



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

“Regulación de la expresión del gen que codifica para la chaperona de histonas ASF1 por el factor CTCF en *Drosophila melanogaster*”

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
Maestra en Ciencias

PRESENTA:  
LB Cristina Muñoz Olivos

Dr. Félix Recillas Targa  
[Instituto de Fisiología Celular](#)

MIEMBROS DEL COMITÉ TUTOR

Dra. Rosa Estela Navarro González  
[Instituto de Fisiología Celular](#)  
Dr. Juan Miranda Ríos  
[Instituto de Investigaciones Biomédicas](#)

Ciudad de México. Marzo, 2017.



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.

*A MIS PADRES*

*Por su ejemplo, por su apoyo  
incondicional y por ser  
mi inspiración.*

*Gracias.*

## AGRADECIMIENTOS INSTITUCIONALES

Este trabajo fue realizado en el Instituto de Fisiología Celular (IFC), en el departamento de Genética Molecular de la Universidad Nacional Autónoma de México (UNAM) bajo la tutoría del Dr. Félix Recillas Targa, con los apoyos de la Dirección General de Asuntos de Personal Académico (DGAPA-PAPIIT) 220503 y de Fronteras de la Ciencia 290.

Los estudios de maestría de Cristina Muñoz Olivos, con número de CVU 515013609, se realizaron con el apoyo de la Beca CONACYT Nacional para posgrados y con el Programa de Apoyo para Estudios del Posgrado (PAEP).

Se agradece y reconoce al comité tutor de este proyecto, constituido por:

Dr. Félix Recillas Targa	Instituto de Fisiología Celular
Dra. Rosa Estela Navarro González	Instituto de Fisiología Celular
Dr. Juan Miranda Ríos	Instituto de Investigaciones Biomédicas

Se agradece el apoyo y la excelente asesoría técnica de la Biol. Georgina Guerrero Avendaño. Asimismo se reconoce el trabajo de la laboratorista Gianelli Cortés González.

Adicionalmente se reconoce la asesoría académica de todos y cada uno de los miembros del laboratorio del Dr. Félix Recillas Targa, en particular al Lic. Rodrigo Gacel Arzate Mejía y al M. en C. Francisco Javier Carmona Aldana.

Se reconoce también el apoyo de la Unidad de Biología Molecular del Instituto de Fisiología Celular en las secuenciaciones de plásmidos realizadas en este proyecto.

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

Dra. Diana María Escalante Alcalde	Instituto de Fisiología Celular
Dra. María Soledad Funes Argüello	Instituto de Fisiología Celular
Dra. Lorena Aguilar Arnal	Instituto de Investigaciones Biomédicas
Dra. Martha Verónica Vázquez Laslop	Instituto de Biotecnología
Dr. Enrique Alejandro Reynaud Garza	Instituto de Biotecnología

## AGRADECIMIENTOS

Quisiera agradecer al Dr. Félix por brindarme la oportunidad de estar en este laboratorio y haberme permitido formar parte de un grupo de personas tan increíbles. Gracias por todas sus enseñanzas y consejos. Para mí su perseverancia, su pasión y dedicación a la ciencia son un gran ejemplo que me llenan de orgullo y admiración.

A mis papás por que siempre han sido mi motor y guía, porque con su ejemplo y apoyo he alcanzado muchas metas en mi vida y no han dejado que me pierda en el camino. A ti papá, por todo tu cariño, porque siempre has creído en mí y me has enseñado a tener fortaleza y perseverancia, enseñándome a levantar en los momentos difíciles. A ti mamá, por tu amor infinito, por tus invaluable consejos y porque me impulsas a ser siempre una mejor persona.

A Pily y a Angie por su amistad, por todas esas noches de alegrías, risas y locuras, por brindarme su apoyo en las buenas y en las malas, por todo lo que me han ayudado a crecer y porque junto a ustedes he aprendido que el ser familia no tiene nada que ver con compartir ADN.

A mi hermano Amauri, porque siempre has creído en mí y has estado dándome tu apoyo incondicional. Porque siempre me escuchas y me das consejos muy locos que hacen que mi vida sea más divertida.

Muchas gracias Geito, porque siempre me has cuidado y has estado al pendiente de mi, tanto en lo laboral como en lo personal. Porque con tu ternura, tus palabras y tus consejos siempre me diste ánimos. En ti me llevo una gran y querida amiga.

A Giane, por toda la ayuda y amistad que me brindaste, sin ti no habría logrado tantas cosas. Gracias por todas esas pláticas y risas que hicieron mi estancia muy divertida, no pude haberme sentado en un lugar mejor, junto a ti me la pase increíblemente bien. También gracias por tus cuidados y consejos.

A los chicos del laboratorio, que además de ser excelentes compañeros, se volvieron grandes amigos. A Gustavo, Hober, Carlos, Tania, Josué, Ceci, Josuecito, Karin, Rodri, Paco, Edgar y Jesús. Gracias a todos y a cada uno de ustedes. Gracias por su solidaridad, sus enseñanzas, sus consejos y más que nada por su amistad y por compartir conmigo momentos increíbles, llenos de risas y diversión. Definitivamente creo que no pude haber estado con un mejor equipo. En particular gracias a Rodri y a Paco, por haber hecho este proyecto posible, porque con paciencia me enseñaron y me brindaron sus conocimientos, además de que siempre me guiaron y orientaron. Gracias chicos por haber sido tan maravillosos conmigo.

# Contenido

1. RESUMEN.....	1
ABREVIATURAS .....	2
2. INTRODUCCIÓN .....	3
1.1 ANTECEDENTES GENERALES.....	3
Organización del genoma al interior del núcleo.....	3
Las chaperonas de histonas. ....	6
Papel de las chaperonas de histonas en el ensamblaje de la cromatina.....	7
El factor de Anti-silenciamiento 1 (ASF1). ....	9
Proteínas de unión a secuencias tipo “insulator”. ....	14
Funciones del factor CTCF. ....	16
1.2 ANTECEDENTES PARTICULARES .....	19
Papel de CTCF durante el desarrollo. ....	20
El factor nuclear CTCF está enriquecido en la región intergénica 5’ de <i>asf1</i> , lo que indica que podría estar participando en la regulación de su expresión. ....	21
Técnica CRISPR-Cas9 .....	22
3. PLANTEAMIENTO DEL PROBLEMA.....	24
4. HIPÓTESIS .....	25
5. OBJETIVOS.....	25
OBJETIVO GENERAL .....	25
OBJETIVOS PARTICULARES.....	25
6. MATERIALES Y MÉTODOS .....	26
Cultivo Celular .....	26
Oligonucleótidos para los ensayos de retardo de movilidad electroforética (EMSA). ....	26
Alineamiento de oligonucleótidos para los ensayos de retardo. ....	27
Marcaje de la sonda con fósforo radioactivo ( <sup>32</sup> P).....	27
Purificación de la sonda marcada.....	27
Cuantificación de la sonda. ....	28

Extracción de proteínas nucleares. ....	29
Cuantificación de proteínas totales. ....	29
Preparación de las muestras para los ensayos de retardo. ....	30
Anticuerpo.....	30
Plásmidos. ....	30
Diseño de los oligonucleótidos del RNA guía para la técnica CRISPR-Cas9.....	32
Alineamiento y fosforilación de oligonucleótidos .....	32
Digestión y desfosforilación del plásmido pAc-sgRNA-Cas9.....	33
Purificación del plásmido pAc-sgRNA-Cas9. ....	33
Ligación de oligonucleótidos en el vector pAc-sgRNA-Cas9. ....	33
Transformación. ....	33
Purificación de los plásmidos.....	34
Transfección para células S2R+.....	34
Extracción de ADN genómico por fenol-cloroformo. ....	35
Selección y dilución clonal. ....	35
Secuenciación.....	35
Extracción de ARN total. ....	36
Comprobación de la integridad del ARN.....	36
Reacción de transcripción reversa.....	37
Reacción en cadena de la polimerasa cuantativa (qPCR). ....	37
Tabla de oligonucleótidos .....	37
7. RESULTADOS.....	39
Unión de CTCF a la región intergénica 5' de <i>asfI</i> .....	39
Generación de la mutación del motivo de unión 1 a CTCF. ....	44
La eliminación de parte del motivo 1 de CTCF impide la unión de este factor.....	47
Regulación de la expresión de <i>asfI</i> por el factor CTCF. ....	49
8. DISCUSIÓN.....	52
9. CONCLUSIONES .....	57

10. PERSPECTIVAS ..... 57

11. REFERENCIAS ..... 59



# 1. RESUMEN

La chaperona de histonas ASF1 es una proteína multifuncional que ha sido implicada en el ensamblaje de cromatina dependiente e independiente de la replicación, en la reparación del ADN, en el remplazo de histonas, en la transcripción génica, en el reclutamiento de complejos co-represores y en el deslizamiento y la relocalización de nucleosomas en regiones promotoras. Sin embargo poco se sabe acerca de su regulación transcripcional. A partir de datos de ChIP-seq publicados en 2009 por el grupo de Victor Corces, en nuestro laboratorio se identificó que el factor nuclear CTCF está enriquecido río arriba del gen *asf1*, en el genoma de *Drosophila melanogaster* (*D. melanogaster*). Ya que CTCF se ha visto involucrado en diversos procesos regulatorios de activación y represión transcripcional, en la formación y mantenimiento de asas de cromatina y en la organización tridimensional del genoma, decidimos corroborar el papel que tiene CTCF en la regulación de *asf1*. En este proyecto demostramos que CTCF reconoce y puede unirse a un motivo de ADN que se encuentra en la región intergénica 5' de *asf1* y cuando dicho motivo se muta, la unión se elimina casi en su totalidad. En una población de la línea celular S2R+ enriquecida con una mutante que ha eliminado parte de este motivo de unión a CTCF, se incrementa la expresión *asf1* y de *l(3)76BDM*, que se encuentra río arriba de *asf1*. Con esto se demuestra que el factor nuclear CTCF participa en la regulación transcripcional de *asf1*, un gen involucrado en el mantenimiento global de la cromatina, indispensable para el desarrollo, supervivencia e implicado en vías de señalización.

