, 1995; Sutherland *et al*, 1997; 1999; Blackwood y Kadonaga, 1998; Fiering *et al*, 2000).

El dominio α -globina de mamíferos contiene genes que están dispuestos en el orden en el que se expresan durante el desarrollo. En ratones transgénicos, estos genes necesitan de un sitio hipersensible al corte por la DNasa I ($\alpha HS-40$) que les confiere un alto nivel de expresión. Para examinar el efecto del promotor α -globina y el *enhancer* $\alpha HS-40$ sobre la expresión de *lacZ*, se generaron ratones transgénicos con estos elementos de regulación. La tinción histoquímica para β -galactosidasa en eritrocitos revela que un número extremadamente pequeño de eritrocitos embrionarios expresan al gen reportero en las líneas que llevan la construcción que contiene sólo al promotor del gen α -globina (αP). En contraste, las líneas que portan la

construcción que contiene el promotor del gen α -globina y el *enhancer* α HS-40 (α PE) muestran que el porcentaje de eritrocitos que expresan β -Gal en la etapa embrionaria es alto. El nivel de expresión en estos eritrocitos no es mayor que en los eritrocitos de las líneas que sólo llevan el promotor α , con lo cual se favorece el modelo binario (Fig. 19) (Sutherland *et al*, 1997). También se ha observado que el *enhancer* SV40 funciona de acuerdo al modelo binario en células de mamífero (Walters *et al*, 1995).

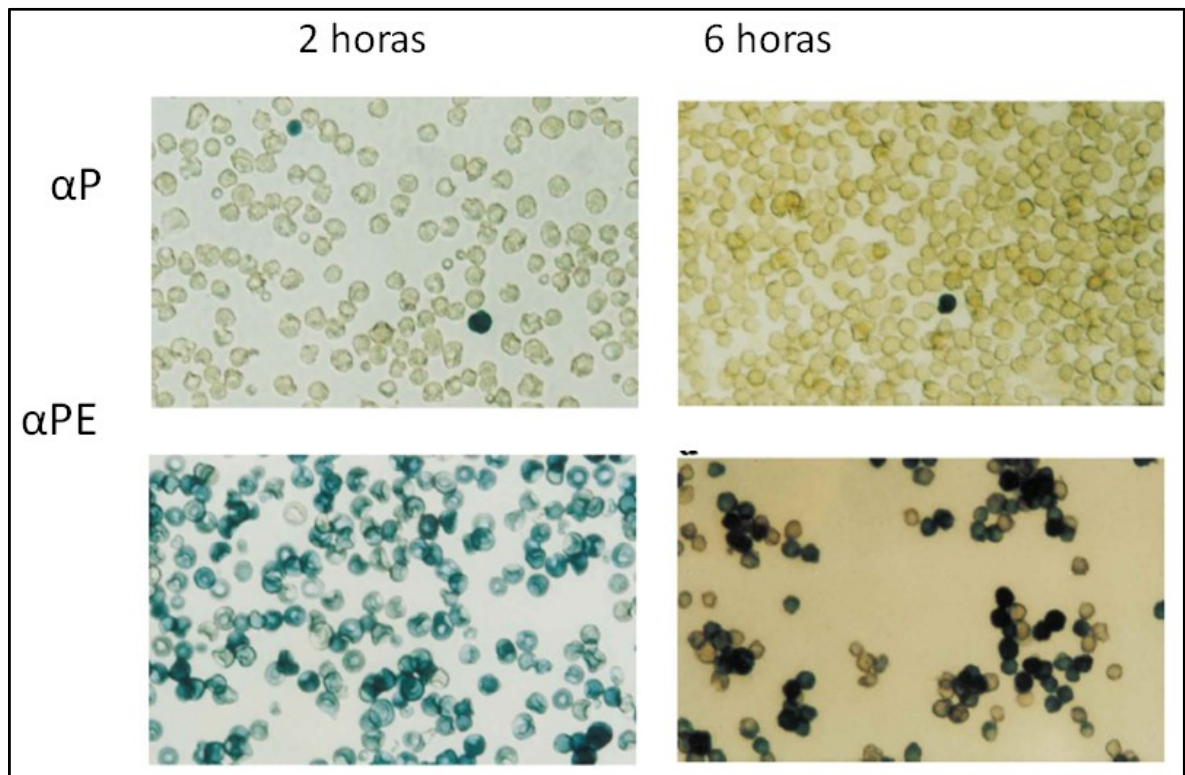


Figura 19. Tinción con X-gal de eritrocitos de ratones transgénicos (11,5 dpc) en presencia del promotor α P y promotor-*enhancer* α PE). Las células se fotografiaron después de incubaciones cronometradas de 2 y 6 horas. En presencia del promotor se observa que aproximadamente el 1% de las células expresan β -Gal. Después de 6 horas de incubación la línea transgénica en presencia del promotor-*enhancer* α muestra que el 98% de las células expresan la enzima β -Gal de acuerdo al modelo binario (Tomado de Sutherland *et al*, 1997).

IV.3.3 Implicaciones de los modelos progresivo y binario de acción de los *enhancers*

La visión más popular relacionada con la actividad de los *enhancers* es que los factores transcripcionales que se unen a ellos generan interacciones con otras proteínas que finalmente incrementan los niveles de expresión de su gen blanco alterando su tasa

transcripcional. Esta visión no toma en cuenta que el DNA está altamente estructurado en forma de cromatina al interior del núcleo y como se ha mencionado anteriormente es una estructura dinámica donde hay zonas de mayor o menor accesibilidad (Kress *et al*, 2010; Qureshi y Mehler, 2010). Es posible que los *enhancers* tengan la propiedad de evitar el silenciamiento de su gen blanco mediante su relocalización hacia zonas transcripcionalmente activas (Sutherland *et al*, 2007; Fiering *et al*, 2000; Chakalova *et al*, 2005). Lo anterior implicaría que la principal función de los *enhancers* sería asegurar la expresión génica en los linajes apropiados en vez de incrementar los niveles de transcripción (Fiering *et al*, 2000).

IV.4 Super-enhancers

El complejo Mediador facilita el reclutamiento de la RNA pol II a los *enhancers* vía los factores de transcripción (Kornberg, 2005; Malik y Roeder, 2010) y es un elemento esencial para el mantenimiento de las células troncales y su diferenciación (Ito *et al*, 2000; Kagey *et al*, 2010; Risley *et al*, 2010). Con base en la cantidad de complejo Mediador unido al genoma, recientemente se identificaron grandes regiones que abarcan hasta 50 kb y muestran un gran enriquecimiento de este complejo denominados *super-enhancers*. Los *super-enhancers*, se asocian fuertemente a la RNA pol II, H3K27ac y están enriquecidos en sitios de unión a factores transcripcionales altamente específicos de tejido pero no con genes constitutivos (Whyte *et al*, 2013; Pott y Lieb, 2015). De forma interesante, el 93% de los pares promotor-*super-enhancers* identificados en base a su frecuencia de interacción se produce dentro de un mismo dominio topológico (TAD) definido por Hi-C (Whyte *et al*, 2013). Teniendo en cuenta los altos niveles de complejo Mediador sobre los *super-enhancers*, estos podrían funcionar como centros de conectividad genómica donde las redes reguladoras apropiadas se organizan estructuralmente en fábricas transcripcionales para coordinar la expresión de los genes (Mousavi *et al*, 2014). Una característica fundamental asociada con el complejo Mediador es que su composición es dinámica. En los mamíferos, el complejo Mediador consta de 26 subunidades (Allen y Taadjes, 2015). La subunidad CDK8 por ejemplo se asocia de forma reversible con el Mediador; cuando la subunidad CDK8 forma parte del complejo cambia su estructura y evita su interacción con la RNA pol II. Alternativamente, las subunidades del

mediador parecen tener roles distintivos en diferentes tipos de células (Balamotis *et al*, 2009). Además, el complejo Mediador interactúa con proteínas estructurales tales como las cohesinas de manera tejida específica (Kagey *et al*, 2010). A pesar de estas evidencias, aún es necesario entender en detalle los posibles mecanismos regulados por complejo Mediador relacionados con el establecimiento y/o mantenimiento del destino celular.

IV.5 RNA pol II y transcripción de *enhancers*

Los *enhancers* activos, además de las marcas de cromatina características con las que se asocian están ocupados por la RNA pol II (De Santa *et al*, 2010). Inicialmente se pensó que podría ser el resultado de la proximidad física con los promotores a los que regulaban, sin embargo, el análisis detallado de los *enhancers* reveló que también son sitios de transcripción (Mousavi *et al*, 2014). Recientemente se ha reportado la transcripción asociada con elementos *enhancer* en diversos contextos celulares como cultivos neuronales primarios, células miogénicas, macrófagos de ratón y células de cáncer de mama (Kim *et al*, 2010; Mousavi *et al*, 2013; Lam *et al*, 2013; Li *et al*, 2013). Se ha propuesto que en células miogénicas los *enhancer* RNAs (eRNAs) regulan la accesibilidad de la maquinaria transcripcional del locus que codifica para reguladores maestros de la miogénesis como MyoD y Miogenina, lo que implica un impacto en la programación celular (Mousavi *et al*, 2013). Utilizando cultivos primarios de neuronas en un sistema *in vitro* bajo estímulos de cloruro de potasio para despolarizar membranas, se identificaron aproximadamente ~12,000 *enhancers* con base en su capacidad de ser transcritos. Estos elementos de regulación, además presentan un claro incremento en la unión de la acetiltransferasa CBP (Kim *et al*, 2010). Bajo estas condiciones, se observó que la asociación de la RNA pol II a los *enhancers* coincide con la síntesis bidireccional de transcritos de RNA de tamaño menor a 2 kb (eRNAs), los cuales tienen Cap en su extremo 5' pero carecen de señales de poli-adenilación en su extremo 3'. La presencia de eRNAs correlaciona fuertemente con la expresión de mRNAs de genes cercanos y con la adquisición de la marca de histonas H3K27ac pero en ausencia de la marca represiva H3K27me3 (Djebali *et al*, 2012; Kaikkonen *et al*, 2013).

De forma consistente, el uso de RNAi específicos que abaten los niveles de eRNAs también reducen los niveles de expresión de los genes codificantes asociados (Kim C T *et al*, 2010; Lam *et al*, 2013). En células de cáncer de mama la unión del receptor de estrógenos α (ER- α) a los *enhancers* genera un incremento global en los niveles de transcripción de los eRNAs (Li *et al*, 2013). El tratamiento de estas células con estradiol induce un incremento en la unión de cohesinas (RAD21 y SMC3) sobre los *enhancers*. Estas evidencias han permitido proponer que una posible función asociada con los eRNAs es la estabilización de asas cromatínicas que favorecen la expresión de genes (Li *et al*, 2013).

El examen detallado de los RNAs nacientes ha permitido identificar una arquitectura común en regiones genómicas donde inicia la síntesis de RNAs. Esta arquitectura se caracteriza por la formación de un *NDR* y la simetría de dos nucleosomas que delimitan pares de transcripción divergentes presentes en promotores y *enhancers*. Adicionalmente, éstos análisis sugieren que las modificaciones post-traduccionales en las histonas solamente reflejan diferencias en los niveles de transcripción pero no pueden diferenciar entre elementos de regulación como promotores y *enhancers* (Core *et al*, 2014). En relación con esta idea, recientemente se ha hallado un grupo de elementos reguladores en *cis* con rasgos dinámicos (*cREDS*) que poseen actividad de promotor y *enhancer* cuando su actividad se mide usando genes reporteros. Los *cREDS* poseen modificaciones post-traduccionales en las histonas propias de *enhancers* activos como H3K27ac y H3K4me1 en un tejido particular, pero en otros tejidos presentan H3K4me3 en vez de H3K4me1 (Leung *et al*, 2015). Sin duda el trabajo futuro ahondará en la naturaleza de estas evidencias pero es posible que los *enhancers* sean un tipo de elementos de regulación con funciones y mecanismos diversos.

IV.6 RNAs que actúan como potenciadores de la transcripción

La mayor parte (70-80%) del genoma de los mamíferos tiene el potencial de ser transcrito en RNAs no codificantes (*ncRNAs*) (De Santa *et al*, 2010; Natoli y Andrau, 2012; Mousavi *et al*, 2014). Datos experimentales recientes apoyan el hecho de que ciertos RNAs intergénicos no codificantes largos tienen funciones importantes en la regulación transcripcional (Khalil *et al*, 2009). Los *lncRNAs* son transcritos de más de 200 nucleótidos que en la mayoría de los casos

reflejan las mismas características en la cromatina que los genes codificantes pero sin contener un marco de lectura abierto funcional (ORF) (Orom *et al*, 2010). Este tipo de RNAs largos han sido implicados en procesos celulares como la inactivación del cromosoma X (Heard y Disteche, 2006). Usando como modelo de estudio queratinocitos humanos primarios, se observó la inducción de RNAs no codificantes largos que responden a las señales de diferenciación celular (Orom *et al*, 2010). Este tipo de transcritos se asocian con las marcas en la histona H3K4me3 y en la H3K36me3, su transcripción procede en una sola dirección, se poli-adenilan, muestran conservación en su secuencia y actúan como *enhancers* de los genes vecinos estimulando su transcripción (Orom *et al*, 2010). Al abatir sus niveles mediante RNAi, los niveles de expresión de sus genes blanco disminuyen sustancialmente. Este tipo de transcritos actúan estimulando la transcripción y se comportan de manera similar a como lo hacen los *enhancers* en *cis* localizados en la secuencia del DNA. En cuanto al posible mecanismo molecular que estimula la expresión génica a través de *lncRNAs*, se ha observado que este tipo de transcritos interactúa con el complejo Mediador favoreciendo las interacciones de la cromatina de largo alcance entre dos dominios distales (Lai *et al*, 2013). Alternativamente, los *lncRNAs* pueden interaccionar con los complejos de remodelación de la cromatina y dirigirlos a regiones específicas del genoma contribuyendo al proceso de transcripción (Dinger *et al*, 2008). Sin embargo, estos RNAs podrían tener otras funciones muy distintas en otros contextos y es posible que aún no hayan sido establecidas.

Los datos experimentales muestran la existencia de diferentes tipos de RNAs asociados a los *enhancers*. El mecanismo propuesto con los eRNAs mediante su asociación con cohesinas permite estabilizar el complejo de inicio de la transcripción y con ello, favorecer la expresión de genes. Una alternativa a este modelo surge de los RNAs largos no codificantes que a través de su estructura secundaria podrían funcionar como factores de andamiaje que permiten el reclutamiento de otros complejos remodeladores de cromatina necesarios para la transcripción (Orom *et al*, 2010). El RNA puede ser en un futuro próximo la pista que nos lleve a un mejor entendimiento del comportamiento tan diverso y flexible que presentan los *enhancers*.

En resumen, todos los modelos de actividad de los *enhancers* discutidos previamente, muestran nuestro interés por comprender la manera en que regulan la expresión de los genes. Sin embargo, también son el resultado de nuestra incapacidad para comprender si existe un mecanismo de acción general para los *enhancers* o bien pueden ser considerados de manera independiente y específica. En consecuencia, sigue siendo necesario realizar esfuerzos que nos permitan comprender los mecanismos moleculares del funcionamiento de los *enhancers*.

V. Modelo de estudio y antecedentes particulares

Desde hace algunas décadas, diversos grupos de investigación, incluyendo el nuestro, se han interesado en describir y caracterizar los dominios génicos α - y β -globina de pollo. El principal objetivo ha sido entender cómo la cromatina y proteínas asociadas a su remodelamiento dirigen la expresión correcta y regulada de los genes globina durante la diferenciación eritroide y el desarrollo. La enorme cantidad de información generada acerca de estos dominios los sitúa como un modelo paradigmático de estudio que contribuye no solo al entendimiento de la regulación de los genes sino al planteamiento de nuevos modelos.

V. 1 Dominio multigénico α -globina de pollo

El dominio α -globina de pollo comprende una región genómica de alrededor de 50 kb localizada en el minicromosoma 14 compuesta de un gen embrionario π y dos adultos α^D y α^A (Valdes-Quezada *et al*, 2013). A lo largo del dominio se han identificado varios sitios de hipersensibilidad a la DNasa I que coinciden con elementos reguladores que permiten expresar los genes en forma diferencial durante el desarrollo y la diferenciación eritroide (Fig. 20) (Furlán-Magaril *et al*, 2011).

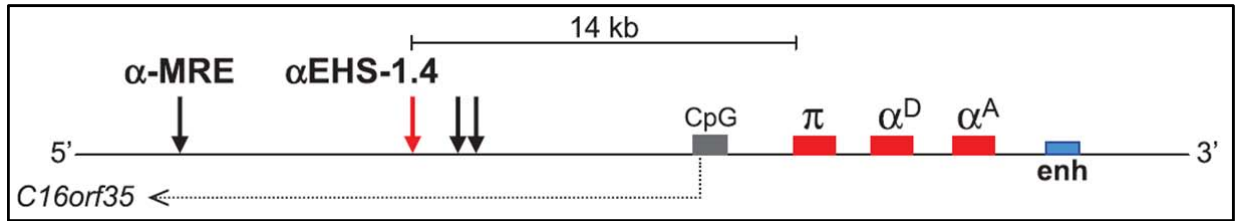


Figura 20. Elementos reguladores y estructurales de la cromatina del dominio α -globina de pollo. Las flechas verticales localizadas en el costado 5' del dominio representan sitios de hipersensibilidad a la DNasa I (DHS) que correlacionan con el elemento α -MRE y el *insulator* α EHS1.4. Se muestra el gen embrionario π , y los dos genes adultos α^D y α^A . Aproximadamente 1 kb río abajo del gen α^A se encuentra el *enhancer* 3' α -globina del dominio. También se muestra el gen *C16orf35* que se transcribe en sentido opuesto a los genes globina (Furlán Magaril *et al*, 2011).

V.2 Factores transcripcionales relacionados con la expresión de los genes globina

En un intento por entender cómo se establecen los programas transcripcionales que permiten a las células troncales comprometerse hacia un linaje celular y diferenciarse, muchos grupos de investigación han centrado sus esfuerzos por entender qué factores transcripcionales determinan la expresión de los genes globina. En seguida, se describen algunos factores de transcripción que permiten la expresión de los genes globina en diferentes organismos. Utilizando células de ratón que representan todas las etapas de la diferenciación eritroide se encontró que la expresión de los genes globina comienza en los progenitores hematopoyéticos multipotentes donde el factor transcripcional GATA-2 interacciona con regiones regulatorias del dominio α -globina (Anguita *et al*, 2004). En etapas subsecuentes, el factor de transcripción GATA-2 es sustituido por GATA-1, que a su vez atrae otras proteínas incluyendo SCL, LMO2, E2A, Ldb-1 y NF-E2 (Anguita *et al*, 2004). En conjunto, los complejos multi-proteicos inducen cambios en la estructura de la cromatina aumentando la acetilación y la metilación en las histonas a escala del dominio. El último evento en la activación implica el reclutamiento de la RNA pol II hacia los promotores de los genes y comienza su transcripción (Anguita *et al*, 2004). Se ha observado que el factor transcripcional GATA-1 requiere de la interacción del co-factor eritroide específico FOG-1 (*Friend of GATA*) para llevar a cabo sus funciones (Tsang *et al*, 1997; Hong *et al*, 2005; Miccio y Blobel, 2010). La proteína FOG-1 puede interactuar con TACC3, CtBP-2 y NuRD por lo que puede activar o reprimir la expresión

de sus genes blanco (Miccio y Blobel, 2010). Su asociación con NuRD se requiere para la completa expresión de los genes β -globina durante la eritropoyesis definitiva (Miccio y Blobel, 2010). Se ha demostrado que el factor transcripcional GATA-1 puede ser acetilado por CBP/p300 incrementando su afinidad por el DNA (Huo y Zhang, 2005). GATA-1 además puede interactuar con el factor PU.1 durante la diferenciación eritroide (Cantor y Orkin, 2005). En nuestro grupo de investigación se demostró que la proteína GATA-1 interacciona con el factor de transcripción bivalente YY1 y modula el funcionamiento del *enhancer* situado en el extremo 3' del dominio α -globina de pollo (Rincón-Arano *et al*, 2005). El factor EKLF (factor eritroide similar a krüppel) tiene efectos sobre la organización de la cromatina durante la diferenciación eritroide gracias al reclutamiento de acetiltransferasas de histonas y es necesario para la formación de un dominio activo de cromatina (Drissen *et al*, 2004; Patrinos *et al*, 2004). BRG1, la subunidad catalítica del complejo SWI/SNF tiene la capacidad de formar complejos proteicos con los factores transcripcionales GATA-1 y EKLF. En células de hígado fetal murino, el complejo formado por BRG1, GATA-1 y EKLF es necesario para la formación de una asa de cromatina entre el elemento α -MRE y el promotor $\alpha 2$ (Kim *et al*, 2009). NF-E2 es un heterodímero formado por la subunidad ubicua NF-E2p18 y la subunidad hematopoyética específica NF-E2p45 (Anguita *et al*, 2004). Este factor se une a la secuencia consenso TGCTGA(G/C)TCA que se encuentra presente en elementos reguladores eritroides específicos. Su función es activar mediante complejos remodeladores de cromatina y participa en la transferencia de la RNA pol II a los promotores de los genes globina (Huo y Zhang, 2005). LDB1 es una proteína nuclear de expresión ubicua que no se une al DNA pero forma parte de un complejo proteico en células eritroides de ratón con las proteínas Lmo2, TAL1, GATA1 y E2A (Meier *et al*, 2006). En el dominio β de mamíferos se ha demostrado que Ldb1/NLI es responsable de regular las interacciones de larga distancia entre el *LCR* y su promotor blanco (Fig. 14; Song *et al*, 2007). El factor multifuncional CTCF también está directamente involucrado en la formación de asas de cromatina en el dominio α - y β -globina (Splinter *et al*, 2006; Valdes-Quezada *et al*, 2013). El factor de transcripción RUNX1 se une a la secuencia consenso (T/C)G(T/C)GGT(T/C) (van Riel *et al*, 2012) y puede formar heterodímeros con la proteína CBF β incrementando su afinidad por el DNA (Sasaki *et al*, 1996; van Riel *et al*, 2012).

El factor RUNX1 es parte del complejo proteico formado por LDB1, GATA1, TAL1, y EtO2 (Meier *et al*, 2006) y tiene un papel esencial en las células troncales hematopoyéticas porque actúa como un represor de genes eritroides a través de su unión con la proteína represora MYEF2 (van Riel *et al*, 2012). En resumen, la transcripción de los genes globina durante el desarrollo y la diferenciación eritroide se regula a través de la interrelación entre elementos reguladores presentes en el DNA y una combinatoria de factores transcripcionales generales y eritroide específicos.

V.3 Elementos reguladores asociados al dominio α -globina de pollo

En la región 5' no codificante del dominio α -globina de pollo se han identificado varios sitios hipersensibles al corte por la DNasa I que incluyen un *LCR* conocido como elemento de regulación principal o α -MRE (Vyas *et al*, 1995). Aproximadamente 14 kb río abajo del α -MRE se localiza el elemento α EHS-1.4 que corresponde a un *insulator* cuya actividad es dependiente de CTCF. Este elemento se ubica dentro de un intrón del gen *C16orf35* que se transcribe en sentido opuesto a los genes globina (Furlán-Magaril *et al*, 2011). Posteriormente se localiza una isla CpG que se encuentra a ~ 4 kb río arriba del gen π . Hacia el extremo 3' del dominio también se ha identificado un elemento regulador eritroide específico situado a una distancia de aproximadamente 1 kb del gen α^A y que interactúa con los factores transcripcionales YY1 y GATA-1 (Fig. 21) (Rincón-Arano *et al*, 2005; Escamilla-Del-Arenal y Recillas Targa, 2008). Recientemente nuestro grupo ha reportado que el gen embrionario π y el *enhancer* distal 3' están en contacto físico en etapas tempranas del desarrollo. En etapas posteriores estos contactos se pierden, el gen embrionario π se silencia por metilación del DNA y posiblemente la proteína CTCF regula las interacciones de larga distancia entre elementos reguladores del dominio (Valdes-Quezada *et al*, 2013).

La comparación de las secuencias de DNA entre pato y pollo del extremo 3' del dominio α -globina, permitió localizar inicialmente 4 sitios de reconocimiento para el factor eritroide específico GATA-1 (Recillas-Targa *et al*, 1993). Mediante transfecciones transitorias en eritroblastos de pollo fue posible identificar dos elementos de regulación adyacentes con funciones opuestas. Los elementos denominados silenciador y *enhancer* se localizan

aproximadamente a 450 pb río abajo del gen adulto α^A (Fig. 21) (Recillas Targa *et al*, 1993). El silenciador posee al menos tres sitios de unión para factores nucleares pero su identidad aún se desconoce. Estos sitios se definieron en ensayos de protección a la DNasa I (*Footprinting*) y se denominan SF1, SF2 y SF3 (Recillas Targa *et al*, 1993). El *enhancer* posee sitios de unión para los factores eritroides YY1 y GATA-1, entre otros.

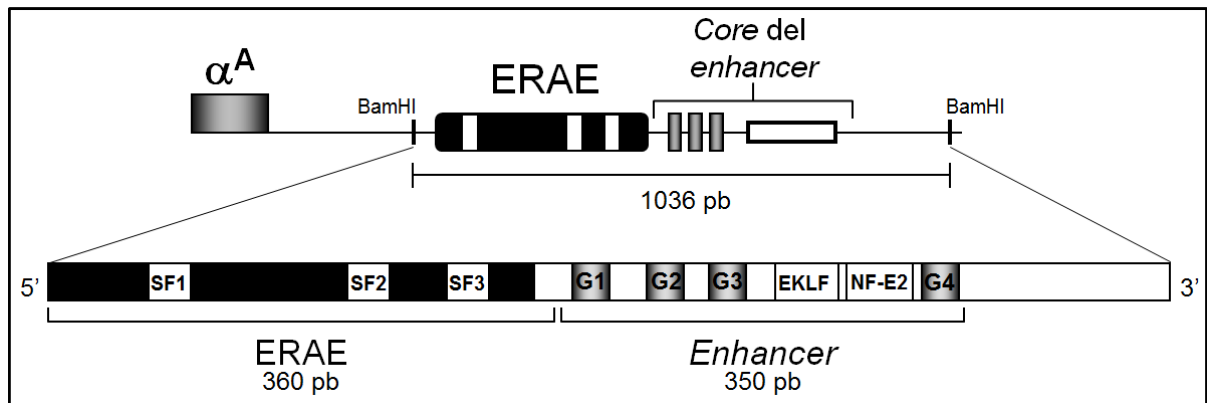


Figura 21. Esquema de la región reguladora 3' α -globina de pollo. Se muestra el elemento ERAE, inicialmente llamado “silenciador” y el *core* del *enhancer*, localizados río abajo del gen α^A -globina de pollo. En mayor aumento se muestra la localización detallada de las secuencias de unión a factores nucleares delimitada por sitios de restricción BamHI. En ensayos de protección a la DNasa I se determinó que el ERAE posee tres sitios de unión a factores nucleares denominados SF1, SF2 y SF3 (Recillas Targa *et al*, 1993). Sobre el *enhancer* se identificaron sitios de unión para los factores transcripcionales EKLf, NF-E2 y cuatro sitios de unión al factor eritroide GATA-1 (representado en la imagen por G1, G2, G3 y G4) (Knezetic y Felsenfeld, 1989; Rincón-Arango *et al*, 2005).

V.4 Organización de la cromatina de los elementos reguladores ERAE-*enhancer* del dominio α -globina de pollo

Para entender mejor la función de los elementos reguladores del dominio 3' α -globina de pollo, se estudió la organización de la cromatina asociada a esta región genómica durante el desarrollo del pollo y la diferenciación eritroide. Los ensayos de accesibilidad de la cromatina al corte por la DNasa I (no se muestran) y el mapeo del posicionamiento nucleosomal revelaron la aparición progresiva de una conformación eritroide específica en el *core* del *enhancer* (Fig. 22). Esta estructura se caracteriza por la presencia de dos sitios hipersensibles al corte por la MNasa (flechas verticales) y una baja densidad nucleosomal o *NDR*, delimitada

por dos nucleosomas (Fig. 21, cilindros punteados). Uno de los nucleosomas se localiza en la región del silenciador (ERAE) y el otro en la región regulatoria de 120 pb. La estructura de la cromatina se remodela gradualmente alcanzando su máxima accesibilidad cuando el *enhancer* es completamente activo en eritrocitos circulantes de 10 días (10dRBC) (Fig. 22).

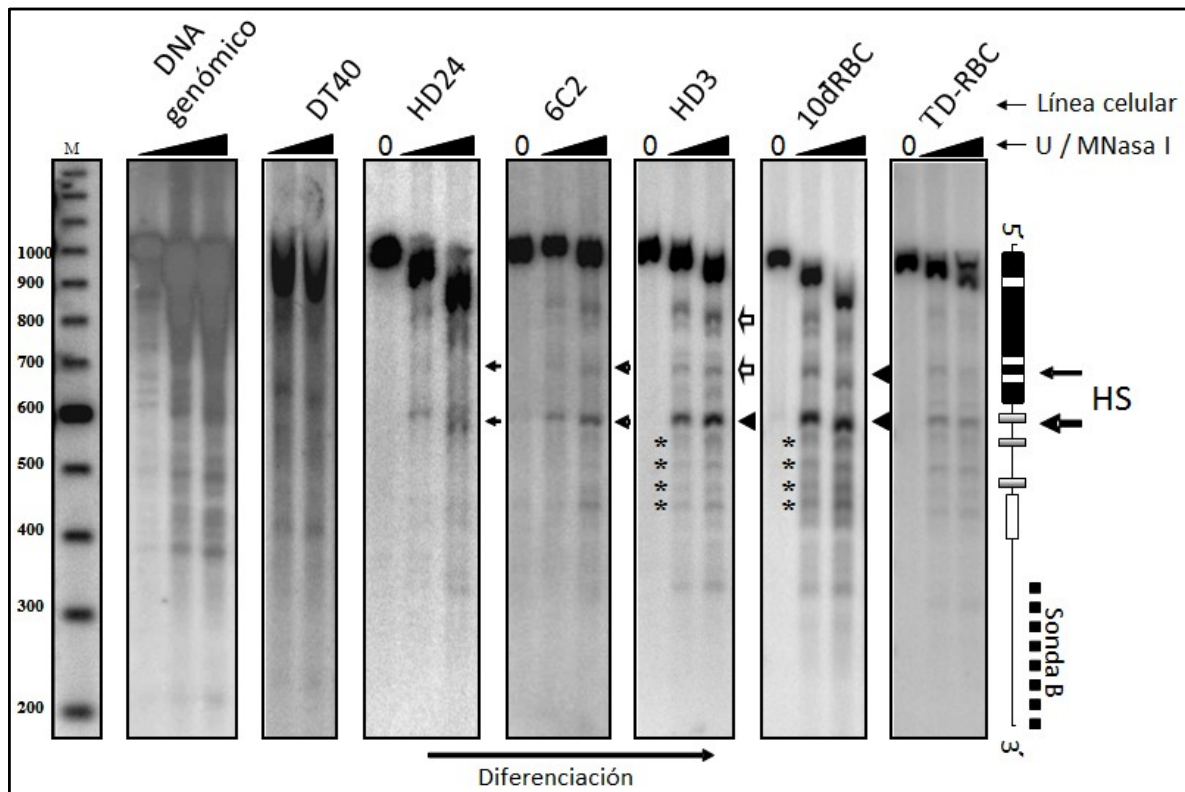


Figura 22. Posicionamiento de nucleosomas determinado por digestión enzimática con la endonucleasa MNasa durante la diferenciación eritroide. Se muestran los patrones de digestión del DNA en distintas etapas de la diferenciación eritroide, desde las células menos diferenciadas (HD24 y 6C2), hasta las más diferenciadas (HD3, eritrocitos circulantes de 10 y 14 días). Los controles negativos incluyen DNA genómico y la línea celular linfóide DT-40. Desde las primeras etapas de la diferenciación eritroide hay dos sitios de hipersensibilidad a la MNasa que aumentan su intensidad hasta el nivel más alto en las células circulantes de 10 días (flechas negras). La posición de estas hipersensibilidades se localiza entre los elementos de regulación ERAE (“silenciador”) y *enhancer*. A partir de eritroblastos de pollo o células HD3 aparecen más sitios (asteriscos) de corte por la enzima sugiriendo una organización de la cromatina más abierta. Las flechas horizontales abiertas reflejan el posicionamiento constante de un nucleosoma río arriba del *enhancer*. Para mayor claridad se muestran los sitios de mayor hipersensibilidad al corte por la endonucleasa en la región ERAE-*enhancer* en una representación esquemática (lado derecho de la imagen) (Tomado de Escamilla-Del-Arenal y Recillas Targa, 2008).

Para determinar si los sitios de unión al factor GATA-1 presentes en el *core* del *enhancer* son importantes para generar la estructura de la cromatina caracterizada previamente, se usó un sistema experimental que permite comparar la actividad de distintos transgenes en el mismo sitio de integración genómica. Usando este sistema fue posible recapitular la estructura endógena de la región regulatoria 3' del dominio, en contraste, las mutaciones puntuales del primer y segundo sitio de unión al factor eritroide GATA-1 muestran cambios en la estructura de la cromatina y pérdida de la actividad transcripcional del gen reportero (Fig. 23) (Escamilla-del-Arenal y Recillas Targa, 2008).

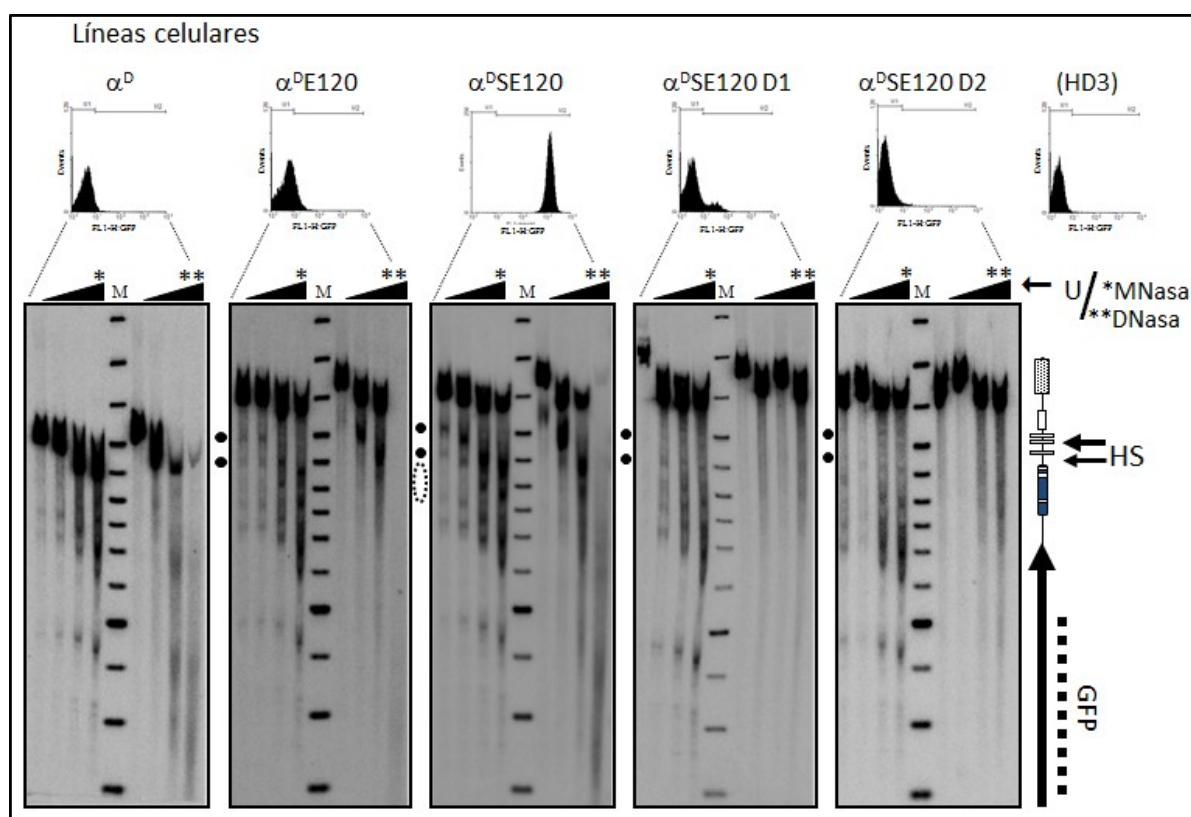


Figura 23. Posicionamiento de nucleosomas determinado por digestión enzimática con la endonucleasa MNasa y DNasa I en el mismo sitio de integración para diversos transgenes que contienen distintas secuencias regulatorias. En la parte superior se muestran los perfiles de expresión de las líneas transgénicas generadas donde sólo la construcción que contiene al ERAE-*enhancer* ($\alpha^D SE120$) expresa al gen reportero *GFP*. En la parte inferior se muestran los patrones de digestión del DNA de todos los vectores. Todas las construcciones tienen el promotor α^D , el $\alpha^D E120$ tiene además al *enhancer*, el $\alpha^D SE120$ tiene al ERAE-*enhancer* y los otros dos plásmidos tienen mutado el primer ($\alpha^D SE120 \Delta 1$) o segundo ($\alpha^D SE120 \Delta 2$) sitio de unión al factor transcripcional GATA-1. Los círculos negros muestran las posiciones de los dos

sitios de hipersensibilidad a la MNasa. El óvalo de puntos muestra el NDR (Tomado de Escamilla-Del-Arenal y Recillas Targa, 2008).