## **ABREVIATURAS**

**ADN** – Ácido desoxirribonucleico.

**ARN** – Ácido ribonucleico

**ARNm** - ARN mensajero

**ASF1** – Factor de Anti-silenciamiento 1

**Cas9** – Proteína 9 asociada a CRISPR

**CIP**– Fosfatasa alcalina de intestino de becerro

**ChIP** - Inmunoprecipitación de cromatina

**ChIP-chip** – ChIP acoplado a microarreglo

**ChIP-seq** – Secuenciación de ChIP

**CP190** – Proteína centrosomal 190

**CRISPR** - Repeticiones Palindrómicas Cortas, Agrupadas y Regularmente Inter-espaciadas

**crRNA** – ARN CRISPR

**CTCF** – Factor de unión a CCCTC

**EMSA** – Ensayo de retardo de la movilidad electroforética

**H3K4me3** – Trimetilación de la lisina 4 de la histona H3

**l(3)76BDm** - Letal 3809

**PAM** – Motivo adyacente al espaciador

**PCR** – Reacción en cadena de la polimerasa

**PNK** - Polinucleótido 5'-hydroxil-cinasa

**qPCR** – PCR cuantitativa

**RP49** – Proteína ribosomal 49

**sgRNA** – RNA guía

**tracrRNA** – Molécula transactivadora de crRNA

**TSS** – Sitio de inicio de la transcripción

## 2. INTRODUCCIÓN

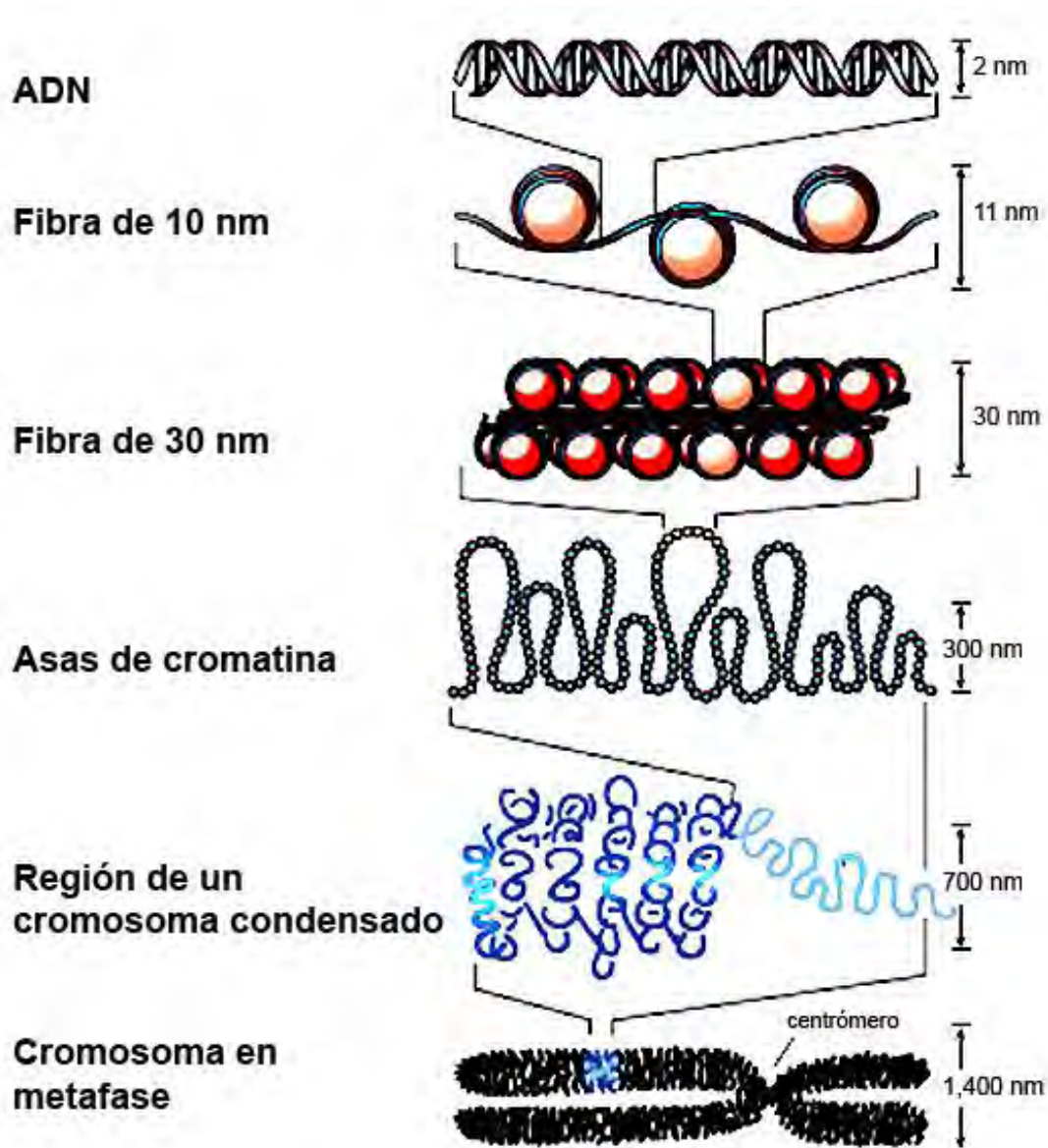
### 1.1 ANTECEDENTES GENERALES

#### **Organización del genoma al interior del núcleo.**

El genoma eucarionte se encuentra altamente estructurado al interior del núcleo en forma de cromatina, siendo su unidad fundamental el nucleosoma. Cada nucleosoma está conformado por 147 pares de bases de ADN enrolladas alrededor de dos tetrámeros de las histonas H2A, H2B, H3 y H4. A su vez, cada uno de estos nucleosomas se conecta con otros por medio de un segmento de unión de ADN (Felsenfeld y Groudine, 2003; Swaminathan et al., 2005). Asimismo, la cadena de polinucleosomas se organiza formando arreglos más complejos, por lo que la cromatina puede formar dominios y estructuras tridimensionales dinámicas (**Figura 1**; Felsenfeld y Groudine, 2003).

La organización de la cromatina debe ser finamente regulada ya que su estructura afecta la estabilidad y la expresión de la información genética. La formación de los nucleosomas permite que el ADN sea acomodado dentro del núcleo e influye en la accesibilidad de componentes celulares, como los factores transcripcionales y la ARN polimerasa al ADN (Ramachandran y Henikoff, 2015).

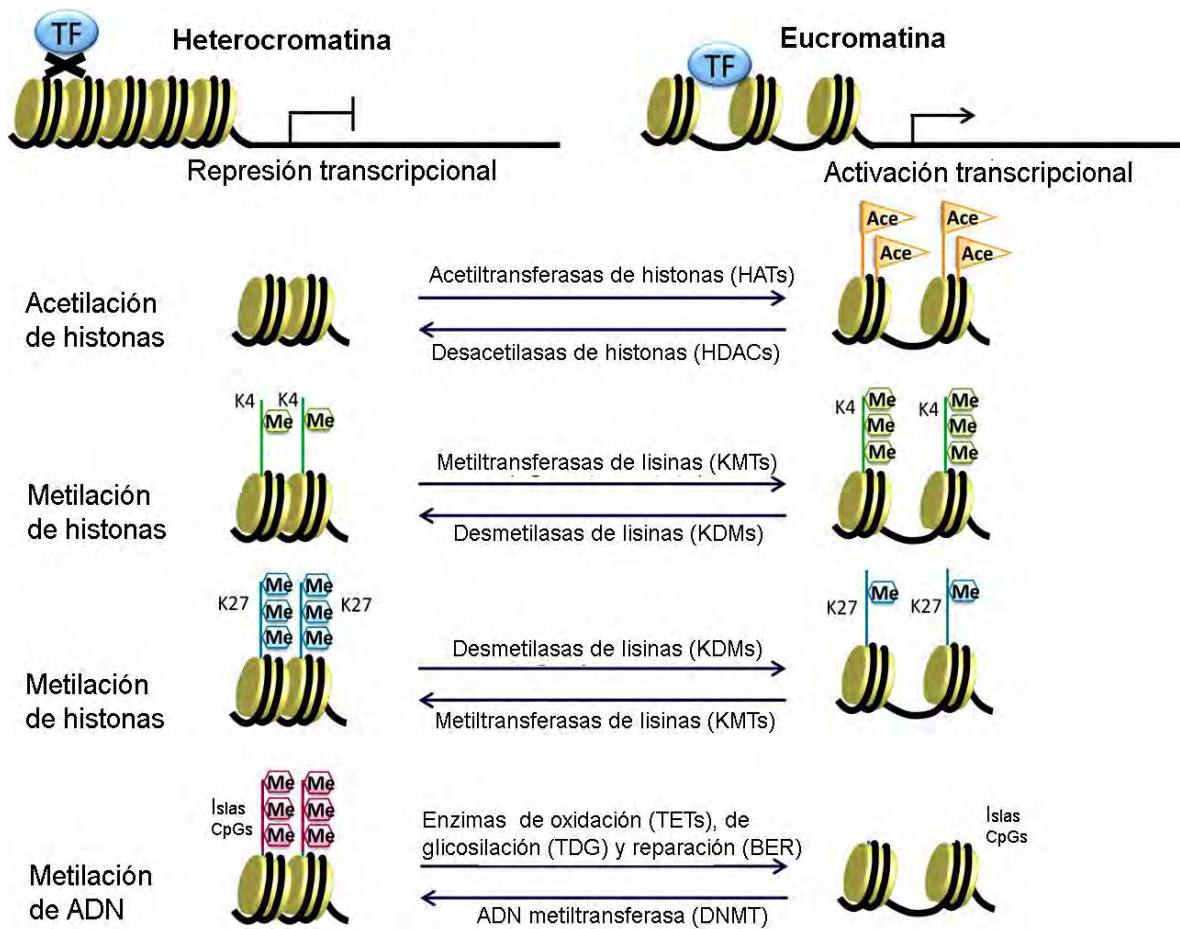
Dependiendo del nivel de compactación, la cromatina puede clasificarse en dos tipos: 1) Eucromatina o cromatina accesible a la maquinaria transcripcional y 2) Heterocromatina o cromatina compacta poco accesible, que suele ser inactiva transcripcionalmente (Felsenfeld y Groudine, 2003). Asimismo, hay dos tipos de heterocromatina en las células: A) la heterocromatina constitutiva; que es mantenida establemente en elementos repetidos, centrómeros, telómeros y transposones y B) la heterocromatina facultativa, que es transitoria y es finamente regulada, según sea el estado fisiológico de la célula. Ya que estos estados de la cromatina pueden cambiar, pueden propagarse a regiones vecinas y pueden afectar la expresión de genes, es importante establecer límites entre cada uno de los dominios cromatínicos. Las secuencias específicas en el genoma y la unión de proteínas efectoras, permiten la formación y la regulación de dichos dominios (Brahmachari y Jain, 2013).