De forma inesperada, estos experimentos mostraron también que en el contexto de la cromatina la presencia del elemento llamado “silenciador” es necesaria para inducir altos niveles de transcripción del gen reportero *GFP*, así que en este trabajo nos interesamos por entender la función del elemento silenciador que renombramos “elemento río arriba del *enhancer*” o ERAE.

VI. Planteamiento del problema

Actualmente se acepta que la regulación de la expresión génica no depende únicamente de los elementos promotores sino que se complementa en gran medida con otros elementos distales donde el contexto tridimensional también juega un papel regulatorio. En este sentido, los *enhancers* tienen una función preponderante ayudando a regular la expresión específica de un gen o un grupo de genes. Aunque existe un gran número de estudios a escala global del genoma, poco se entiende acerca de los mecanismos moleculares que permiten la actividad de los *enhancers* y menos aún de la regulación de su actividad en un entorno cromatínico. Por ello, surge el interés por entender si la estructura de la cromatina es un factor determinante en la actividad asociada con el *enhancer* 3' α -globina de pollo. Actualmente contamos con una descripción detallada de la estructura de la cromatina de esta región pero se desconoce si el posicionamiento nucleosomal y la presencia del elemento ERAE influyen en su actividad.

VII. Hipótesis

La caracterización de la estructura de la cromatina en la región reguladora del dominio 3' α -globina de pollo se definió previamente por la presencia de dos sitios de hipersensibilidad a la nucleasa MNasa entre el ERAE y el *enhancer*, que además se encuentra delimitada por dos nucleosomas. La accesibilidad de la estructura de la cromatina es mayor durante la diferenciación eritroide y coincide con la actividad del elemento *enhancer*. Por lo que proponemos que la estructura de la cromatina asociada con el elemento ERAE-*enhancer* situado en el costado 3' del dominio α -globina de pollo, tiene un papel fundamental en su

capacidad para estimular la actividad transcripcional de su gen blanco durante el desarrollo y la diferenciación eritroide.

VIII. Objetivo general

Determinar si la estructura de la cromatina regula la actividad potenciadora a través del elemento ERAE-*Enhancer* localizado en el costado 3' no codificante del dominio α -globina de pollo.

IX. Objetivos particulares

IX.1 Entender el mecanismo por el cual el elemento ERAE-*enhancer* localizado en el costado 3' no codificante del dominio α -globina de pollo incrementa los niveles de transcripción y determinar si se ajusta al modelos binario y/o progresivo de acción de los *enhancers*.

IX.2 Caracterizar la función del elemento río arriba del *enhancer* (ERAE) en la regulación del *enhancer* 3' α -globina de pollo.

IX.3 Evaluar la influencia de la estructura de la cromatina en la actividad del *enhancer* situado en el extremo 3' no-codificante del dominio α -globina de pollo.

IX.4 Determinar si los cambios asociados con la estructura de la cromatina modifican la unión del factor transcripcional CTCF en la región 3' no codificante del dominio α -globina de pollo.

X. Resultados

X.1 Generación de líneas estables con los plásmidos que contienen al *enhancer*, ERAE-*enhancer* y ERAE.

Para investigar la contribución del ERAE (antes llamado silenciador) en la actividad del *enhancer* 3' del dominio α -globina de pollo se generaron líneas transgénicas estables en células de pollo llamadas HD3 o LSCHD3, las cuales están arrestadas en el estadio de eritroblasto. Las secuencias de DNA correspondientes al *enhancer* 3', ERAE-*enhancer* y ERAE

se clonaron en el plásmido pG α D3 en presencia del promotor α^D . Todas las construcciones tienen como gen reportero a la proteína verde fluorescente (*GFP*) permitiéndonos evaluar su actividad mediante citometría de flujo o FACS. Los perfiles de expresión obtenidos por FACS provenientes de poblaciones clonales pueden ser inactivas, activas o bifásicas (es decir, mezclas de células activas e inactivas) (Fig. 24).

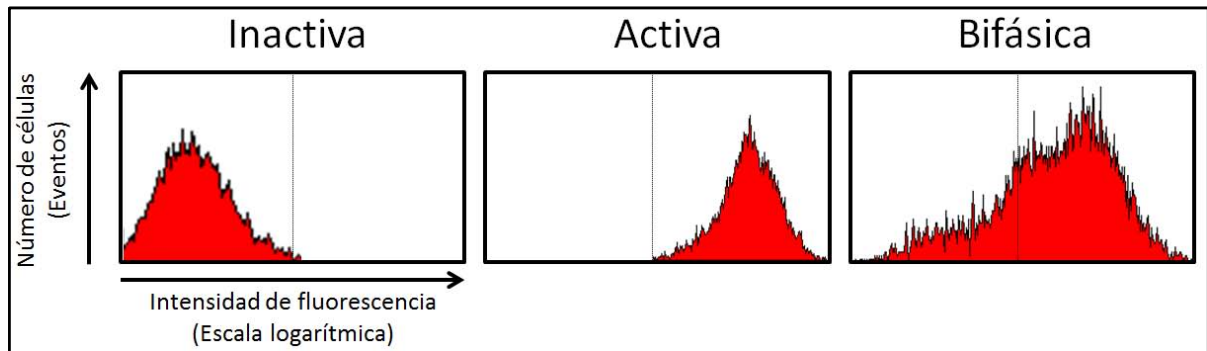


Figura 24. Perfiles de expresión obtenidos por citometría de flujo provenientes de poblaciones clonales. Se muestran tres tipos de perfiles de expresión que pueden corresponder a poblaciones activas, inactivas o bifásicas. El eje de las abscisas corresponde a la intensidad promedio de fluorescencia y su escala es logarítmica. El eje de las ordenadas corresponde al número de células o eventos analizados. Mediante el uso de ésta metodología es posible analizar dos parámetros en forma simultánea: 1) El porcentaje de células fluorescentes y 2) La intensidad media de fluorescencia. Cabe mencionar que ambos parámetros son relevantes para diferenciar entre los modelos binario y progresivo de acción de los *enhancers*.

Los plásmidos que incluyen a los elementos de regulación fueron transfectados de manera estable en la línea celular HD3. Después de 48 horas de realizada la transfección, las células se transfirieron a una matriz de celulosa (*Methocel*) en presencia de un marcador de selección que corresponde a la droga G-418 (o Geneticina, a una concentración de 0,9 mg/ml; ésta cantidad ha demostrado ser eficiente para impedir el crecimiento de células no transfectadas, Furlán-Magaril *et al*, 2011). Después de 2-3 semanas de selección se aislaron clonas individuales y fueron transferidas a pozos pequeños para facilitar su expansión. La integración de una sola copia de cada transgén en el genoma de la célula receptora fue demostrada por ensayos de *Southern blot* y las líneas generadas fueron seleccionadas en base a su integridad y peso molecular (Fig. 25).

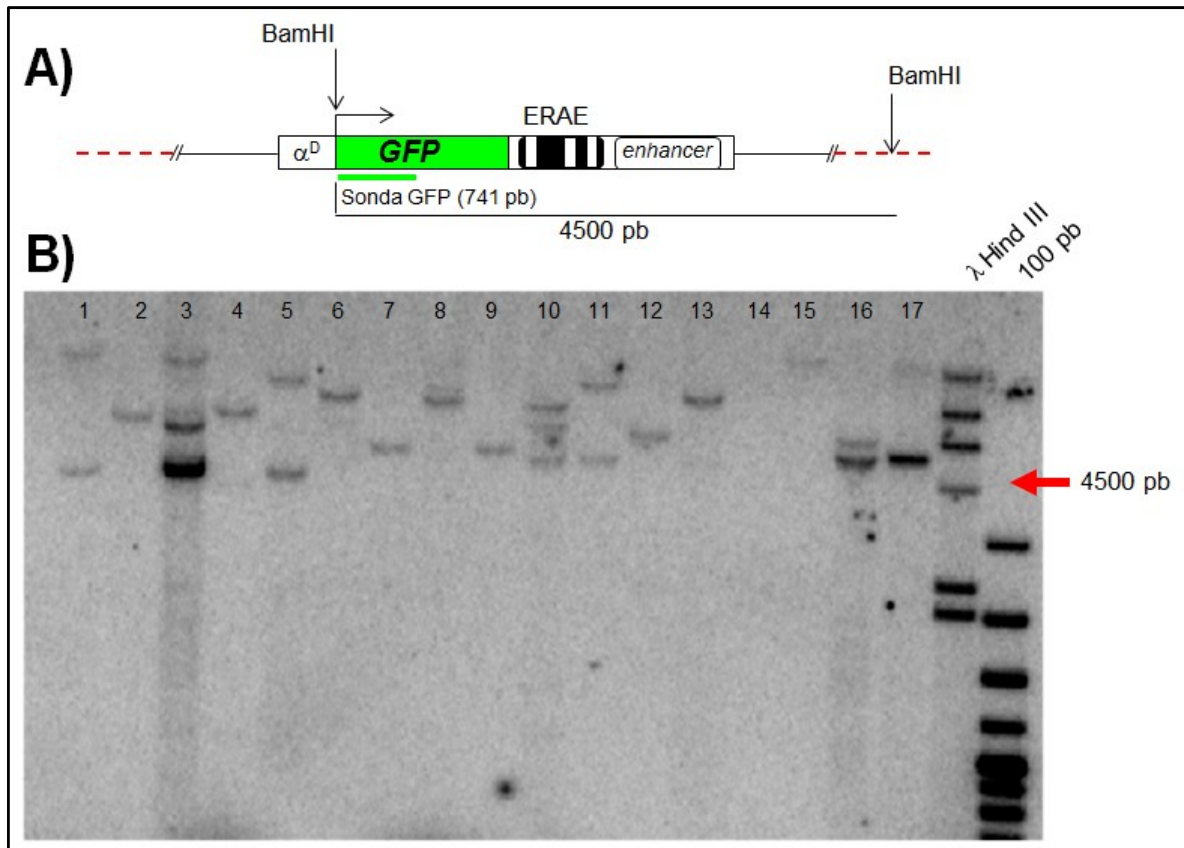


Figura 25. Integración de una sola copia de cada transgén en el genoma de las células HD3 confirmado por ensayos de *Southern blot*. **A)** Representación gráfica del transgén que contiene el gen reportero *GFP* y las secuencias reguladoras correspondientes al ERAE-*enhancer*, las flechas indican el sitio de restricción de la enzima BamHI dentro del plásmido y fuera del él (línea punteada roja). **B)** Autorradiografía representativa de un ensayo de Southern Blot. El DNA genómico de cada línea transgénica (1 a 17) se sometió a digestión con la enzima Bam HI para determinar el número de copias de cada transgén y su integridad. Para revelar el patrón de bandas, la membrana se hibridó con una sonda de 741 pb de DNA marcada radiactivamente con ^{32}P . La sonda reconoce específicamente la región *GFP* del plásmido. Los datos que incluimos en el análisis provienen de clonas dónde se observa una sola banda en cada carril (1-17) con un peso mayor o igual a 4750 pb pues representa una línea transgénica clonal donde el transgén se encuentra integrado una sola vez y de manera intacta (sin deleciones) en el genoma de las células HD3.

X.2 El ERAE disminuye la expresión del gen reportero *GFP* en el contexto α -globina de pollo

Cada una de las líneas estables generadas en células HD3 se analizó mediante citometría de flujo para determinar la expresión de cada transgén en presencia del promotor α^D -globina de pollo y las secuencias reguladoras de interés. En la Fig. 26 se muestra el porcentaje (%) de

células fluorescentes definidas por FACS, lo cual refleja la actividad transcripcional del gen reportero para cada transgén obtenido a partir de 10 sitios de inserción diferentes (n = 10). Un pequeño porcentaje (12%) de células positivas a la transcripción del gen reportero *GFP* disminuyen cuando el ERAE está en presencia del *enhancer* 3' α -globina de pollo (Fig. 26, comparar 54% vs 42%).

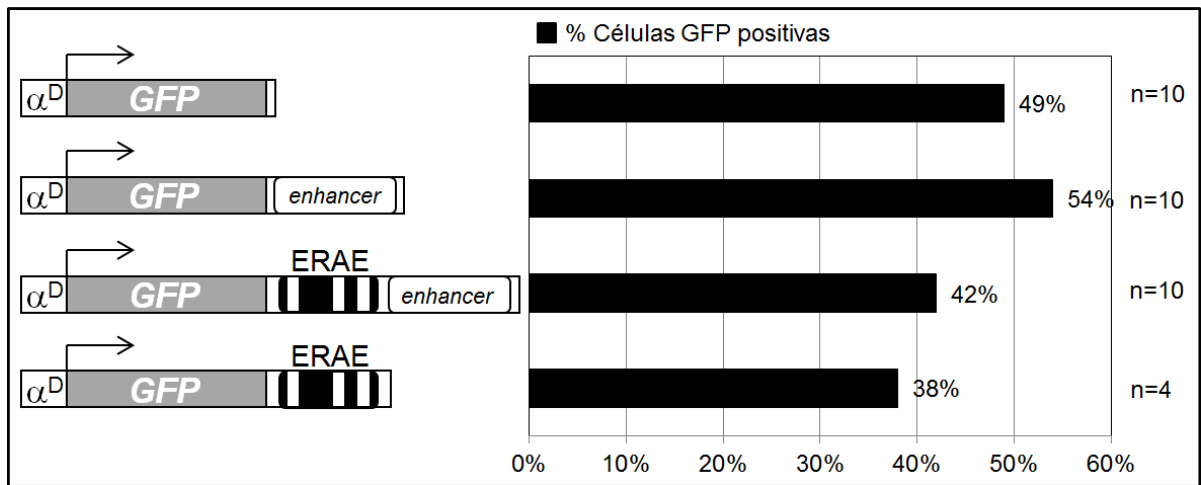


Figura 26. La gráfica muestra el porcentaje (%) de células fluorescentes para cada transgén obtenido por citometría de flujo. Del lado izquierdo se muestran los transgenes que se insertaron en copia única y que fueron seleccionados en base a su integridad mediante ensayos de *Southern blot*. El número de clones independientes se muestra del lado derecho de la gráfica como (n).

Los datos anteriores sugieren que la presencia del ERAE reduce la proporción (%) de células activas a la transcripción sobre la actividad del *enhancer* 3' α -globina de pollo y sobre el promotor α^D -globina de pollo independientemente del sitio de inserción genómica del transgén.

Al analizar la intensidad promedio de fluorescencia de las células que expresan el gen reportero de cada transgén, se observa una drástica reducción en presencia del ERAE (Fig. 27). Por lo tanto, el ERAE modula principalmente la frecuencia de transcripción del gen reportero *GFP*, de acuerdo con el modelo progresivo de acción de los *enhancers* (modelo binario) (Blackwood y Kadonaga, 1998; Fiering *et al*, 2000; Boulger y Groudine, 2011).

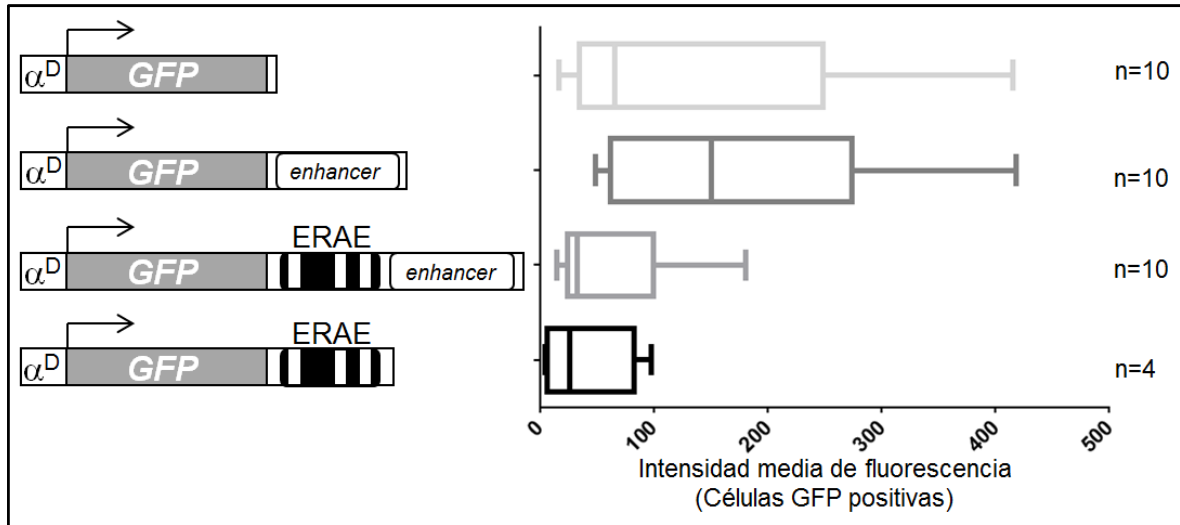


Figura 27. Intensidad media de fluorescencia para cada transgén definido por citometría de flujo. La gráfica de cajas y brazos muestra la intensidad promedio de fluorescencia emitida por el gen reportero *GFP* para cada transgén procedente de líneas estables que contienen copias individuales de cada transgen en células HD3. El número de clones independientes se muestra del lado derecho de la gráfica como (n) y representan 10 sitios de inserción diferentes.

Las fluctuaciones en el número de moléculas de RNA mensajero varían en el tiempo y entre células; estos cambios pueden ser expresados en términos estadísticos como la varianza dividida entre el promedio de la expresión génica, este parámetro se denomina factor Fano (Brown *et al*, 2013). Un valor alto de factor Fano correlaciona con el modelo progresivo de acción de los *enhancers* pues implica una mayor dispersión de los datos mientras que un valor bajo correlaciona mejor con el modelo binario. Para determinar la influencia de los elementos de regulación en la modulación de la actividad transcripcional, decidimos utilizar el factor Fano. Encontramos un valor menor en presencia del ERAE-*enhancer* (factor Fano=53.42) que en presencia del *enhancer* (factor Fano=88.33) (Ver Tabla 1). Este resultado sugiere que el *enhancer* se ajusta al modelo progresivo de actividad de los *enhancers* y la presencia del ERAE regula esta actividad en el contexto eritroide de las células HD3 de pollo.

En resumen, se observa un efecto negativo sobre la transcripción del gen reportero en presencia del ERAE que afecta principalmente la tasa de inicio de la transcripción y apoya el modelo progresivo de actividad de los *enhancers* (Figura 26 y 27). Una implicación que deriva de estos resultados es que la principal función del elemento ERAE es generar una estructura

de cromatina que reduzca la variabilidad en la expresión de gen reportero, a pesar del ambiente cromatínico en el que se encuentre.

X.3 La actividad del ERAE es contexto y tejido dependiente

Con el fin de entender qué tan general es la actividad del ERAE, reemplazamos las secuencias correspondientes al promotor y *enhancer* α -globina del pollo por el promotor y *enhancer* viral SV40, en presencia o ausencia del ERAE. La actividad de estas secuencias reguladoras se evaluó a través de la expresión del gen reportero *GFP* a partir de clonas individuales de cada transgén en la línea celular HD3 y al igual que la serie de experimentos anteriores se determinó el número e integridad de las inserciones por ensayos de Southern blot.

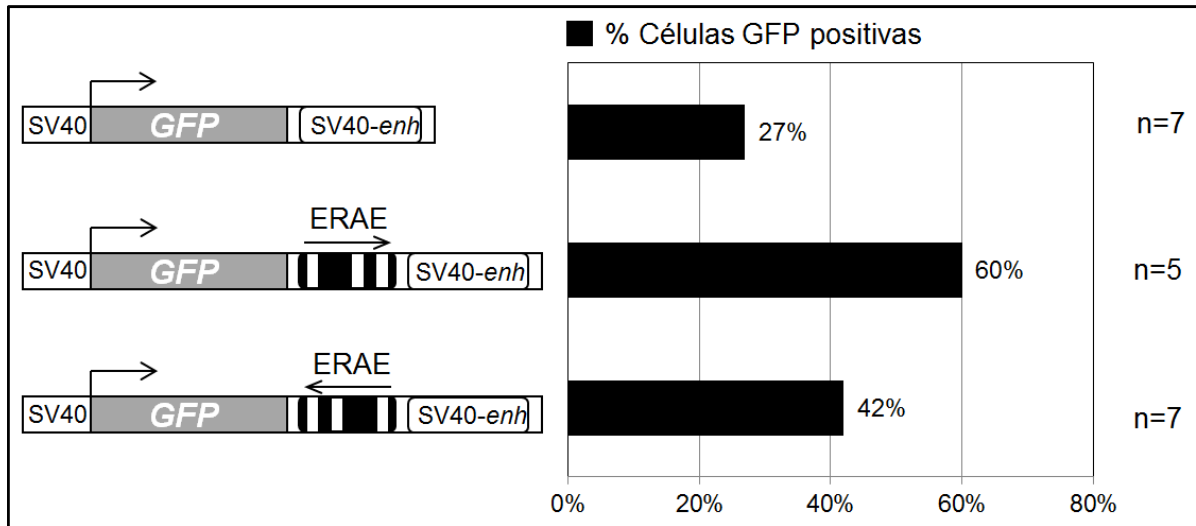


Figura 28. Actividad del elemento ERAE en un sistema heterólogo de regulación. La gráfica muestra el porcentaje (%) de células *GFP* positivas en presencia del ERAE, y el promotor y *enhancer* viral SV40. Independientemente de la orientación del ERAE; en ambos casos se observa un incremento en el porcentaje de células *GFP* positivas.

En contraste con los resultados anteriores se observó un incremento pronunciado en el número de células positivas a la transcripción de la *GFP* al ensayar el transgén que contiene al ERAE, independientemente de su orientación (Fig. 28).

Al analizar al ERAE en la orientación 3'-5' se observa una reducción en el promedio de la intensidad media de fluorescencia (Fig. 29).

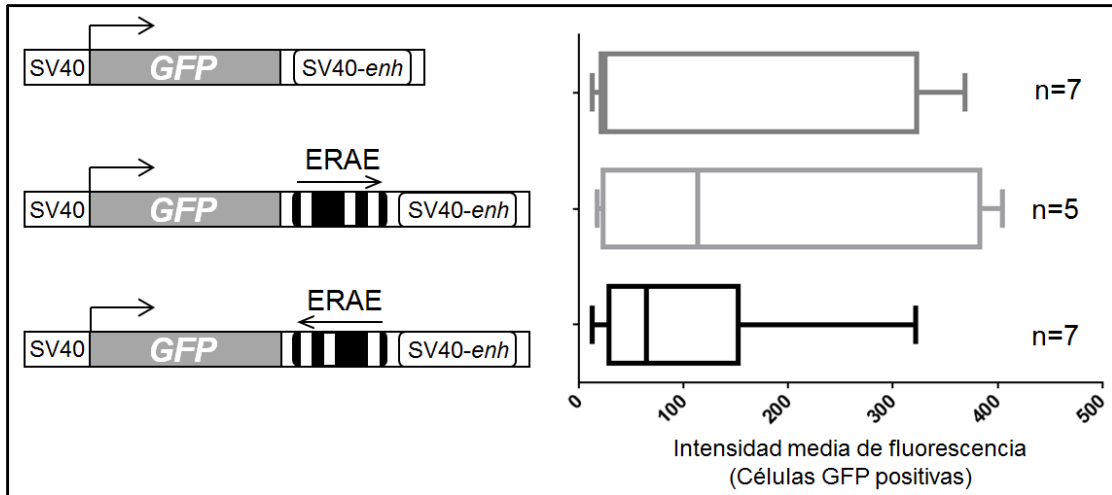


Figura 29. Las líneas celulares con sus correspondientes transgenes (lado izquierdo) se evaluaron en términos de su intensidad media de fluorescencia. La gráfica de cajas y brazos muestra la intensidad media de fluorescencia emitida por el gen reportero para cada transgén, el efecto negativo sobre éste parámetro sólo se observa cuando el ERAE está en orientación (3'-5').

De manera alternativa decidimos plantearnos la misma pregunta pero ahora analizando la actividad del ERAE en el contexto de un promotor mínimo (que contiene únicamente una caja TATA y dos sitios de unión al factor de transcripción Sp1) en células HD3 (Fig. 30). La contribución del ERAE se evaluó a través de la expresión del gen reportero *GFP* en clonas unicopia de cada uno de los transgenes. En este caso observamos un efecto positivo sobre la transcripción del gen reportero *GFP* en presencia del ERAE ya que aumenta la proporción de células activas a la transcripción, pero también aumenta la variabilidad de la intensidad media de fluorescencia. En este contexto, es posible que el promotor no sea un elemento dominante y por lo tanto le permite al ERAE modular la expresión del gen reportero no sólo en forma binaria sino también progresiva.

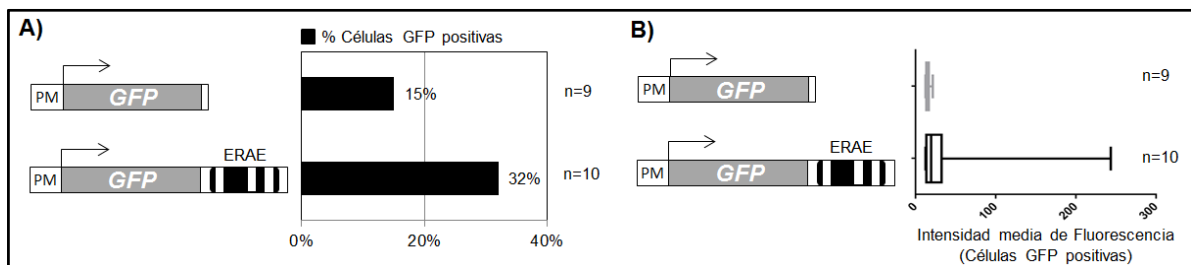


Figura 30. Efecto del ERAE sobre un promotor mínimo. **A)** La presencia del ERAE aumenta la

proporción de células que expresan el gen reportero *GFP*. **B)** La variabilidad de la intensidad media de fluorescencia de la gráfica de caja y brazos aumenta en presencia del ERAE.

Por otra parte, las construcciones que contienen las secuencias de SV40-ERAE se transfectaron en las células HD3 y se analizaron grupos de células (*pools*) resistentes a la droga G-418 en vez de clones individuales (Fig. 31). En ausencia del ERAE se observó un perfil de citometría característico de la actividad de un *enhancer*, con una actividad dispersa del gen reportero que va de intermedia a muy alta (Fig. 31). En concordancia con nuestros resultados previos (Fig. 27), la presencia del elemento ERAE disminuyó la intensidad de fluorescencia de la *GFP* y se redujo el perfil de distribución de las células positivas para *GFP* (observe el eje de las abscisas entre las diferentes construcciones en la Fig. 31), es decir, disminuyó la variabilidad en la expresión del gen reportero *GFP*.

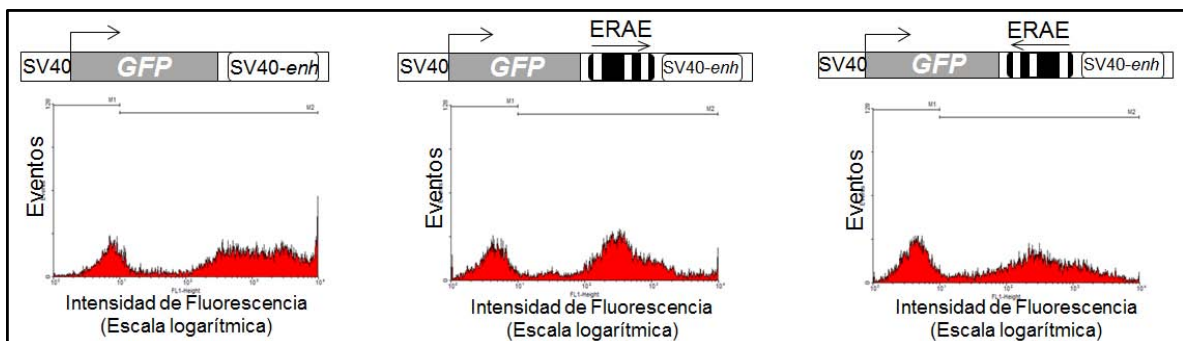


Figura 31. Perfiles de citometría representativos de un grupo de células transfectadas con cada construcción después de la selección con el antibiótico G-418. En la parte superior se muestra cada transgén con sus elementos regulatorios y debajo su perfil de expresión. En ausencia del ERAE los niveles de transcripción son más altos (eje de las abscisas) en comparación con las poblaciones que contienen este elemento de regulación.

Los altos niveles de expresión del gen reportero del transgén que tiene al promotor-*enhancer* SV40 son comparables con datos obtenidos por microscopía de fluorescencia en clones unicopia que contienen la secuencias reguladoras en un contexto eritroide (Fig. 32). En todos los casos hay aproximadamente el mismo número de células, sin embargo la mayor intensidad de fluorescencia se observa cuando el vector contiene únicamente al elemento *enhancer* 3' α -globina de pollo y el promotor α^D . Esta intensidad promedio de fluorescencia se reduce drásticamente en presencia del ERAE. Lo anterior es una vez más consistente con el modelo progresivo de acción de los *enhancers* en presencia del ERAE.

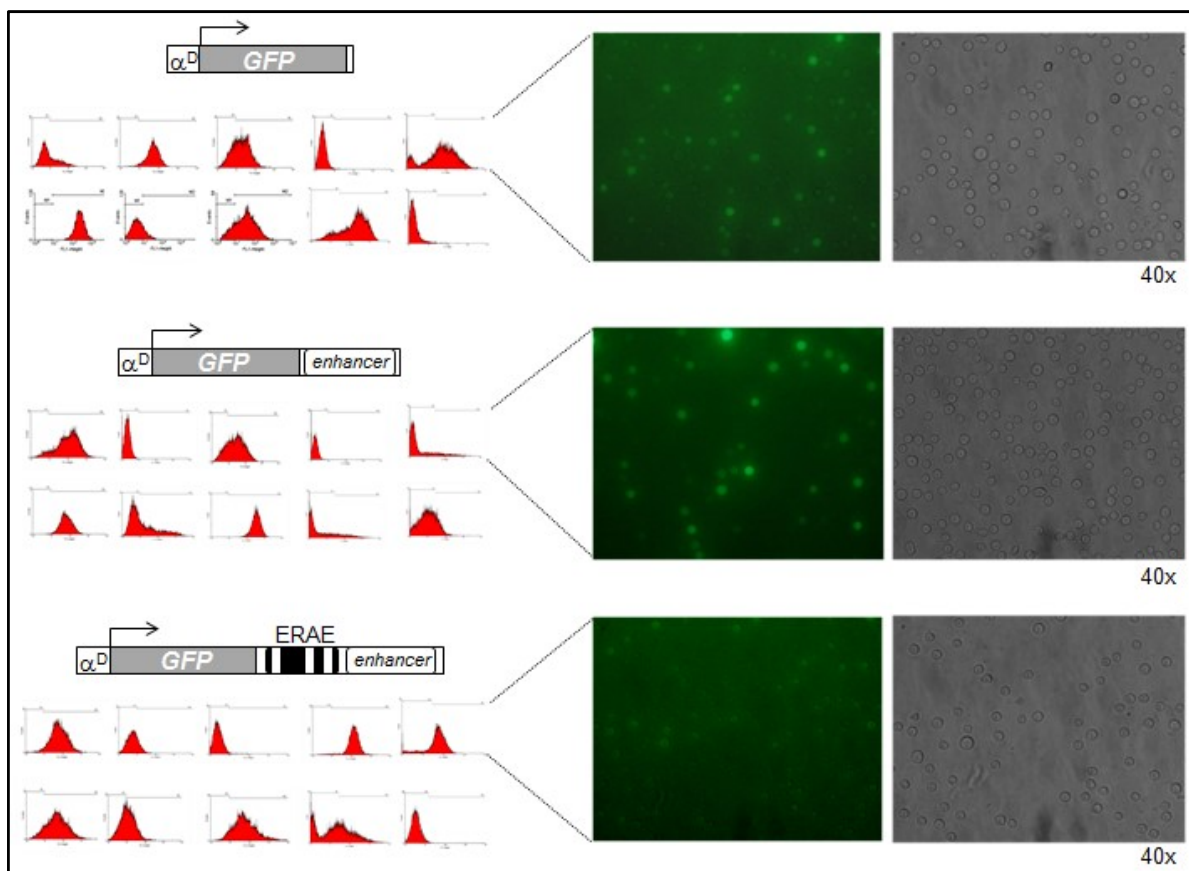


Figura 32. Microscopía de fluorescencia representativa de una clona unicopia transfectada con los transgenes que contienen el promotor α^D y los elementos de regulación eritroides *enhancer* y ERAE-*enhancer*. A simple vista se observa una gran variación en la intensidad de fluorescencia proveniente de cada población clonal, lo que sugiere que las células provenientes de un mismo transgén no contienen las mismas cantidades de proteína *GFP* en un mismo momento. En campo claro se muestran el total de células analizadas.

Cuando analizamos la actividad transcripcional del ERAE en presencia de secuencias heterólogas observamos que su actividad depende del tipo de promotor al cual regula (Figs. 26 y 28). Esto nos llevó a preguntarnos el efecto que tendría el ERAE en otras líneas celulares donde la combinación de factores transcripcionales no es eritroide específica. Para abordar esta pregunta, decidimos clonar el ERAE en ambas orientaciones respecto al promotor y *enhancer* viral SV40 en presencia del gen reportero *Luciferasa*. Las transfecciones transitorias de estos transgenes muestran que la expresión del gen reportero varía de acuerdo al tipo celular analizado, por ejemplo, en la línea celular linfóide se observa un claro efecto negativo sobre la transcripción mientras que en fibroblastos primarios de pollo la actividad del gen

reportero *Luciferasa* depende de la orientación en que está clonado el ERAE (Fig. 33). En conclusión, los datos experimentales sugieren que la actividad del ERAE es contexto y tejido dependiente.