**Figura 1. Niveles de organización del genoma.** El ADN al enrollarse alrededor de un octámero de histonas da lugar a la unidad más básica de la cromatina: el nucleosoma. En los niveles siguientes de organización, los nucleosomas forman estructuras más complejas, llegando al nivel de compactación máximo: el cromosoma metafásico (tomada y modificada de Felsenfeld y Groudine, 2003).

Las diferentes variantes de las histonas influyen las propiedades físicas y la dinámica de los nucleosomas. Por estas razones los estados cromatínicos son regulados mediante diversos mecanismos que alteran la estructura, la estabilidad y la composición de estos nucleosomas (Weber y Henikoff, 2014). Las modificaciones tanto sobre las histonas como sobre el ADN generan “marcas epigenéticas” que facilitan o restringen la accesibilidad al ADN y son reconocidas por factores regulatorios específicos (**Figura 2**; Berger, 2007). Dentro de los diferentes factores que participan en la modulación de la cromatina, están las

enzimas que modifican post-traduccionalmente a las histonas. Estas enzimas pueden generar modificaciones covalentes en las histonas como la acetilación, la metilación, la ubiquitinación y la fosforilación, entre otras. Además, la estructura de los nucleosomas es blanco de la acción de los remodeladores de cromatina dependientes de ATP, responsables de la movilización de los nucleosomas y de las chaperonas de histonas, encargadas de la inserción y remoción de las diferentes variantes de las histonas (Moshkin et al., 2009; Weber y Henikoff, 2014).



**Figura 2. Esquema de la regulación de la estructura de la cromatina.** La heterocromatina impide el acceso a factores transcripcionales (TF), mientras que la eucromatina es accesible para la activación transcripcional. Dichas estructuras son moduladas por la acción de diversos factores que modifican el ADN, las histonas y la densidad nucleosomal. K, lisina; Me, metilación; Ace, acetilación; Islas CpGs; islas de Citosina-Guanina (tomada y modificada de Ohtani y Dimmeler, 2011).

## **Las chaperonas de histonas.**

Las chaperonas de histonas son proteínas que se asocian selectivamente a diferentes tipos de histonas en cuanto estas últimas son sintetizadas. Las chaperonas participan en la oligomerización, el importe nuclear, el ensamblaje y la localización de las histonas en el genoma (Hammond et al., 2017). Además, colaboran en la asociación específica de las histonas con el ADN durante los diversos procesos celulares y modulan las modificaciones post-traduccionales de las histonas, lo que afecta potencialmente la expresión de genes (Avvakumov et al., 2011; De Koning et al., 2007).

Las chaperonas de histonas pueden ser clasificadas con base a su selectividad. Aquellas que reconocen preferencialmente los dímeros H2A-H2B y las que reconocen los dímeros H3-H4. Asimismo, todas ellas reconocen variantes particulares de cada histona. Algunas chaperonas se unen específicamente a aquellas variantes que se expresan en la fase S del ciclo celular o pueden reconocer a las variantes de remplazo que se expresan de manera constitutiva en bajos niveles y son independientes de la fase del ciclo celular (De Koning et al., 2007).

Dentro de las chaperonas que tienen mayor afinidad por las histonas H2A-H2B, están NAP1 (“Nucleosome Assembly Protein-1”), Nucleoplasmina, Nucleofosmina, TAF-I $\beta$  (“Template Activating Factor I $\beta$ ”), Nucleolina, VSP75 (“Variant-specific Surface Protein”) y FACT (“Facilitates Chromatin Transcription”). Las que se unen a los dímeros H3-H4 son ASF1 (“Anti-silencing Factor 1”), NASP (“Nuclear Autoantigenic Sperm Protein”), HIRA (“Histone cell cycle Regulation-defective homolog A”), CAF-1 (“Chromatin Assembly Factor-1”), Rtt106 (“Regulator of Ty1 Transposition”) y Spt6 (“Histone Chaperone Suppressor of Ty6”) (Burgess y Zhang, 2013; De Koning et al., 2007).

Es de suma relevancia que la unión de las chaperonas de histonas sea selectiva y suceda de manera espacio-temporal. Esto permite que exista un fino control para el ensamblaje y desensamblaje de la cromatina, dependiendo de los procesos celulares que estén llevándose a cabo. La unión de las chaperonas a las histonas se facilita en parte por interacciones electroestáticas entre las histonas que son básicas, con los fragmentos ácidos de algunas de las chaperonas; sin embargo, la afinidad depende en gran manera de las características estructurales de cada chaperona, como la existencia de  $\alpha$ -hélices y estructuras

$\beta$ , así como de las modificaciones post-traduccionales que presenten estos factores (De Koning et al., 2007).

Poco se sabe acerca de cómo se regulan estas chaperonas transcripcionalmente y cuáles factores de unión a la cromatina influyen en el control de dicha regulación. Esto es relevante debido al importante papel que juegan las chaperonas en los diversos procesos celulares. Nosotros tenemos particular interés en la chaperona de histonas ASF1 (ver más adelante).

### **Papel de las chaperonas de histonas en el ensamblaje de la cromatina.**

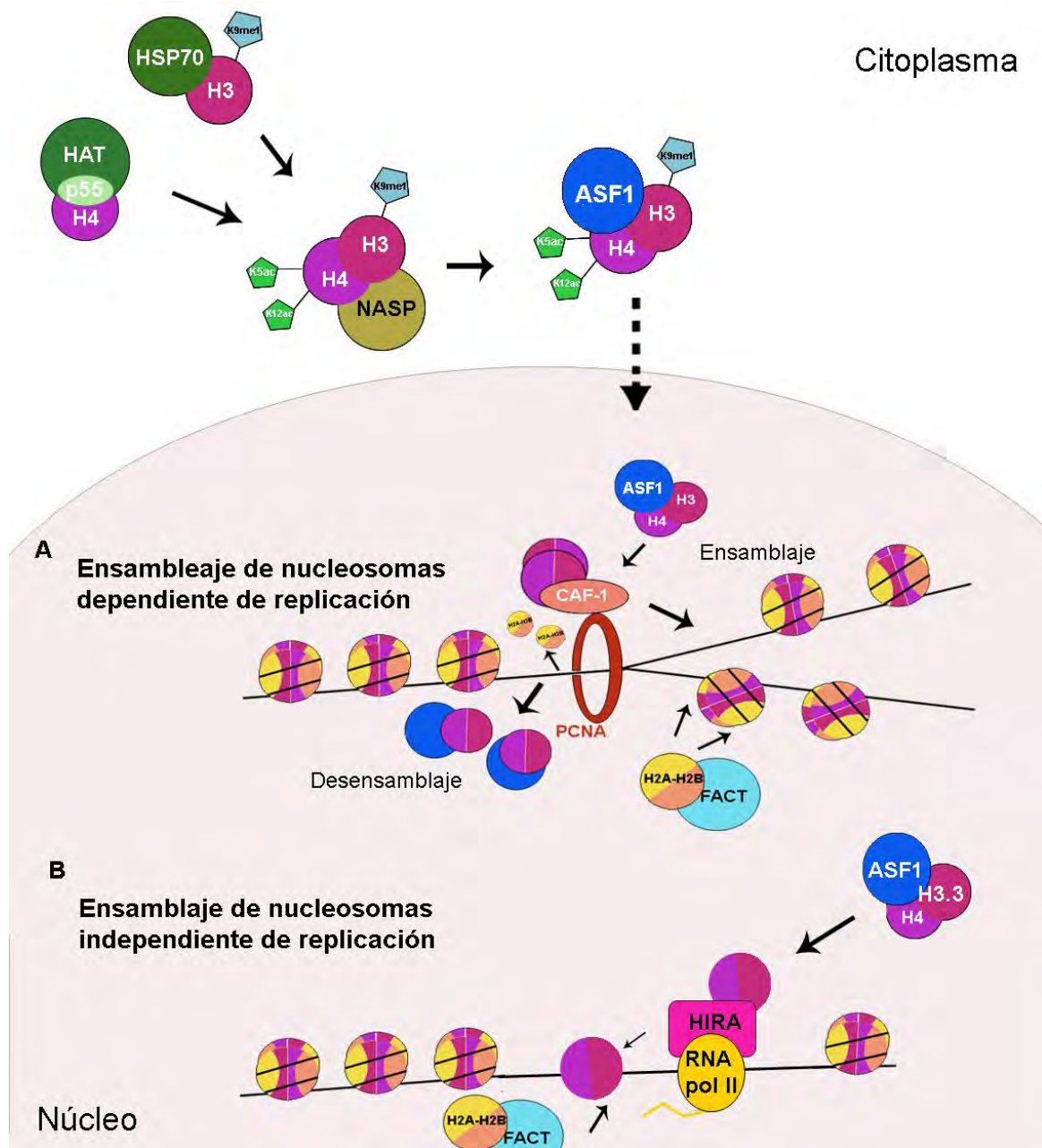
Las chaperonas de histonas participan en cada uno de los pasos en donde las histonas se ven involucradas. Sin embargo, aún falta por conocer parte de los mecanismos, la participación y las interacciones que tienen las chaperonas con otros factores de unión a cromatina (Fan et al., 2015).

El ensamblaje de los nucleosomas puede ser dependiente e independiente de la replicación y numerosos estudios han dado indicios de cómo se lleva a cabo la formación de los nucleosomas y de la cromatina (Eitoku et al., 2008; Green et al., 2005).

El ensamblaje de los nucleosomas se da de manera general, en dos pasos: 1) Dos dímeros de las histonas H3-H4 o en su forma de tetrámero se depositan en el ADN y 2) dos dímeros de las histonas H2A-H2B se unen a cada lado del tetrámero (H3-H4)<sub>2</sub> (**Figura 3**). El desensamblaje se da básicamente con los mismos pasos, pero de manera inversa (Eitoku et al., 2008).

Después de la replicación, las histonas recién sintetizadas se pliegan y se asocian para formar los dímeros, H2A-H2B y H3-H4. Posteriormente, el dímero H3-H4 debe unirse a la chaperona de histonas NASP (“Nuclear Autoantigenic Sperm Protein”) para evitar la degradación. En esta etapa, las histonas del dímero H3-H4 adquieren marcas (modificaciones post-traduccionales) necesarias para el reconocimiento de otros factores que participan en su inserción en el ADN. Las marcas propias de histonas recién sintetizadas son la di-acetilación de la histona H4 en las lisinas 5 y 12 (H4K5 y H4K12) y la monometilación de la lisina 9 de la histona 3 (H3K9me1) (Campos et al., 2010). A continuación, la chaperona de histonas

ASF1 (“Factor de Anti-silenciamiento 1”) se une al dímero de histonas formando un complejo estable y junto con karioferina 4 se transportan al interior del núcleo.



**Figura 3. Ensamblaje de los nucleosomas de maneras dependiente e independiente de la replicación.** Los dímeros H3-H4 después de su síntesis, deben asociarse a diferentes chaperonas de histonas que les brindarán estabilidad y permitirán su asociación con en el ADN. En los procesos dependientes e independientes de la replicación, la chaperona de histonas ASF1 participa en la inserción de dichos dímeros, así como en su remoción del nucleosoma. ASF1 reconoce en el citoplasma a las histonas recién sintetizadas gracias a la presencia de la marca de histonas H3K9me1, lo que permite su importación al núcleo. Sin embargo, dentro del núcleo, ASF1 también es capaz de reciclar dicho dímero. Una vez que los dímeros H3-H4 son depositados en el ADN, se da la asociación con los dímeros H2A-H2B y se forma finalmente el nucleosoma.



En la formación del nucleosoma dependiente de replicación, ASF1 se une al dímero conformado por la histona H4 y a una de las variantes de la histona H3, H3.1 o H3.2. Posteriormente, ASF1 transfiere este dímero a la chaperona de histonas CAF-1, para permitir la formación del tetrasoma (H3-H4)<sub>2</sub> y su inserción en la cromatina. A su vez, CAF-1 interacciona con la proteína PCNA (“Proliferating Cell Nuclear Antigen”) y se ubica detrás de la helicasa en el ADN en replicación. En la formación del nucleosoma independiente de la replicación, ASF1 se une a las histonas H3.3-H4, las transfiere a la chaperona de histonas HIRA y esta última es capaz de depositar al tetrámero en el ADN (Campos et al., 2010; Campos et al., 2015; Eitoku et al., 2008).

Una vez que el tetrámero (H3-H4)<sub>2</sub> se ha depositado en el ADN, un dímero de las histonas H2A-H2B se deposita a cada lado del tetrámero. Las chaperonas de histonas que interaccionan con estos últimos dímeros son FACT y NAP1. Sin embargo aún no se ha aclarado si una de ellas o ambas depositan al dímero H2A-H2B en el ADN (Ransom et al., 2010).

Finalmente, la última histona en depositarse es H1. Esta se une al ADN de unión entre dos nucleosomas y no interacciona directamente con las otras histonas. En mamíferos existen 11 isoformas y mientras que algunas de ellas tienen un papel en la represión transcripcional y en la compactación de la cromatina, en *D. melanogaster* sólo existe un subtipo y está asociado con represión transcripcional (Parseghian, 2015).

En resumen, durante el ensamblaje de la cromatina, todas las histonas tanto las recién sintetizadas como las recicladas, deben ser reconocidas por las chaperonas. En una serie de pasos, las chaperonas depositan en primera instancia los tetrámeros de histonas (H3-H4)<sub>2</sub> y posteriormente dos dímeros de las histonas H2A-H2B que al asociarse al ADN dan lugar a la formación de los nucleosomas e influyen así en el establecimiento de dominios cromatínicos.

### **El factor de Anti-silenciamiento 1 (ASF1).**

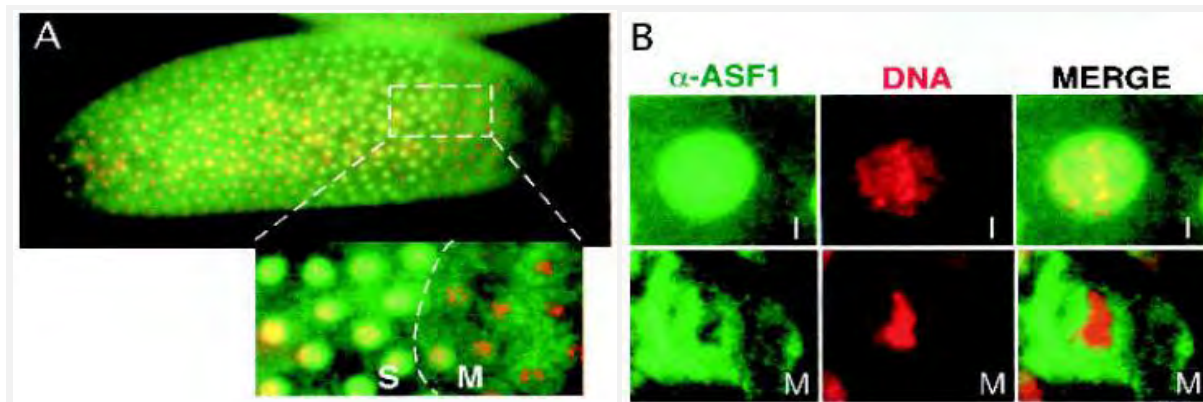
Entre las chaperonas de histonas, ASF1 es la más conservada entre los eucariontes (Natsume et al., 2007). Esta chaperona tiene un papel muy importante en el ensamblaje de la cromatina, ya que está involucrada en todos los procesos donde participan las histonas H3 y H4. Estos procesos incluyen los dependientes y los independientes de la replicación celular, la

reparación del ADN, el remplazo de histonas, la transcripción génica, el reclutamiento de complejos co-represores y el deslizamiento y relocalización de nucleosomas en regiones promotoras (Burgess y Zhang, 2013; De Koning et al., 2007; Goodfellow et al., 2007; Moshkin et al., 2009).

En *D. melanogaster* existe un solo gen de *asf1* que codifica dos variantes del transcrito, la isoforma A y la isoforma B. La isoforma A contiene 1505 nucleótidos y la isoforma B contiene 1129 nucleótidos (carece de 376 nucleótidos en el extremo 3' en comparación con la variante A. Ambas isoformas se traducen en una proteína que contiene los mismos aminoácidos (Emmert, 2014). La distribución de las variantes de ASF1 no han sido descritas. En los vertebrados, las plantas y en *Caenorhabditis elegans* (*C. elegans*) existen dos parálogos a los que se les denomina ASF1a y ASF1b. Ambos son similares en su estructura primaria, sin embargo, la distribución de sus ARNs mensajeros (ARNm) es distinta. Mientras que el ARNm de ASF1a es ubicuo, el de ASF1b se encuentra principalmente en el timo, los testículos, el intestino delgado, el colon y las células germinales (Eitoku et al., 2008; Messiaen et al., 2016).

ASF1 está compuesta por dos dominios, el dominio amino-terminal y el carboxilo-terminal. El dominio amino-terminal está conservado entre mamíferos, *D. melanogaster* y levaduras. Este dominio forma una estructura en lámina  $\beta$  que le permite interactuar con las porciones carboxilo-terminal de las histonas H3 y H4 o con las chaperonas de histonas HIRA y CAF-1, aunque no puede unirse al dímero y a las chaperonas al mismo tiempo. Además, el complejo ASF1 disocia al tetrámero (H3-H4)<sub>2</sub> y promueve el desensamblaje de los nucleosomas, debido a que ASF1 interactúa con las porciones carboxilo-terminal de las histonas, regiones necesarias para la interacción entre los dos dímeros de H3-H4 (De Koning et al., 2007; Eitoku et al., 2008; Moshkin et al., 2002; Natsume et al., 2007). ASF1 contiene también una porción carboxilo terminal divergente, que permanece expuesta al solvente y provee una interfase para la interacción con otros factores que participan en diversos procesos que involucran al ADN (Eitoku et al., 2008).