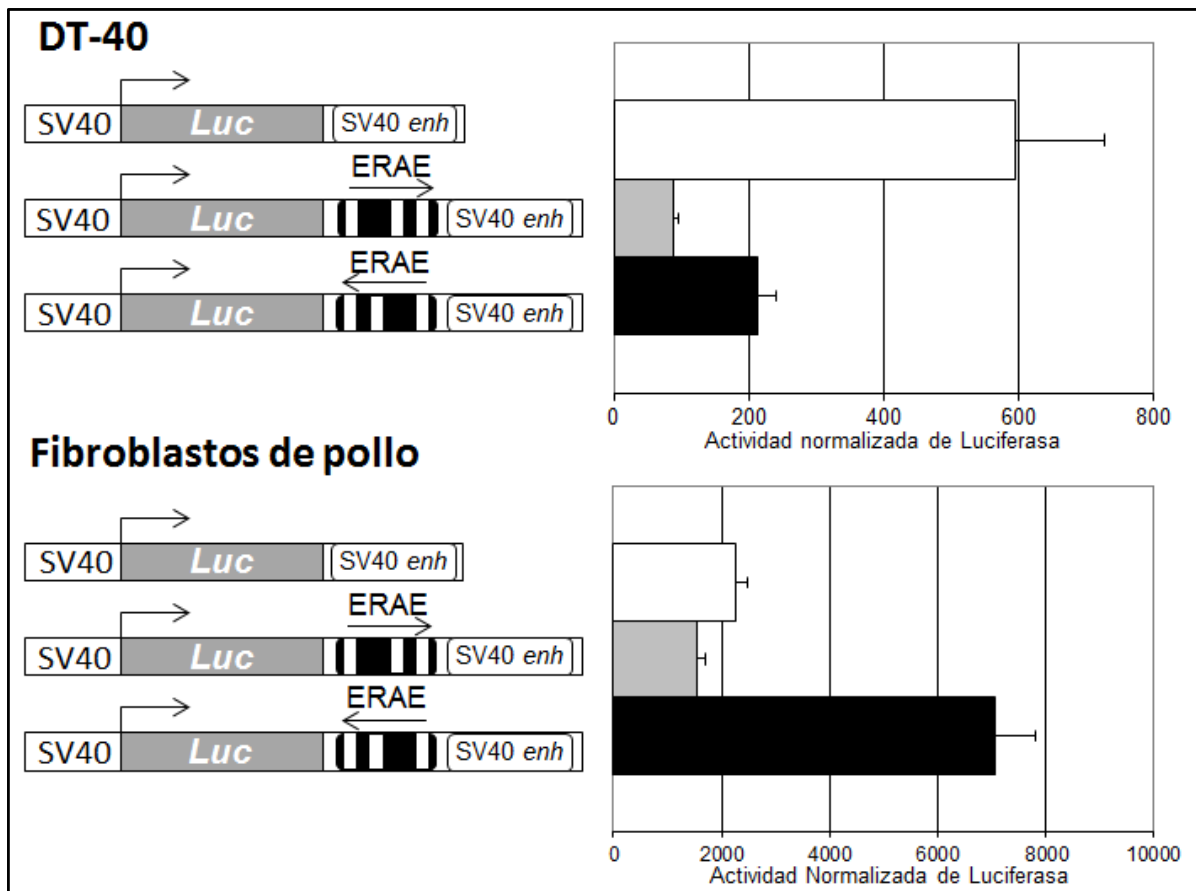


Figura 33. La actividad del ERAE en un contexto heterólogo es específica de tejido. Las gráficas muestran la actividad normalizada de la *Luciferasa* que se utilizó como gen reportero bajo la influencia del promotor-*enhancer* viral SV40 en presencia o ausencia del ERAE clonado en dirección 5'-3' (→) y 3'-5' (←).

X.4 El factor de transcripción RUNX-1 interacciona *in vitro* con la secuencia ERAE

El ERAE fue definido inicialmente como un elemento silenciador en un contexto no cromatinizado midiendo la actividad del gen reportero *Luciferasa* en transfecciones transitorias (Recillas Targa *et al*, 1993). En el fragmento de DNA de 360 pb que comprende el ERAE se definió la presencia de al menos tres secuencias de unión a factores nucleares

llamados SF1, SF2 y SF3 (Fig. 21 y 34) (Recillas Targa *et al*, 1993). La secuencia que corresponde al sitio SF1 contiene una secuencia repetida directa (R1 y R2, que se indican con flechas horizontales de color azul) (Fig. 34). Adyacente a la segunda repetida identificamos en este trabajo un sitio de unión al factor transcripcional RUNX-1 cuyo consenso corresponde a (T/C)G(T/C)GGT(T/C) (van Riel *et al*, 2012). En la Fig. 33 el sitio de unión a RUNX1 se muestra subrayado y con letras de color azul que además coincide con la repetida 2 (R2) (Fig. 34).

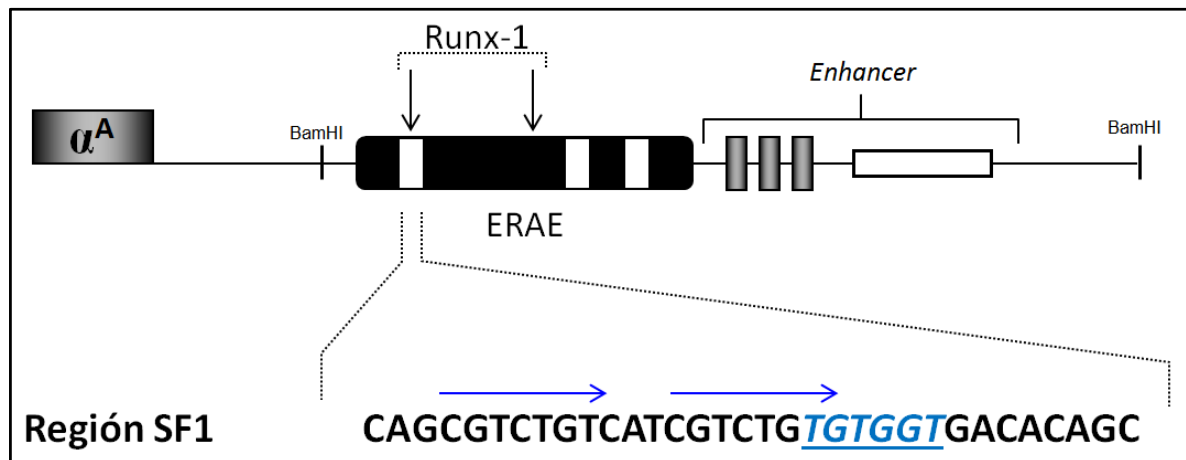


Figura 34. Esquema gráfico de la región genómica correspondiente al gen adulto α^A y la región de regulación 3' del dominio α -globina de pollo. Se muestra parte de la secuencia de DNA correspondiente a la región SF1 que incluye dos secuencias repetidas directas (Flechas horizontales azules, R1 y R2). La R2 coincide con el sitio de unión al factor RUNX-1 (se muestra con letras azules).

La proteína RUNX-1 ha sido previamente descrita como parte del complejo proteico formado por LDB1, GATA-1, TAL1, y EtO2 en células hematopoyéticas (Meier *et al*, 2006), pero su actividad está en función de los co-factores con los que interactúa (Blyth *et al*, 2005; Pencovich, *et al*, 2011). Se ha demostrado además que los factores LDB1 y GATA-1 interaccionan *in vivo* con proteínas estructurales como las cohesinas por lo que RUNX-1 podría formar parte de un complejo que favorece los contactos de larga distancia entre promotores y *enhancers* que son necesarios para la activación de los genes (Smith y Shilatifard, 2014). En este trabajo demostramos la interacción *in vitro* del factor transcripcional RUNX-1 a la región del ERAE mediante ensayos de retardo de movilidad electroforética (EMSA) (Fig. 35B) y súper-retardo (Fig. 36B).

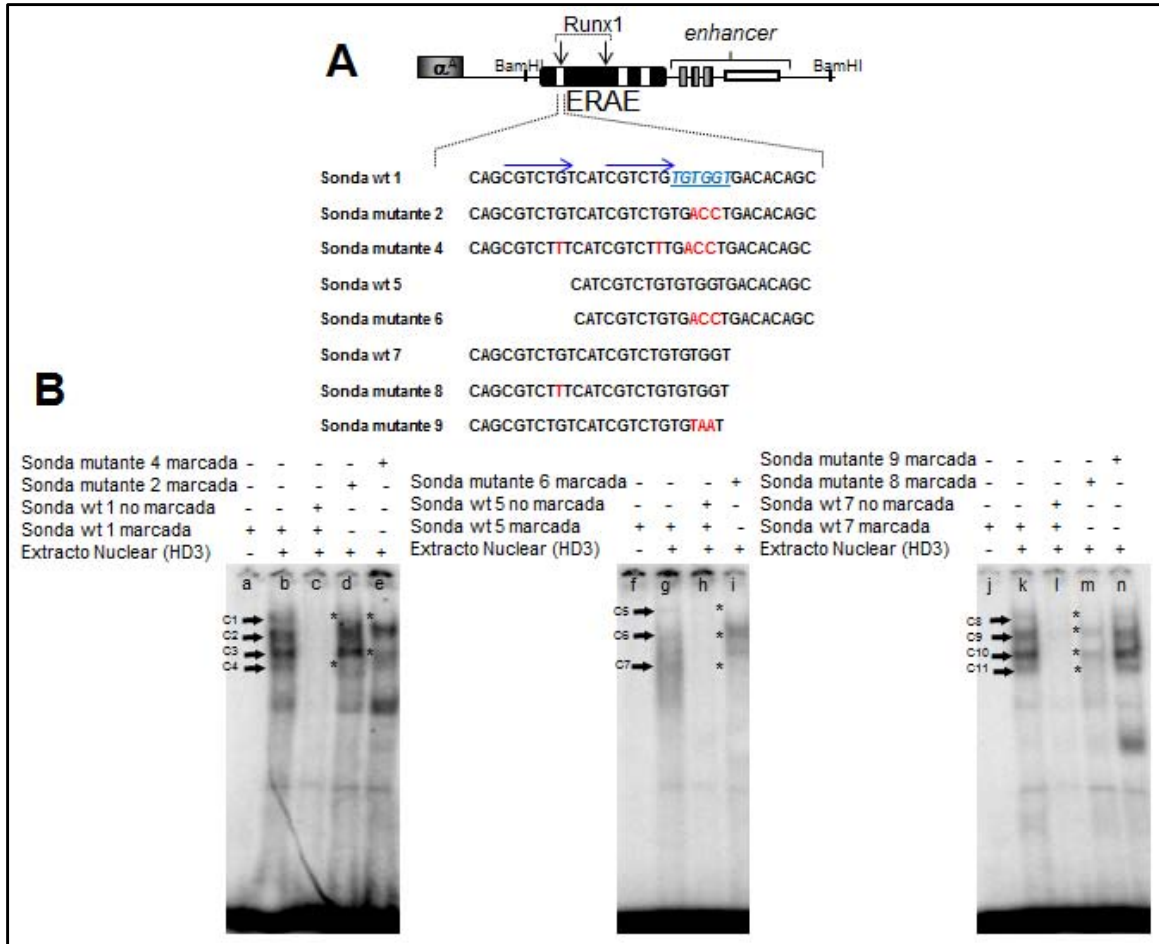


Figura 35. Ensayo de movilidad electroforética de la secuencia genómica correspondiente a SF1 localizado en el extremo 5' del ERAE. **A)** Representación gráfica de la región genómica correspondiente al gen adulto α^A y la región de regulación 3' del dominio α -globina de pollo. Las flechas verticales negras indican los sitios potenciales de unión al factor transcripcional RUNX-1. Las flechas horizontales azules indican la posición de la secuencia repetida directa en la región SF1. Únicamente se muestran las secuencias *forward* utilizadas como sondas para realizar el ensayo. Con letras rojas se indican los nucleótidos que fueron intercambiados para generar sondas mutantes. **B)** Ensayo de retardo de movilidad electroforética o EMSA. En cada carril (a-n) se muestra la sonda radiactiva usada y la formación de complejos proteicos con extractos nucleares de células HD3. Los asteriscos localizados en el lado izquierdo de cada carril se usan cuando los complejos han desaparecido usando las sondas mutantes.

Un aspecto relevante del experimento de EMSA es que las mutaciones en la secuencia consenso TGTGGT para el factor RUNX-1 (sondas mutantes 2 y 6), así como las mutaciones en las secuencias repetidas directas R1 y R2 (sondas mutantes 4 y 8) afectan la unión de complejos proteicos sobre la secuencia SF1. En los ensayos de retardo, los complejos 1 y 4 (C1

y C4) dependen de la integridad del sitio de unión al factor transcripcional RUNX-1 (comparar carriles b, d y e de la Fig. 35B). El complejo C3 parece depender en su totalidad de la secuencia repetida directa 1 (R1) (comparar carriles b y e de la Fig. 35B). La sonda que incluye la secuencia repetida directa 2 (R2) intacta pero el sitio de unión a RUNX-1 mutado pierde el patrón de unión de todos los complejos (comparar carriles g e i de la Fig. 35B). Los complejos C8 y C11, dependen totalmente de la integridad de la repetida directa R1 (comparar carriles k y m de la Fig. 35B) pero también afectan la unión de los complejos C9 y C10 (ver carriles k y m en la Fig. 35B). En conclusión, las mutaciones que afectan la integridad de los sitios R1, R2 y RUNX-1 afectan diferencialmente la unión de los complejos proteicos que interactúan con estas secuencias.

Para los ensayos de súper-retardo se utilizó un anticuerpo que reconoce al factor RUNX-1 (ab92336 de *Abcam*) y se usó como control positivo la región promotora del gen *FOS* de humano (Fig. 36A), esta secuencia de DNA se obtuvo a partir de un análisis *in vivo* mediante ensayos de ChIP-seq donde se reportó la unión del factor transcripcional RUNX-1 en la línea celular K562 (Pencovich *et al*, 2011). El ensayo de súper-retardo confirma que los complejos formados (C1 a C4 de la Fig. 36A) en presencia del anticuerpo que reconoce al factor RUNX-1 con la sonda del promotor *FOS* retardan su migración (comparar carriles b y c de la Fig. 36A). Adicionalmente, el complejo de menor peso molecular (C5) desaparece. Alternativamente se utilizaron un par de anticuerpos no relacionados como controles negativos (anti-IgG y anti-CTCF) (Fig. 36A). Estos resultados sugieren que la proteína RUNX-1 interactúa específicamente con el promotor *FOS* de humano *in vitro*.

Al usar la sonda del sitio SF1 (que incluye R1, R2 y el sitio de unión a RUNX1), se observa que el complejo con el peso molecular más bajo (C8) desaparece en presencia del anticuerpo anti-RUNX-1. El ensayo muestra que el complejo C6 migra ligeramente retardado y el complejo C7 casi desaparece (comparar carriles i y j de la Fig. 36B). Aunque se requieren pruebas adicionales para demostrar la unión *in vivo* del factor transcripcional RUNX-1 a la secuencia del ERAE como una inmunoprecipitación de cromatina (ChIP), una hipótesis que surge de estos resultados es que la unión del factor RUNX-1 al ERAE podría contribuir con la formación

de un complejo proteico que permite la comunicación a distancia con su promotor blanco en el contexto del dominio α -globina de pollo.

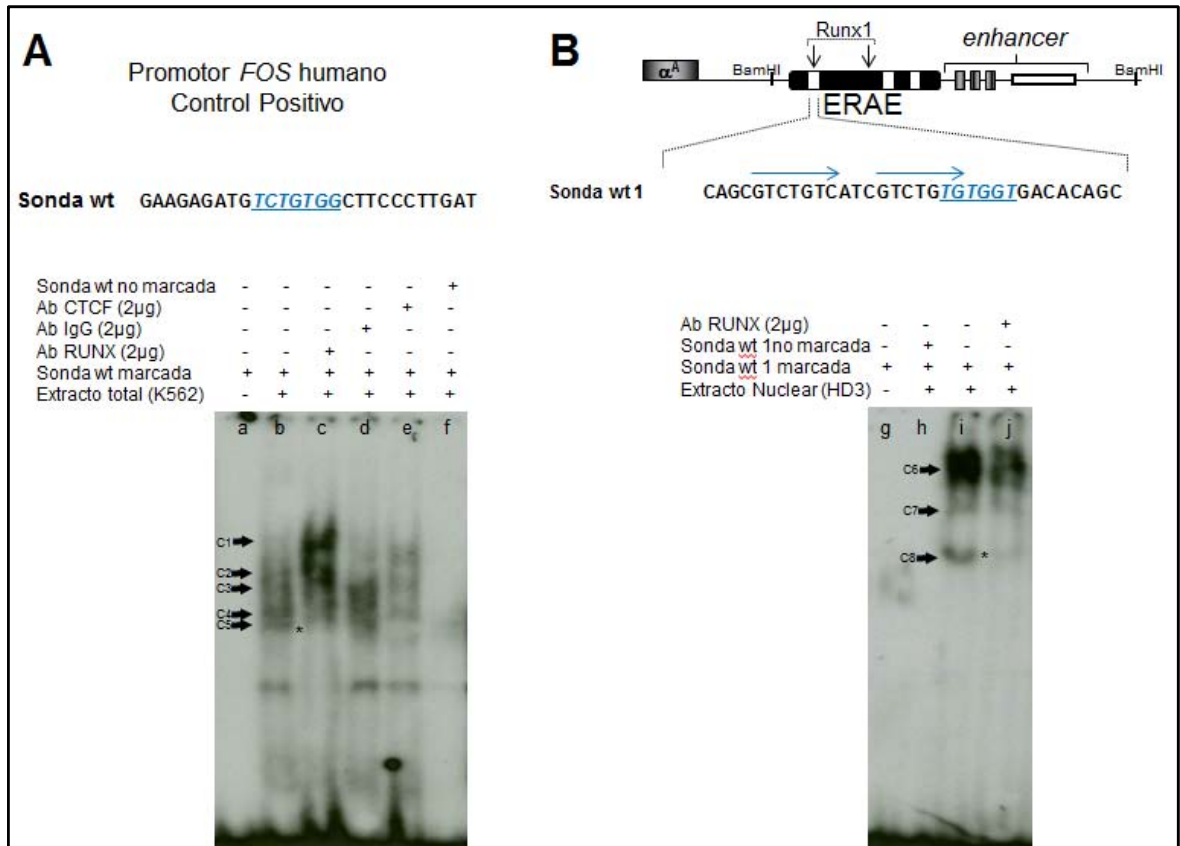


Figura 36. El factor de transcripción RUNX-1 forma complejos proteicos *in vitro* con la secuencia del ERAE. **A)** Ensayo de súper retardo usando la sonda del control positivo que corresponde al promotor *FOS* de humanos (Pencovich *et al*, 2011) y extractos totales provenientes de células K562. Únicamente se muestran las secuencias *forward* utilizadas como sondas para realizar el ensayo. El sitio de unión al factor RUNX-1 se muestra con letras azules. Los complejos formados por la interacción proteína-DNA se muestran como C1 a C5. El asterisco localizado en el lado izquierdo del carril c muestra que el complejo nucleoproteico ha desaparecido usando el anticuerpo anti-RUNX-1. **B)** Ensayo de súper-retardo usando las sondas de la región SF1 del ERAE. Las flechas verticales negras indican los sitios potenciales de unión al factor transcripcional RUNX-1. Las flechas horizontales azules indican la posición de las secuencias repetidas directas (R1 y R2) en la región SF1. En cada carril se muestra la sonda usada marcada radiactivamente con 32 P. La formación de complejos proteicos se indica desde el C6 al C8. Los extractos nucleares usados provienen de células HD3. El asterisco localizado en el lado izquierdo del carril j muestra que el complejo nucleoproteico ha desaparecido usando el anticuerpo anti-RUNX-1.

X.5 La integridad del ERAE es necesaria para la actividad del *enhancer 3'α-globina* de pollo

Para determinar qué región (es) a nivel de la secuencia del ERAE son relevantes para modular la actividad del *enhancer 3'α-globina* de pollo, generamos nuevos plásmidos con deleciones graduales del extremo 5' del ERAE. La actividad transcripcional del gen reportero *GFP* se evaluó mediante citometría de flujo en clonas individuales de cada transgén en las línea celular HD3 cuyo número e integridad de las inserciones se verificó por ensayos de *Southern blot*. La eliminación gradual de la secuencias llevó a una reducción en el número de células que expresan el gen reportero (Fig. 37), lo que implica que la integridad del ERAE es necesaria para evitar el silenciamiento del transgen.

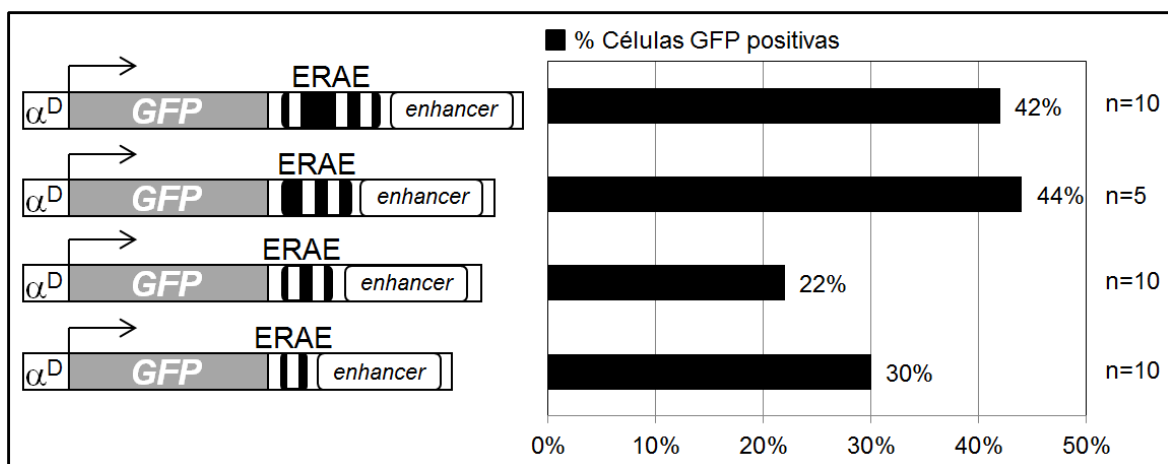


Figura 37. Eliminación secuencial del elemento ERAE en presencia del *enhancer 3'α-globina* de pollo y el promotor α^D . Las deleciones consideran la posición de los sitios de unión SF1, SF2 y SF3. Se aislaron clonas estables de cada transgén y se evaluó su actividad mediante la expresión de *GFP*. La gráfica muestra el porcentaje (%) de células fluorescentes para cada transgen donde las últimas dos deleciones afectan negativamente el porcentaje de células activas a la transcripción.

Asimismo, la reducción en el porcentaje de células activas a la transcripción correlaciona con una mayor variabilidad en la expresión del gen reportero *GFP* (Fig. 38).

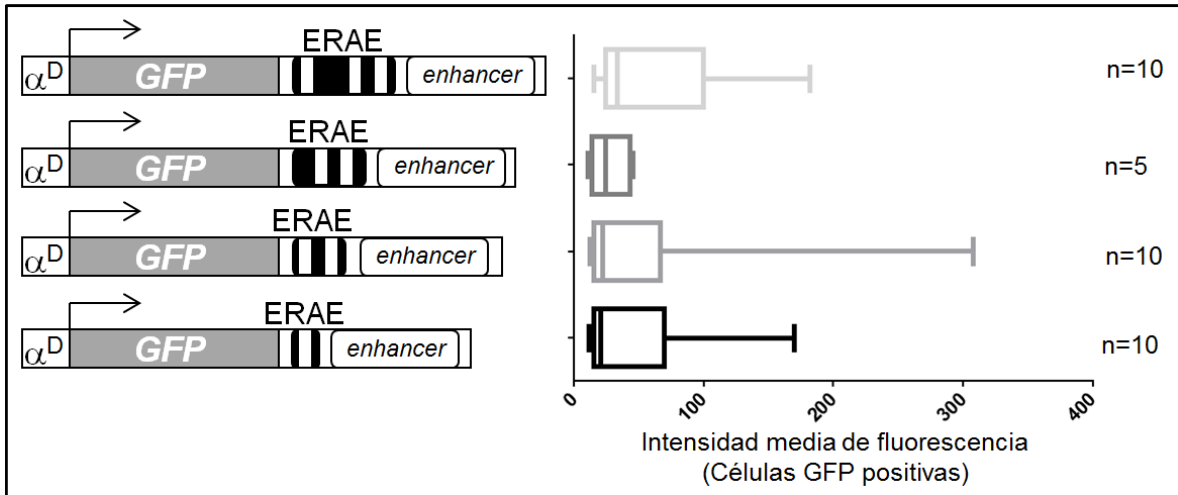


Figura 38. La integridad a nivel de la secuencia del ERAE es necesaria para reducir la variabilidad en la expresión del gen reportero. En la gráfica de caja y brazos se muestra la intensidad promedio de fluorescencia para cada transgén.

Un análisis detallado de la estructura de la cromatina durante el desarrollo del embrión de pollo y la diferenciación eritroide permitió definir una estructura cromatínica particular en la región ERAE-*enhancer* 3' α -globina (Fig. 39).

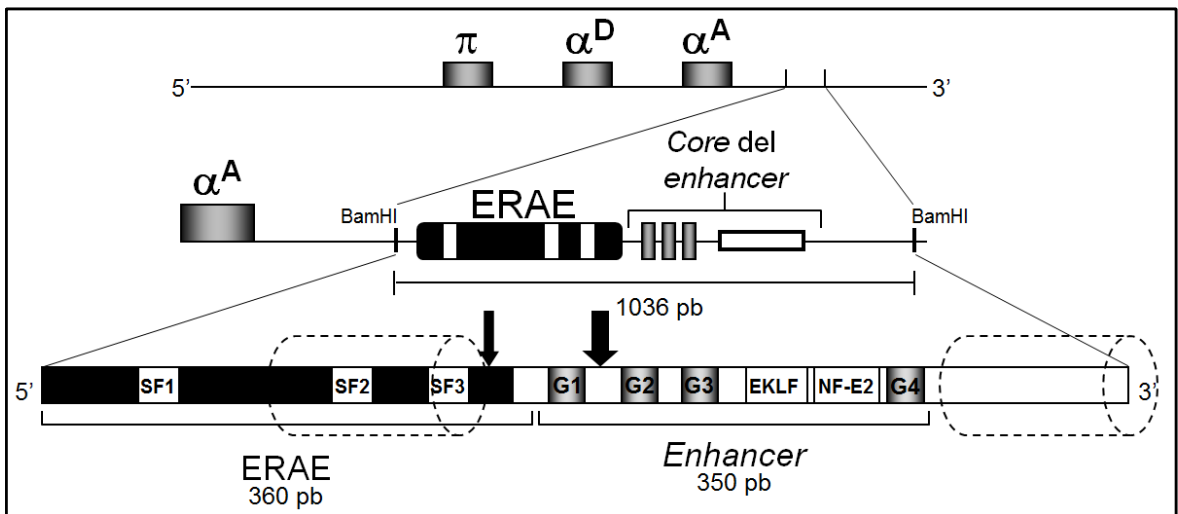


Figura 39. Esquema del dominio α -globina de pollo. Se muestran el elemento ERAE y el *core* del *enhancer*, además de la localización detallada de las secuencias de unión a factores nucleares. La región ERAE-*enhancer* está situada a unos 450 pb río abajo del gen α^A -globina de pollo. Las flechas verticales y su espesor representan la posición e intensidad de los sitios hipersensibles al corte por la MNasa previamente definidos. También se muestra la posición de dos nucleosomas (cilindros punteados) que enmarcan el *core* del *enhancer* generando un

NDR. Este posicionamiento nucleosomal se observa en células HD3 y eritrocitos circulantes de 10 días de desarrollo (10dRBC) (Escamilla-Del-Arenal y Recillas-Targa, 2008).

Esta estructura está definida por dos sitios hipersensibles al corte por la MNasa que correlacionan con la actividad del *enhancer*. Ambas hipersensibilidades se encuentran separadas a una distancia de 100 pb (+/-20 pb) cada una de la otra (Figura 39; Escamilla-Del-Arenal y Recillas-Targa, 2008). Esta distancia no es suficiente para permitir el posicionamiento de un nucleosoma, por lo tanto, decidimos alterar la organización de la estructura de la cromatina para determinar si esto afecta la función del *enhancer*.

X.6 La perturbación de la marca epigenética que se caracteriza por dos sitios de corte a la enzima MNasa y el posicionamiento nucleosomal entre el ERAE y *enhancer* genera la pérdida de la actividad transcripcional del gen reportero

Nuestra hipótesis propone que los dos sitios hipersensibles al corte por la MNasa (Figura 39, flechas verticales) y la posición que ocupan los nucleosomas (Figura 39, cilindros punteados) son característicos del estado activo del *enhancer*. Para abordar esta pregunta decidimos modificar esta organización introduciendo 25, 50 y 75 pb de DNA no relacionado (proveniente del fago lambda) entre los dos sitios de corte por la enzima MNasa (Figura 40).

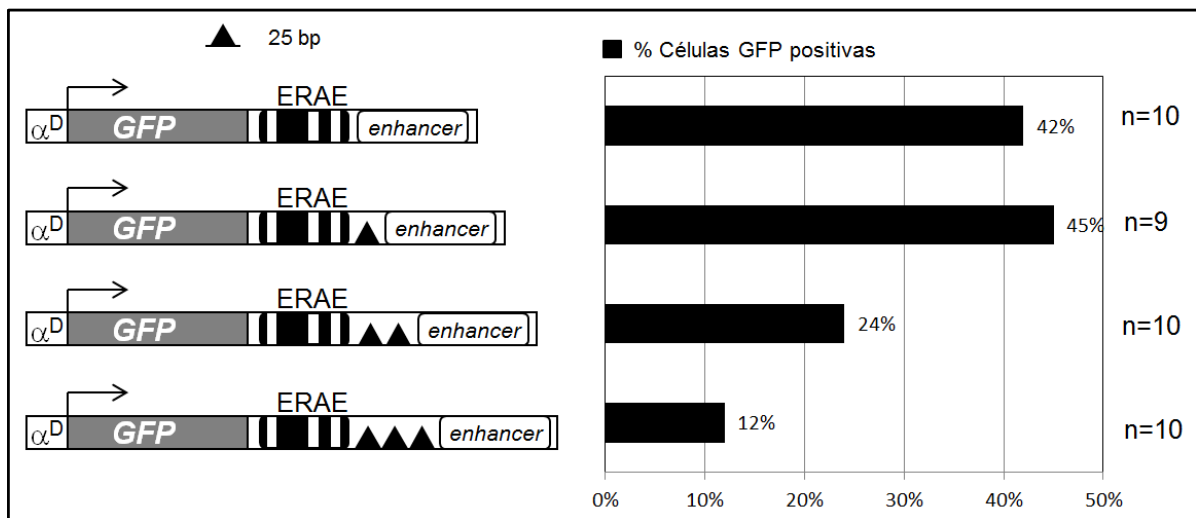


Figura 40. Pérdida de la actividad del *enhancer* mediante la introducción de secuencias no relacionadas entre los sitios hipersensibles al corte por la MNasa que co-localizan entre los elementos ERAE-*enhancer*. La gráfica muestra el porcentaje (%) de células fluorescentes

obtenido para cada construcción. Cuanto mayor es la cantidad de DNA insertado, mayor es la pérdida en la actividad transcripcional del gen reportero. Cada triángulo representa un incremento en la secuencia de 25 pb.

Cuando la distancia entre el ERAE-*enhancer* incrementa en 25 pb, no se observa ningún efecto negativo sobre la actividad transcripcional del gen reportero. En contraste, cuando se incorporaron 50 y 75 pb, hay una disminución proporcional en el porcentaje de células activas a la transcripción (15% y 30%, respectivamente en la Fig. 40). Estos datos sugieren que la actividad estimuladora del *enhancer* se ve afectada por posibles cambios asociados con la estructura de la cromatina.

X.7 El incremento en la distancia entre las dos hipersensibilidades al corte por la enzima MNasa entre el ERAE-*enhancer* 3' α -globina de pollo afecta el posicionamiento de los nucleosomas

La separación de los dos sitios hipersensibles al corte por la MNasa afecta la actividad ERAE-*enhancer* reduciendo significativamente el porcentaje de células activas a la transcripción (Fig. 41). Sin embargo, desconocíamos si la pérdida en la actividad se debía a un cambio en el posicionamiento de los nucleosomas sobre estos elementos de regulación. Para detectar posibles cambios asociados con la estructura de la cromatina se realizaron ensayos de accesibilidad a la MNasa en dos líneas celulares transgénicas. Una de ellas contiene el transgén con la región ERAE-*enhancer* sin modificaciones en distancia (wt) (Fig. 41 A) y la otra línea celular contiene el transgén con una inserción de 75 pb de DNA (Fig. 41 B). Usando una sonda que hibrida con la secuencia *GFP* en la línea wt se reveló el patrón de accesibilidad esperado, similar al endógeno (Fig. 41 A). En contraste, en la línea celular que llevan la inserción de 75 pb encontramos una drástica re-localización del nucleosoma localizado sobre el ERAE (Fig. 41 B). En conclusión, la posición de los nucleosomas que delimitan al *core* del *enhancer* y la formación de los dos sitios hipersensibles al corte por nucleasa MNasa entre el ERAE-*enhancer* son características epigenéticas necesarias para mantener la actividad del *enhancer* 3' α -globina de pollo.

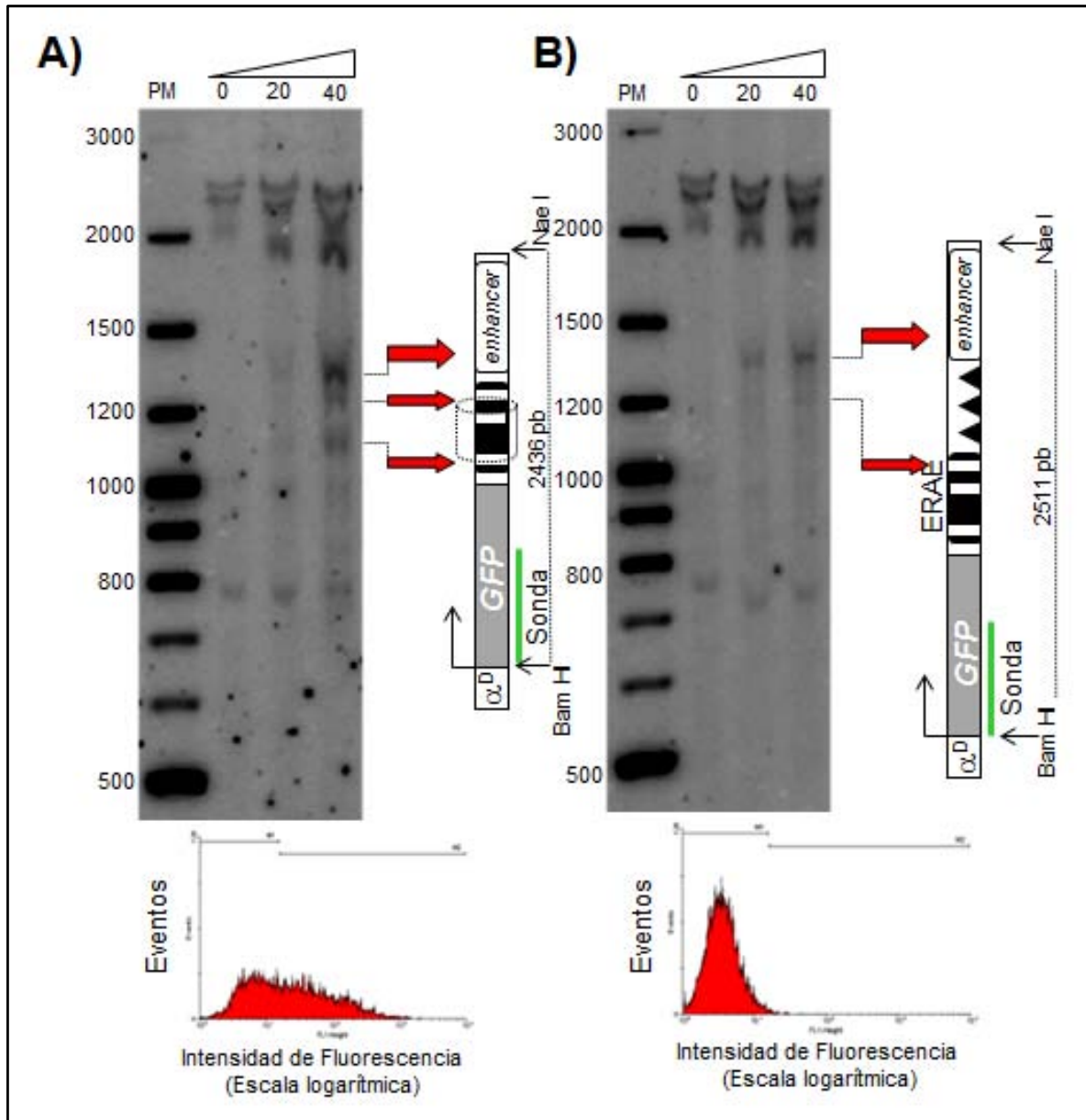


Figura 41. El incremento de la distancia entre los dos sitios hipersensibles al corte por la MNasa altera la posición que ocupan los nucleosomas en la secuencia ERAE. **A)** Patrón de accesibilidad al corte con la enzima MNasa en cantidades crecientes (0=no digestión, 20 y 40 Unidades de MNasa) en la línea estable que contiene la región ERAE-*enhancer* (wt). En la parte inferior se muestra el perfil de expresión obtenido por citometría de flujo correspondiente a una clona activa. **B)** Patrón de accesibilidad al corte con la enzima MNasa en cantidades crecientes (0=no digestión, 20 y 40 Unidades de MNasa) de la clona con un incremento en la distancia (75 pb) entre los dos sitios hipersensibles al corte por la MNasa entre ERAE y *enhancer*. Esta clona no expresa el transgén (ver perfil de expresión obtenido por citometría de flujo) y correlaciona con un patrón de accesibilidad distinto al de la clona wt. Ambas membranas se incubaron al mismo tiempo con una sonda que hibrida en la región

correspondiente a la *GFP* en cada transgén. En dos experimentos independientes se obtuvieron resultados similares.

X.8 El incremento en la distancia entre las dos hipersensibilidades al corte por la enzima MNasa entre el ERAE-*enhancer* 3 ' α -globina de pollo no afecta la unión del factor transcripcional CTCF

La perturbación de la estructura de la cromatina asociada a los elementos de regulación podría alterar la unión de factores transcripcionales sobre la región en estudio. Para demostrar si se afecta la unión de CTCF en la clona mutante decidimos evaluar su presencia mediante inmunoprecipitación de la cromatina pues ya se ha reportado que CTCF es un factor importante relacionado con la formación de asas de cromatina en los genes globina (Valdes-Quezada *et al*, 2013). Encontramos que CTCF puede unirse a la región ERAE-*enhancer* en ambas clonas wt y mutante (Fig. 42). Con éste resultado, favorecemos la idea de que la pérdida en la actividad transcripcional del gen reportero en la clona mutante de 75 pb es compatible con cambios en la estructura de la cromatina que no alteran la unión de CTCF, sin embargo, en el futuro será pertinente evaluar si otros TFs pierden o no su capacidad de unirse al DNA.

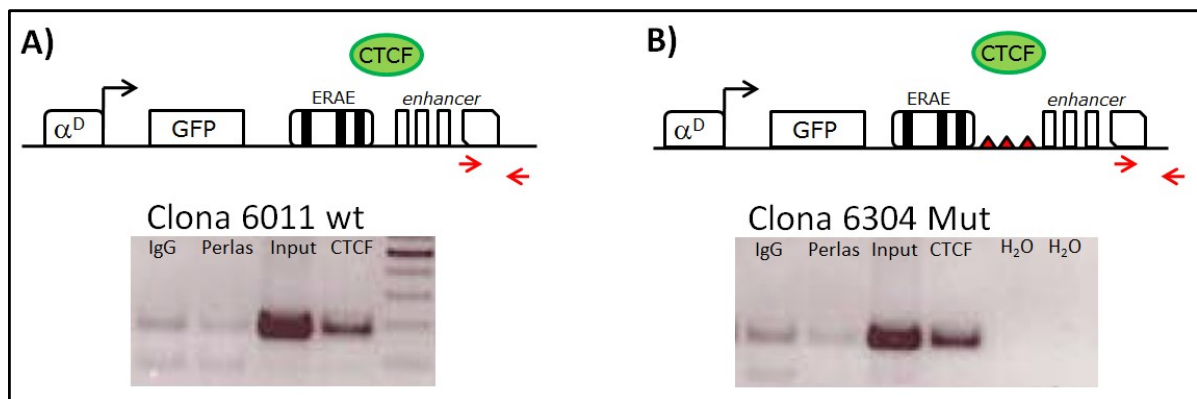


Figura 42. Ensayo de inmunoprecipitación de la cromatina para evaluar la presencia de CTCF en la región ERAE-*enhancer*. **A)** El factor transcripcional CTCF se une a la región ERAE-*enhancer* en una clona wt. **B)** El factor transcripcional CTCF se une a la región ERAE-*enhancer* en una clona mutante con 75 pb de distancia entre ambos elementos de regulación. Las flechas rojas indican la posición de los oligonucleótidos utilizados para amplificar la región genómica por PCR de punto final. Cada triángulo rojo representa un incremento en la secuencia de 25 pb y en verde se muestra la proteína CTCF.

X.9 Modelo de la regulación de la actividad potenciadora mediante el posicionamiento de nucleosomas

El posicionamiento nucleosomal sobre el ERAE y *enhancer 3'* α -globina de pollo correlaciona con la máxima actividad del *enhancer* (Fig. 22 y 41 A), sin embargo, la inserción de 75 pb de DNA entre ambos elementos de regulación es suficiente para alterar la estructura tridimensional y el posicionamiento de los nucleosomas presentes en la región regulatoria ERAE-*enhancer*. En nuestro análisis de accesibilidad de la cromatina se observó que en la construcción mutante, el ERAE parece estar desprovisto de nucleosomas, lo que sugiere que los factores de transcripción como CTCF y GATA-1 podrían unirse y reclutar factores de remodelación de la cromatina dependientes de ATP (Fig. 43). Otra posibilidad es que un nucleosoma se posicione entre la intersección de ambos sitios de hipersensibles al corte por la MNasa ocultando los sitios de unión para factores nucleares que impiden la comunicación entre el *enhancer* y su promotor blanco. Sin embargo, no favorecemos esta última hipótesis porque en el ensayo de accesibilidad observamos una banda débil correspondiente al posicionamiento de un nucleosoma sobre el ERAE, lo que sugiere que este es un evento infrecuente dentro de la población celular analizada.

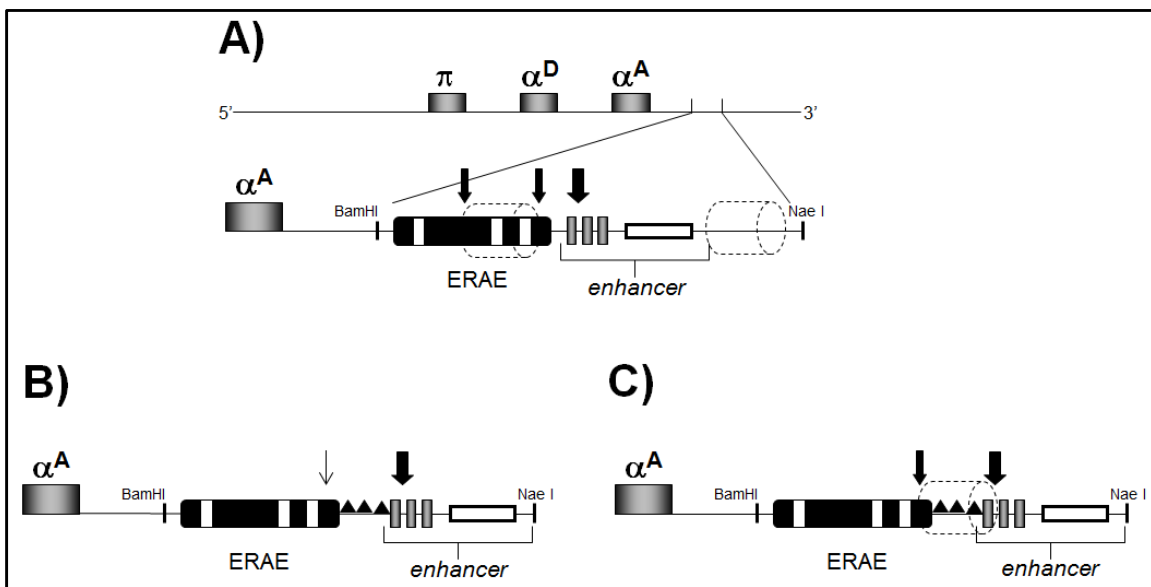


Figura 43. Modelo de la regulación de la actividad potenciadora mediante el posicionamiento de nucleosomas. **A)** Esquema gráfico de la región ERAE-*enhancer* del dominio α -globina de pollo. Las flechas verticales y su espesor representan la posición y la intensidad de los sitios

hipersensibles al corte por la MNasa definidos anteriormente (Escamilla-Del-Arenal y Recillas-Targa, 2008). Los cilindros punteados representan el posicionamiento de los nucleosomas sobre la región ERAE-*enhancer*. Esta organización de la cromatina correlaciona con la actividad máxima del *enhancer*. **B)** Modelo de la región ERAE-*enhancer* con el incremento de distancia de 75 pb donde se favorece una región libre de nucleosomas y la pérdida del nucleosoma sobre el ERAE. **C)** Modelo de la región ERAE-*enhancer* donde se favorece la ganancia de un nucleosoma entre los dos sitios de hipersensibilidad a la MNasa ocultando los sitios de unión a factores transcripcionales, evitando la comunicación con su promotor blanco.

Finalmente, proponemos que el posicionamiento de un nucleosoma sobre el elemento ERAE y la presencia de dos sitios hipersensibles al corte por la MNasa separados por 100 (+/- 20) pb son características locales de la cromatina necesarias en la actividad del *enhancer* 3' α -globina de pollo en el contexto de la cromatina.

XI. Discusión

Los *enhancers* se definen clásicamente como secuencias con una organización modular con motivos de unión a factores transcripcionales ubicuos y tejido específicos. Pueden estar situados en regiones genómicas intragénicas e intergénicas y su actividad es independiente de su ubicación y distancia respecto al sitio de inicio de la transcripción (Blackwood y Kadonaga, 1998; Bulger y Groudine, 2011; Smith y Shilatifard, 2014). Recientemente se han propuesto nuevas maneras de identificar a los *enhancers* basados en la combinatoria de marcas en las histonas asociadas a estas regiones genómicas. Por ejemplo, se ha propuesto que las marcas en las histonas como la H3K27ac y H3K4me1 son características de estos elementos de regulación. Las secuenciaciones masivas de genomas completos en combinación con ensayos de captura conformacional de cromosomas (3C) han permitido determinar que los *enhancers* interaccionan físicamente con sus genes blanco (Marsman y Horsfield, 2012; De Laat y Duboule, 2013). De forma inesperada se ha reportado que los *enhancers* pueden interactuar con RNAs largos no codificantes para facilitar los contactos a larga distancia con su promotor blanco (Orom y Shiekhattar, 2013). Una nueva clase de transcritos pequeños provenientes de regiones potenciadoras denominados *enhancer* RNAs (eRNA) han sido identificados y se ha propuesto que su presencia correlaciona con la actividad de genes codificantes cercanos, al parecer esta nueva clase de RNAs no codificantes participan en los mecanismos de acción de los *enhancers* (Mousavi *et al*, 2013). A pesar de los avances en la comprensión de los mecanismos que permiten a los *enhancers* funcionar, aún se desconoce el efecto de la organización local de la estructura de la cromatina relacionada con ellos.

En este trabajo se describe la función de una secuencia reguladora denominada ERAE que se localiza adyacente al *enhancer 3' α-globina* de pollo. Los resultados muestran que el ERAE es un componente integral del *enhancer 3' α-globina*. Mediante análisis de accesibilidad al corte por la nucleasa MNasa, observamos que la configuración de la cromatina asociada con el ERAE es crítica para la actividad de *enhancer 3' α-globina* de pollo (Fig. 41). La formación de dos sitios de hipersensibilidad al corte por la MNasa y el posicionamiento de un nucleosoma sobre la región del ERAE son necesarios para la actividad del *enhancer*. La alteración en la

configuración de la región correspondiente al ERAE-*enhancer* mediante la inserción de DNA que separa los dos sitios hipersensibles al corte por la MNasa causa la pérdida del nucleosoma 5' sobre el ERAE y reduce el potencial de *trans*-activación del *enhancer* (Fig. 40 y 41). Por lo tanto, no sólo las interacciones a larga distancia, sino también los cambios en la estructura de la cromatina local afectan la actividad de los *enhancers*.

Los datos de dispersión calculados por medio del Factor Fano (Tabla 1) favorecen la hipótesis del modelo progresivo de acción del *enhancers* 3' dado que en presencia del ERAE disminuye la variabilidad en la expresión del gen reportero. Esto sugiere que el ERAE regula principalmente la frecuencia de transcripción y/o genera una estructura de cromatina adecuada independientemente del sitio cromatínico y se refleja en una menor dispersión de los datos. En relación con lo último, determinamos que la actividad transcripcional correlaciona con una estructura de la cromatina que reduce la variabilidad en la expresión del gen reportero, a través del posicionamiento de un nucleosoma sobre la región ERAE.

Aunque el principal efecto del ERAE es el de modular la frecuencia de activación de su gen blanco en el contexto eritroide (de acuerdo al modelo progresivo), también reduce modestamente el porcentaje de células activas a la transcripción (modelo binario). Lo anterior sugiere que el ERAE también afecta una porción de células activas. Si bien es cierto que el porcentaje es modesto (12%), los datos podrían ser relevantes ya que provienen de sitios genómicos de integración independientes (Fig. 26). Esta hipótesis se apoya con los experimentos de deleciones del ERAE y del incremento en la distancia entre ambos sitios hipersensibles al corte por la MNasa pues en ambos casos hay una pérdida del porcentaje de células fluorescentes. Si nuestra hipótesis es correcta, entonces la presencia de un nucleosoma localizado sobre el ERAE podría ser el responsable de regular la transición entre ambos modelos de actividad de los *enhancers*. Actualmente, los *enhancers* siguen siendo validados funcionalmente mediante la identificación de las secuencias que presentan la mayor actividad transcripcional en sistemas que usan genes reporteros, sin embargo las secuencias adyacentes como el ERAE, con frecuencia son omitidas. La identificación del ERAE, que contribuye con la organización local de la cromatina asociada a la actividad del *enhancer*

sugiere que las regiones adyacentes deben ser tomadas en consideración en el análisis de actividad de los *enhancers*, sobre todo en el contexto cromatínico. Tal vez la marca de histona H3K4me1, que se extiende sobre los *enhancers* podría ser utilizada para identificar elementos reguladores como el ERAE en escala genómica o más recientemente utilizando a los eRNAs.

El ERAE-*enhancer* adquiere una estructura cromatínica particular cuyos sitios hipersensibles al corte por la MNasa aparecen progresivamente en las células rojas circulantes durante el desarrollo del pollo alcanzando la mayor accesibilidad de la cromatina en el día 10, cuando el *enhancer* es completamente activo (Escamilla Del Arenal y Recillas Targa, 2008). Desconocemos el mecanismo que permite la remodelación progresiva de la cromatina, pero varios factores podrían estar involucrados. El factor nuclear eritroide específico GATA-1, y los dos sitios de unión situados en cada lado de la hipersensibilidad más prominente entre el ERAE y el *enhancer* son críticos para la correcta actividad de estos elementos de regulación (Escamilla Del Arenal y Recillas Targa, 2007). Además, BRG1, la subunidad catalítica del complejo de SWI/SNF dependiente de ATP es capaz de interactuar con GATA-1 y EKLF, ambos presentes en el *enhancer* 3' α -globina de pollo (Xu *et al*, 2006; Kim *et al*, 2009). Otra posibilidad es el reclutamiento del complejo remodelador de cromatina NuRD a través de GATA-1 y FOG-1 (Miccio *et al*, 2010). Por lo tanto, los dos sitios hipersensibles al corte por la MNasa pueden reflejar la formación de un gran complejo que favorece la liberación de nucleosomas situados sobre el *enhancer* que permiten su actividad máxima.

Por otra parte, los nucleosomas que delimitan el *enhancer* podrían actuar como puntos de anclaje para atraer al complejo de remodelación dependiente de ATP como RSC (Hartley y Madhani, 2009). Este escenario puede complementarse con la incorporación de la variante de histona H2A.Z, que suele estar presente en nucleosomas asociados con promotores y *enhancers* de genes transcripcionalmente activos (Hartley y Madhani, 2009; Barski *et. al*, 2007). Por lo tanto, la perturbación de la estructura de la cromatina en la región ERAE-*enhancer* a través de secuencias de DNA provenientes del genoma del fago lambda modifica su estructura tridimensional. Los cambios en el posicionamiento nucleosomal sobre el ERAE quizá están relacionados con la alteración de la estructura tri-dimensional del complejo

nucleo-proteico anclado al *enhancer* necesario para activar la expresión de su gen blanco. Aunque la presencia de nucleosomas suele correlacionar con silenciamiento epigenético, en algunas circunstancias es necesario compactar el DNA a través de los nucleosomas para acercar sitios distales de unión a factores de transcripción (Lu *et al*, 1995; Stünkel *et al*, 1997; Gossett y Lieb 2012). En nuestro análisis de accesibilidad a la cromatina se observó que el ERAE parece libre de nucleosomas, lo que sugiere que la unión de factores de transcripción como CTCF (Fig. 42) y la remodelación de la cromatina aún están presentes.

Hemos demostrado la flexibilidad funcional asociada con la región del ERAE (Fig. 28, Fig. 33) así como la unión *in vitro* del factor transcripcional RUNX-1 (Fig.36). Curiosamente, ésta familia de factores de transcripción se han asociado con la estimulación o el silenciamiento de sus genes blanco en función de los co-factores con los que interactúan (Blyth *et al*, 2005; Pencovich *et al*, 2011) en diferentes tejidos (Pencovich *et al*, 2011). RUNX1 participa en la formación de complejos proteicos formados por LDB1, GATA-1 y TAL1, entre otros (Meier *et al*, 2006; van Riel *et al*, 2012). Estas interacciones pueden ser relevantes para los contactos de larga distancia entre el *enhancer* y el gen al que regulan. Por último, no podemos descartar la posibilidad de que se requiere la configuración descrita del ERAE-*enhancer* para interactuar de manera eficiente con el promotor α^D . Uno de los próximos desafíos será descubrir el orden jerárquico de los eventos que permiten el ERAE-*enhancer trans*-activar a sus genes blanco durante la diferenciación eritroide y el desarrollo del pollo.

XII. Conclusiones

- ✓ La cromatina modula la actividad de los elementos regulatorios que se encuentran en el extremo 5' del dominio α -globina de pollo.
- ✓ El ERAE regula la actividad del *enhancer* 3' α -globina de pollo, disminuyendo la variabilidad de la transcripción del gen reportero *GFP* independiente del sitio de inserción del transgén.
- ✓ La actividad del ERAE es contexto y tejido dependiente.
- ✓ El factor transcripcional RUNX-1 se une *in vitro* a la región SF1 presente en el ERAE.
- ✓ Se necesita la integridad de la secuencia del elemento de ERAE para mantener la actividad del *enhancer* 3' α -globina de pollo.
- ✓ La alteración de la estructura de la cromatina caracterizada por dos sitios hipersensibles al corte por MNasa tiene un efecto negativo sobre la actividad del *enhancer* localizado en el extremo 3' del dominio α -globina de pollo.
- ✓ La introducción de secuencias no relacionadas entre ambos sitios de corte por la MNasa alteran el posicionamiento del nucleosoma que se localiza sobre la región ERAE y esto a su vez, correlaciona con la pérdida de la actividad del *enhancer* 3' α -globina de pollo.
- ✓ La introducción de secuencias no relacionadas entre ambos sitios de corte a la MNasa no afecta la capacidad de unión del factor transcripcional CTCF sobre la región ERAE-*enhancer* 3' α -globina de pollo.

XIII. Perspectivas

XIII.1 Demostrar la unión *in vivo* del factor transcripcional RUNX-1 en la región de ERAE-*enhancer*.

XIII.2 Caracterizar la función de la proteína RUNX-1 en la función del *enhancer*.

XIII.3 Determinar la interacción del factor RUNX-1 con factores ubicuos y eritroide específicos en el dominio α -globina de pollo.

XIII.4 Evaluar la unión de otros factores transcripcionales como GATA-1 y RUNX-1 en la clona mutante con incremento de 75 pb de distancia entre ERAE-*Enhancer*.

XIII.5 Determinar el orden jerárquico de los cambios que ocurren en la estructura de la cromatina en la región ERAE-*Enhancer* y que permiten *trans*-activar a su(s) gene(s) blanco.

XIV. Bibliografía

- 1) Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell*. 103(4), 667-78 (2000)
- 2) Ahmad K, Henikoff S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell*. 9, 1191-1200 (2002)
- 3) Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol*. 16(3):155-66 (2015)
- 4) Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, Higgs DR. Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J*. 23(14), 2841-52 (2004)
- 5) Antalis TM, Costelloe E, Muddiman J, Ogbourne S, Donnan K. Regulation of the plasminogen activator inhibitor type-2 gene in monocytes: localization of an upstream transcriptional silencer. *Blood*. 88(10), 3686-97 (1996)
- 6) Arnosti DN, Kulkarni MM. Transcriptional enhancers: Flexible billboards or intelligent enhanceosomes. *J Cell Biochem*. 94, 890–898 (2005)
- 7) Balamotis MA, Pennella MA, Stevens JL, Wasyluk B, Belmont AS, Berk AJ. Complexity in transcription control at the activation domain-mediator interface. *Sci Signal*. 2(69):ra20 (2009)
- 8) Banerji J, Banerji J, Olson L, Schaffner W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell*. 33, 729-740 (1983)
- 9) Banerji J, Rusconi S, Schaffner W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell*. 27, 299–308 (1981)
- 10) Barberis A, Petrascheck M. Transcription Activation in Eukaryotic Cells. *Encyclopedia of Life Sciences* (2003)

- 11) Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High resolution profiling of histone methylation in the genome. *Cell*. 129(4), 823-37 (2007)
- 12) Beato M, Eisefeld K. Transcription factor access to chromatin. *Nucleic Acids Res*. 25(18):3559-63 (1997)
- 13) Becker PB, Hörz W. ATP-dependent nucleosome remodeling. *Annu Rev Biochem*. 71,247-73 (2002)
- 14) Berger SL. Histone modifications in transcriptional regulation. *Curr opin genet dev*. 12,142-148 (2002)
- 15) Berger SL. The complex language of chromatin regulation during transcription. *Nature*. 447(7143), 407-412 (2007)
- 16) Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 16(1), 6-21 (2002)
- 17) Biterge B, Schneider R. Histone variants: key players of chromatin. *Cell Tissue Res*. 356(3), 457-66 (2014)
- 18) Blackwood EM, Kadonaga J T. Going the distance: a current view of enhancer action. *Science*. 281(5373), 60-63 (1998)
- 19) Blyth K, Cameron ER, Neil JC. The RUNX genes: gain or loss of function in cancer. *Nat Rev Cancer*. 5(5), 376-387 (2005)
- 20) Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, Ghavi-Helm Y, Wilczynski B, Riddell A, Furlong EE. Tissue specific analysis of chromatin states identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet*. 44(2), 148-156 (2012)
- 21) Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. High resolution mapping and characterization of open chromatin across the genome. *Cell*. 132(2), 311-322 (2009)
- 22) Brown CR, Mao C, Falkovskaia E, Jurica MS, Boeger H. Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biol*. 11, e1001621 (2013)
- 23) Bulger M, Groudine M. Functional and mechanistic diversity of distal transcriptional enhancers. *Cell*. 144(3) 327-339 (2011)

- 24) Bulger M, Groudine M. Looping versus linking: toward a model for long gene activation. *Genes Dev.* 13(19) 2465-2477 (1999)
- 25) Bulger M, Groudine M. TRAPping enhancer function. *Nat Genet.* 32 (4), 555-556 (2002)
- 26) Bulger M, van Doorninck J H, Saitoh N, Telling A, Farrell C, Bender MA, Felsenfeld G, Axel R, Groudine M. Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes. *Proc Natl Acad Sci USA.* 96(9), 5129-34 (1999)
- 27) Calo E, Wysocka J. Modification of enhancer chromatin: What, how and why? *Mol. Cell.* 49, 825-837 (2013)
- 28) Cantor AB, Orkin SH. Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. *Semin Cell Dev Biol.* 16(1) 117-128 (2005)
- 29) Chakalova L, Debrand E, Mitchell JA, Osborne CS, Fraser P. Replication and transcription: shaping the landscape of the genome. *Nat Rev Genet.* 6, 669-677 (2005)
- 30) Cho H, Orphanides G, Sun X, Yang XJ, Ogryzko V, Lees E, Nakatani Y, Reinberg D. A human RNA polymerase II complex containing factors that modify chromatin structure. *Mol Cell Biol.* 18(9), 5355-5363 (1998)
- 31) Clapier CR, Cairns BR. The biology of chromatin remodeling complexes. *Annu Rev Biochem.* 78, 273-304 (2009)
- 32) Core LJ, Martins AL, Dank CG, Waters CT, Siepel A, Lis JT. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet.* 46, 1311-1320 (2014)
- 33) Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet.* 2(4), 292-301 (2001)
- 34) Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA.* 107(50), 21931-21936 (2010)

- 35) Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA. Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol.* 145(6), 1119-31 (1999)
- 36) de Bruijn MF, Speck NA. Core-binding factors in hematopoiesis and immune function. *Oncogene.* 23(24),4238-48 (2004)
- 37) de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature.* 502(7472), 499-506 (2013)
- 38) De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* 8(5), e1000384 (2010)
- 39) Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science.* 295 (5558), 1306-11 (2002)
- 40) de la Serna IL, Ohkawa Y, Imbalzano AN. Chromatin remodeling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nat Rev Genet.* 7(6) 461-473 (2006)
- 41) Demers C, Chaturvedi CP, Ranish JA, Juban G, Lai P, Morle F, Aebersold R, Dilworth FJ, Groudine M, Brand M. Activator-mediated recruitment of the MLL2 methyltransferase complex to the α -globin locus. *Mol Cell.* 27, 573-584 (2007)
- 42) Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Solda G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC, Mattick JS. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res,* 18(9):1433-45 (2008)
- 43) Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 485(7398), 376-80 (2012)
- 44) Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell

- K, Bell I, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR. Landscape of transcription in human cells. *Nature*. 489 (7414), 101-8 (2012)
- 45) Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 16(10), 1299-309 (2006)
- 46) Drissen R, Palstra RJ, Gillemans N, Splinter E, Grosveld F, Philipsen S, de Laat W. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev*. 18(20), 2485-90 (2004)
- 47) ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, *et al*. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 447(7146), 799-816 (2007)
- 48) Escamilla-Del-Arenal M, Recillas-Targa F. GATA-1 modulates the chromatin structure and activity of the chicken α -globin 3' enhancer. *Mol Cell Biol*. 28(2), 575-586 (2007)
- 49) Fanucchi S, Shibayama Y, Burd S, Weinberg MS, Mhlanga MM. Chromosomal contact permits transcription between coregulated genes. *Cell*. 155(3), 606-20 (2013)
- 50) Felsenfeld G, Boyes J, Chung J, Clark D, Studitsky V. Chromatin structure and gene expression. *Proc Natl Acad Sci USA*. 93, 9384-9388 (1996)
- 51) Felsenfeld G, Groudine M. Controlling the double helix. *Nature*. 421, 448-453 (2003)

- 52) Feng Q, Zhang Y. The MeCP1 complex represses transcription through preferential binding, remodeling and deacetylating methylated nucleosomes. *Genes Dev.* 15(7), 827-32 (2001)
- 53) Fiering S, Whitelaw E, Martin DI. To be or not to be active: the stochastic nature of enhancer action. *Bioessays.* 4, 381-387 (2000)
- 54) Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I, Leonhardt H. Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One*, 6:e21306 (2011)
- 55) Furlan-Magaril M, Rebollar E, Guerrero G, Fernández A, Moltó E, González-Buendía E, Cantero M, Montoliu L, Recillas-Targa F. An insulator embedded in the chicken α -globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic Acids Res.* 39(1), 89-103 (2011)
- 56) Gaffney DJ, McVicker G, Pai AA, Fondufe-Mittendorf YN, Lewellen N, Michelini K, Widom J, Gilad Y, Pritchard JK. Controls of nucleosome positioning in the human genome. *PLoS Genet.* 8(11), e1003036 (2012)
- 57) Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet.* 7(9), 703-13 (2006)
- 58) Gossett AJ, Lieb JD. In vivo effects of histone H3 depletion on nucleosome occupancy and position in *Saccharomyces cerevisiae*. *PLoS Genet.* 8 e1002771 (2012)
- 59) Green MR. Eukaryotic transcription activation: Right on target. *Mol Cell.* 18, 399-402 (2005)
- 60) Greive SJ, von Hippel PH. Thinking quantitatively about transcriptional regulation. *Nat Rev Mol Cell Biol.* 6, 221-232 (2005)
- 61) Grosveld F, Antoniou M, Berry M, De Boer E, Dillon N, Ellis J, Fraser P, Hanscombe O, Hurst J, Imam A, Lindenbaum M, Philipsen S, Pruzina S, Strouboulis J, Raguz-Bolognesi S, Talbot D. The regulation of human globin gene switching. *Philos Trans R Soc Lond B Biol Sci.* 339(1288), 183-91 (1993)
- 62) Grosveld F. Activation by locus control regions? *Curr Opin Genet Dev.* 9(2), 152-7(1999)

- 63) Grosveld, F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell*. 51, 975–985 (1987)
- 64) Guarente L. UAS and Enhancers: Common mechanism of transcriptional activation in yeast and mammals. *Cell*. 52(3), 303-305 (1988)
- 65) Guy LG, Kothary R, Wall L. Position effects in mice carrying a lacZ transgene in cis with the β -globin LCR can be explained by a graded model. *Nucleic Acids Res*. 25(21) 4400-4407 (1997)
- 66) Hake SB, Garcia BA, Kauer M, Baker SP, Shabanowitz J, Hunt DF, Allis CD. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc Natl Acad Sci USA*. 102, 6344-6349 (2005)
- 67) Hartley PD, Madhani HD. Mechanisms that specify promoter nucleosome location and identity. *Cell*. 137(3), 445-58 (2009)
- 68) He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, Zhang Y, Xu K, Ni M, Lupien M, Mieczkowski P, Lieb JD, Zhao K, Brown M, Liu XS. Nucleosome dynamics define transcriptional enhancers. *Nat Genet*, 42(4):343-7 (2010)
- 69) Heard E, Disteché CM. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev*. 20(14), 1848-67 (2006)
- 70) Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 39(3), 311-318 (2007)
- 71) Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type specific gene expression. *Nature*. 459(7243), 108-112 (2009)
- 72) Henikoff S. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet*. 9(1) 15-26 (2008)

- 73) Herz HM, Mohan M, Garruss AS, Liang K, Takahashi YH, Mickey K, Voets O, Verrijzer CP, Shilatifard A. Enhancer-associated H3K4 monomethylation by Trithorax-related, the *Drosophila* homolog of mammalian Mll3/Mll4. *Genes Dev.* 26(23), 2604-20 (2012)
- 74) Ho Y, Elefant F, Cooke N, Liebhaber S. A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. *Mol Cell.* 9(2), 291-302 (2002)
- 75) Ho Y, Elefant F, Liebhaber SA, Cooke NE. Locus control region plays an active role in long-range gene activation. *Mol Cell.* 23(3), 365-375 (2006)
- 76) Hon GC, Song CX, Du T, Jin F, Selvaraj S, Lee AY, Yen CA, Ye Z, Mao SQ, Wang BA, Kuan S, Edsall LE, Zhao BS, Xu GL, He C, Ren B. 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. *Mol Cell.* S1097-2765(14)00681-9 (2014)
- 77) Hong W, Nakazawa M, Chen YY, Kori R, Vakoc CR, Rakowski C, Blobel GA. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J.* 24(3) 2367-78 (2005)
- 78) Hughes AL, Rando OJ. Mechanisms underlying nucleosome positioning in vivo. *Annu Rev Biophys.* 43,41-63 (2014)
- 79) Huo X, Zhang J. Important roles of reversible acetylation in the function of hematopoietic transcription factors. *J Cell Mol Med.* 9(1) 103-112 (2005)
- 80) Ito M, Yuan CX, Okano HJ, Darnell RB, Roeder RG. Involvement of the TRAP220 component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol Cell.* 5(4), 683-93 (2000)
- 81) Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* 33 Suppl, 245-54 (2003)
- 82) Jansen A, Verstrepen KJ. Nucleosome positioning in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev.* 75(2), 301-20 (2011)

- 83) Jiang C, Pugh F. Nucleosome positioning: advances through genomics. *Nat Rev Genet.* 10, 161-172 (2009)
- 84) Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. H3.3/H2A.Z double variant-containing nucleosomes mark nucleosome free regions of active promoters and other regulatory regions. *Nat Genet.* 41(8) 941-945 (2009)
- 85) Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* 21(12) 1519-1529 (2007)
- 86) Juven-Gershon T, Hsu JY, Theisen JW, Kadonaga JT. The RNA polymerase II core promoter: the gateway to transcription. *Curr Opin Cell Biol.* 3,253-259 (2008)
- 87) Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA. Mediator and cohesin connect gene expression and chromatin architecture. *Nature.* 467(7314), 430-5 (2010)
- 88) Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, Chun HB, Tough DF, Prinjha RK, Benner C, Glass CK. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell.* 51:310-25 (2013)
- 89) Kamakaka R. Silencers and locus control regions: opposite sides of the same coin. *Trends Biochem Sci.* 22(4), 124-128 (1997)
- 90) Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci.* 106(28), 11667-72 (2009)
- 91) Khoury G, Gruss P. Enhancer elements. *Cell.* 33, 313-314 (1983)
- 92) Kim SI, Bresnick EH, Bultman SJ. BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nucleic Acids Res.* 37(18), 6019-27 (2009)
- 93) Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. *Nature.* 465(7295), 182-187 (2010)

- 94) Kircher M, Kelso J. High-throughput DNA sequencing-concepts and limitations. *Bioessays*. 32(6):524-36 (2010)
- 95) Knezetic JA, Felsenfeld G. Identification and characterization of a chicken α -globin enhancer. *Mol Cell Biol*. 9, 893-901 (1989)
- 96) Ko MS, Nakauchi H, Takahashi N. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J*. 9, 2835-2842 (1990)
- 97) Kornberg RD. Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci*. 30(5), 235-9 (2005)
- 98) Kouzarides T. Chromatin modifications and their function. *Cell*. 128,693-705 (2007)
- 99) Kress C, Ballester M, Devinoy E, Rijnkels M. Epigenetic modifications in 3D: nuclear organization of the differentiating mammary epithelial cell. *J Mammary Gland Biol Neoplasia*. 15(1),73-83 (2010)
- 100) Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, Richards DP, Wu X, Emili A, Hughes TR, Buratowski S, Greenblatt JF. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell*. 12(6), 1565-76 (2003)
- 101) Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhhattar R. Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature*, 494(7438):497-501 (2013)
- 102) Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, Lee CY, Watt A, Grossman TR, Rosenfeld MG, Evans RM, Glass CK. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature*. 498(7455), 511-5 (2013).
- 103) Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet*. 8(2), 104-15 2007

- 104) Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. A high resolution atlas of nucleosome occupancy in yeast. *Nature Genetics* 39, 1235-1244 (2008)
- 105) Leung D, Jung I, Rajagopal N, Schmitt A, Selvaraj S, Lee AY, Yen CA, Lin S, Lin Y, Qiu Y, Xie W, Yue F, Hariharan M, Ray P, Kuan S, Edsall L, Yang H, Chi NC, Zhang MQ, Ecker JR, Ren B. Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature*. 518(7539):350-4 (2015)
- 106) Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell*. 128(4), 707-719 (2007)
- 107) Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature*. 498, 516-520 (2013)
- 108) Lorch Y, Griesenbeck J, Boeger H, Maier-Davis B, Kornberg RD. Selective removal of promoter nucleosomes by the RSC chromatin-remodelling complex. *Nat Struct Mol Biol*. 18(8), 881-885 (2011)
- 109) Lu Q, Wallrath LL, Elgin SC. The role of a positioned nucleosome at the *Drosophila melanogaster* hsp 26 promoter. *EMBO J*. 14, 4738-4746 (1995)
- 110) Malik S, Roeder RG. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet*. 11(11), 761-72 (2010)
- 111) Marsman J, Horsfield JA. Long distance relationship: enhancer-promoter and dynamic gene transcription. *Biochim Biophys Acta*. 1819(11-12), 1217-1227 (2012)
- 112) Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet*. 7, 29-59 (2006)
- 113) Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I, Pugh BF. Nucleosome organization in the *Drosophila* genome. *Nature*. 453, 358-364 (2008)
- 114) Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, Gering M, Patient R, Hostert A, Grosveld F. Novel binding partners of Lbd1 are required for haematopoietic development. *Development*. 133(24), 4913-23 (2006)

- 115) Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmc enriched within active genes and accessible chromatin in the nervous system. *Cell*. 151:1417-30 (2012)
- 116) Merika M, Williams AJ, Chen G, Collins T, Thanos D. Recruitment of CBP/p300 by the INF β enhanceosome is required for synergistic activation of transcription. *Mol Cell*. 2, 277-287(1998).
- 117) Merika M, Thanos D. Enhanceosomes. *Curr Opin Genet Dev*. 11(2), 205-8 (2001)
- 118) Miccio A, Blobel GA. Role of the GATA-1/FOG-1/NurD pathway in the expression of human beta-like globin genes. *Mol Cell Biol*. 30(14), 3460-70 (2010)
- 119) Misteli T. Beyond the sequence: cellular organization of genome function. *Cell*. 128(4), 787-800 (2007)
- 120) Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartorelli V. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol Cell*. 51(5),606-17 (2013)
- 121) Mousavi K, Zare H, Koulis M, Sartorelli V. The emerging roles of eRNAs in transcriptional regulatory networks. *RNA Biol*. 11(2), 106-10 (2014)
- 122) Natesan S, Gilman MZ. DNA bending and orientation-dependent function of YY1 in the c-fos promoter. *Genes Dev*. 7(12B), 2497-509 (1993)
- 123) Natoli G, Andrau JC. Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet*. 46, 1-19 (2012)
- 124) Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Blüthgen N, Dekker J, Heard E. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485(7398), 381-5 (2012)
- 125) van Berkum NL, Lieberman-Aiden E, Williams L, Imakaev M, Gnirke A, Mirny LA, Dekker J, Lander ES. Hi-C: A Method to Study the Three-dimensional Architecture of Genomes. *J Vis Exp*. 39, 1869 (2010).