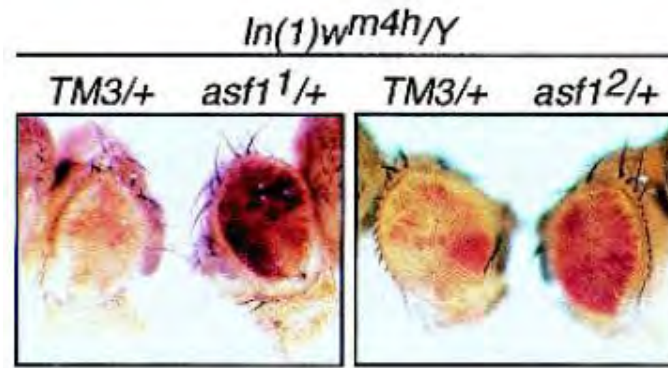
En *D. melanogaster*, ASF1 se encuentra tanto en el citoplasma como en la cromatina y su localización es dependiente del ciclo celular (**Figura 4**). Esta chaperona se localiza principalmente en el núcleo durante la fase S y en el citoplasma durante la mitosis (Moshkin et al., 2002).



**Figura 4. Localización intracelular de la proteína ASF1 en la embriogénesis de *D. melanogaster*.** En verde ASF1 y en rojo ADN. A) En estas inmunotinciones de embriones de *D. melanogaster* en una etapa temprana, se observa que ASF1 tiene una localización nuclear en la fase S del ciclo celular y citoplasmática durante la mitosis. B) Localización de la proteína ASF1 durante la interfase y la mitosis en embriones celularizados (tomada de Moshkin et al., 2002).

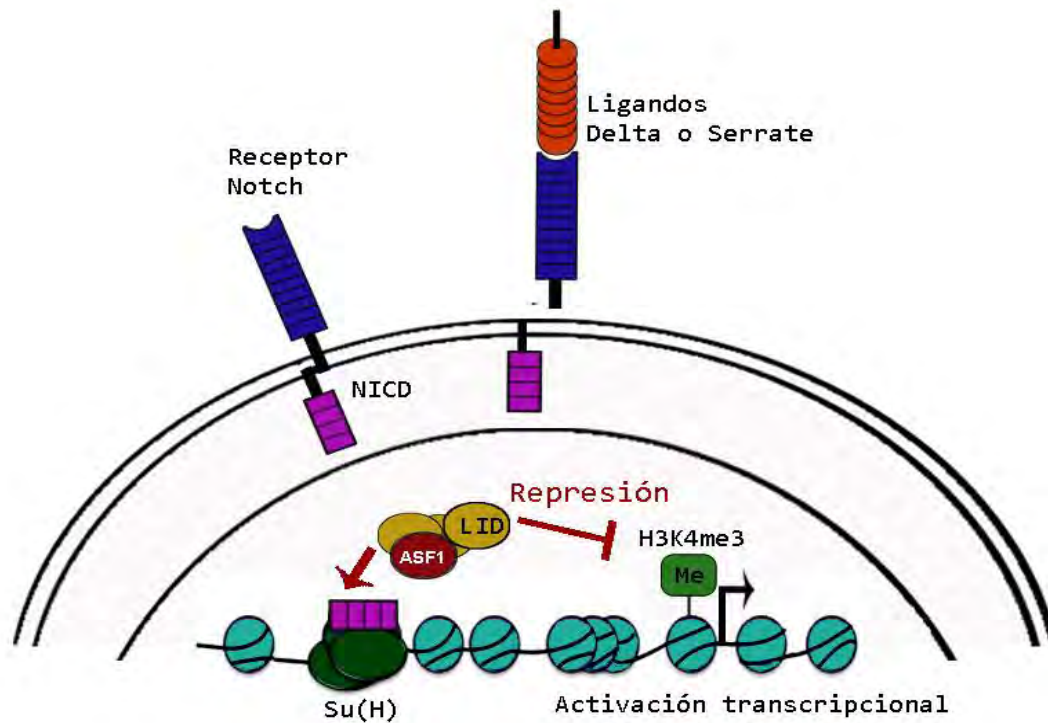
A través del estudio de diversos organismos y líneas celulares, incluyendo las líneas celulares de *Drosophila*, la línea DT40 de pollo (linfoblastos), las células germinales de ratón y células pluripotentes humanas, se conocen algunos procesos donde interviene ASF1. La pérdida de ASF1 provoca que las células permanezcan en la fase S del ciclo celular, impidiendo la supervivencia y la diferenciación celular (Gonzalez-Muñoz et al., 2014; Messiaen et al., 2016; Moshkin et al., 2002; Sanematsu et al., 2006). Los individuos que carecen de la proteína funcional ASF1 en *D. melanogaster* mueren en etapas embrionarias o larvarias. Los individuos que son mutantes heterocigotos de esta chaperona, tienen una supresión del efecto de variegación por posición, es decir, presentan apertura de la cromatina en regiones de heterocromatina centroméricas. Esto indica que ASF1 también podría estar participando en la formación y mantenimiento de la heterocromatina (**Figura 5**; Moshkin et al., 2002).

Algunos ensayos *in vitro* sugieren que ASF1 tendría un papel en la regulación de la transcripción de genes específicos, ya que la ARN PolIII y los factores transcripcionales TAFII250 y TFIID pueden reclutar a ASF1 a regiones determinadas en el genoma. Asimismo, ASF1 es capaz de formar un complejo con la subunidad MCM2 de la helicasa replicativa y con las histonas H3-H4. Por lo que se propone que ASF1 juega un papel directo durante la replicación del ADN, al promover el desensamblaje de los nucleosomas y permitir el paso de la horquilla de replicación (Huang et al., 2015).



**Figura 5. ASF1 en el mantenimiento de regiones de heterocromatina.** A la izquierda de cada panel, se encuentran individuos adultos machos control, que tienen clonas celulares donde el gen *white* (*w* que da pigmentación roja en el ojo) se encuentra invertido en una región de heterocromatina pericentrica en el cromosoma X. Debido a que *w* se silencia por la propagación estocástica de la heterocromatina de regiones vecinas en cada célula, se produce el efecto de variegación. A la derecha de cada panel se encuentran machos heterocigotos con mutaciones para ASF1 (*asf1*<sup>1</sup> y *asf1*<sup>2</sup> tienen deletados 126 y 8 aminoácidos respectivamente). Puede observarse que los ojos de estos animales son casi en su totalidad rojos, indicando que *w* fue des-reprimido, una característica de la supresión del efecto de variegación por posición (tomada de Moshkin et al., 2002).

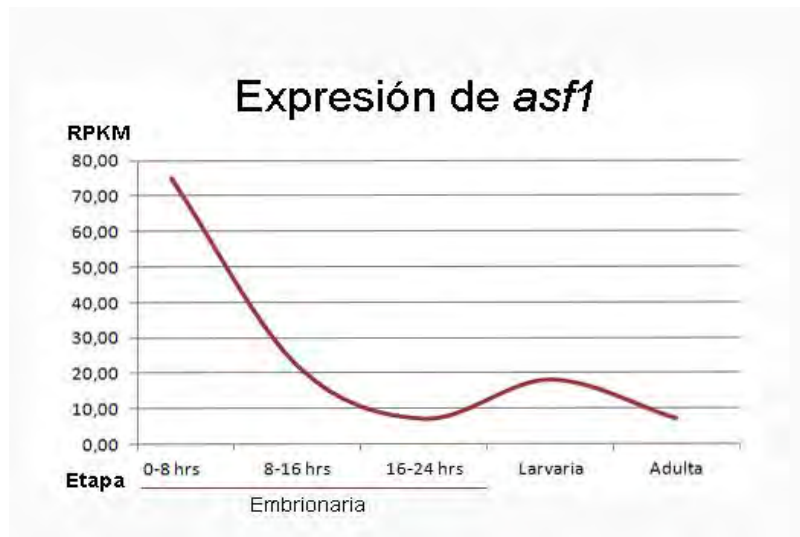
Uno de los pocos ejemplos de la regulación que ejerce ASF1 sobre la transcripción de genes específicos, se demuestra con la evidencia de que esta chaperona es un co-regulador negativo de la vía de Notch en *Drosophila* (Goodfellow et al., 2007). La vía de Notch regula procesos celulares que ocurren tanto en la embriogénesis como en la adultez, promueve la diferenciación hacia un linaje celular específico, genera inhibición lateral célula-célula y contribuye a la formación de patrones en diversos tejidos (San Juan et al., 2012). De manera normal, cuando un estímulo activa al receptor Notch, el dominio intracelular (NICD) del receptor se escinde y se une al factor transcripcional Su(H) (“Suppressor of Hairless”). Ambos se transportan al núcleo, donde su unión al ADN genera una respuesta transcripcional rápida que favorece la expresión de genes blanco de esta vía (Artavanis-Tsakonas et al., 1999). Sin embargo, ASF1 regula negativamente la transcripción de estos genes, ya que se une a Su(H) y recluta un complejo co-represor, dentro del cual se encuentra una desmetilasa de la histona H3K4me<sub>3</sub>, marca asociada a activación transcripcional (**Figura 6**; Moshkin et al., 2009). El papel de ASF1 en la vía de Notch resalta su importancia durante los procesos de desarrollo, proliferación, sobrevivencia celular y apoptosis, ya que esta vía participa en todos ellos (San Juan et al., 2012).



**Figura 6. Participación de ASF1 en la regulación de genes específicos (genes blanco de la vía de Notch).** Con la activación del receptor Notch, su dominio intracelular (NICD) se trasloca al núcleo y junto al factor transcripcional “Suppressor of Hairless” (Su(H)), permiten la transcripción de algunos genes blanco de esta vía. ASF1 puede unirse físicamente con Su(H) y reclutar un complejo co-represor, el cual contiene a LID, una desmetilasa de H3K4me3, lo que reprime la transcripción de dichos genes (Moshkin et al., 2009).