- 126) Ogbourne S, Antalis T. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem.* 331 (1), 1-14 (1998)
- 127) Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet.* 12(4), 283-93 (2011)
- 128) Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R. Long noncoding RNAs with enhancer-like function in human cells. *Cell.* 143, 46-58 (2010)
- 129) Orom UA, Shiekhattar R. Long-noncoding RNAs usher in a new era in the biology of enhancers. *Cell.* 154(6), 1190-3 (2013)
- 130) Ozonov EA, van Nimwegen E. Nucleosome free regions in yeast promoters result from competitive binding of transcription factors that interact with chromatin modifiers. *PLoS Comput Biol.* 9(8), e1003181 (2013)
- 131) Panne D. The enhanceosome. *Curr opin struct biol.* 18, 236-242 (2008)
- 132) Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell.* 144(2), 200-13 (2011)
- 133) Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, Skotte J, Wutz A, Porse B, Jensen ON, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* 38, 4958–4969 (2010)
- 134) Patrinos GP, de Krom M, de Boer E, Langeveld A, Imam AM, Strouboulis J, de Laat W, Grosveld FG. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev.* 18, 1495–1509 (2004)
- 135) Patrusky B. Gene enhancers. *Mosaic.* 17(3), 36-44 (1986)
- 136) Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, Imbert J, Andrau JC, Ferrier P, Spicuglia S. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J.* 30(20), 4198-210 (2011)

- 137) Pencovich N, Jaschek R, Tanay A, Groner Y. Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. *Blood*. 117, e1-e14 (2011)
- 138) Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, Bland MJ, Wagstaff W, Dalton S, McDevitt TC, Sen R, Dekker J, Taylor J, Corces VG. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*. 153(6), 1281-95 (2013)
- 139) Plank JL, Dean A. Enhancer function: mechanistic and genome-wide insights come together. *Cell*. 55 (1), 5-14 (2014)
- 140) Pott S, Lieb JD. What are super-enhancers? *Nat Genet*, 47, 8-12 (2015)
- 141) Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. Locus Control Regions. *Blood*. 100(9), 3076-3086 (2002)
- 142) Qureshi IA, Mehler MF. Impact of nuclear organization and dynamics on epigenetic regulation in the central nervous system: implications for neurological disease states. *Ann N Y Acad Sci*. 1204, E20-37 (2010)
- 143) Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 470(7333), 279-283 (2011)
- 144) Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159, 1665-1680 (2014)
- 145) Targa FR, de Moura Gallo CV, Huesca M, Scherrer K, Marcaud L. Silencer and enhancer elements located at the 3' side of the chicken and duck α globin encoding gene domains. *Gene*. 129, 229-237 (1993)
- 146) Recillas-Targa F, Escamilla-del-Arenal M. Participación de la estructura de la cromatina en la regulación de la expresión génica. *Mensaje bioquímico*. 23, 173-200 (2004)

- 147) Rincón-Arano H, Valadez-Graham V, Guerrero G, Escamilla-Del-Arenal M, Recillas-Targa F. YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side α -Globin Enhancer Activity. *J Mol Biol.* 349(5), 961–975 (2005)
- 148) Risley MD, Clowes C, Yu M, Mitchell K, Hentges KE. The Mediator complex protein Med31 is required for embryonic growth and cell proliferation during mammalian development. *Dev Biol.* 342(2), 146-56 (2010)
- 149) Robertson KD. DNA methylation and human disease. *Nat Rev Genet.* 6(8), 597-610 (2005)
- 150) Roh TY, Cuddapah S, Zhao K. Active chromatin domains are defined by acetylation islands revealed by genome wide mapping. *Genes Dev.* 19(5), 542-552 (2005)
- 151) Roh TY, Wei G, Farrell CM, Zhao K. Genome wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res.* 17(1), 74-81 (2007)
- 152) Rönnerblad M, Andersson R, Olofsson T, Douagi I, Karimi M, Lehmann S, Hoof I, de Hoon M, Itoh M, Nagao-Sato S, Kawaji H, Lassmann T, Carninci P, Hayashizaki Y, Forrest AR, Sandelin A, Ekwall K, Arner E, Lennartsson A; FANTOM consortium. Analysis of the DNA methylome and transcriptome in granulopoiesis reveals timed changes and dynamic enhancer methylation. *Blood.* 123(17), e79-89 (2014)
- 153) Sandaltzopoulos R, Becker PB. Heat shock increases the reinitiation rate from potentiated chromatin templates. *Mol Cell Biol.* 18(1), 361-367 (1998)
- 154) Sasaki K, Yagi H, Bronson RT, Tominaga K, Matsunashi T, Deguchi K, Tani Y, Kishimoto T, Komori T. Absence of fetal liver hematopoiesis in mice deficient in transcriptional coactivator core binding factor beta. *Proc Natl Acad Sci USA.* 93, 12359-12363 (1996)
- 155) Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Pereira CF, Fisher AG, Adams DJ, Flicek P, Crawford GE, Laframboise T, Tesar P, Wei CL, Scacheri PC. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet.* 6(7), e1001023 (2010)

- 156) Schöler HR, Gruss P. Cell type-specific transcriptional enhancement in vitro requires the presence of trans-acting factors. Hans R. Schöler and Peter Gruss. The *EMBO J.* 4(11), 3305-3013 (1985)
- 157) Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell.* 128(4), 735-45 (2007)
- 158) Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, Zhang K, Zhang Y. Genome-wide analysis reveals TET-and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell.* 153, 1-15 (2013)
- 159) Shones DE *et al.* Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. Dynamic regulation of nucleosome positioning in the human genome. *Cell.* 132, 887-898 (2008)
- 160) Skene PJ, Henikoff S. Histone variants in pluripotency and disease. *Development.* 140(12), 2513-24 (2013)
- 161) Smith CL, Peterson CL. ATP-dependent chromatin remodeling. *Curr Top Dev Biol.* 65,115-48 (2005)
- 162) Smith E, Shilatifard A. Enhancer biology and enhanceropathies. *Nat Struct Mol Biol.* 21(3), 210-219 (2014)
- 163) Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu GL, Jin P, He C. Genome-wide profiling of 5-Formylcytosine reveals its roles in epigenetic priming. *Cell.* 153, 1-14 (2013)
- 164) Song SH, Hou C, Dean A. A positive role for NLI/Ldb1 in long range B-globin locus control region function. *Mol Cell.* 28(5), 810-822 (2007)
- 165) Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 20(17), 2349-54 (2006)
- 166) Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, Münzel M, Wagner M, Müller M, Khan F, Eberl HC, Mensinga A, Brinkman AB, Lephikov K, Müller U, Walter J,

- Boelens R, van Ingen H, Leonhardt H, Carell T, Vermeulen M. Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives. *Cell*, 152:1146-59 (2013)
- 167) Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schübeler D. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*. 480(7378), 490-5 (2011)
- 168) Steffen PA, Ringrose L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat Rev Mol Cell Biol*. 15(5), 340-56 (2014)
- 169) Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 403(6765), 41-5 (2000)
- 170) Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol*. 12(6), R54 (2011)
- 171) Stünkel W, Kober I, Seifart KH. A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene. *Mol Cell Biol*. 17, 4397-4405 (1997)
- 172) Sun FL, Elgin SC. Putting boundaries on silence. *Cell*. 99(5), 459-62 (1999)
- 173) Sutherland HG, Martin DI, Whitelaw E. A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. *Mol Cell Biol*. 3, 1607-1614 (1997)
- 174) Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*. 116(1), 51-61 (2004)
- 175) Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Höfer T, Rippe K. Genome-wide nucleosome positioning during embryonic stem cell development. *Nat Struct Mol Biol*, 19(11):1185-92 (2012).
- 176) Thanos D, Maniatis T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell*. 83(7), 1091-100 (1995)

- 177) Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutuyavin T, Lajoie B, Lee BK, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, Navas PA, Stamatoyannopoulos G, Iyer VR, Lieb JD, Sunyaev SR, Akey JM, Sabo PJ, Kaul R, Furey TS, Dekker J, Crawford GE, Stamatoyannopoulos JA. The accessible chromatin landscape of the human genome. *Nature*. 489(7414), 75-82 (2012)
- 178) Tie F, Banerjee R, Conrad PA, Scacheri PC, Harte PJ. Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. *Mol Cell Biol*. 32(12), 2323-34 (2012)
- 179) Tie F, Banerjee R, Saiakhova AR, Howard B, Monteith KE, Scacheri PC, Cosgrove MS, Harte PJ. Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. *Development*. 141(5), 1129-39 (2014)
- 180) Tillo D, Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Field Y, Lieb JD, Widom J, Segal E, Hughes TR. High nucleosome occupancy is encoded at human regulatory sequences. *PLoS One*. 5(2), e9129 (2010)
- 181) Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*. 10(6), 1453-65 (2002)
- 182) Tolstorukov MY, Kharchenko PV, Goldman JA, Kingston RE, Park PJ. Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. *Genome Res*. 19(6), 967-77 (2009)
- 183) Tsang AP, Visvader JE, Turner CA, Fujiwara Y, Yu C, Weiss MJ, Crossley M, Orkin SH. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*. 90(1), 109-19 (1997)

- 184) Vakoc CR *et al.* Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, Groudine M, Weiss MJ, Dekker J, Blobel GA. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell*. 17(3), 453-62 (2005)
- 185) Valdes-Quezada C, Arriaga-Canon C, Fonseca-Guzmán, Guerrero G, Recillas-Targa F. CTCF demarcates chicken embryonic α -globin gene autonomous silencing and contributes to adult stage-specific gene expression. *Epigenetics*. 8(8), 827-38 (2013)
- 186) van Berkum NL, Lieberman-Aiden E, Williams L, Imakaev M, Gnirke A, Mirny LA, Dekker J, Lander ES. Hi-C: a method to study the three-dimensional architecture of genomes. *J Vis Exp*. 6;(39) (2010)
- 187) van Riel B, Pakozdi T, Brouwer R, Monteiro R, Tuladhar K, Franke V, Bryne JC, Jorna R, Rijkers EJ, van Ijcken W, Andrieu-Soler C, Demmers J, Patient R, Soler E, Lenhard B, Grosveld F. A novel complex, RUNX1-MYEF2, represses hematopoietic genes in erythroid cells. *Mol Cell Biol*. 32(19), 3814-22 (2012)
- 188) Vignali M, Hassan AH, Neely KE, Workman JL. ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol*. 20(6), 1899-910 (2000)
- 189) Vinci MC. Sensing the environment: Epigenetic regulation of gene expression. *J Physic Chem Biophysic*. S3, 001 (2011)
- 190) Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. *Nature*. 461(7261), 199-205 (2009)
- 191) Vyas P, Vickers MA, Picketts DJ, Higgs DR. Conservation of position and sequence of a novel, widely expressed gene containing the major human α -globin regulatory element. *Genomics*. 29, 679–689 (1995)
- 192) Walker MD, Edlund T, Boulet AM, Rutter WJ. Cell specific expression controlled by the 5' flanking region of insulin and chymotrypsin genes. *Nature*. 306, 557-561 (1983)
- 193) Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, Martin DI. Enhancers increase the probability but not the level of gene expression. *Proc Natl Acad Sci USA*. 92(15), 7125-7129 (1995)

- 194) Wang Q, Carroll JS, Brown M. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol Cell.* 19(5), 631-642 (2005)
- 195) Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet.* 40(7), 897-903 (2008)
- 196) West AG, Fraser P. Remote control of gene expression. *Hum Mol Genet.* 14(1), 101-111 (2005)
- 197) Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 153(2), 307-19 (2013)
- 198) Xu Z, Meng X, Cai Y, Koury MJ, Brandt SJ. Recruitment of the SWI/SNF protein Brg1 by a multiprotein complex effects transcriptional repression in murine erythroid progenitors. *Biochem J.* 399(2), 297-304 (2006)
- 199) Yasui D, Miyano M, Cai S, Varga-Weisz P, Kohwi-Shigematsu T. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*, 419(6907):641-5 (2002)
- 200) Yie J, Senger K, Thanos D. Mechanism by which the IFN- β enhanceosome activates transcription. *Proc Natl Acad Sci USA.* 96(23), 13108-13113 (1999)
- 201) Yoshimura FK, Davison B, Chaffin K. Murine leukemia virus long terminal repeat sequences can enhance gene activity in a cell-type-specific manner. *Mol Cell Biol.* 5(10), 2832-2835 (1985)
- 202) Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min JH, Jin P, Ren B, He C. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell.* 149(6), 1368-80 (2012)
- 203) Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science.* 309(5734), 626-30 (2005)
- 204) Yuh CH, Bolouri H, Davidson EH. Genomic cis-regulatory logic: Experimental and computational analysis of sea urchin gene. *Science.* 279 (5358), 1896-1902 (1998)

- 205) Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell-Res*, 21(4), 564-578 (2011)
- 206) Zabidi MA, Arnold CD, Schernhuber K, Pagai M, Rath M, Frank O, Stark A. Enhancer core promoter specificity separates developmental and housekeeping gene regulation. *Nature*. 518(7540), 556-559 (2015)
- 207) Zentner GE, Henikoff S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol*. 20(3), 259-266 (2013)
- 208) Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res*. 21(8), 1273-1283 (2011)
- 209) Zhang Y, Moqtaderi Z, Rattner BP, Euskirchen G, Snyder M, Kadonaga JT, Liu XS, Struhl K. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol*. 16(8), 847-853 (2009)
- 210) Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*. 12(1), 7-18 (2011)

XV. Anexos

XV.1 Materiales y métodos

XV.1.1 Plásmidos y construcciones

Los sub-fragmentos de DNA que contienen la región 3' del dominio α -globina de pollo se amplificaron por PCR a partir de DNA genómico de eritrocitos circulantes de 10 días de desarrollo de embriones de *G. gallus* (Alpes, Puebla, México) clonado en el vector pG α D3 que contiene *GFP* como gen reportero bajo el control del promotor α^D . La secuencia ERAE se obtuvo usando los siguientes cebadores: FSil - PstI : 5' - GGCAACGCTGCAGCACTCCTTACCCCATGA - 3' y 5' - FSilPstI : GGCAACGCTGCAGCACACAGCACCCAGCA - 3' . Las escisiones del ERAE se obtuvieron utilizando los siguientes cebadores y se sub-clonaron en el vector pG α D3 : ASIL1 : 5'CGCGGATCCCATGCACTCCTTACCCCATG -3' , DSil1.2 : 5'-GCGGGATCCCAACAGGAACCATCAGCACTTGGG -3', DSil2 : 5' - GCGGGATCCCTGCTCTTTGGGCTATGACCCTC - 3' y A120P4 : 5' - CGCGGATCCGACGTGGGCAGCAGATAGCCTCG -3' Los transgenes mutantes que aumentan la distancia entre ERAE y *enhancer* se obtuvieron mediante la creación de una mutación (generada por PCR con el cebador: FMIDSE : 5'-CGCGGACGTCAGTCACAGATCATGGTTACGACGTCGCGC - 3' y RMIDSE : 5' - GCGCGACGTCGTAACCATGATCTGTGACTGACGTCCGGGGGGCG - 3') que incorpora un nuevo sitio de restricción AatII donde se clonaron uno, dos o tres copias de 25 pares de bases de DNA proveniente del genoma del fago lambda. El plásmido pG α D3SE mutante se secuenció para confirmar la mutación y la secuencia insertada es: 5' - CAGTCACAGATCATGGTTACGACGT - 3' .

XV.1.2 Cultivo celular

Las células HD3 derivan de pollo (*Gallus gallus*) y corresponden a eritroblastos transformados por virus, sensibles a temperatura. Estas células se cultivaron a 37°C en *Dulbecco Modified*

Eagles Medium (DMEM, GIBCO) suplementado con 10 % de suero fetal bovino, 2 % de suero de pollo y antibióticos (100 U *Penicillin* / 0.1 mg estreptomina /ml) en una incubadora con CO₂ (Boyes y Felsenfeld, 1996; Rincón-Arano *et al*, 2005)

XV.1.3 Transfecciones transitorias y estables

Las células HD3 se lavaron dos veces con buffer de fosfato salino (PBS) y se resuspendieron en medio DMEM. Posteriormente se contaron en una cámara de Nuebauer y se usaron 1×10^6 células HD3, que se transfirieron a una placa de 6 pozos y se transfectaron con 1 μ g (1 μ g/ μ l) del plásmido correspondiente, además de 2 μ l de *Lipofectamine*[®] 2000 como el protocolo indica. Después de 6 horas, se añadieron 4 ml de medio no selectivo a las células transfectadas. Después de 48 horas de recuperación, las células se utilizaron para medir la actividad de luciferasa en el caso de ensayo de transfección transitoria o bien fueron transferidas a un medio cuya matriz de celulosa contiene G-418 (a una concentración de 0.9mg/ml) para la selección de clonas estables (Rincón-Arano *et al*, 2005; Furlan-Magaril *et al*, 2011).

XV.1.4 Aislamiento de clonas

Después de 48 horas de transfección y recuperación de células en medio sin antibiótico, aproximadamente 5×10^5 (500 μ l) de las células transfectadas se transfirieron a una matriz de celulosa (*Methocel*, Fluka) que contiene 0.9 mg/ml de G-418 (geneticina, Calbiochem), una cantidad que previamente ha demostrado matar todas las células no transfectadas para aislar colonias individuales que fueron seleccionadas después de 2-3 semanas para su posterior expansión en medio líquido adicionado con antibiótico durante la primer semana, luego se retiró y se llevaron a cabo los siguientes experimentos (Furlan-Magaril *et al*. 2011).

XV.1.5 Citometría de flujo

Se obtuvo 1 ml de células provenientes de una misma línea transgénica a una confluencia de 70-90%. Posteriormente se centrifugaron a 3000 rpm y se retiró el sobrenadante. Las células se resuspendieron en 500 μ l del buffer FACS FLOW[®]. En el citómetro BDFacscalibur[®] los

parámetros tamaño y granulosidad de las células se ajustaron usando como blanco las células HD3 wt. Los datos correspondientes a la intensidad media de fluorescencia y el porcentaje de células fluorescentes se analizaron con el programa WinMidi®.

XV.1.6 Southern blot

El DNA genómico se obtuvo después de la incubación con RNasa (0.125 mg/ml de RNasa A) durante 1 hora a 37° C y posteriormente se incubó con 0.5 mg / ml de proteinasa K y SDS al 0.5% durante la noche. Se digirieron 10 µg de cada muestra de DNA con las endonucleasas de restricción correspondientes (2 U/µg) que se indican en cada experimento particular. Después de la digestión, las muestras corrieron en geles de agarosa al 1 % a 110 V durante 90 minutos en TBE 1X. El DNA se transfirió a una membrana (Amersham Hybond-XL, GE Healthcare) en buffer de NaOH 0.5 M durante toda la noche. La sonda corresponde a un fragmento de 741 pb de DNA de la secuencia del gen *GFP* obtenido del plásmido pEGFP-1 mediante una restricción enzimática BamHI/NotI. La sonda se marcó radiactivamente con Random Primers (sistema de marcaje de DNA RadPrime, Invitrogen). La hibridación se realizó a 65° C durante la noche en 30 % de SSC 20X, Denhart y 1 % de SDS, pH 7,2. Dos lavados se realizaron a 65° C en 1 % de SDS y 10 % de SSC 20X o 1 % de SSC 20X (Boyes y Felsenfeld, 1996).

XV.1.7 Microscopía de Fluorescencia

Las células HD3 transfectadas con los plásmidos correspondientes que expresan a la proteína *GFP* hace posible que puedan ser observadas *in vivo* al microscopio de fluorescencia sin ningún tratamiento adicional. 24 horas antes de observar las células en el microscopio confocal FV10, las células HD3 se lavaron dos veces con medio DMEM. Posteriormente las células se resuspendieron y contaron en una cámara de Neubauer. Se usaron 1×10^6 células transfectadas para cada plásmido y se transfirieron a una placa de 6 pozos. Al día siguiente las células se transportaron a la unidad de microscopía del IFC envueltas en papel aluminio y condiciones de esterilidad para observarse en el microscopio de fluorescencia. Las fotografías se tomaron en aumento de 40X.

XV.1.8 Ensayo de movilidad electroforética (EMSA) y súper-retardo

Células HD3 se lavaron dos veces con PBS y se resuspendieron 3.5×10^7 células en 2 ml de buffer de NI (Tris - HCl 1 M pH 7.5, sacarosa 900 mM, HEPES 150 mM, KCl 900 mM, MgCl 40 mM de DTT 1 M, NaCl 450 mM) más 2 ml de buffer de lisis durante 1 minuto para obtener núcleos aislados (se monitorizó la integridad de los núcleos al añadir NP-40). La lisis de núcleos se realizó utilizando RIPA (1 % de NP-40) y la concentración de proteína se determinó por el ensayo de Bradford. Los oligonucleótidos utilizados para este ensayo fueron radiactivamente marcados con ^{32}P usando PNK (Biolabs). Para cada ensayo se utilizaron 15-20 μg de proteína nuclear y se incubó con poli (di-dC) como un competidor no específico durante 15 minutos en hielo y 15 minutos a 25°C ; a continuación, las sondas correspondientes ($1\mu\text{l} = 20000$ cpm) fueron añadidas e incubadas en hielo 15 minutos, y 15 minutos a temperatura ambiente. Después de las incubaciones, las muestras se cargaron en un gel de acrilamida al 5 % sin urea y se aplicó 150 V durante 105 minutos usando TBE 0.25X (Valadez-Graham *et al*, 2004).

XV.1.9 Digestión con la endonucleasa MNasa

Las células HD3 se lavaron dos veces con PBS y se resuspendieron 3.5×10^7 células en 2 ml de buffer NI (Tris-HCl 1 M pH 7.5, sacarosa 900 mM, HEPES 150 mM, KCl 900 mM, MgCl 40 mM de DTT 1 M, NaCl 450 mM) más 2 ml de buffer de lisis durante 1 minuto para obtener núcleos aislados (NI más 0.5 % de NP- 40, en este paso se monitorizó la integridad de los núcleos). Los núcleos se lavaron dos veces con 10 ml de NI y se resuspendió en 300 μl de NI. Para las digestiones con nucleasa micrococcal (MNasa, Worthington), 100 μl de núcleos se utilizaron para cada reacción con CaCl_2 (2 μl , 0,3 M), se añadieron NI y enzima MNasa a un volumen final de 300 μl . Los núcleos se digirieron con 0, 20 y 40 U de MNasa, respectivamente, durante 5 min a 25°C . Las digestiones se detuvieron mediante la adición de EDTA a 10 mM, seguido por la adición de RNasa (0.125 mg/ml de RNasa A) durante 1 hora a 37°C y se incubaron posteriormente con 0.5 mg/ml de proteinasa K y SDS al 0.5 % durante la noche. El DNA se extrajo con fenol-cloroformo y una vez con cloroformo. El DNA se precipitó con etanol y se resuspendió en H_2O . Se digirieron 50 μg de cada muestra de DNA con BamHI seguido de la

restricción con la enzima NaeI (Boyes y Felsenfeld, 1996, Escamilla-Del-Arenal y Recillas-Targa F, 2007).

XV.1.10 Inmunoprecipitación de cromatina

Las células HD3 se obtuvieron a partir de cultivos celulares con un 80-90% de confluencia. Se lavaron dos veces con PBS centrifugando a 1200 rpm durante 5 min. Se utilizaron 10×10^6 células (cantidad suficiente para precipitar con 10 anticuerpos diferentes) y se resuspendieron en 30 ml de PBS. Se agregaron 3 ml de solución de formaldehído al 11 % (Formaldehído al 11 %, 0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM Hepes) con movimiento constante durante 10 minutos. A continuación se añade glicina (2.5 mM) a una concentración final del 125 mM durante 5 min para detener el entrecruzamiento, se agitan los tubos suavemente y se centrifugan durante 5 min a 1200 rpm a 4°C. Lavar 2 veces las células con PBS frío centrifugando cada vez a 1200 rpm por 5 min. El pellet de células se resuspende en 1 ml de buffer de lisis (1 % SDS, 5 mM EDTA, 50 mM de TRIS-HCl pH 8.1) con inhibidores de proteasas (para preparar 5 ml usar 1 µl Leupeptina [1 mg/ml], 0.1 µl Aprotinina [10 mg/ml], 50 µl Butirato de Sodio [1 M], 50 µl PMSF [10 mg/ml], 5 µl Bensamidina [100 mM], 5 µl Pepstatina [1 mg/ml]). A continuación, fragmenta (sonicar) la muestra hasta obtener el tamaño de fragmentos deseada (300-400 pb). Centrifugar muestra a 4°C por 5 min a 11000 rpm y recuperar el sobrenadante. Guardar una alícuota de 50 µl como Input. Hacer una dilución 1:10 de la muestra con buffer de dilución (1 % Tritón X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl) con inhibidores de proteasas (para preparar 20 ml usar 4 µl Leupeptina [1 mg/ml], 0.4 µl Aprotinina [10 mg/ml], 200 µl Butirato de Sodio [1 M], 200 µl PMSF [10 mg/ml], 20 µl Bensamidina [100 mM], 20 µl Pepstatina [1 mg/ml]). Usar 1 ml de cromatina soluble en un tubo eppendorf para cada inmunoprecipitación. Agregar el anticuerpo específico e incubar durante toda la noche a 4°C en agitación constante. Agregar 30µl de perlas de sefarosa A/G (0.2 g/100 µl) previamente bloqueadas con BSA (100 µg/ml) y esperma de salmón (500 µg/ml) e incubar durante 2-4 horas en agitación constante a 4°C. Se recuperan las perlas y se lavan secuencialmente por 10 min en agitación constante a 4° en los buffers Paro I (0.1 % SDS, 1 % Tritón X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), Paro II (0.1 % SDS, 1 % Tritón

X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl) y Paro III (0.25 M LiCl, 1 % NP-40, 1 % Desoxicolato, 1 mM EDTA, 10 mM Tris-HCl pH 8.1) centrifugando cada vez 1 minuto a 2000 rpm a 4°C. Hacer 2 lavados con buffer TE (10 mM Tris-HCl, pH 7.4-8.0, 1 mM EDTA pH8) centrifugando cada vez 1 minuto a 2000 rpm a 4°C. Agregar 150 µl de solución de lavado (0.1 M Bicarbonato de Sodio, 1 % SDS) a las perlas. Añadir proteinasa K (20 µg/µl) y RNasa (5 µg/µl) a cada muestra. Incubar las muestras a 37°C durante 1 hora, posteriormente incrementar la temperatura a 65°C durante 6 horas para revertir el entrecruzamiento. Extraer el DNA y eluir en 50µl de agua. Realizar las PCRs con primers específicos.

XV.2 Tabla 1. Factor Fano

Promotor α^D	Intensidad media de Fluorescencia (Células GFP positivas)	Potenciador	Intensidad media de Fluorescencia (Células GFP positivas)	ERAP-Potenciador	Intensidad media de Fluorescencia (Células GFP positivas)
1	222.72	1	256.72	1	27.22
2	365.49	2	237.25	2	78.11
3	66.68	3	53.14	3	181.89
4	67.22	4	48.71	4	34.33
5	43.59	5	325.88	5	32.99
6	40.6	6	88.16	6	15.23
7	17.65	7	212.65	7	44.37
8	325.88	8	64.89	8	30.67
9	17.09	9	418.33	9	17.09
10	64.89	10	65.17	10	164.6
Promedio	123.181	Promedio	177.09	Promedio	62.65
DS	124.489689	DS	125.07133	DS	57.852626
VAR	15497.6827	VAR	15642.8376	VAR	3346.92634
Factor Fano (VAR/Promedio)	125.812282	FF	88.3326986	FF	53.4226072



A regulatory element affects the activity and chromatin structure of the chicken α -globin 3' enhancer



Estela García-González, Félix Recillas-Targa *

Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Ciudad de México, DF, Mexico

ARTICLE INFO

Article history:

Received 15 May 2014

Received in revised form 9 September 2014

Accepted 10 September 2014

Available online 17 September 2014

Keywords:

Globin genes

Transcription

Enhancer

Chromatin

Nucleosome

Transcription factor

ABSTRACT

Gene promoters are frequently insufficient to drive the spatiotemporal patterns of gene expression during cell differentiation and organism development. Enhancers convey these properties through diverse mechanisms, including long-distance interactions with target promoters via their association with specific transcription factors. Despite unprecedented progress in the knowledge of enhancer mechanisms of action, there are still many unanswered questions. In particular, the contribution of an enhancer's local chromatin configuration to its mechanism of action is not completely understood. Here we describe a novel regulatory element, the Upstream Enhancer Element (UEE), which modulates the activity of the chicken α -globin 3' enhancer by regulating its chromatin structure, specifically by positioning a nucleosome upstream of the core enhancer. This element binds nuclear factors and confers a more restricted activation on the α -globin 3' enhancer, suggesting a progressive/rheostatic model for enhancer activity. Our results suggest that the UEE activity contributes to the positioning of a nucleosome that is necessary for the α -globin 3' enhancer activation.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Transcription activation by distant regulatory elements is a characteristic of many developmentally regulated gene families [1]. Notably, many metazoan genes are influenced by numerous enhancers that are capable of increasing transcriptional events at target genes, despite being located large distances away [2]. Recent advances in epigenomic profiling and genome-wide sequencing technologies have brought to light enhancer-specific chromatin signatures, such as enrichment of the H3K4me1 or H3K27ac histone modifications, the presence of the histone acetyltransferases CBP and p300, as well as evidence supporting the hypothesis that enhancers come in close proximity to their target genes with collaboration of structural protein complexes [1–4]. However, additional mechanisms of enhancer function, such as local nucleosome positioning in association with chromatin remodelers and tissue specific transcription factors, are less understood.

The notion that enhancers have the ability to increase transcriptional levels, either by promoting the number of transcription events in specific population of cells or by increasing the total number of transcriptionally active cells, has been addressed by two models [5]. In the binary model the gene can be in the “on” or “off” state, and the presence of an enhancer stimulates the active state, thereby increasing the probability that more cells will become transcriptionally active. In the

progressive model the RNA polymerase machinery loads more frequently on the same DNA template, thereby increasing the number of RNA molecules transcribed from the same gene, but not increasing the number of cells activating the gene [6].

Extensive studies of the α -globin locus have contributed to a better understanding of enhancer function and regulation during erythroid cell differentiation [7]. Indeed, our group previously described the α -globin 3' enhancer defined as a *cis*-acting DNA sequence that increase transcription accompanied by a “silencer” element [8–10]. In recent years we have re-analyzed these regulatory elements and found that in its genomic context the silencer is necessary to induce the expression of a reporter gene and seems to constitute an integral component of the α -globin 3' enhancer [9]. We determined the positioning of nucleosomes over this genomic region and found two prominent erythroid-specific micrococcal nuclease hypersensitive sites (MNase-HSs). In addition, we described two well-positioned nucleosomes that shield the binding sequences for the ubiquitous and erythroid-specific transcription factor GATA-1, and which are also correlated with an enrichment of histone acetylation [9]. Consequently, chromatin remodeling and generation of nucleosome-free regions are coincident with maximal enhancer activity [9,11].

Based on the novel activity of the initially named “silencer” element, now referred to in this work as the Upstream Enhancer Element (UEE), we show that this element takes part in the overall enhancer activity. Here, we interrogated the role of the UEE in the chicken α -globin 3' enhancer activity and chromatin structure. In randomly integrated genomic contexts we assessed the role of the UEE in α -globin 3' enhancer activity by employing *GFP* as a reporter gene. Our results show that

* Corresponding author at: Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México, DF 04510, Mexico. Tel.: +52 55 56 22 56 74; fax: +52 55 56 22 56 30.

E-mail address: frecilla@ifc.unam.mx (F. Recillas-Targa).

the UEE activity seems to be context-dependent and has a negative effect on transcription, reducing the number of transcriptionally active cells and especially the fluorescence intensity (burst size) in erythroid cells. Sequential upstream deletion of the 360 bp UEE element causes a drastic reduction in the number of cells expressing *GFP*. Finally, changes in the distance between the two MNase-HSs by insertion of unrelated sequences also cause a loss of enhancer activity by affecting the positioning of the 5' nucleosome that shields the enhancer. In summary, we describe a sequence that is needed for the α -globin 3' enhancer activity of the chicken α -globin locus.

2. Materials and methods

2.1. Plasmid constructs

DNA sub-fragments containing the chicken 3' α -globin regulatory region (UEE-enhancer) were PCR-amplified from 10-day-old embryonic erythrocyte (10dRBC) genomic DNA of *Gallus gallus* (Alpes, Puebla, México), and cloned in pG α D3 containing *GFP* as a reporter gene under the control of the α^D promoter. UEE sequence was obtained using the following primers: FSil-PstI: 5'-GGCAACGCTGCAGCACTCCTACCCCATGA-3' and FSilPstI: 5'-GGCAACGCTGCAGCACAGCACCA GCA-3'. Deletions were obtained using the following primers and sub-cloned into the pG α D3 construct: ASil1: 5'CGCGGATCCCATGCACCTCTACCCCATG-3', DSil1.2: 5'-GC GGGATCCCAACGGAACCATCAGCACTTGGG-3', DSil2: 5'-GCGGGATCCCTG CTCCTTGGGCTATGACCTC-3' and A120P4: 5'-CGCGGATCCGACGTGGGCGAG CAGATAGCCTCG-3'. The mutant transgenes that increase the distance between the UEE and enhancer were obtained by creating a mutation (generated by PCR with the primer: FMIDSE: 5'-CGCGGACGTACGACAGATCATGGTTACGACGTCCGCG-3' and RMIDSE: 5'-GCGCGACGTGTAACCATGATCTGTGACTGACGTCCGGGGG GCG-3') that incorporates a novel AatII restriction site in which one, two or three copies of the 25 bp lambda-DNA were cloned. The mutant plasmid pG α D3SE was sequenced to confirm the mutation, and the inserted sequence is: 5'-CAGTCACAGATCATGGT TACGACGT-3'.

2.2. Cell culture

HD3 cells are derived from chicken (*G. gallus*) and correspond to an erythroblast virus-transformed and temperature-sensitive cell line. They were grown at 37 °C in Dulbecco's Modified Eagles Medium (DMEM, GIBCO) supplemented with 10% bovine fetal serum, 2% chicken serum and antibiotics (100 U Penicillin/0.1 mg Streptomycin/ml) in a CO₂ incubator [8,12].

2.3. Transfections

HD3 cells were washed twice with phosphate buffered saline (PBS) and resuspended in DMEM media. 1×10^6 HD3 cells were then transferred to a 6-well plate and transfected with 1 μ g (1 μ g/ μ l) of corresponding plasmid plus 2 μ l of Lipofectamine® 2000, as indicated by the protocol. After 6 h, 4 ml of non-selective medium was added to transfected cells. Following 48 h of cell recovery, luciferase activity was measured in the case of transient transfection assays, or else the cells were transferred to a cellulose matrix media containing G-418 Sulfate for selection of stable clones [8,13].

2.4. Cell clone isolation

After 48 h of cell recovery, approximately 5×10^5 (500 μ l) of transfected cells was transferred to a cellulose matrix (Methocel, FLuka) containing 0.9 mg/ml of G-418 (Geneticin, antibiotic from Calbiochem), an amount previously shown to kill all of the non-transfected cells. Isolated individual colonies were picked after 2–3 weeks and expanded in liquid medium containing the antibiotic to perform subsequent experiments [13].

2.5. Southern blot

Genomic DNA was obtained after incubation with RNase (0.125 mg/ml RNase A) for 1 h at 37 °C, and incubated subsequently with 0.5 mg/ml proteinase K and 0.5% SDS overnight. Ten microgram of each DNA sample was digested with the appropriate restriction endonuclease (2 U/ μ g) as indicated for the particular experiment. Following digestion, samples were loaded in 1% agarose gels at 110 V for 90 min in 1 \times TBE. The DNA was transferred to a membrane (Amersham Hybond-XL, GE Healthcare) in 0.5 M NaOH overnight. The probe is a 741 bp GFP DNA fragment obtained from the pEGFP-1 plasmid by BamHI and NotI enzyme digestion. It was radiolabeled by random primers (RadPrime DNA labeling system, Invitrogen). Hybridization of the two experimental conditions was performed in the same membrane under the same conditions at 65 °C overnight in 30% of 20 \times SSC, Denhardt's reagent and 1% SDS, pH 7.2. Two washes were performed at 65 °C in 1% SDS and 10% of 20 \times SSC or 1% of 20 \times SSC [12].

2.6. Electrophoretic mobility shift assay (EMSA) and super-shift assay

HD3 cells were washed twice with PBS and resuspended 3.5×10^7 cells in 2 ml of NI buffer (1 M Tris-HCl pH 7.5, 900 mM Sucrose, 150 mM HEPES, 900 mM KCl, 40 mM MgCl, 1 M DTT, 450 mM NaCl) plus 2 ml of lysis buffer (NI plus 0.5% NP-40, in this step nuclei integrity was monitored) for 1 min to obtain isolated nuclei. Nucleus lysis was performed using RIPA (1% NP-40) and protein concentration was determined by the Bradford assay. Oligonucleotides used for this assay were annealed and radioactively marked using PNK (BioLabs). Fifteen micrograms of total nuclear protein was incubated with Poly (dI-dC) as a non-specific competitor for 15 min on ice and 15 min at 25 °C, then an appropriate probe (1 μ l = 20,000 cpm) was added and incubated on ice for 15 min, and then for 15 min at room temperature. After incubations, samples were loaded into a 5% acrylamide gel without urea and applied 150 V for 105 min using 0.25 \times TBE [14]. For the super-shift we used 2 μ g of each antibody tested (anti-RUNX1 / AML1 + RUNX3 + RUNX2 [EPR3099] from Abcam (ab92336), IgG antibody from Santa Cruz Biotechnology (sc-2028) and CTCF antibody [anti-cCTCF₈₆₋₂₃₃] against chicken obtained as described previously [14]. The positive control sequence was obtained from a reported RUNX1 ChIP-seq assay in the K562 cell line; this sequence is located in the *FOS* gene proximal promoter region [15]. The same conditions used to perform EMSA assays were used for super-shift experiments, except that samples were loaded into a 5% acrylamide gel without urea and applied 120 V for 4 h at 4 °C in 0.25 \times TBE buffer.

2.7. MNase digestion

HD3 cells were washed twice with PBS and resuspended 3.5×10^7 cells in 2 ml of NI buffer (Tris-HCl 1 M pH 7.5, Sucrose 900 mM, HEPES 150 mM, KCl 900 mM, MgCl 40 mM DTT 1 M, NaCl 450 mM) plus 2 ml of lysis buffer (NI plus 0.5% NP-40, in this step nuclei integrity was monitored) for 1 min to obtain isolated nuclei. Nuclei were washed twice with 10 ml of NI buffer and resuspended in 300 μ l of NI buffer. For digestions with micrococcal nuclease (MNase, Worthington), 100 μ l of nuclei was used for each reaction and CaCl₂ (2 μ l, 0.3 M), NI buffer and MNase were added to a final volume of 300 μ l. Nuclei were digested with 0, 20 and 40 U of MNase, respectively, for 5 min at 25 °C. The digestions were stopped by addition of EDTA to 10 mM, followed by addition of RNase (0.125 mg/ml RNase A) for 1 h at 37 °C, and incubated subsequently with 0.5 mg/ml proteinase K and 0.5% SDS overnight. The DNA was extracted with phenol-chloroform and once with chloroform. DNA was precipitated with ethanol and resuspended in H₂O. 50 μ g of each DNA sample was digested with BamHI followed by NaeI digestion [9, 12].

2.8. Fano factor calculation

Fano factor measures the fluctuations in gene expression in terms of variance over the mean of gene expression [16]. To perform the Fano factor calculation we obtained the mean fluorescence intensity by flow cytometry measurements for the population of cells that express the *GFP* reporter gene for every single genetic copy of each transgene. We divided the variance over the mean fluorescence to determine the Fano factor [16].

3. Results

3.1. The Upstream Enhancer Element diminishes *GFP* reporter expression in the α -globin context

To determine the contribution of the UEE to 3' enhancer activity, we generated single genetic copies of each transgene containing the α -globin 3' enhancer, the UEE adjacent to the α -globin 3' enhancer, and the UEE alone under the control of the chicken α^D -globin gene

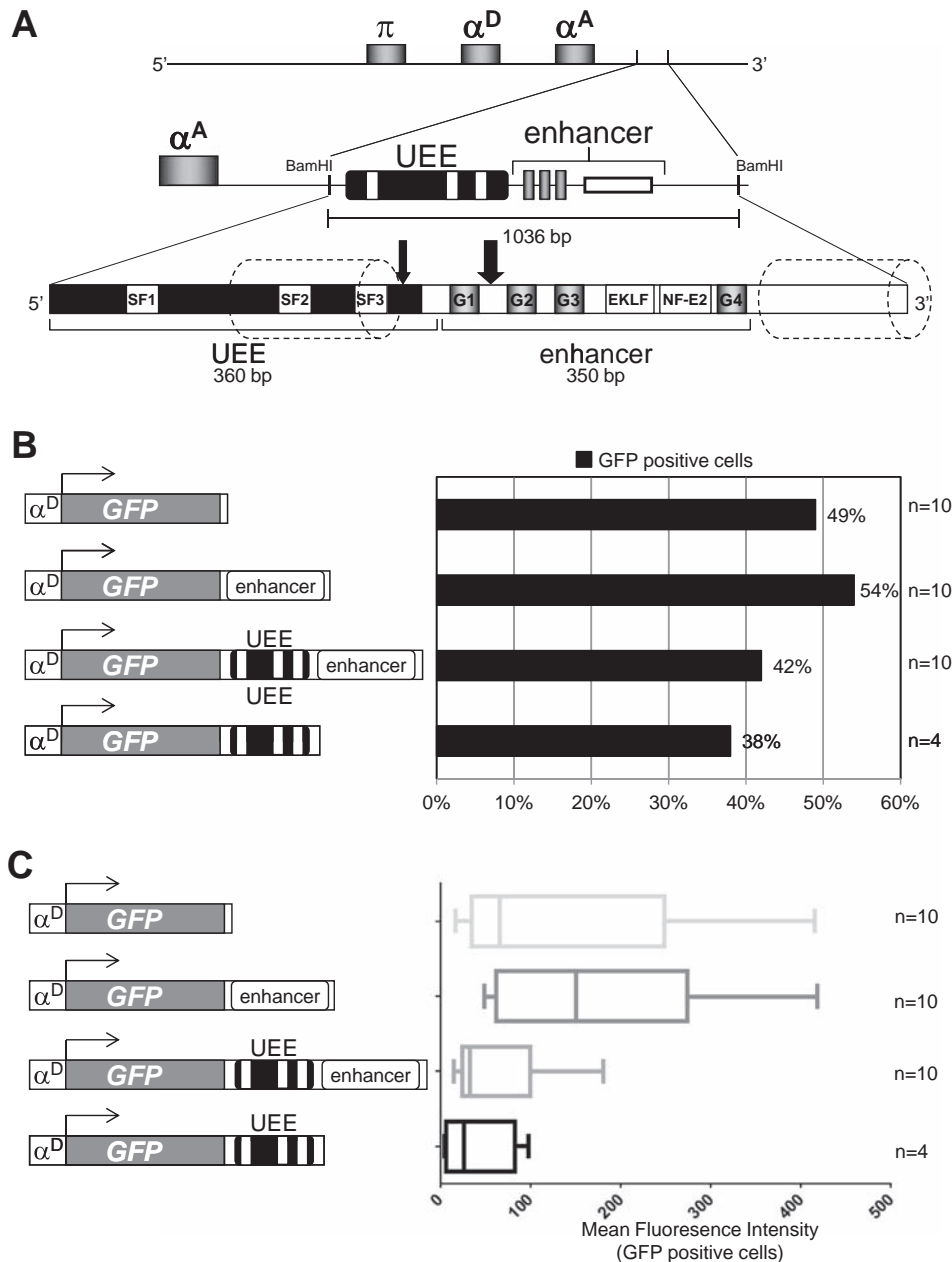


Fig. 1. Contribution of the Upstream Enhancer Element (UEE) to the chicken α -globin 3' enhancer. (A) Schematic of the chicken α -globin gene domain, the Upstream Enhancer Element (UEE) and the core enhancer, and the detailed localization of nuclear factor binding sequences. The UEE-enhancer region is located about 400 bp downstream of the chicken adult α^A gene. The UEE possesses three binding sites, SF1, SF2 and SF3 for nuclear factors defined previously by DNase footprinting [10]. Based on their binding sequences these factors appear not to be erythroid-specific and their identity remains to be determined [10]. Four GATA-1 binding sites are present in the core enhancer as well as Sp1/EKLF, NF-E2 and YY1 binding sites [8,9]. The vertical arrows and their thickness represent the position and intensity, respectively, of the MNase hypersensitive sites defined previously. In addition (dotted lines) we show the position of two well-positioned nucleosomes defined in HD3 cells and 10-day embryonic red blood cells [9]. Of note, at these stages two nucleosomes located between these two fixed nucleosomes are evicted [9]. (B) Stable lines containing single copies of each construct generated in erythroblast HD3 cells. The graph shows the percentage (%) of fluorescent cells defined by flow cytometry. Single genetic copies of each transgene were randomly integrated and selected on the basis of their integrity and verified by Southern blotting (Supplementary Fig. S1). (C) The same HD3 cell lines with their corresponding transgenes were evaluated in terms of the *GFP* mean fluorescence intensity.

promoter in the temperature-sensitive erythroblast cell line LSCCHD3, also known as HD3 cells (Supplementary Fig. S1). The activity of the regulatory sequences was assessed using the *GFP* reporter gene. Fluorescence histograms were obtained from flow cytometry measurements with the number of *GFP*-positive cells and their corresponding mean fluorescence intensity (Supplementary Fig. S2). We found a negative impact on the α -globin promoter and α -globin 3' enhancer in the presence of the UEE. The number of *GFP*-positive cells decreased slightly, by 12 percentage points when the UEE is next to the α -globin 3' enhancer (Fig. 1B), but a pronounced effect on the fluorescence intensity was observed (Fig. 1C). Therefore, the UEE mainly modulates the activity of the α -globin 3' enhancer by reducing the burst size according to the progressive/rheostatic model of enhancer activity hypothesis, rather than

affecting the burst frequency of transcription (binary model of enhancer activity) [6]. Consistently, the calculation of data dispersion using the Fano factor in the presence of the UEE linked to the enhancer is lower (Fano factor = 53.42) compared to that obtained from the enhancer alone (Fano factor = 88.33) (Supplementary Table 1) [16]. This result suggests that in the presence of the UEE, the α -globin 3' enhancer behaves as described in the enhancer progressive/rheostatic model in the α -globin context [5].

3.2. UEE activity is context- and tissue-dependent

In order to understand how general the activity of the UEE is, we replaced the promoter and enhancer α -globin sequences with the viral

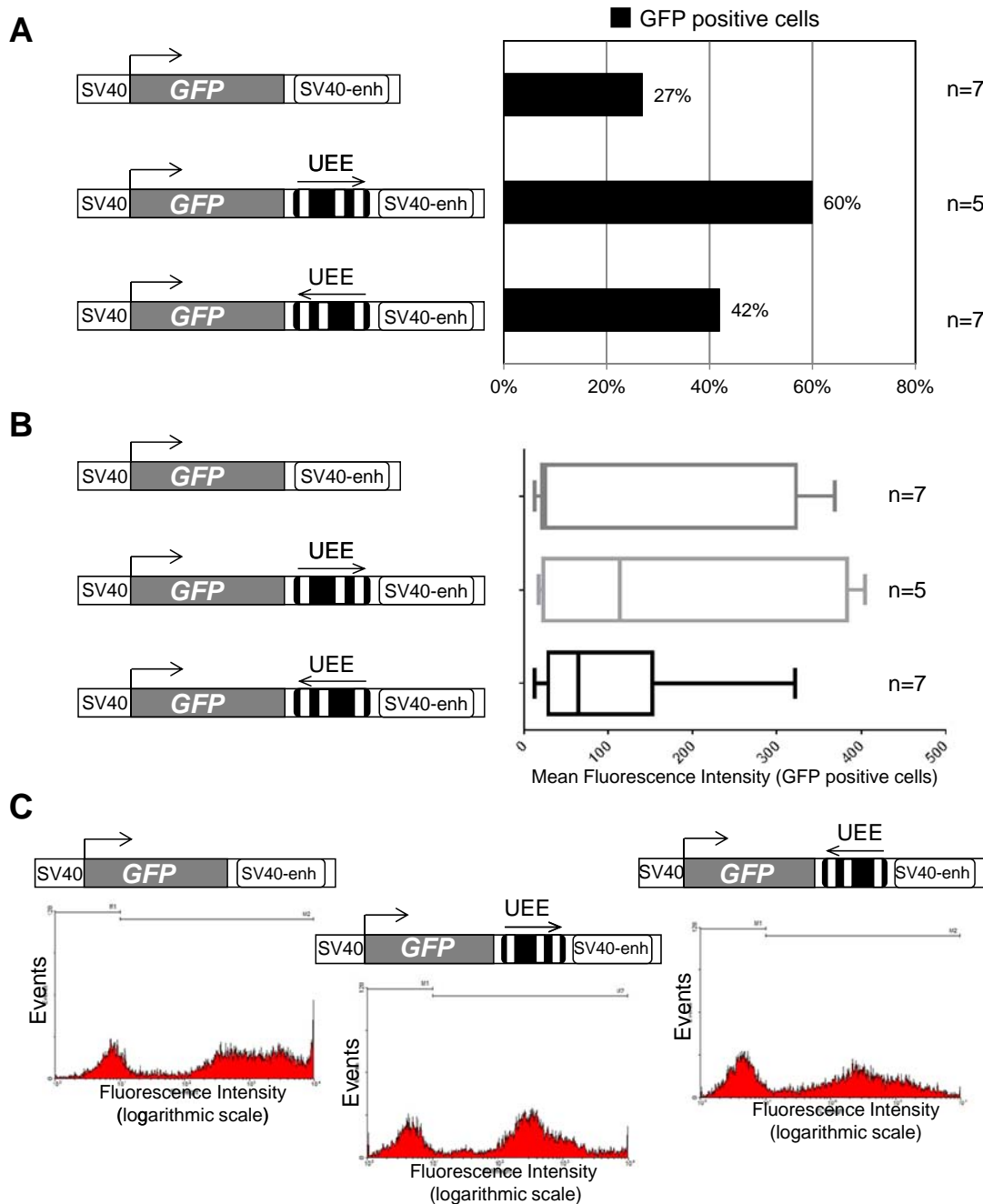


Fig. 2. Activity of the Upstream Enhancer Element (UEE) in a heterologous system. (A) Single genetic copies of each G418 drug-resistant transgene construct expressing the *GFP* gene under the control of the viral SV40 promoter–enhancer. (B) The same HD3 cell lines with their corresponding transgenes were evaluated in terms of the *GFP* mean fluorescence intensity. (C) Representative flow cytometry profile of cellular pools transfected with each construct after antibiotic G418 selection is shown. It is relevant to note that in the absence of the UEE there is a large dispersion of the fluorescence intensity in the HD3 stable transformed cell population.

SV40 promoter–enhancer pair, in the presence or absence of the UEE (Fig. 2A). The activity of these regulatory sequences was assessed through the expression of *GFP* in single genetic copies of each transgene in the HD3 cell line. In contrast to our previous results (Fig. 1), we observed a pronounced increment in the number of GFP-positive cells containing the UEE independently of its orientation relative to the transcriptional start site (TSS). Similar data was observed when placing a minimal promoter in the presence of the UEE (Supplementary Fig. S3A and S3B). A reduction in the mean fluorescence intensity was observed only when the UEE was placed in reverse orientation (Fig. 2B). In an alternative assay, each construct containing the SV40–UEE sequences was transfected in HD3 cells and pools of cells resistant to G418 were selected instead of single clones (Fig. 2C). In the absence of the UEE we observed a fluorescent cytometry profile characteristic of the activity of an enhancer, with a diverse activity ranging from an intermediate to very high reporter gene activity (Fig. 2C). In agreement with our previous results (Fig. 1C), the presence of the UEE element decreased the *GFP* fluorescence intensity and narrowed the distribution profile of GFP-positive cells (Fig. 2C). Interestingly, the assessment of promoter–enhancer SV40 elements in the presence of the UEE in transient transfection assays was tissue-specific (Supplementary Fig. S3C). In summary, in a heterologous context the UEE also reduces transcription variability that decrease *GFP* expression levels, and it seems to be tissue-dependent.

3.3. The integrity of the UEE is necessary for 3' enhancer activity

The UEE was initially defined as a silencer element in a non-chromatinized context [10]. In the 360 bp DNA fragment comprising the UEE we defined the presence of at least three nuclear factor-binding DNA motifs named SF1, SF2 and SF3 (Fig. 1A), with SF1 being the most prominent binding site, as shown by *in vitro* DNase I footprinting [10]. Intriguingly, none of the binding motifs correspond

to the known erythroid nuclear factors. The SF1 binding site is comprised of a direct repeat that is required for SF1 nuclear factor binding. We have also identified a novel RUNX1 binding site adjacent to the second direct repeat. Thus, through EMSAs and super-shift assays (Supplementary Figs. S4 and S5), we showed the *in vitro* binding of Runx using probes corresponding to the SF1 binding site. One possibility is that the RUNX transcription factor can contribute to the chromatin configuration of the UEE, but further work is required to understand such a contribution. To test the requirement of the UEE for 3' enhancer function, we generated deletions of the UEE and tested them as single-copy integrants in HD3 cells. Elimination of sequences from the 5' end of the UEE led to a reduction in the number of cells expressing the reporter gene with an increment in the fluorescence intensity (Fig. 3). This may reflect that mutants are unable to adopt the proper chromatin configuration over the UEE.

3.4. Perturbation of two MNase-HSs in the UEE and 3' enhancer affects enhancer activity

Our previously published data suggested that a construct containing the enhancer without the UEE cannot activate *GFP* expression [9]. Therefore, we asked whether chromatin structure associated with UEE–enhancer has a role in its overall activity. Two erythrocyte-specific nuclease hypersensitive sites (HSs) were identified, which appear when the enhancer is fully active during chicken embryo development and erythroid differentiation [9]. These two HSs are separated by 100 (+/– 20) bp, which are not sufficient to position a nucleosome. We hypothesized that this chromatin arrangement, which is characteristic of the active state of the enhancer, affects the function of the UEE and enhancer elements. To test this hypothesis, we perturbed the chromatin arrangement by introducing 25, 50 and 75 bp of unrelated lambda phage DNA between the two HSs in our transgene constructs (Fig. 4A). We found that adding 25 bp between the two MNase-HSs

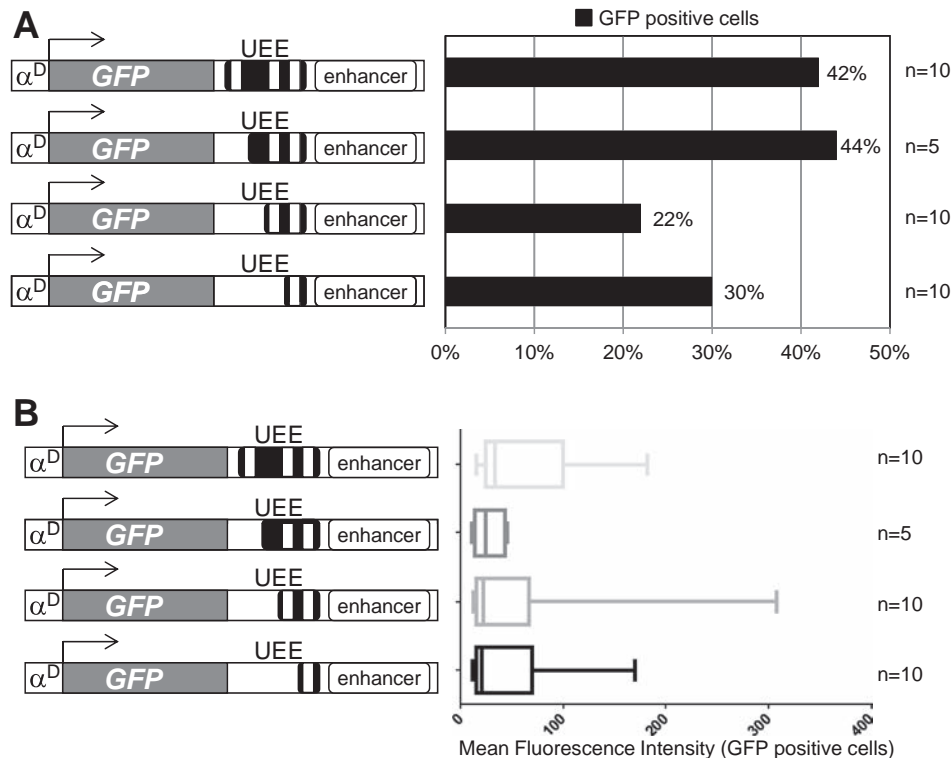


Fig. 3. Sequential deletion of the Upstream Enhancer Element in the presence of the α -globin 3' core enhancer affects its activity. Taking advantage of the position of the SF1 to SF3 nuclear factor binding sites (black boxes) we generated three different deletions that shortened the UEE. Individual HD3 cells stable clones were isolated and single-copy integrants selected. (A) The graphs show the percentage (%) of fluorescent cells and (B) the mean fluorescence intensity defined by flow cytometry.

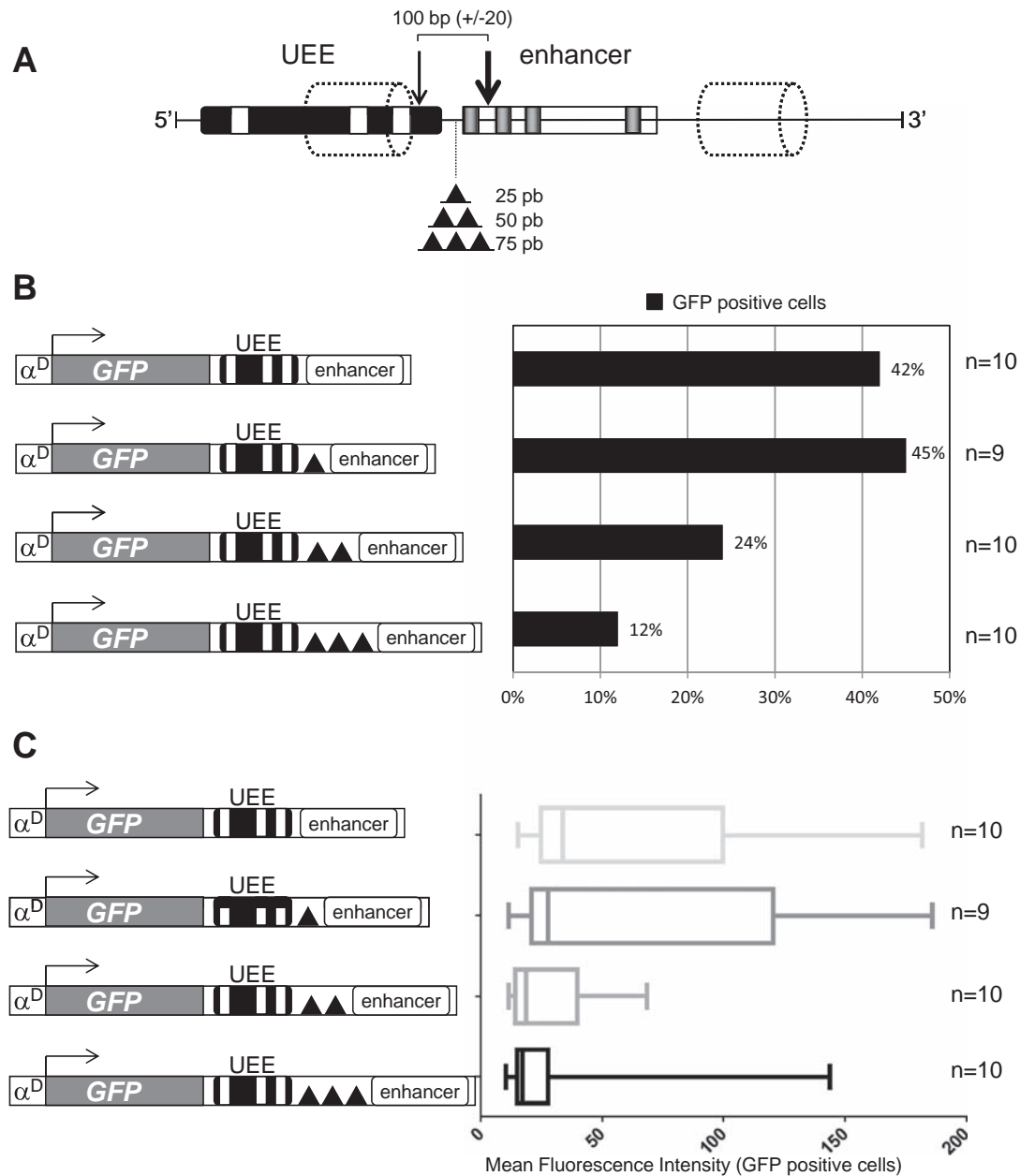


Fig. 4. Increasing the distance between the two MNase-HSs affects UEE-enhancer activity. (A) Schematic of the location of the two MNase-HSs (separated by 100 (+/-20) bp) and the point mutation that allows the introduction of a new AatII restriction site to insert 25, 50 and 75 bp of unrelated lambda-DNA between the two MNase-HSs. A new set of individual stable clones were generated and single-copy integrants were selected in HD3 cells. (B) As shown previously, the graph shows the percentage (%) of fluorescent cells and (C) the mean fluorescence intensity defined by flow cytometry.

did not reduce transgene activity. In contrast, the addition of 50 and 75 bp of lambda DNA decreased the number of GFP-positive cells by 18% and 30%, respectively (Fig. 4B). These results suggest that nucleosome positioning over the UEE and the formation of the MNase-HSs are epigenetic features that influence α -globin 3' enhancer activity.

3.5. Increasing the distance between the two MNase-HSs in the UEE and 3' enhancer perturbs nucleosome positioning

Separation of the two MNase-HSs affected the UEE-enhancer activity (Fig. 4), however, whether this effect is associated with nucleosome repositioning in the UEE-enhancer elements was not clear. We therefore performed MNase assays to determine the nucleosome positioning in two independent HD3 stably transfected cell lines, one carrying the unaltered transgene (Fig. 5A), and the other carrying the transgene

containing a 75 bp insertion of unrelated DNA (Fig. 5B). Using a probe that hybridizes to the GFP sequence, we found the two expected prominent MNase-HSs in cells carrying the unaltered transgene, as previously seen in endogenous chromatin and in transgene sequences (Fig. 5A). In contrast, in the HD3 cell lines carrying the 75 bp insertion we found a de-localization of the nucleosome located 5' of the main HSs and over the UEE (Fig. 5B). The chromatin accessibility analysis shows that the UEE appears nucleosome-free but is unable to affect the reporter gene activity, thus suggesting that binding of transcription factors like GATA-1 and CTCF (data not shown) is still present and can recruit ATP-chromatin remodeling complexes (Fig. 5B and Supplementary Fig. S6). This interpretation is further supported by our previous ChIP assays and PCR amplification with DNA associated with mono- and dinucleosomes isolated by sucrose gradient as a measure of MNase sensitivity and nucleosome depletion [9]. Another possibility is that the

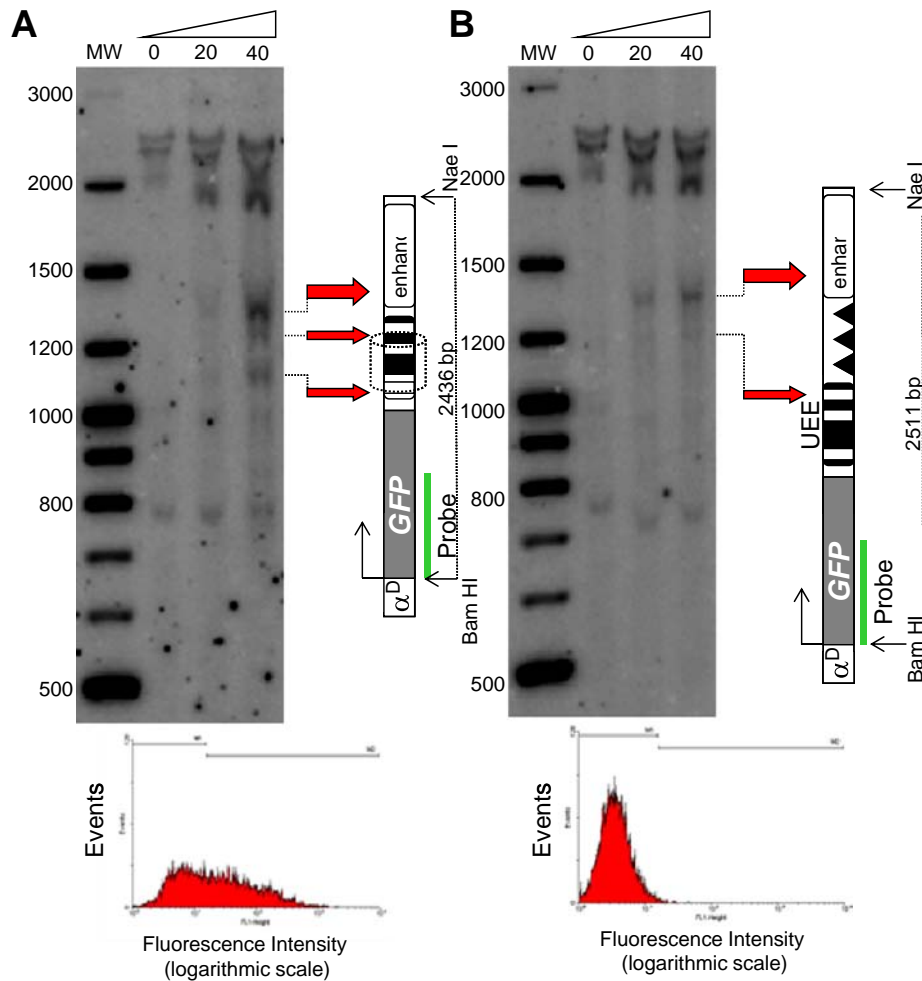


Fig. 5. Increasing the distance between the two MNase-HSs delocalizes the nucleosomes over the UEE-enhancer. (A) Southern blot showing no-digestion (0 Units) and MNase digestion (20 Units and 40 Units) of a control stable line. A probe corresponding to GFP sequences was used for the Southern blot. A schematic of the MNase-HS bands is shown along with the flow cytometry profile of this particular stable line. (B) The same as in (A) but here we chose a stable line with the highest increment in distance (75 bp) between the two MNase-HSs. Of note, transgene expression is lost (see flow cytometry graph). This is one representative set of experiments out of the two independent clones analyzed. Similar results were obtained from both set of experiments.

positioning of a nucleosome between the intersection of both MNase-HSs occludes the regulatory binding sites, thus preventing nuclear factor binding and communication with the target promoter, although we do not favor this hypothesis because in the performed accessibility assay we observed a faint band suggesting that this is an infrequent event in the cell population (Fig. 5 and Supplementary Fig. S6). In summary, we propose that the positioning of a nucleosome over the UEE element and the presence of two MNase-HSs separated by 100 (\pm 20) bp are local chromatin features required for the activity of the chicken α -globin 3' enhancer.

4. Discussion

Enhancers are classically defined as sequences with a modular organization of binding motifs for ubiquitous and tissue-specific transcription factors. They can be located in intragenic and intergenic genomic regions, and their activity is independent of their location and distance with respect to their target gene [1,2,5]. In recent years a large amount of data has been generated to propose novel ways of identifying enhancers. For example, the definition of characteristic histone marks like the H3K27ac and H3K4me1 modifications, and the development of high-throughput methods demonstrating long-distance mechanisms of enhancer action [17,18]. Unexpectedly, enhancers are also colluding with long non-coding RNAs to facilitate contacts over long distances

with their target promoter [19]. More recently, RNAs transcribed from regulatory elements like enhancers, known as enhancer RNAs (eRNAs), represent a novel class of non-coding transcripts that participate in enhancer mechanisms of action [20]. Despite the progress in understanding the mechanisms of long-range enhancer function, less is known about the effect of local chromatin organization in the activation potential of enhancers in a cell population.

Here, we describe a regulatory sequence of the chicken α -globin 3' enhancer, the UEE, located 5' of the core enhancer. Our results confirm that the UEE is a component of the α -globin 3' enhancer. In addition, the UEE-enhancer chromatin configuration is critical for the α -globin enhancer activity (Figs. 4 and 5). Formation of two MNase-HSs and the eviction of two nucleosomes positioned over the enhancer are required for full activity of the enhancer, which is flanked by two firmly positioned nucleosomes (Fig. 5). Perturbing the configuration of the UEE-enhancer by separating the two MNase-HSs caused loss of the 5' well-positioned nucleosome and affected the activation potential of the enhancer (Fig. 5B and Supplementary Fig. S6). Therefore, the activity of enhancers is affected by not only long-distance interactions, but also by the local chromatin structure.

Our dispersion data calculated through the Fano factor favors the hypothesis of the progressive model of enhancer action in the presence of the UEE. This suggests that the UEE mainly regulates the burst size instead of regulating the frequency of transcription in the α -globin

context. However, UEE presence also reduces the percentage of transcriptionally active cells, suggesting that it also regulates the activation frequency of the number of DNA templates or cells. Even when the percentage is modest (12%), this data could be relevant since it comes from different and independent genomic sites of transgene insertion. If our interpretation is correct, then the presence of a nucleosome could be the switch that regulates the transition between both models of enhancer activity.

At present, enhancers are validated through the identification of sequences with the highest activity in reporter systems, but adjacent sequences are frequently omitted. The identification of the UEE, which contributes to the local chromatin organization and enhancer activity, suggests that adjacent regions should be considered when analyzing enhancer activity *in vivo*. Perhaps the H3K4me1 modification, which spans over identified enhancers and their surrounding sequences, could be used to identify additional enhancer regulatory elements in genome-wide studies.

The UEE-enhancer acquires a particular epigenetic signature through the formation of two MNase-HSs [9]. Interestingly these two MNase-HSs appear progressively during chicken embryonic red blood cell development, reaching a peak of chromatin accessibility at day 10, when the enhancer is fully active, and coincident with eviction of the two nucleosomes positioned over the enhancer [9]. How this progressive chromatin remodeling occurs is not clear, but several factors might be involved. The erythroid-specific nuclear factor GATA-1, and the two GATA-1 binding sites located on each side of the 3'-most MNase-HS, are critical for activity of the 3' enhancer [9]. In addition, BRG1, the catalytic subunit of the SWI/SNF ATP-dependent chromatin remodeling complex is able to interact with GATA-1 and EKLF, both of which are present in the 3' enhancer [21,22]. Another possibility is the recruitment of the chromatin remodeling complex NuRD through the GATA-1 and FOG-1 factors [23]. Furthermore, the SWI/SNF complex can regulate transcription by nucleosome eviction over promoter elements [11,24]. Therefore, the two MNase-HSs may reflect the formation of a large complex, favoring the release of nucleosomes located over the enhancer and allowing maximal activity.