A pesar de que la función y las interacciones de la chaperona de histonas han sido ampliamente estudiadas, aún queda mucho por dilucidar. No existe información sobre los mecanismos que regulan la expresión de *asf1*; sin embargo, los patrones de expresión del ARNm varían en las diferentes etapas del desarrollo de *D. melanogaster*. Por lo tanto, la regulación de la transcripción de *asf1* debe ser finamente regulada de manera espacio-temporal (**Figura 7**).

Datos de ChIP y ChIP-chip del genoma completo de *D. melanogaster*, publicados por el grupo de Corces, indican que las proteínas de unión a “insulators” CTCF y CP190 están enriquecidas río arriba del sitio de inicio de la transcripción de *asf1* (Bushey et al., 2009). Estos datos sugieren que CTCF y CP190 podrían estar participando en la regulación de la transcripción de *asf1*.



**Figura 7. Niveles de expresión de *asf1* durante el desarrollo de *D. melanogaster*.** Este gráfico representa los niveles de ARNm de *asf1* en *D. melanogaster*, donde puede observarse que la expresión varía dependiendo la etapa del desarrollo. Los niveles de expresión, representados en RPKMs (lecturas por kilobase por cada millón de lecturas mapeadas) se obtuvieron en la base de datos Flybase a partir de ensayos de RNAseq (Emmert, 2014).

En conclusión, ASF1 es una proteína multifuncional, indispensable para la formación adecuada y el mantenimiento de dominios de cromatina. Además la presencia y actividad de ASF1 sobre sus blancos está definida por su ubicación citoplasmática o nuclear y por su interacción con otros factores. Por lo tanto, la expresión de esta chaperona de histonas debe ser finamente regulada, dependiendo del tejido y etapa del desarrollo, para que los distintos procesos celulares ocurran de manera correcta.

### **Proteínas de unión a secuencias tipo “insulator”.**

Los factores transcripcionales participan en la regulación de la expresión espacio-temporal de los genes al unirse a regiones regulatorias en el ADN y contribuyen a la formación de interacciones tridimensionales. De estos factores, las proteínas de unión a “insulators” permiten la formación de asas cromatínicas y generan la división del genoma en dominios regulatorios (Negre et al., 2010).

Los “insulators” han sido caracterizados tanto en *D. melanogaster* como en mamíferos y levaduras. Estos son secuencias de ADN reconocidas por proteínas específicas, que al unirse ejercen su función. Funcionalmente los “insulators” se definen por tres características:

1) proveen una barrera que impide la propagación de heterocromatina a dominios vecinos, para prevenir la expresión aberrante de genes; 2) regulan interacciones de regiones regulatorias como potenciadores, silenciadores y promotores, impidiendo o facilitando el mantenimiento de estados génicos transcripcionalmente activos; 3) tienen una función arquitectónica en el genoma al permitir la formación de asas cromatínicas y mediando interacciones inter- e intracromosómicas (Herold et al., 2012).

*Drosophila* ha sido un organismo adecuado para el estudio de las propiedades de las proteínas de unión a “insulators” ya que con la manipulación de sistemas de recombinación y diversas mutantes ha sido posible obtener diferentes combinaciones de regiones regulatorias en la misma posición genómica, lo que permite estudiar el papel de un elemento regulatorio en particular (Kyrchanova y Georgiev, 2014). Entre los primeros “insulators” que fueron identificados en *D. melanogaster* se encuentran los que están en los complejos de genes homeóticos *Bithorax* (BX-C) (Barges et al., 2000) y *Antennapedia* (ANT-C) (Belozarov et al., 2003), los “insulators” que enmarcan la región promotora del gen *Notch* (Vazquez y Schedl, 2000), los “insulators” *scs* y *scs'* que limitan el dominio donde se encuentran dos genes que codifican para las proteínas de choque térmico Hsp70 (Kellum y Schedl, 1992), así como los que enmarcan los retrotransposones *gypsy* e *Idefix* (Kyrchanova y Georgiev, 2014).

Diversos estudios de análisis de genoma completo en *Drosophila*, indican que las proteínas de unión a “insulators” pueden localizarse en miles de sitios en el genoma, gracias a la presencia de motivos en el ADN que han sido caracterizados por ensayos de ChIP-seq. Además, estos “insulators” interaccionan entre ellos mediante diferentes combinatorias, así como con distintos promotores y factores transcripcionales generales, como lo es TAF3 (Kyrchanova y Georgiev, 2014).

Hasta la fecha se han caracterizado alrededor de 10 proteínas de unión a secuencias “insulators” en *D. melanogaster* denominadas: Su (Hw) (“Suppressor of Hairy wing”), mod (mdg4) (“modifier of mdg4”), BEAF 32 (“Boundary Element Associated Factor of 32 kDa”), Zw5 (“Zeste white 5”), GAF (“GAGA factor”), Pita, ZIPIC (“Zinc Finger Protein-Interacting with CP190”), Ibf1 (“Insulator Binding Factor 1”), Ibf2 (“Insulator Binding Factor 2”), el complejo Elba1/Elba2/Elba3 (“Early Boundary Activity DNA Binding Factor) y CTCF (“CCCTC-binding Factor”). Todas ellas pueden ser encontradas en asociación con CP190 (Centrosomal protein 190), la cual no se une directamente al ADN, pero se le ha descrito como una molécula que permite la interacción a larga distancia entre estos elementos

“insulator” y provee estabilidad a las interacciones (Cuartero et al., 2014; Maksimenko et al., 2015).

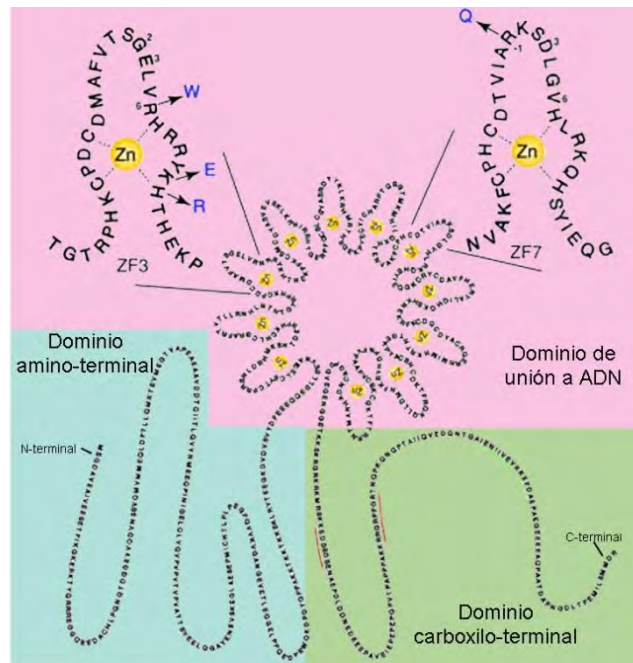
Por lo tanto, la actividad de las proteínas de unión a “insulators” es indispensable para que se regule el contacto entre diversas regiones regulatorias en el genoma, para la estructuración tridimensional de éste y para el establecimiento de dominios transcripcionales. Dentro de las proteínas de unión a “insulators”, la proteína CTCF es una de las mejor caracterizadas.

### **Funciones del factor CTCF.**

Dentro de las moléculas de unión a “insulators”, CTCF es la única que se encuentra tanto en *D. melanogaster* como en vertebrados (Herold et al., 2012). El hecho de que CTCF es una proteína muy conservada, con un papel importante en la regulación de genes y en la organización del genoma, la ha hecho un buen candidato para evaluar cómo los cambios en las secuencias del ADN inducen la unión funcional de este factor con consecuencias en la regulación génica (Ni et al., 2012).

El dominio de CTCF que está más conservado entre las diferentes especies, es el dominio de unión al ADN, que contiene 11 dedos de zinc, con los que reconoce sitios blanco en el genoma. CTCF puede unirse tanto a regiones intragénicas, como a intergénicas o cerca de sitios de los inicio de la transcripción (TSS) en el ADN (Holohan et al., 2007). CTCF reconoce distintas regiones en el genoma, debido a que la combinatoria de los dedos de zinc que participan en dicha unión es dependiente de la secuencia del ADN a la que se está uniendo (Ohlsson et al., 2001). Sin embargo las porciones amino- y carboxilo-terminales difieren mucho más entre las especies, lo que probablemente les confiera especificidad para asociarse a los diferentes factores. La porción amino-terminal de CTCF contiene un dominio de multimerización, por medio del cual forma complejos diméricos o tetraméricos (Bonchuk et al., 2015). La porción carboxilo terminal contiene 200 aminoácidos con los que interacciona con otras proteínas, como CP190. Además esta porción, le confiere parcialmente su actividad “insulator” en ciertas regiones cromosómicas (Bonchuk et al., 2015; Holohan et al., 2007). En la **Figura 8** se muestra la estructura de la proteína CTCF.



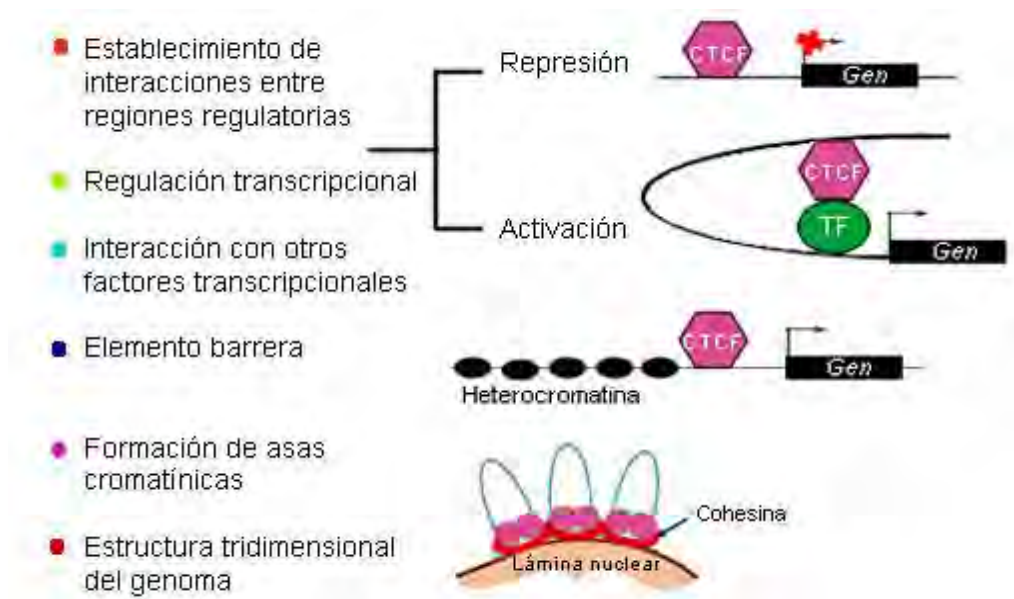


**Figura 8. Estructura proteica de CTCF.** La proteína de unión a “insulators” CTCF está conformada por tres dominios: 1) el dominio de unión a ADN, que contiene 11 dedos de zinc. Está sumamente conservado en diferentes organismos y reconoce motivos en el ADN al cual se une, 2) el dominio amino-terminal sirve para que CTCF pueda multimerizarse y contiene sitios de ADP-ribosilación, modificación post-traduccional que le brinda sus propiedades como “insulator” en numerosos sitios en el genoma (Yu et al., 2004), 3) el dominio carboxilo-terminal que participa en la asociación a otras moléculas como ARN (Saldaña-Meyer et al., 2014) y proteínas (tomada y modificada de Ohlsson et al., 2001).

El factor CTCF regula a diferentes niveles la transcripción de genes. (Herold et al., 2012). Diversos mecanismos de regulación ocurren cuando CTCF se une a sus sitios blanco, ya que estos sitios son elementos de respuesta involucrados no solo en la activación transcripcional, sino en la represión (**Figura 9**). Su unión interfiere con la metilación del ADN y por lo tanto inhibe la represión transcripcional. También puede reclutar y se puede asociar a otros factores que generarán las diferentes respuestas (Ohlsson et al., 2001).

Con los ensayos de captura conformacional de cromosomas (3C), se determina la estructura tridimensional de la cromatina y se identifican las regiones de ADN que se contactan a través de complejos proteicos. A través de esta metodología se sabe que CTCF, mediante su asociación a Cohesinas en mamíferos y CP190 en *Drosophila*, no solo participa en la formación de interacciones a corta distancia, sino que interviene en interacciones a larga distancia mediante la formación de asas cromatínicas. En las distintas líneas celulares, entre el

50 y 80% de los sitios que son ocupados por CTCF, son ocupados también por Cohesina (Ong y Corces, 2014). Estas propiedades permiten el establecimiento de dominios topológicos, la organización tridimensional del genoma y la regulación de la expresión de genes, tanto individuales como de familias de genes (Ghirlando y Felsenfeld, 2016). Se ha demostrado también, que en mamíferos, la mayoría de las asas cromatínicas formadas por CTCF dependen de que los motivos de unión a CTCF se encuentren orientados en una manera convergente (de Wit et al., 2015).



**Figura 9. Funciones del factor CTCF.** CTCF es un factor clave para la estructuración tridimensional del genoma al reconocer sitios blanco en el genoma. Esta involucrado en la activación y represión transcripcional, ya que puede permitir o evitar interacciones entre distintas regiones regulatorias al crear asas cromatínicas. Impide la propagación de heterocromatina y establece a su vez, dominios transcripcionales activos o silentes. (tomada y modificada de Liyanage et al., 2014)

Otros procesos en los que está involucrado CTCF, son la inactivación del cromosoma X, impronta génica, “splicing” alternativo y recombinación V(D)J. Por estas razones, se ha especulado que CTCF es un organizador maestro de la cromatina, ya que controla diversos procesos celulares, como la transcripción, la replicación, la reparación y establece programas específicos para cada línea celular (Ong y Corces, 2014).

Cabe recalcar que CTCF ocupa sitios en el genoma de manera específica y dependiente, tanto de la etapa del desarrollo como del tejido de origen. A pesar de que existen sitios conservados, aun no queda totalmente claro cómo se recluta este factor y cómo su unión y las interacciones que promueve son altamente estables (Lupiáñez et al., 2015). Una gran variedad de otras proteínas pueden reclutarse a los sitios de unión de CTCF y pueden fungir como co-factores de este factor. Por ejemplo, TFII-I, el cual ayuda a estabilizar a CTCF en algunas regiones promotoras, el factor TAF3 que participa en la formación de asas junto a CTCF y CP190, además de proteínas a las que se une en loci específicos, como YY1, Kaiso, CHD8, PARP1, Maz, JunD, ZNF143, Prdm5 y Nucleofosmina (Ong y Corces, 2014).

Es importante destacar, que CTCF no solo se une a secuencias en el ADN y a proteínas, sino que también se une a ARNs. Esta unión a ARNs se da gracias a que CTCF contiene una región de unión a ARN (RBR), conformada por los dedos de zinc 10 y 11 y parte su porción carboxilo-terminal. Un ejemplo, es el reconocimiento y unión de CTCF al ARN anti sentido WRAP53, que regula a *p53* (Saldaña-Meyer et al., 2014).

En conclusión, CTCF actúa como un regulador maestro del genoma gracias a que modula la estructura de la cromatina y regula la activación y represión transcripcional en los múltiples y diversos procesos celulares, mediante su interacción con otros elementos.

## 1.2 ANTECEDENTES PARTICULARES

Por ensayos de ChIP-seq en la línea celular Kc (línea celular neuronal derivada de embriones en fase tardía de *D. melanogaster*), realizados por el grupo de Corces, se identificaron los sitios en el genoma donde se une CTCF. En este estudio, definieron que el 49% de los sitios de unión a CTCF se encuentran en las regiones intergénicas, en el extremo 5' (a 200-300 pares de bases río arriba) de los sitios de inicio de la transcripción de genes involucrados principalmente en procesos del desarrollo. Esto sugiere que CTCF tiene un papel en la regulación de genes individuales específicos (**Tabla 1**; Bushey et al., 2009).

**Tabla 1.**

	Gene ontology analysis for biological process of genes with an insulator protein localization site within 200 bp of the TSS		
	P-value		
	Genes with Su(Hw)	Genes with dCTCF	Genes with BEAF
Developmental process	>0.1	3.70E-05	>0.1
Multicellular organismal development	>0.1	2.03E-05	>0.1
Anatomical structure development	>0.1	1.88E-04	>0.1
Imaginal disc development	>0.1	1.50E-04	>0.1
Regulation of compound eye photoreceptor development	>0.1	1.64E-03	>0.1
Cell cycle	>0.1	2.80E-07	4.10E-10
RNA metabolic process	>0.1	>0.1	2.10E-11
Nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process	>0.1	>0.1	2.20E-11
Biopolymer metabolic process	>0.1	>0.1	2.50E-10
Cellular metabolic process	>0.1	>0.1	8.70E-05
Primary metabolic process	>0.1	>0.1	4.90E-04

**Tabla 1. Procesos biológicos de los genes a los que se asocian las proteínas “insulator”.** La unión de las proteínas “insulator” CTCF y BEAF a 200 pares de bases del sitio de inicio de la transcripción (TSS) de genes, muestra que cada una de estas proteínas está involucrada en la regulación de procesos distintos. CTCF se une principalmente a genes asociados al desarrollo, mientras que BEAF se asocia a procesos metabólicos (Bushey et al., 2009).

### **Papel de CTCF durante el desarrollo.**

Se sabe que CTCF es vital para la supervivencia de los diferentes organismos donde se ha estudiado, sin embargo aun no se conocen los mecanismos con los que contribuye al desarrollo de los mismos. Cuando se abole la expresión de *CTCF* en *D. melanogaster*, los individuos se desarrollan hasta la etapa de pupa, pero no llevan a cabo la metamorfosis. Esto indica que a pesar de que existen más proteínas de unión a “insulators”, estas no pueden suplir la función de CTCF (Mohan et al., 2007). En mamíferos, la ausencia de la proteína funcional de CTCF es letal en etapas tempranas del desarrollo embrionario, previo a la implantación (Moore et al., 2012). Tanto en *Drosophila* como en mamíferos, la eliminación de sitios de unión a CTCF específicos alteran, de manera espacio-temporal, los patrones de expresión de los genes. Esta expresión aberrante de genes provoca malformaciones, cambios homeóticos y retraso en el desarrollo (Holohan et al., 2007; Lupiáñez et al., 2015; Mohan et al., 2007). Cabe mencionar, que datos obtenidos en nuestro laboratorio indican que CTCF también es indispensable para el desarrollo de *Danio rerio* (pez cebra). Al utilizar morfollinos que impiden la traducción de CTCF, se afecta el desarrollo de todo el organismo, incluyendo el sistema nervioso lo que genera una letalidad embrionaria temprana (Carmona-Aldana, resultados no publicados).

Por lo tanto, la actividad adecuada de CTCF, es indispensable para el desarrollo, la formación de patrones y la supervivencia de los diversos organismos donde está presente, gracias a que este factor controla la regulación de numerosos genes, mediante la formación de dominios transcripcionales y participa en todos procesos celulares.