Complementing this scenario, the firmly positioned nucleosomes shielding each side of the core enhancer may also affect enhancer activity (Fig. 1A). Such nucleosomes may act as anchor points to attract the ATP-dependent chromatin remodeling complex RSC [25]. This can be complemented by the incorporation of the histone variant H2A.Z, which is commonly present in nucleosomes associated with promoters and enhancers of transcriptionally active genes [25,26]. Thus, perturbation of the chromatin structure of the α -globin 3' enhancer through unrelated DNA sequences can alter its tri-dimensional structure, and in particular, the de-localization of the well-positioned nucleosome over the UEE could disable the nucleoprotein complex anchored to the enhancer necessary to activate gene expression. For example, in some circumstances DNA compaction through nucleosomes is necessary to bring distal transcription factor binding sites closer [27–29]. In our chromatin accessibility analysis we observed that the UEE appears nucleosome-free, suggesting that binding of transcription factors (like CTCF, data not shown) and chromatin remodeling are still present (Fig. 5B and Supplementary Fig. S6).

We have shown the functional flexibility associated with the UEE region and the potential binding of the RUNX factor (Figs. 1, 2 and Supplementary Fig. S3). Interestingly RUNX transcription factors are associated with both stimulation and silencing of target genes depending on its associated partners in a tissue-dependent manner [15,30]. RUNX1 participates in the formation of complexes composed of LDB1, GATA-1 and TAL1 among other proteins [31,32]. These interactions could be relevant for long-distance contacts with target promoters since GATA-1, LDB1, GATA-1 and cohesin proteins have been demonstrated to support *in vivo* enhancer-promoter communication [2]. Finally, we cannot discard the possibility that the described configuration of the UEE-enhancer is required for the adequate assembly of nuclear factors needed for long-

distance contacts with their target promoters, like the interaction between EKLF/GATA-1/FOG-1 with BRG1, the chromatin remodeling protein ATRX or even CTCF [33–36]. Thus, an upcoming challenge will be to uncover the hierarchical order of events that gives the UEE-enhancer the capacity to activate its target genes during erythroid differentiation and chicken development.

Acknowledgments

We thank Catherine Farrell, Paul Delgado-Olguín and Rodrigo Arzate-Mejía for their comments on the manuscript and suggestions. We acknowledge the technical assistance of Georgina Guerrero Avendaño and Fernando Suaste Olmos. This work was supported by the DGAPA, UNAM (IN209403, IN203811 and IN201114), and CONACyT (42653-Q, 128464 and 220503) to FR-T. Ph.D. fellowship from CONACyT and Dirección General de Estudios de Posgrado-Universidad Nacional Autónoma de México (DGEP) to EG-G. Additional support was provided by the PhD Graduate Program, “Doctorado en Ciencias Biomédicas”, to the Instituto de Fisiología Celular and the Universidad Nacional Autónoma de México.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbarm.2014.09.009>.

References

- [1] M. Bulger, M. Groudine, Functional and mechanistic diversity of distal transcription enhancers, *Cell* 144 (2011) 327–339.
- [2] E. Smith, A. Shilatfard, Enhancer biology and enhanceropathies, *Nat. Struct. Mol. Biol.* 21 (2014) 210–219.
- [3] G.A. Maston, S.G. Landt, M. Snyder, M.R. Green, Characterization of enhancer function from genome-wide analyses, *Annu. Rev. Genomics Hum. Genet.* 13 (2012) 29–57.
- [4] A. Rada-Iglesias, R. Bajpai, T. Swigut, S.A. Brugmann, R.A. Flynn, J. Wysocka, A unique chromatin signature uncovers early developmental enhancers in humans, *Nature* 470 (2011) 279–285.
- [5] E.M. Blackwood, J.T. Kadonaga, Going the distance: a current view of enhancer action, *Science* 281 (1998) 60–63.
- [6] S. Fiering, E. Whitelaw, D.I.K. Martin, To be or not to be active: the stochastic nature of enhancer action, *BioEssays* 22 (2000) 381–387.
- [7] J.A. Knezetic, G. Felsenfeld, Identification and characterization of a chicken α -globin enhancer, *Mol. Cell. Biol.* 9 (1989) 893–901.
- [8] H. Rincón-Arango, V. Valadez-Graham, G. Guerrero, M. Escamilla-Del-Arenal, F. Recillas-Targa, YY1 and GATA-1 interaction modulate the chicken 3'-side α -globin enhancer activity, *J. Mol. Biol.* 349 (2005) 961–975.
- [9] M. Escamilla-Del-Arenal, F. Recillas-Targa, GATA-1 modulates the chromatin structure and activity of the chicken α -globin 3' enhancer, *Mol. Cell. Biol.* 28 (2008) 575–586.
- [10] F. Recillas-Targa, C.V. De Moura Gallo, M. Huesca, K. Scherrer, L. Marcaud, Silencer and enhancer elements located at the 3'-side of the chicken and duck α -globin-encoding gene domain, *Gene* 129 (1993) 229–237.
- [11] C. Jiang, B.F. Pugh, Nucleosome positioning and gene regulation: advances through genomics, *Nat. Rev. Genet.* 10 (2009) 161–172.
- [12] J. Boyes, G. Felsenfeld, Tissue specific factors additively increase the probability of the all-or-none formation of a hypersensitive site, *EMBO J.* 15 (1996) 2496–2507.
- [13] M. Furlan-Magaril, E. Rebollar, G. Guerrero, A. Fernández, E. Moltó, E. González-Buendía, M. et al., An insulator embedded in the chicken α -globin locus regulates chromatin domain configuration and differential gene expression, *Nucleic Acids Res.* 39 (2011) 89–103.
- [14] V. Valadez-Graham, S.V. Razin, F. Recillas-Targa, CTCF-dependent enhancer blockers at the upstream region of the chicken α -globin gene domain, *Nucleic Acids Res.* 32 (2004) 1354–1362.
- [15] N. Pencovich, R. Jaschek, A. Tanay, Y. Groner, Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models, *Blood* 117 (2011) e1–e14.
- [16] C.R. Brown, C. Mao, E. Falkovskaia, M.S. Jurica, H. Boeger, Linking stochastic fluctuations in chromatin structure and gene expression, *PLoS Biol.* 11 (2013) e1001621.
- [17] J. Marsman, J.A. Horsfield, Long distance relationships: enhancer-promoter communication and dynamic gene transcription, *Biochem. Biophys. Acta* 1819 (2012) 1217–1227.
- [18] W. de Laat, D. Duboule, Topology of mammalian developmental enhancers and their regulatory landscapes, *Nature* 502 (2013) 499–506.
- [19] U.A. Orom, R. Shiekhattar, Long noncoding RNAs usher in a new era in the biology of enhancers, *Cell* 154 (2013) 1190–1193.

- [20] K. Mousavi, H. Zare, S. Dell'Orso, L. Grontved, G. Gutierrez-Cruz, A. Derfoul, et al., eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci, *Mol. Cell* 51 (2013) 606–617.
- [21] Z. Xu, X. Meng, Y. Cai, M.J. Koury, S.J. Brandt, Recruitment of the SWI/SNF protein Brg1 by a multiprotein complex effects transcriptional repression in murine erythroid progenitors, *Biochem. J.* 399 (2006) 297–304.
- [22] S.I. Kim, E.H. Bresnick, S.J. Bultman, BRG1 directly regulates nucleosome structure and chromatin looping of the α globin locus to activate transcription, *Nucleic Acids Res.* 37 (2009) 6019–6027.
- [23] A. Miccio, G.A. Blobel, Role of the GATA-1/FOG-1/NuRD pathway in the expression of human β -like globin genes, *Mol. Cell Biol.* 30 (2010) 3460–3470.
- [24] M.L. Dechassa, A. Sabri, S. Pondugula, S.R. Kassabov, N. Chatterjee, M.P. Klade, B. Bartholomew, SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes, *Mol. Cell* 38 (2010) 590–602.
- [25] P.D. Hartley, H.D. Madhani, Mechanisms that specify promoter nucleosome location and identity, *Cell* 137 (2009) 445–458.
- [26] A. Barski, S. Cuddapah, K. Cui, T.Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, High-resolution profiling of histone methylations in the human genome, *Cell* 129 (2007) 823–837.
- [27] Q. Lu, L.L. Wallrath, S.C.R. Elgin, The role of a positioned nucleosome at the *Drosophila melanogaster* hsp26 promoter, *EMBO J.* 14 (1995) 4738–4746.
- [28] W. Stünkel, I. Kober, K.H. Seifart, A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene, *Mol. Cell Biol.* 17 (1997) 4397–4405.
- [29] A.J. Gossett, J.D. Lieb, *In vivo* effects of histone H3 depletion on nucleosome occupancy and position in *Saccharomyces cerevisiae*, *PLoS Genet.* 8 (2012) e1002771.
- [30] K. Blyth, E.R. Cameron, J.C. Neil, The RUNX genes: gain or loss of function in cancer, *Nat. Rev. Cancer* 5 (2005) 376–387.
- [31] N. Meier, S. Krpic, P. Rodriguez, J. Strouboulis, M. Monti, J. Krijgsveld, M. Gering, R. Patient, A. Hostert, F. Grosveld, Novel binding partners of Lbd1 are required for haematopoietic development, *Development* 133 (2006) 4913–4924.
- [32] B. van Riel, T. Pakozdi, R. Brouwer, R. Monteiro, K. Tuladhar, V. France, et al., A novel complex RUNX-MYEF2, represses hematopoietic genes in erythroid cells, *Mol. Cell Biol.* 32 (2012) 3814–3822.
- [33] R. Drissen, R.J. Palstra, N. Gillemans, E. Splinter, F. Grosveld, S. Philipsen, et al., The active spatial organization of the β -globin locus requires the transcription factor EKLF, *Genes Dev.* 18 (2004) 2485–2490.
- [34] C.R. Vakoc, D.L. Letting, N. Gheldof, T. Sawado, M.A. Bender, M. Groudine, et al., Proximity among distant regulatory elements at the β -globin locus requires GATA-1 and FOG-1, *Mol. Cell* 17 (2005) 453–462.
- [35] E. Splinter, H. Heath, J. Kooren, R.J. Palstra, P. Klous, F. Grosveld, et al., CTCF mediates long-range chromatin looping and local histone modification in the β -globin locus, *Genes Dev.* 20 (2006) 2349–2354.
- [36] K. Ratnakumar, E. Bernstein, ATRX: the case of a peculiar chromatin remodeler, *Epigenetics* 8 (2013) 3–9.



Chromatin remodeling effects on enhancer activity

Estela García-González¹ · Martín Escamilla-Del-Arenal² · Rodrigo Arzate-Mejía¹ · Félix Recillas-Targa¹

Received: 8 October 2015 / Revised: 4 March 2016 / Accepted: 14 March 2016
© Springer International Publishing 2016

Abstract During organism development, a diversity of cell types emerges with disparate, yet stable profiles of gene expression with distinctive cellular functions. In addition to gene promoters, the genome contains enhancer regulatory sequences, which are implicated in cellular specialization by facilitating cell-type and tissue-specific gene expression. Enhancers are DNA binding elements characterized by highly sophisticated and various mechanisms of action allowing for the specific interaction of general and tissue-specific transcription factors (TFs). However, eukaryotic organisms package their genetic material into chromatin, generating a physical barrier for TFs to interact with their cognate sequences. The ability of TFs to bind DNA regulatory elements is also modulated by changes in the chromatin structure, including histone modifications, histone variants, ATP-dependent chromatin remodeling, and the methylation status of DNA. Furthermore, it has recently been revealed that enhancer sequences are also transcribed into a set of enhancer RNAs with regulatory potential. These interdependent processes act in the context of a complex network of chromatin interactions, which together contributes to a renewed vision of how gene activation is coordinated in a cell-type-dependent manner. In this review, we describe the interplay between

genetic and epigenetic aspects associated with enhancers and discuss their possible roles on enhancer function.

Keywords Enhancer · Promoter · Histone modifications · Chromatin · CTCF · Topological associating domains · Non-coding RNAs · Enhancer RNA

Introduction

It was recognized almost 35 years ago that enhancers are important DNA regulatory elements that eukaryotic cells use to generate specificity by controlling differential programs of gene expression. Next-generation sequencing methodologies such as ChIP-seq and RNA-seq have enabled us to integrate chromatin and transcriptional states and explore the dynamic regulation of gene expression during cellular differentiation and development. In particular, enhancers have emerged as critical players in distinguishing transcriptional states that show a high degree of variation between different cell types. Enhancer DNA regulatory elements were first described in monkey tumor virus studies [1, 2]. The simian virus SV40 enhancer consists of a 72 base pairs (bp)-long repeated sequence whose deletion reduces the viral protein levels expressed in early stages of infection, and as a result, abolishing virus viability [1]. The existence of enhancers with similar stimulatory characteristics was subsequently reported in eukaryotic genomes [3–7]. The first cellular enhancer was found within an intron of the mouse immunoglobulin heavy chain gene whose stimulatory activity depended on a subset of nuclear factors with high cell-type specificity [3, 4, 8]. In this review, we will cover the most relevant chromatin-associated aspects of enhancer functions, including recent discoveries involving long non-coding RNAs (lncRNAs) and enhancer RNAs (eRNAs).

✉ Félix Recillas-Targa
frecilla@ifc.unam.mx

¹ Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Circuito Exterior S/N, Ciudad Universitaria, C.P. 04510 Mexico City, México

² Department of Biochemistry and Molecular Biophysics, Mortimer B. Zuckerman Mind Brain and Behavior Institute, Columbia University, New York City, NY 10027, USA

How do enhancers affect transcription?

The way in which enhancers stimulate transcription is a central question that remains poorly understood. As enhancers were characterized based on their property to increase the transcriptional levels of target genes, the amount of gene product or the number of cells that activate transcription must be important. Therefore, two models have been suggested to explain their function: the binary model and the progressive or rheostatic model [9]. The binary model for enhancer function proposes that enhancers increase the probability that a higher portion of cells activate transcription at a given locus within a cell population [9–13]. In the progressive model, enhancers increase the number of RNA molecules transcribed from the target gene, but not the number of cells that initiate transcription [14]. It is currently unsolved if these models of enhancer function can be generalized or even if alternative models exist.

Transcription factors as mediators of enhancer activity

A central characteristic of enhancers is to function as TF-binding platforms that affect transcription by direct stimulation of their target promoters, located a few kilobases (kb) up to megabases away [15–18]. Although many enhancers are found in intergenic regions, they can also be localized within protein coding genes, in particular introns [19, 20]. Enhancer sequences are commonly considered to be 200–500 bp in length where DNA-binding sites for multiple TFs are clustered [18]. Many lineage-specific TFs and their recruitment to binding sites within enhancers govern lineage-specific gene transcription [21–23]. Two models for TF binding at enhancers have been proposed [24]. In the “enhanceosome” model, the DNA sequence acts as a scaffold for the ordered and cooperative binding of TFs to form a protein complex that can activate transcription. In this scenario, enhancer activity emerges from a network of interactions and its action is lost just by the absence of one protein [25, 26]. In contrast, in the “billboard” model, TF binding is independent from each other and TFs do not act as a single unit [24].

TFs typically have short recognition sequences (4–10 nucleotides), often with a high degree of degeneracy, which increases the probability to find their target sequences in the genome. For example, the predicted estrogen receptor recognition sequence occurs more than one million times in the human genome, but only 10,000–16,000 binding sites are occupied in a human cell line [20, 27, 28]. It is unclear why a major proportion of binding sites are not occupied. For example, DNA packaging into nucleosomes could block the accessibility to

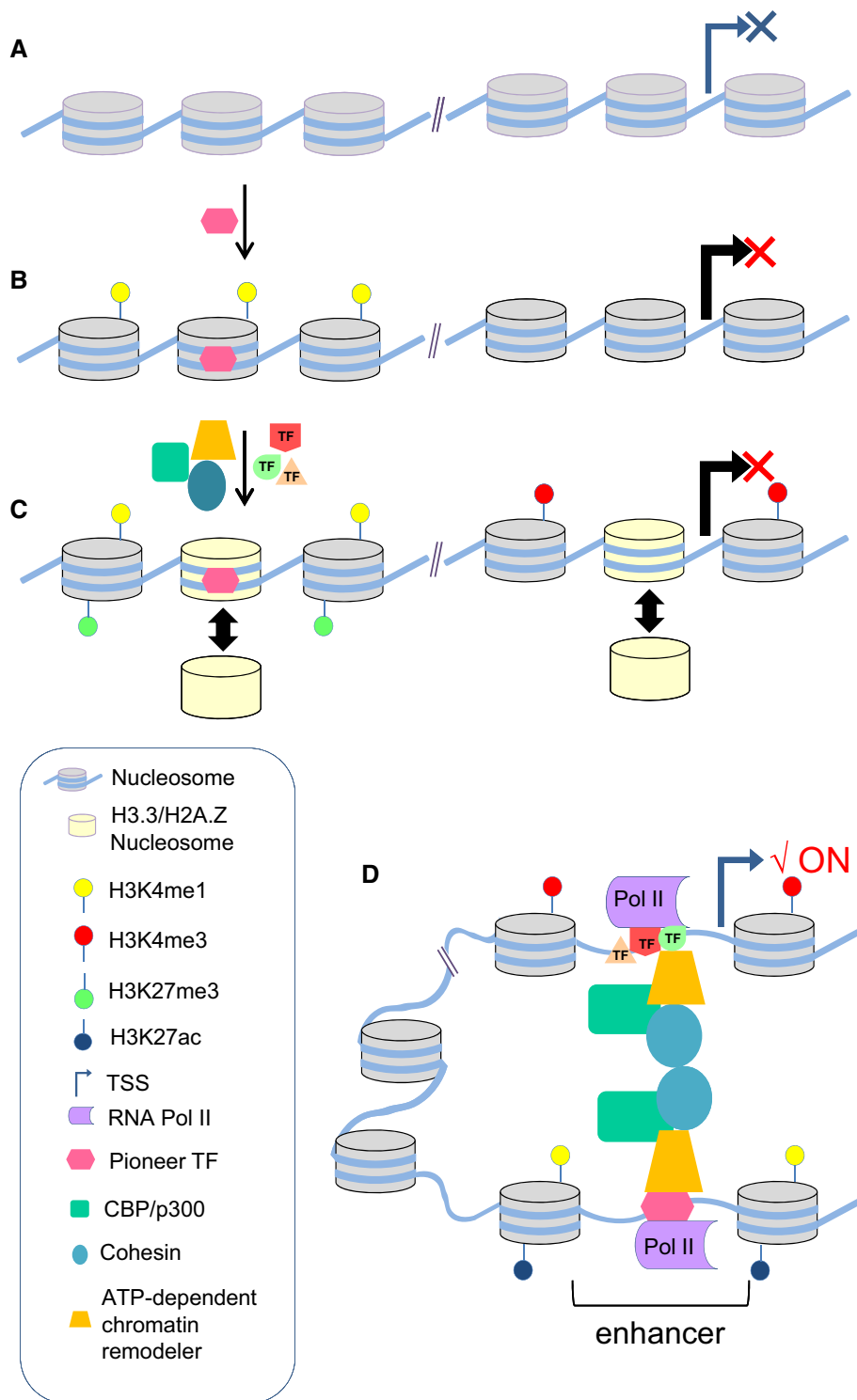
their binding sites [29]. However, clustering of binding sites contributes to overcome the chromatin barrier by increasing the cooperativity between TFs like in the “enhanceosome” model [18, 21, 25, 28, 30]. Cooperative binding was observed using reconstituted nucleosome cores and DNA harboring different combinations of binding sites for different and unrelated TFs including GAL4, USF, and NF- κ B. The binding of the first TF stimulates the binding of the second by up to two orders of magnitude [31]. This evidence supports the idea that cooperative TF binding to chromatinized DNA could significantly increase the affinity of each factor for nucleosomal DNA at regulatory elements in the genome and thereby overcoming the natural chromatin barrier (Fig. 1) [21–23].

DNA accessibility differs between cell types and changes dynamically during organism development, cell differentiation and in response to external and internal cell stimuli [32, 33]. Such changes in DNA accessibility are believed to be mediated by regulatory proteins, including a special class of TFs called pioneer factors (Fig. 1). Pioneer or nucleation factors, such as FOXA, PU.1, OCT4, and GATA-1, initially bind their consensus sequence within the context of the chromatin, at the level of the solenoid or 30 nm fiber facilitating the binding of additional TFs and co-factors to further open the chromatin in response to hormones or other differentiation signals [23, 29].

The induction of pluripotent stem cells (iPS) from somatic cells through the overexpression of pluripotency TFs contributed to the finding that some TFs are able to recognize their DNA binding sites within silent chromatin [21, 23, 34, 35]. Pluripotency factors including OCT4, SOX2, and KLF4 preferentially recognize partial motifs displayed on the nucleosome surface [23]. The majority (70 %) of pluripotency TF binding initiates gene activation by first binding enhancer sequences of genes that promote reprogramming [23, 35–38]. Pioneer factors are hypothesized to direct lineage specification by changing chromatin states at their binding sites. One well-studied example is OCT4, a TF involved in establishing and maintaining the pluripotent state. In one study, OCT4 was shown to change the chromatin state at the *MYOD1* locus. Overexpression of OCT4 resulted in its binding to the enhancer of *MYOD1*, and a switch from H3K4me1 to H3K4me3 at the *MYOD1* promoter. Interestingly, the H3K27me3 mark, which was present before OCT4 overexpression, remained [39]. This bivalent promoter state has been described in embryonic stem cells and could be mediated in part by OCT4 [36].

It is important to consider that TF binding does not always imply direct stimulation of gene expression. For example, in the case of pioneer factors that can access DNA where other factors cannot, binding is primarily required to initiate additional chromatin remodeling events that will prepare the chromatin template for the binding of

Fig. 1 Chromatin dynamics at enhancers and promoters during gene activation. **a** Silent chromatin characterized by a regular pattern of nucleosomes positioned over regulatory elements including the enhancer and its target promoter. **b** Initial binding of a Pioneer TF within nucleosomes first to the enhancer facilitates the binding of additional TFs and co-factors to open the chromatin through histone modifications. **c** Additional TFs and the action of ATP-dependent chromatin remodeling complexes in response to hormones or other differentiation signals facilitate the recruitment of the transcription machinery to both the promoter and the enhancer. Histone variant exchange at regulatory elements may contribute to chromatin relaxation. **d** Physical contacts between enhancer and its target promoter occurs when the gene is transcribed. Under this condition, the enhancer is characterized by its association with high levels of the H3K27ac histone modification



other cell-type-specific TFs (Fig. 1) [39]. The ability of TFs to activate transcription on chromatin templates is also dependent on the recruitment of coactivator proteins [21, 30, 40, 41]. By definition, coactivators lack sequence-specific DNA binding competency, but they function as histone modifiers or by recruiting complexes with histone

modifier capabilities (i.e., histone acetyltransferases), ATP-dependent chromatin remodelers, or mediators of long-range chromatin contacts [42–44]. Therefore, pioneer TFs and their regulatory partners can selectively bind to the site where an enhancer will be remodeled in order to be fully activated in a regulated and highly specific manner.

Histone modifications at enhancers

The histone H3 lysine 4 mono-methylation is present in nucleosomes linked to enhancer elements

Systematic genome-wide studies of histone post-translational modifications have revealed new insights into transcriptional regulation [33, 45, 46]. In particular, H3K4me1 was the first histone modification globally linked to distal regulatory regions [19] in a tissue-specific manner [33]. Comparably, histone H3K4me3 mark was predominantly enriched at gene promoter regions [19]. Nevertheless, a clear-cut discrimination between enhancers and promoters cannot be made based on their histone H3K4 methylation status as histone H3K4me2 or H3K4me3 marks have been also detected at active enhancers bound by the RNA Pol II [47–51].

In mouse and human embryonic stem cells, enhancers directing promoter activity of genes linked to developmental processes are pre-labeled by H3K4me1. This leads to the hypothesis that this modification participates in enhancer priming [18, 52–54]. Histone H3K4me1, unlike H3K27ac or RNA Pol II, appears to persist at enhancers even after their disengagement from promoters [55–57]. Once incorporated, H3K4me1 can be read by chromodomain-containing acetyltransferases like TIP60, which catalyzes acetylation of H2A at Lys5 [58, 59]. In addition, histone H3K4me1 contributes to maintain a permissive chromatin state by repelling the interaction of proteins that recognize unmethylated H3 lysine 4, BHC80 (CoREST-LSD1 complex) and de novo DNA methylation complex [60, 61]. Thus, H3K4 methylation may protect distal regulatory elements from being targeted by DNA methylation and keep them in a poised state, until ready for activation with the proper signal. Nevertheless, the presence of H3K4me1 at enhancers in *Drosophila melanogaster*, an organism lacking Dnmt3 homolog and low levels of DNA methylation [62], suggests that other mechanisms may be at play, potentially involving direct recognition of H3K4me1.

Histone acetyltransferases occupancy at enhancers

In line with the positive correlation between histone acetylation and gene expression [48, 63], genome wide binding of enzymes with histone acetyltransferase activity represents a central aspect in enhancer function [52, 64]. CBP/p300 proteins are the most studied co-factors with intrinsic histone acetyltransferase (HATs) activity [65]. Genomic occupancy of p300/CBP during development in humans shows that 95 % of p300 in vivo binding is found at promoter distal regions [52, 65]. Consistent with these data, p300 binding is an accurate predictor of in vivo

enhancers in the developing mouse embryo. Additionally, most p300 bound regions coincide with DNase I hypersensitive sites (DHS) and active gene expression during development [64]. CBP and p300 HATs acetylate over 70 proteins, including themselves, however, histone 3 lysine 18 (H3K18) and H3K27 are their major in vivo histone targets [66–68]. One role for CBP/p300 enhancer binding may be to integrate metabolic information of the levels of acetyl coenzyme A (which acts as donor of acetyl groups) [69]. For example, in yeast where the Gcn5p/SAGA complex catalyzes the acetylation of histones at promoters of genes important for growth according to increased levels of acetyl coenzyme A [70]. Alternatively, CBP/p300 may recruit RNA Pol II to enhancers marked with H3K4me1 to transcribe the enhancer itself and generate a special class of transcripts, collectively called eRNAs, which have been implicated in the mechanism of action of enhancers (Fig. 2) [71, 72].

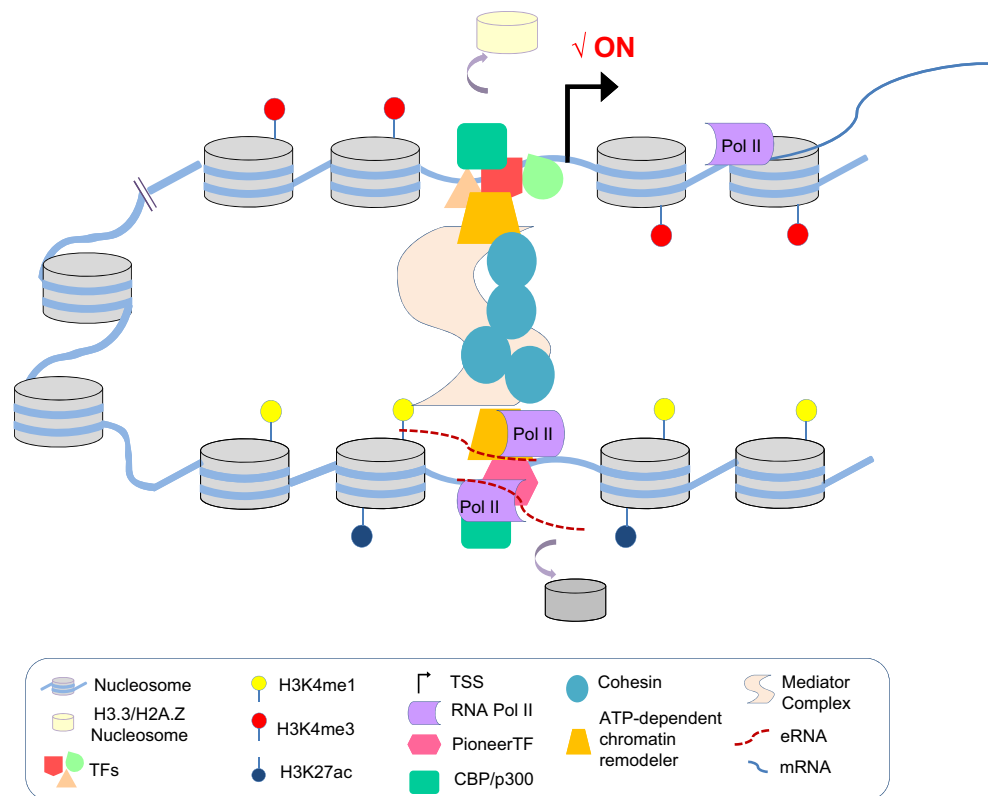
Apart from CBP/p300, other HATs have been found to interact with enhancers [72]. A possible reason for binding of different HATs to enhancer regions could be a differential recruitment of cofactors or TFs. Moreover, TFs could be potentially modified by HATs at target enhancers affecting their activity or protein interactions [73]. In conclusion, histone acetyltransferases are widely associated with enhancer elements, which argue in favor of a functional role for acetylation in enhancer function.

Histone H3K27ac and H3K27me3 marks demarcate active and poised enhancers

Despite a cell-type-specific genome-wide correlation between enhancers and the presence of H3K4me1 histone modification or the p300 coactivator complex [19, 45], the use of these parameters does not accurately predict enhancer activity on target promoters. Therefore, histone chromatin analysis at enhancer sequences during ESC differentiation revealed a clear correlation between the H3K27ac histone mark and an increased gene activity of proximal genes; in contrast, the presence of H3K27me3 showed the opposite trend [52–54, 56]. To some extent, the presence of acetylation at many enhancers may attenuate nucleosome stability [25] or improve chromatin accessibility. Since acetylation reduces the affinity of histones for negatively charged DNA, it may help TFs to access their binding sites more efficiently [74].

The different combinations of histone marks at enhancers on a genome-wide scale in different cell types permitted the classification of poised enhancers containing H3K4me1 and H3K27me3 histone marks and active enhancers associated with H3K4me1 and H3K27ac histone marks [52, 53]. Detailed analysis revealed that poised enhancers are lineage-specific enhancers that become

Fig. 2 Gene transcription is coupled to non-coding transcription at enhancers. Active enhancers are transcribed in a cell-type-specific manner producing bidirectional eRNAs. eRNA transcription coincides with transcription of target gene, high levels of H3K27ac, and a common architecture of transcriptional initiation characterized by nucleosome depleted region formation. eRNAs can exert their regulatory action through the stabilization of chromatin loops



activated upon differentiation [52]. Nucleosome remodeling accompanies changes in the H3K27ac and H3K4me1 histone marks at active enhancers (Fig. 1). At the time of TF binding, active enhancers exhibit single peak of RNA Pol II enrichment at the center. On the contrary, poised enhancers show a single peak of H3K4me1 and H3K27me3 histone marks without RNA Pol II occupancy [53, 56]. As previously mentioned, poised enhancers contain the H3K27me3 mark instead of H3K27ac. Remarkably, in *D. melanogaster*, the H3K27me3 histone demethylase UTX and the chromatin remodeler Brm directly bind to CBP/p300 and both *brm* mutations and UTX knockdown reduce H3K27ac levels at enhancer regions [75]. This provides a mechanism that couples H3K27 demethylation and the presence/action of a histone acetyltransferase (CBP) in order to switch from the inactive to the active enhancer status.

A limited set of modifications including H3K4me1/3, H3K27ac, and H3K27me3 are ubiquitously utilized to map and demarcate active and poised enhancers and promoters. Yet, these modifications represent only a fraction of the full repertoire of histone modifications at enhancers and promoters. Evidence is still lacking as to whether histone marks such as H3K4me1 or H3K27ac are sufficient, necessary, or even mechanistically involved in the activity of these regulatory elements [76, 77]. Both marks have fairly high turnover rates, and it is thus likely that neither is

inherited across mitosis nor is instructive for future cellular generations [78].

Remodeling the chromatin structure to access enhancer sequences

Chromatin remodeling by changing nucleosome density or position at enhancers is critical to generate a regulatory environment typically “devoid” of nucleosomes [also known as nucleosome depleted regions (NDR)] such that the DNA is accessible to be bound by TFs. In fact, chromatin remodeling has been widely associated with *cis*-regulatory elements in the genome such as promoters and enhancers [79–81]. At least one-third of the identified DHSs in 125 cell types were tissue-specific, located distally from annotated promoters [81], highly enriched for the H3K27ac histone mark, and associated with active enhancer sequences [33].

Nucleosome positioning and eviction at enhancers through chromatin remodeling complexes

For years, an intense debate has developed in terms of the role of DNA sequence in nucleosomal positioning. Sequence analysis of enhancers responsive to androgen stimulation show enrichment for an AT-rich motif (AA/

TT/TA/AT) at NDRs that is absent at nucleosome-enriched regions. In contrast, GC-rich motifs show the opposite tendency [82]. Similar observations of nucleotide composition at NDR of regulatory elements like enhancers have been reported in yeast and *D. melanogaster* [83–85]. Despite these observations, there is no clear consensus about the role of those sequences in nucleosome positioning and chromatin remodeling.

In contrast, ATP-dependent chromatin remodeling complexes directly contribute to nucleosomal positioning. Unlike TFs, these complexes do not bind directly to the DNA but contain protein domains that recognize other proteins or covalent histone modifications that assist them to associate to specific regions of the chromatin [86]. Importantly, in humans there is evidence that functional enhancers are associated with nucleosome remodeling [87]. SATB1 (special AT-rich sequence binding 1) is a protein found predominantly in thymocytes that recruits the CHRAC and ACF nucleosome mobilizing complexes within the interleukin-2 receptor gene (*IL-2R α*), in particular, toward the exon 1 and the SBS700 region of the gene, repressing its expression through changes in chromatin accessibility [88]. Local alteration in the pattern of cleavage by micrococcal nuclease is evident between wild-type and SATB1 null thymocyte nuclei possibly due to changes in nucleosome distribution over the enhancer [88].

Furthermore, the recruitment of the ATP-dependent chromatin remodeling SWI/SNF (Smarca4/BRG1) complex to enhancers is critical to initiate and establish the transcriptional program that promotes oligodendrocyte differentiation and subsequent myelination of the central nervous system (CNS) in humans [89]. The oligodendrocyte-lineage determination factor (Olig2) recruits the chromatin remodeling factor Brg1 to enhancer elements that exhibit enrichment in H3K27ac and that control oligodendrocyte differentiation [52, 89]. A Brg1 ATPase mutant domain inhibits myelin gene expression, suggesting that ATPase activity is essential for Brg1 function to promote myelination programs by chromatin remodeling of Olig2-enhancers [89].

Another ATP-dependent chromatin remodeling enzyme implicated in the regulation of enhancer activity in mammals is CHD7, a member of the Chromodomain Helicase DNA-binding protein family. Genetic mutations of *CHD7* are a cause of CHARGE syndrome [86, 90]. In mouse embryonic stem cells (mESCs), CHD7 co-localizes on a genome-wide level with loci exhibiting high levels of p300, H3K4me1, DHSs, and binding of master regulators like OCT4, SOX2, and NANOG, suggesting its association with enhancer sequences [91]. Significant correlation between CHD7 occupancy and reduced gene expression suggests that CHD7 functions to limit the expression of a subset of ESC-specific genes [91]. However, whether

CHD7 changes the chromatin structure associated with enhancers remains to be explored.

Nucleosome positioning changes according to gene activity, and is an important element of transcriptional regulation [45]. Recent genome-wide data confirm that DNA accessibility is highly dynamic and changes during lineage specification, and that particular patterns of nucleosome positioning along enhancer sequences emerge during gene activation [33, 45, 82, 92].

In prostate cancer cells, the androgen receptor primarily binds to enhancer sequences [82]. Before androgen receptor activation, androgen receptor binding loci are already marked with two well-positioned H3K4me2-containing nucleosomes that flank the binding site, along with a well-positioned H2A.Z/H3K4me2 containing nucleosome occluding the binding site. After androgen receptor activation, the central H3K4me2-modified nucleosome is destabilized and the region between the flanking nucleosome is increased probably reflecting the binding of the androgen receptor that could act as boundary that direct the positioning of nearby nucleosomes [82]. However, it is unknown if the eviction of the central nucleosome is a consequence of active chromatin remodeling or if rather the presence of the H2A.Z histone variant causes nucleosome instability.

Nucleosome positioning is also important at the 3' chicken α -globin enhancer. When the enhancer reaches its highest activity, two flanking nucleosomes are well positioned and two additional nucleosomes, shielding the functional part, are evicted concomitant with the appearance of two micrococcal hypersensitive sites (MHSs). Introduction of unrelated lambda DNA increasing the distance between this two the MHSs affects enhancer activity and nucleosome organization [92, 93]. The local chromatin configuration of the enhancer is dependent on the binding of the erythroid-specific factor GATA-1, probably by recruiting Brg1, since it physically interacts with GATA-1 at the α -globin locus of murine fetal liver cells [94, 95]. The artificial expression of the pioneer factor GATA-1 in HeLa cell line resulted in the formation of DHSs and hyperacetylation of histones within GATA-1 bound regulatory elements. Functional analysis of GATA-1 binding at interaction sites in an erythroid context revealed that GATA-1 plays a major role in defining both chromatin structure and enhancer activity of the 3' chicken α -globin enhancer since point mutations of GATA-1 interaction sites abolish both the DHSs and the enhancer capacity to *trans*-activate a reporter gene [92]. In an attempt to distinguish between the two models of α -globin enhancer activity (binary vs. progressive), we identified a novel regulatory element which modulates the activity of the chicken α -globin enhancer [92, 93]. This element, named upstream enhancer element (UEE), increases the activity of

this enhancer but not the numbers of cells in which the enhancer is active in agreement with the progressive model [93]. We propose that this mode of action is in part facilitated by the fixed positioning of nucleosomes located upstream and downstream of the core enhancer since the UEE is important for positioning the nucleosome located upstream of the enhancer's core [92, 93].

Nucleosome eviction at promoters has also been described, for example upon transcriptional activation of the yeast *PHO5* gene promoter and for the mammalian MMTV LTR promoter regulated by steroid hormones [96, 97]. Therefore, nucleosome eviction is a phenomena that is shared by both enhancers and promoters sequences, which further suggests, along with enrichment of shared histone marks (like H3K4me1), that both regulatory sequences have common mechanisms of action (see later “[Similarities between enhancers and promoters](#)”).

Together, these results suggest that the coordinated repositioning and/or eviction of nucleosomes through chromatin remodeling complexes represent a critical step to render functional enhancers.

DNA hydroxymethylation in enhancer function

An alternative proposed mechanism involved in the regulation of enhancer activity is related to DNA methylation and hydroxymethylation. It has been well established that cytosine DNA methylation (5mC) is an essential epigenetic modification catalyzed by DNA methyltransferases [98, 99]. Genome-wide studies show a negative correlation between DNA methylation and chromatin accessibility [81]. Similarly, TFs and co-activator binding are inversely correlated with DNA methylation [85, 100, 101]. Concordantly, active enhancers are normally depleted of DNA methylation [100, 102].

5mC can be further oxidized by TET proteins through the active demethylation pathway generating 5hmC, 5fC, and 5caC, which can regenerate unmodified cytosines [103, 104]. Importantly, cytosine 5-hydroxymethylation (5hmC), which is the initial intermediate in the enzymatic cytosine demethylation cascade [103, 104], is a novel epigenetic modification that overlaps with histone H3K4me1 and H3K27ac at enhancers during differentiation [105, 106]. A single base resolution map of 5hmC in mouse and human ESCs, revealed that 5hmC is most abundant at both poised and active enhancers [107], rather than at CpG-rich promoters, as was previously suggested [108, 109]. Importantly, 5hmC enrichment at enhancers is characterized by a bimodal distribution flanking TF-binding sites [107].

The role of 5hmC on enhancer function remains to be determined, but one possibility is that a protein directly

recognizes this DNA modification, which then can affect transcriptional activation. Proteins have been identified that preferentially recognize the 5hmC mark. For instance, the peptide product of the *Uhrf2* gene whose expression increases upon differentiation of neural progenitor cells (NPC) [104, 110, 111]. Alternatively, new candidate molecules that can shape the DNA methylation landscape by recruiting both DNMT and TET enzymes to chromatin are long non-coding RNAs [112, 113]. However, further studies are needed to demonstrate that these mechanisms can occur at enhancer regions.

Long non-coding RNA with enhancer-like function

Most of the mammalian genome (70–80 %) has the potential to be transcribed into non-coding RNAs (ncRNAs) [114–116]. Recent experimental data support a role for lncRNAs in transcriptional regulation [117, 118]. lncRNAs have been implicated in cellular processes such as X chromosome inactivation, genomic imprinting, development, cell differentiation, and several pathologies, among many other functions [119, 120]. Using a model of primary human keratinocytes, long non-coding RNAs that respond to cell differentiation signals were found to be upregulated [118]. These non-coding transcripts act by stimulating transcription and behave similarly to classical enhancers. It has been observed that this kind of transcript interacts with the Mediator complex, thereby favoring long-range chromatin interactions between two loci [121]. As an alternative molecular mechanism to stimulate gene expression, lncRNAs might produce a more permissive chromatin environment at regulatory regions by direct recruitment of chromatin remodeling complexes [122]. Therefore, particular attention needs to be paid to define whether lncRNAs themselves possess or contribute to enhancer function and if the Mediator complex as a critical player in these interdependent interactions.

Non-coding transcription at enhancers, a novel regulatory mechanism

Detailed genome-wide analysis by RNA-seq and Global Run-On sequencing (GRO-seq) assays at high resolution, together with the data accumulated by the ENCODE Project reveals that a majority of active enhancers are transcribed in a cell-type-specific manner and produce relatively short transcripts (200–300 nucleotides), known as enhancer RNAs or eRNAs [45, 49, 51, 71, 114, 123–126]. This type of transcript has been observed in different cellular contexts including primary neuronal cultures, myogenic cells, mouse macrophages, and breast cancer

cells among many others [71, 124–127]. eRNAs exhibit a 5' cap and are generally not spliced or polyadenylated. Interestingly, the majority of enhancer transcription is bidirectional (Fig. 2) [128]. Apparently, there is a tight correlation between eRNA expression and the transcription of nearby genes, even though eRNA relative abundance is frequently low [128]. In line with this, transcription of eRNAs is positively correlated with an acquisition of active histone marks at enhancers, particularly H3K27ac, and the absence of the repressive histone H3K27me3 mark (Fig. 2) [129, 130].

An important question is how eRNAs become activated as specifically regulated transcription units and, of course, which mechanisms define the precise initiation site of enhancer transcription. Additionally, an aspect that remains unanswered is if eRNA represent a general feature of enhancer function, and if so, which is the hierarchy of eRNA transcription during enhancer action.

eRNAs have been implicated in different processes. For example, they participate in the formation of enhancer-promoter contacts in collaboration with the cohesin complex (RAD21 and SMC3) (Fig. 2) [125] and they can interact with the negative elongation factor (NELF) complex to facilitate the transition of paused RNA polymerase II into productive elongation by a decoy mechanism upon induction of immediate early genes in neurons [44, 131]. Alternatively, in macrophages, co-repressor/histone deacetylase complex NCoR-HDAC3 can be recruited by the nuclear receptors Rev-Erbs at the response elements in enhancers and promoters of target genes to establish a macrophage-specific program of repression. The repressive function of Rev-Erbs is dependent of their ability to inhibit eRNAs transcription, therefore suggesting eRNAs transcription is essential for gene expression of enhancer target genes [124].

In summary, there is increasing evidence that implicates eRNAs in the activity of enhancers. Their mechanism of action remains to be further investigated but there is no doubt that they have a relevant role in transcription regulation. One unexplored function of eRNA may have to do with multiple long-distance interactions that can be mediated by CTCF or even the Mediator complex. Importantly, they may be linked to different diseases, in particular, because a significant number of SNPs are found in enhancer regions perturbing chromatin remodeling activities and/or TFs binding [132].

Similarities between enhancers and promoters

Evidence for enhancer transcription has led the field to compare the nucleosome and chromatin architecture of enhancers and TSSs [51]. Examination of nascent RNA

identified common architecture of transcriptional initiation characterized by NDR formation and the symmetry of two nucleosomes flanking divergent transcription pairs at enhancers suggesting that the same principle applies to both enhancers and promoters (Fig. 2). However, to what extent the chromatin architecture associated with promoters and enhancers accounts for transcription initiation will require further analysis since NDR formation has been also observed at the 3' end of genes [83, 133–135]. Whether a universal architecture for promoters and enhancers is applicable to all loci is still controversially debated [82, 136].

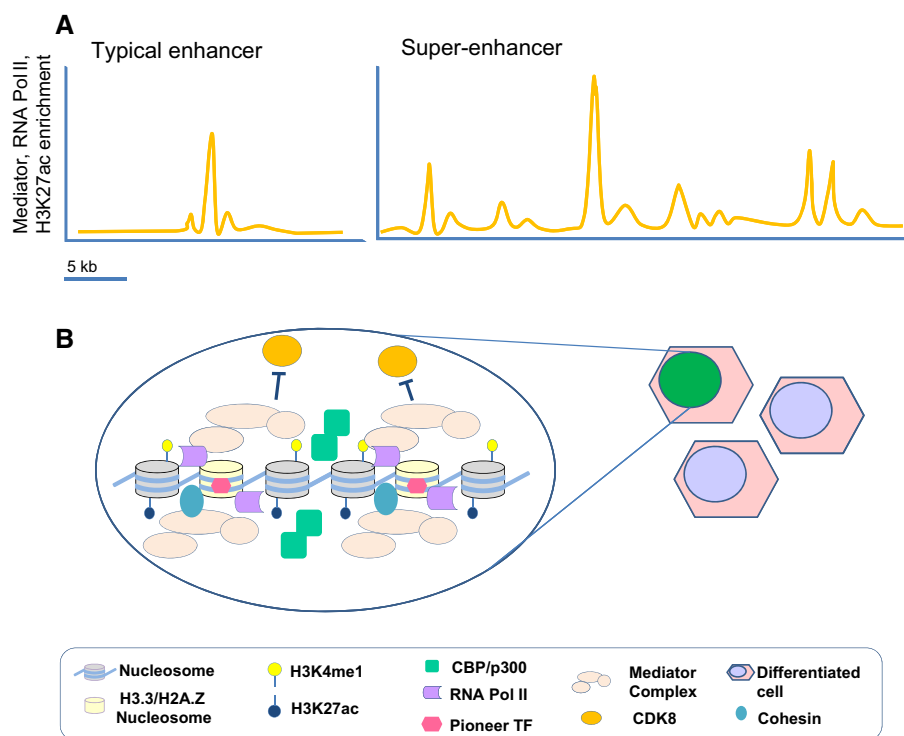
Since the line distinguishing transcription start sites from enhancers has become blurred by the finding that transcription originates from enhancers and promoters, histone modifications seem to not be responsible for differences in transcription levels [51], therefore it is possible that chromatin features cannot distinguish between them (Figs. 1, 2). However, at the functional level, both elements work differently as the excision of the proximal promoter region does not elicit transcription despite the presence of the enhancer [19, 71].

In any case, different regulatory mechanisms remain to be studied. For example, and based in genome-wide scale studies, in particular, with antibodies against a sub-set of histone modification signatures, active enhancer for some cell types and tissues have been found, with active promoters having different signatures. Based on such difference it has been proposed that this is a group of *cis*-regulatory elements with dynamic signatures also named cREDS [137]. This type of regulatory element shows histone marks associated with active enhancers like H3K27ac and H3K4me1 in a particular tissue, but in other cell types, histone modifications shift towards H3K4me3 [137]. In addition, intragenic enhancers can function as promoters of the embedded gene generating abundant, spliced and multi-exonic poly(A)⁺ RNAs but with low protein coding potential and thus far with unknown function [138]. In conclusion, there are many convergent features between promoters and enhancers that deserve a new visualization and strategies to study them.

Super-enhancers coordinate the expression of tissue-specific genes

Several studies have implicated the Mediator complex, which is essential for the maintenance of stem cells and their differentiation [139–141] as a facilitator of the RNA Pol II recruitment to enhancers via TFs (Fig. 3) [142, 143]. Recently, large areas of up to 50 kb showing a great enrichment of Mediator complex components were termed super-enhancers [144]. Despite the difficulty to

Fig. 3 a Comparison of genomic features associated with typical vs. super-enhancers. Super-enhancers are not functionally defined but they harbor chromatin enhancer features like very high enrichment of Mediator complex, H3K27ac, RNA Pol II and TFs, they span large genomic areas, and regulate genes that specify cell identity. **b** Mediator complex binds genomic regions in a tissue-specific manner and may be directly involved in integrating external and internal metabolic signals to modulate gene expression. The dynamic structure and composition of the Mediator complex may be a key feature in regulating the process of transcription since its interaction with the CDK8 subunit prevents its interaction with RNA Pol II



functionally define super-enhancers, they frequently harbor one or multiple (up to six) DHSs, each one with enhancer features including RNA Pol II association, high density of TF binding, and the highest enrichment of H3K27ac histone mark (Fig. 3) [144, 145]. Intriguingly, Mediator subunits seem to have different roles in diverse cell types, and they can interact with structural proteins such as cohesin in a tissue-specific manner [139, 146]. Future work is necessary to unravel the possible mechanisms driving cell fates regulated by the Mediator complex.

Three-dimensional circuitry of enhancer sequences

Regulatory elements like enhancers and promoters interact in space through loop formation and this has important consequences for gene expression [15]. Loop formation is thought to increase the local concentration of TFs or cofactors involved in transcriptional regulation including cohesin or Mediator [139, 147, 148]. Most of the loops occur within topologically associating domains (TADs) in metazoans like *D. melanogaster* and mammals. TADs are megabase-sized chromosomal regions often invariant between tissues. Sequences inside TADs interact more frequently with each other than with sequences outside the TADs [149–152]. These structures have been proposed to create a microenvironment where genes and regulatory sequences can interact with high frequency, and share

histone modifications and patterns of gene expression [149, 152]. For example, *Shh* gene expression depends on an enhancer located 1 Mb away from its target promoter. Both enhancer and promoter are inside the same TAD and contact each other through loop formation [17].

Despite extensive annotation of enhancer sequences through mapping of histone modifications like H3K4me1, H3K27ac, or cofactors like p300 [19, 52, 64, 153, 154] enhancers looping dynamics have not been fully investigated. Using chromosome conformation capture-based (Hi-C) method to map long-range chromatin interactions maps for specific regulatory sequences elucidated a complex network of interactions between all annotated promoters and enhancers was recently elucidated in mouse embryonic stem cells (mESCs) and fetal liver cells (FLC) [154]. Less than half of the promoter-enhancer interactions are conserved between mESCs and FLCs. This also holds for human embryonic stem cells (hESCs) and differentiated human cells suggesting that enhancer looping interactions are heavily rewired during cells differentiation in mouse and human [154–156]. In contrast, in *D. melanogaster*, a small set of developmental enhancers active in mesodermal embryonic cells (6–8 h after egg laying) show no significant change of looping interactions with their target promoters despite a clear induction of gene expression for their associated genes when compared with multipotent cells (3–4 h after egg laying) where mesodermal genes are not expressed [157]. Authors suggest that those enhancer

looping interactions are pre-established very early in development and possibly primed for rapid gene expression since paused RNA Pol II was found at the sites of loop interaction. Thus, in human and mouse cells, differentiation is accompanied by an extensive change of enhancer-promoter looping interactions while in *D. melanogaster* enhancer-promoter looping interactions may be set very early during embryonic development. Whether these differences originate from the small number of enhancer sequences analyzed in *D. melanogaster* or it reflects a biological difference in the dynamics of enhancer-promoter looping interactions, possibly due to differences in early development between vertebrates and invertebrates (*D. melanogaster* embryonic development lasts just 24 h) remains to be determined.

Looping interactions seem to be very different between cell types but little is known about their dynamics as a response to different stimuli in differentiated cells. In humans, a high-resolution genome-wide map of chromatin contacts (Hi-C) in primary fibroblast cells treated and untreated with *TNF alpha* suggest that despite a clear activation of *TNF alpha* response genes, the vast majority of enhancer chromatin contacts with promoters of *TNF alpha* response genes are unchanged [155]. This trend was also evident in other cell types and under different stimuli. This suggests that at least in human differentiated cells, pre-existing looping interactions between enhancer and promoters of genes responsive to signaling inducible TFs are a feature of the 3D regulatory landscape of the genome. Even though, looping interactions seem to be pre-established at enhancers and promoters in differentiated cells the connectivity landscape of these regulatory sequences is very different between cell types [154–156].

The molecular mechanisms by which enhancer and promoter interactions are established are just becoming clear but the CCCTC-binding factor (CTCF) seems to be directly involved. CTCF is a DNA-binding protein best characterized as an insulator associated protein and that has been suggested to be a critical player for genome organization through formation of chromatin loops with cohesin (Figs. 1, 2). Remarkably, 92 % of chromatin loops identified in human cells have CTCF binding sites (CBS) in a convergent orientation [155]. Most of the loci at the anchors of the loops were enhancers and promoters, which suggest that CTCF binding in a convergent orientation, could bring together regulatory elements through loop formation. At the protocadherin (*Pcdh*) gene cluster, the inversion of CBS with a reversed orientation at the *Pcdh* enhancer using the CRISPR-Cas9 system abolished loop formation between enhancer and target promoters in human and mouse neurons [158]. This data strongly support that the convergent orientation of CBS is critical for loop formation.

The use of the CRISPR-Cas9 system for genome editing has opened a whole new era of research and will enable us to test the function of enhancers endogenously by direct removal or disruption of their sequence in vivo. Moreover, this new technology will allow to understand the mechanistic consequences of mutations within enhancer sequences that are associated with certain diseases, e.g., by creating them in animal models [159, 160]. For instance, analysis of human patient-derived cells demonstrated that mutations identified in human limb malformation syndromes affected proper interactions between enhancers and promoters of the WNT6/IHH/EPHA4/PAX3 region mainly by disrupting TAD borders. This highlights the importance of insulation of genes and their regulatory elements, like enhancers, by TADs to ensure a proper program of gene expression, protecting against non-canonical interactions with other regulatory elements [160].

The emerging picture suggests that enhancers act as complex networks that favor three-dimensional contacts with regulatory regions in the genome where context-dependent signaling outcomes may determine cell fate choices.

Conclusions and prospects

Functional specialization of cells and tissue types is vital for all multicellular organisms. This requires cells to respond to developmental and environmental cues by generating specific patterns of gene expression on the basis of an identical set of genetic material. Enhancers are the main regulators that enable cell-type-specific gene expression. To gain such accuracy, enhancers are bound by specific TFs and 3D chromatin architecture-mediating proteins, marked by specific post-translational histone modifications, and generate non-coding transcripts called eRNAs. In addition, high-resolution maps of promoter–enhancer interactions are providing new insights into enhancer function, generating hypotheses which can now be directly tested using genome editing techniques. With these new tools it will be possible in the future to explore, in great detail, how a cell adopts its unique transcriptional identity.

Acknowledgments We acknowledge Karin Meier for critical reading of the manuscript. This work was supported by the DGAPA-PAPIIT, UNAM (IN209403, IN203811 and IN201114), CONACyT (42653-Q, 128464 and 220503) and Fronteras de la Ciencia-2015 (Grant 290) to FR-T, and by a PhD fellowship from CONACyT and Programa de Apoyo a los Estudios del Posgrado (PAEP), UNAM to EG-G and RA-M. Additional support was provided by the Ph.D. Graduate Program, “Doctorado en Ciencias Bioquímicas y Biomédicas”, to the Instituto de Fisiología Celular, Universidad Nacional Autónoma de México.

References

1. Khoury G, Gruss P (1983) Enhancer elements. *Cell* 33:313–314
2. Banerji J, Rusconi S, Schaffner W (1981) Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27:299–308
3. Banerji J, Olson L, Schaffner W (1983) A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33:729–740
4. Gillies SD, Morrison SL, Oi VT, Tonegawa S (1983) A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33:717–728
5. Walker MD, Edlund T, Boulet AM, Rutter WJ (1983) Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. *Nature* 306:557–561
6. Edlund T, Walker MD, Barr PJ, Rutter WJ (1985) Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5' flanking elements. *Science* 230:912–916
7. Goodbourn S, Zinn K, Maniatis T (1985) Human β -interferon gene expression is regulated by an inducible enhancer element. *Cell* 41:509–520
8. Schöler HR, Gruss P (1985) Cell type-specific transcriptional enhancement in vitro requires the presence of *trans*-acting factors. *EMBO J* 4:3305–3313
9. Ko MSH, Nakauchi H, Takahashi N (1990) The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J* 9:2835–2842
10. Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, Martin DIK (1995) Enhancers increase the probability but not the level of gene expression. *Proc Natl Acad Sci USA* 92:7125–7129
11. Sutherland HGE, Martin DIK, Whitelaw E (1997) A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. *Mol Cell Biol* 3:1607–1614
12. Yie J, Senger K, Thanos D (1999) Mechanism by which the IFN- β enhanceosome activates transcription. *Proc Natl Acad Sci USA* 96:13108–13113
13. Sandaltzopoulos R, Becker PB (1998) Heat shock increases the reinitiation rate from potentiated chromatin templates. *Mol Cell Biol* 18:361–367
14. Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K (2012) Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Res* 22:490–503
15. Bulger M, Groudine M (2011) Functional and mechanistic diversity of distal transcription enhancers. *Cell* 144:327–339
16. Ong CT, Corces VG (2011) Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* 12:283–293
17. De Laat W, Duboule D (2013) Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499–506
18. Calo E, Wysocka J (2013) Modification of enhancer chromatin: what, how and why? *Mol Cell* 49:825–837
19. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD et al (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39:311–318
20. Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN et al (2009) ChIP-Seq of ER α and RNA polymerase II defines genes differentially responding to ligands. *EMBO J* 28:1418–1428
21. Soufi A, Donahue G, Zaret KS (2012) Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151:994–1004
22. Spitz F, Furlong EE (2012) Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* 13:613–626
23. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS (2015) Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161:555–568
24. Arnosti DN, Kulkarni MM (2005) Transcriptional enhancers: intelligent enhanceosomes or flexible billboards? *J Cell Biochem* 94:890–898
25. Merika M, Williams AJ, Chen G, Collins T, Thanos D (1998) Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription. *Mol Cell* 1:277–287
26. Panne D (2008) The enhanceosome. *Curr Opin Struct Biol* 18:236–242
27. Vega VB, Lin CY, Lai KS, Kong SL, Xie M, Su X et al (2006) Multiplatform genome-wide identification and modeling of functional human estrogen receptor binding sites. *Genome Biol* 7:R82
28. Joseph R, Orlov YL, Huss M, Sun W, Kong SL, Ukil L et al (2010) Integrative model of genomic factors for determining binding site selection by estrogen receptor- α . *Mol Syst Biol* 6:456
29. Zaret KS, Carroll JS (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25:2227–2241
30. Siersbæk R, Rabiee A, Nielsen R, Sidoli S, Traynor S, Loft A et al (2014) Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Rep* 7:1443–1455
31. Adams CC, Workman JL (1995) Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Mol Cell Biol* 15:1405–1421
32. Bell O, Tiwari VK, Thomä NH, Schübeler D (2011) Determinants and dynamics of genome accessibility. *Nat Rev Genet* 12:554–564
33. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenyk M, Yen A, Heravi-Moussavi A et al (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518:317–330
34. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676
35. Iwafuchi-Doi M, Zaret KS (2014) Pioneer transcription factors in cell reprogramming. *Genes Dev* 28:2679–2692
36. You JS, Kelly TK, De Carvalho DD, Taberlay PC, Liang G, Jones PA (2011) OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. *Proc Natl Acad Sci USA* 108:14497–14502
37. Sammons MA, Zhu J, Drake AM, Berger SL (2014) TP53 engagement with the genome occurs in distinct local chromatin environments via pioneer factor activity. *Genome Res* 25:179–188
38. Smith E, Shilatifard A (2014) Enhancer biology and enhanceropathies. *Nat Struct Mol Biol* 21:210–219
39. Taberlay PC, Kelly TK, Liu CC, You JS, De Carvalho DD, Miranda TB et al (2011) Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* 147:1283–1294
40. Malik S, Roeder RG (2005) Dynamic regulation of Pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* 30:256–263

41. Weake VM, Workman JL (2010) Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet* 11:426–437
42. Vignali M, Hassan AH, Neely KE, Workman JL (2000) ATP-dependent chromatin remodeling complexes. *Mol Cell Biol* 5:1899–1910
43. Näär AM, Lemon BD, Tijan R (2001) Transcriptional coactivator complexes. *Annu Rev Biochem* 70:475–501
44. Becker PB, Hörz W (2002) ATP-dependent nucleosome remodeling. *Annu Rev Biochem* 71:247–273
45. ENCODE Project Consortium et al (2007) Identification and analysis of functional elements in 1 % of the human genome by the ENCODE pilot project. *Nature* 447:799–816
46. Kornberg RD, Lorch Y (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98:285–294
47. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z et al (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837
48. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S et al (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40:897–903
49. Koch F, Andrau JC (2011) Initiating RNA polymerase II and TIPS as hallmarks of enhancer activity and tissue-specificity. *Transcription* 2:263–268
50. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H et al (2011) H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J* 30:4198–4210
51. Core LJ, Martins AL, Dank CG, Waters CT, Siepel A, Lis JT (2014) Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet* 46:1311–1320
52. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470:279–283
53. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ et al (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA* 107:21931–21936
54. Zentner GE, Tesar PJ, Scacheri PC (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* 21:1273–1283
55. Bogdanovic O, Fernandez-Miñán A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, van Kruysbergen I et al (2012) Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. *Genome Res* 22:2043–2053
56. Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N et al (2012) Tissue specific analysis of chromatin states identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet* 44:148–156
57. Whyte WA, Bilodeau S, Orlando DA, Hoke HA, Frampton GM, Foster CT et al (2012) Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* 482:221–225
58. Altaf M, Auger A, Monnet-Saksouk J, Brodeur J, Piquet S, Cramet M et al (2010) NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *J Biol Chem* 285:15966–15977
59. Jeong KW, Kim K, Situ AJ, Ulmer TS, An W, Stallcup MR (2011) Recognition of enhancer element-specific histone methylation by TIP60 in transcriptional activation. *Nat Struct Mol Biol* 18:1358–1365
60. Lan F, Collins RE, De Cegli R, Alpatov R, Horton JR, Shi X et al (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* 448:718–722
61. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* 448:714–717
62. Lyko F, Ramsahoye BA, Jaenisch R (2000) DNA methylation in *Drosophila melanogaster*. *Nature* 408:538–540
63. Cho H, Orphanides G, Sun X, Yang X-J, Ogryzko V, Lees E et al (1998) A human RNA polymerase II complex containing factors that modify chromatin structure. *Mol Cell Biol* 18:5355–5363
64. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A et al (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457:854–858
65. Yao TP, Oh SP, Fuchs M, Zhou ND, Ch'ng LE, Newsome D et al (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* 93:361–372
66. Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A et al (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* 136:3131–3141
67. Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE et al (2011) Distinct roles of GCN5/PCAF-mediated H3K9ac and CB/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* 30:249–262
68. Holmqvist P-H, Mannervik M (2013) Genomic occupancy of the transcriptional co-activators p300 and CBP. *Transcription* 4:18–23
69. Mouchiroud L, Eichner LJ, Shaw RJ, Auwerx J (2014) Transcriptional coregulators: fine-tuning metabolism. *Cell Metab* 20:26–40
70. Cai L, Sutter BM, Li B, Tu BP (2011) Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Mol Cell* 42:426–437
71. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J et al (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465:182–187
72. Krebs AR, Karmodiya K, Lindahl-Allen M, Struhl K, Tora L (2011) SAGA and ATAC histone acetyltransferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell* 44:410–423
73. Daitoku H, Sakamaki J, Fukamizu A (2001) Regulation of FoxO transcription factors by acetylation and protein–protein interactions. *Biochem Biophys Acta* 1813:1954–1960
74. Speranzini V, Pilotto S, Sixma TK, Mattevi A (2016) Touch, act and go: landing and operating on nucleosomes. *EMBO J* 35:376–388
75. Tie F, Banerjee R, Conrad PA, Scacheri PC, Harte PJ (2012) Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. *Mol Cell Biol* 32:2323–2334
76. Hödl M, Basler K (2012) Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr Biol* 22:2253–2257
77. Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J (2013) A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* 339:698–699
78. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR (2012) Dynamics and memory of heterochromatin in living cells. *Cell* 149:1447–1460
79. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z et al (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132:311–322
80. Sabo PJ, Kuehn MS, Thurman R, Johnson BE, Johnson EM, Cao H et al (2006) Genome-scale mapping of DNase I sensitivity in vivo using DNA microarrays. *Nat Methods* 3:511–518
81. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E et al (2012) The accessible chromatin landscape of human genome. *Nature* 489:75–82

82. He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q et al (2010) Nucleosome dynamics define transcriptional enhancers. *Nat Genet* 42:343–347
83. Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ et al (2008) Nucleosome organization in the *Drosophila* genome. *Nature* 453:358–364
84. Iyer V, Struhl K (1995) Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J* 14:2570–2579
85. Anderson JD, Widom J (2001) Poly (dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. *Mol Cell Biol* 21:3830–3839
86. Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. *Annu Rev Biochem* 78:273–304
87. Shones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z et al (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132:887–898
88. Yasui D, Miyano M, Cai S, Varga-Weisz P, Kohwi-Shigematsu T (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature* 419:641–645
89. Yu Y, Chen Y, Kim B, Wang H, Zhao C, He X et al (2013) Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell* 152:248–261
90. Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB et al (2004) Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 36:955–957
91. Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Pereira CF, Fisher AG et al (2010) CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet* 6:e1001023
92. Escamilla-Del-Arenal M, Recillas-Targa F (2008) GATA-1 modulates the chromatin structure and activity of the chicken α -globin 3' enhancer. *Mol Cell Biol* 28:575–586
93. García-González E, Recillas-Targa F (2014) A regulatory element affects the activity and chromatin structure of the chicken α -globin 3' enhancer. *Biochim Biophys Acta* 1839:1233–1241
94. Xu Z, Meng X, Cai Y, Koury MJ, Brandt SJ (2006) Recruitment of the SWI/SNF protein Brg1 by a multiprotein complex effects transcriptional repression in murine erythroid progenitors. *Biochem J* 399:297–304
95. Kim SI, Bresnick EH, Bultman SJ (2009) BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nucleic Acids Res* 37:6019–6027
96. Almer A, Rudolph H, Hinnen A, Hörz W (1986) Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction release additional upstream activating DNA elements. *EMBO J* 5:2689–2696
97. Richard-Foy H, Hager GL (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J* 6:2321–2328
98. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99:247–257
99. Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A et al (2014) DNA methylation dynamics of the human preimplantation embryo. *Nature* 511:611–615
100. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A et al (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480:490–495
101. Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, Vernot B et al (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489:83–90
102. Bock C, Beerman I, Lien W-H, Smith ZD, Gu H, Boyle P, Gnirke A, Fuchs E, Rossi DJ, Meissner A (2012) DNA methylation dynamics during *in vivo* differentiation of blood and skin stem cells. *Mol Cell* 47:633–647
103. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL et al (2013) Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* 153:692–706
104. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C et al (2013) Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* 152:1146–1159
105. Sérandour AA, Avner S, Percevault F, Demay F, Bizot M, Lucchetti-Miganeh C et al (2011) Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res* 21:555–565
106. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol* 12:R54
107. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A et al (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* 149:1368–1380
108. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M et al (2011) Genome wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473:394–397
109. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K et al (2011) Genome wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* 25:679–684
110. Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I et al (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One* 6:e21306
111. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* 151:1417–1430
112. Di Ruscio A, Ebraldize AK, Benoukraf T, Amabile G, Goff LA, Terragni J et al (2013) DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* 503:371–376
113. Arab K, Park YT, Lindroth AM, Schäfer A, Oakes C et al (2014) Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Mol Cell* 55:604–614
114. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK et al (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol* 8:e1000384
115. Natoli G, Andrau JC (2012) Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet* 46:1–19
116. Mousavi K, Zare H, Koulunis M, Sartorelli V (2014) The emerging roles of eRNAs in transcriptional regulatory networks. *RNA Biol* 11:106–110
117. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D et al (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA* 106:11667–11672
118. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G et al (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143:46–58
119. Heard E, Disteché CM (2006) Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 20:1848–1867
120. Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. *Annu Rev Biochem* 81:145–166
121. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA et al (2013) Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* 494:497–501

122. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB et al (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* 18:1433–1445
123. Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT et al (2011) A rapid, extensive and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* 145:622–634
124. Lam MTY, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y et al (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* 498:511–515
125. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY et al (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498:516–520
126. Mousavi K, Zare H, Dell'Orso S, Grontved L, Gutierrez-Cruz G, Derfoul A et al (2013) eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol Cell* 51:606–617
127. Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA et al (2013) eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol Cell* 49:524–535
128. Lam MTY, Li W, Rosenfeld MG, Glass CK (2014) Enhancer RNAs and regulated transcriptional programs. *Trends Biochem Sci* 39:170–182
129. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A et al (2012) Landscape of transcription in human cells. *Nature* 489:101–108
130. Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD et al (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell* 51:310–325
131. Schaukowitz K, Joo JY, Liu X et al (2014) Enhancer RNA facilitates NELF release from immediate early genes. *Mol Cell* 56:29–41
132. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J et al (2014) An atlas of active enhancers across human cell types and tissues. *Nature* 507:455–461
133. Jiang C, Pugh F (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* 10:161–172
134. Jansen A, Verstrepen KJ (2011) Nucleosome positioning in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 75:301–320
135. Huang H, Liu H, Sun X (2013) Nucleosome distribution near the 3' ends of genes in the human genome. *Biosci Biotechnol Biochem* 77:2051–2055
136. Duttke SH, Lacadie SA, Ibrahim MM, Glass CK, Corcoran DL, Benner C, Heinz S, Kadonaga JT, Ohler U (2015) Human promoters are intrinsically directional. *Mol Cell* 57:674–684
137. Leung D, Jung I, Rajagopal N, Schmitt A, Selvaraj S, Lee AY et al (2015) Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature* 518:350–354
138. Kowalczyk MS, Hughes JR, Garrick D, Lynch MD, Sharpe JA, Sloane-Stanley JA et al (2012) Intragenic enhancers act as alternative promoters. *Mol Cell* 45:447–458
139. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL et al (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430–435
140. Risley MD, Clowes C, Yu M, Mitchell K, Hentges KE (2010) The Mediator complex protein Med31 is required for embryonic growth and cell proliferation during mammalian development. *Dev Biol* 342:146–156
141. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH et al (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153:307–319
142. Malik S, Roeder RG (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* 11:761–772
143. Ito T, Ikehara T, Nakagawa T, Kraus WL, Muramatsu M (2000) p300-mediated acetylation facilitates the transfer of histone H2A–H2B dimers from nucleosomes to a histone chaperone. *Genes Dev* 14:1899–1907
144. Pott S, Lieb JD (2015) What are super-enhancers? *Nat Genet* 47:8–12
145. Allen BL, Taatjes DJ (2015) The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* 16:155–166
146. Fiering S, Whitelaw E, Martin DI (2000) To be or not to be active: the stochastic nature of enhancer action. *BioEssays* 4:381–387
147. Plank JL, Dean A (2014) Enhancer function: mechanistic and genome-wide insights come together. *Mol Cell* 55:5–14
148. Heinz S, Romanoski CE, Benner C, Glass CK (2015) The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* 16:144–154
149. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y et al (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376–380
150. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N et al (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485:381–385
151. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M et al (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148:458–472
152. Bouwman BA, de Laat W (2015) Getting the genome in shape: the formation of loops, domain and compartments. *Genome Biol* 16:154
153. Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K et al (2014) Genome-scale functional characterization of *Drosophila* developmental enhancers *in vivo*. *Nature* 512:91–95
154. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L et al (2015) Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet* 47:598–606
155. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY et al (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503:290–294
156. Schoenfelder S, Furlan-Magaril M, Mifsud B, Tavares-Cadete F, Sugar R, Javierre BM et al (2015) The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res* 25:582–597
157. Ghavi-Helm Y, Klein FA, Pakodzi T, Ciglar L, Noordermeer D, Huber W et al (2014) Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512:96–100
158. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU et al (2015) CRISPR inversion of CTCF sites alters genome topology an enhancer/promoter function. *Cell* 162:900–910
159. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E et al (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161:1012–1025
160. Sanyal A, Lajoie BR, Jain G, Dekker J (2012) The long-range interaction landscape of gene promoters. *Nature* 489:109–113