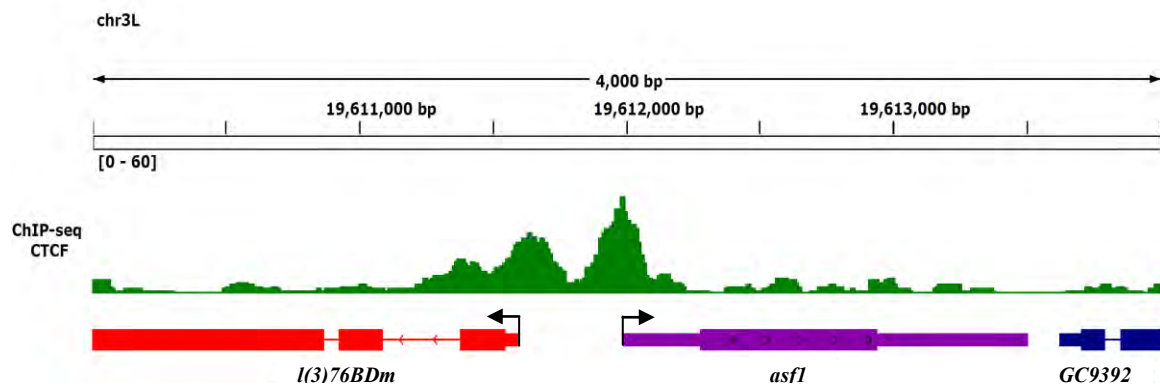
**El factor nuclear CTCF está enriquecido en la región intergénica 5' de *asf1*, lo que indica que podría estar participando en la regulación de su expresión.**

A partir de los datos de ChIP-seq de CTCF obtenidos por Bushey et al, 2009 encontramos, por un análisis de Ontología de genes que CTCF está enriquecido en regiones cercanas al sitio de inicio de la transcripción de genes involucrados en el desarrollo, en el ciclo celular y en vías de señalización. En nuestro laboratorio, indentificamos que entre los genes que están involucrados en estos procesos, *asf1* puede ser regulada por CTCF, ya que este factor está enriquecido en la región intergénica 5' de *asf1*, a comparación de sitios vecinos.

En la **Figura 10**, se muestra un esquema de los datos de enriquecimiento de CTCF en la región intergénica entre *asf1* y *l(3)76BDm*. Este último, es un gen que se encuentra a 379 pares de bases, en una posición antisentido respecto a *asf1*. Se sabe muy poco de este gen, excepto que su ausencia es letal para *D. melanogaster* (Spradling et al., 1999). Además, tomando en cuenta la similitud de la secuencia, la base de datos Flybase predice que *l(3)76BDm* es ortólogo del gen que codifica para TRS85 en *Saccharomyces cerevisiae* (Curators, 2008). TRS85 es una subunidad del complejo TRAPP (“Transport Protein Particle”), que participa en el transporte de vesículas entre el retículo endoplásmico y el aparato de Golgi), sin embargo, no se han realizado más estudios al respecto.

En *D. melanogaster* CTCF se encuentra frecuentemente entre genes transcritos de manera divergente, lo que sugiere que CTCF permite la expresión diferencial de estos genes (Smith et al., 2009). CTCF al unirse a la región intergénica entre *asf1* y *l(3)76BDm*, podría regular la expresión diferencial de ambos genes.

Estos antecedentes nos permiten proponer que CTCF regula la expresión de la chaperona de histonas, mediante su unión a la región intergénica 5' de *asf1*.



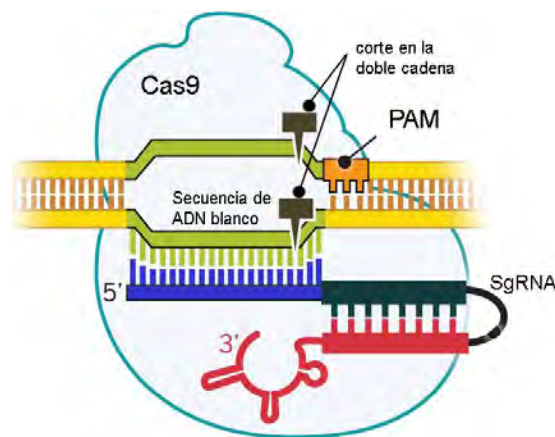
**Figura 10. Enriquecimiento de la señal de CTCF en la región intergénica entre *asf1* y *l(3)76BDm*.** En este esquema se muestran datos obtenidos del “Genome Browser” de la Universidad de California Santa Cruz (UCSC), de ensayos de ChIP-seq realizados en embriones de 0-12 horas de *D. melanogaster* (Bushey et al., 2009). Los gráficos se generaron con el programa IGV (Integrative Genomics Viewer). Se puede observar que existe un enriquecimiento de la señal de CTCF (verde) en la región intergénica entre *asf1* (morado) y *l(3)76BDm* (rojo) a comparación de regiones vecinas en el genoma. Estos datos sugieren que CTCF puede unirse a la posible región reguladora de *asf1* y por lo tanto, participa en la regulación de su transcripción (imagen creada por Arzate-Mejía).

## Técnica CRISPR-Cas9

En este trabajo, se empleó la técnica CRISPR-Cas9, para generar eliminaciones sobre el sitio en que CTCF se une, en la región intergénica 5' de *asf1*.

La técnica CRISPR (“Clustered Regularly Interspaced Short Palindromic Repeats”)-Cas9, es una técnica que permite la edición del genoma de forma muy precisa. Se originó a partir del sistema bacteriano CRISPR-Cas tipo II, el cual provee a las bacterias un sistema inmunitario adaptativo contra ácidos nucleicos provenientes de fagos y plásmidos (Doudna y Charpentier, 2014). El locus CRISPR consiste en un arreglo de secuencias cortas repetidas, intercaladas con espaciadores (20 nucleótidos de ADN adquirido de ácidos nucleicos invasores). Estas secuencias, en una etapa posterior, son transcritas y procesadas para dar lugar a los crRNAs (un solo espaciador flanqueado de una secuencia repetida). Dicho procesamiento y maduración se da por la actividad de la molécula tracrRNA (“trans-activating crRNA”) y la ARNasa III. Para que este sistema sea funcional, el crRNA se combina con la proteína Cas9 (endonucleasa) formando un complejo efector, que reconoce secuencias complementarias a los espaciadores en el ácido nucleico externo a la bacteria, lo cual induce un corte de doble cadena en esta secuencia (Karvelis et al., 2013). Este sistema

ha sido adaptado para la edición del genoma en otros organismos, incluida *Drosophila*, donde el crRNA y el tracrRNA han sido fusionados para crear un solo RNA guía (sgRNA) (**Figura 11**). Mediante el diseño de oligonucleótidos que codifican para el sgRNA de 20 nucleótidos, se dirige a la Cas9 a sitios específicos, siendo el único requisito para su acción, la presencia de los oligonucleótidos NGG (N puede ser cualquier nucleótido, mientras que G es Guanina), también llamada secuencia PAM (“Protospacer Adjacent Motif”), la cual debe encontrarse adyacente al extremo 3’ de los 20 oligonucleótidos en el ADN blanco. Posterior al corte inducido por la endonucleasa, los sistemas de reparación del ADN entran en acción y pueden generar mutaciones: eliminaciones, inserciones o cambios de un solo nucleótido (Ran et al., 2013).



**Figura 11. Sistema de edición genómica CRISPR-Cas9.** Para que este sistema sea funcional se necesita de la endonucleasa Cas9, un RNAguía (sgRNA), que es complementario a 20 nucleótidos de ADN en la secuencia blanco, con los cuales se aparea y la secuencia PAM. El sgRNA lleva a la endonucleasa a sitios específicos en el genoma, donde promueve un corte en la doble cadena del ADN, tres bases río arriba de la secuencia PAM. Posteriormente, por el sistema de reparación del ADN, pueden generarse mutaciones en la secuencia blanco (tomada y modificada de Doudna y Charpentier, 2014).

### 3. PLANTEAMIENTO DEL PROBLEMA

La chaperona de histonas ASF1 participa en los procesos donde las histonas H3-H4 se ven involucradas, en el ensamblaje y desensamblaje de la cromatina, en el remplazo de histonas, en la transcripción génica, en la reparación del ADN, en el reclutamiento de complejos co-represores, en el deslizamiento y relocalización de nucleosomas en regiones promotoras, además de otros procesos celulares importantes (Burgess y Zhang, 2013; De Koning et al., 2007; Goodfellow et al., 2007; Moshkin et al., 2009). Sin embargo, los mecanismos que regulan la expresión de *asf1* aún se desconocen.

Diversos estudios han sugerido que CTCF tiene un papel en la regulación transcripcional de genes involucrados en el desarrollo y en la supervivencia de diversos organismos. Esta proteína de unión a “insulators” juega un papel muy importante en la organización tridimensional de la cromatina y en la regulación transcripcional de genes, mediante su unión a motivos en el ADN. En *D. melanogaster* el 49% de los sitios ocupados por CTCF se encuentran a 200-300 pares de bases río arriba del sitio de inicio de la transcripción de genes involucrados en procesos de desarrollo (Bushey et al., 2009).

En el análisis de nuestro laboratorio de los datos de ChIP-seq de Bushey et al., 2009, se buscaron genes involucrados tanto en procesos de desarrollo como genes de vías de señalización que pudieran ser regulados por CTCF. Encontramos que CTCF se encuentra enriquecido río arriba del sitio de inicio de la transcripción de *asf1*.

En este estudio queremos dilucidar si CTCF tiene un papel en la regulación de la transcripción de *asf1* y del gen *l(3)76BDm*, que se encuentra posicionado de manera divergente a *asf1* a 379 pares de bases río arriba del extremo 5' de *asf1*. Con estos datos pretendemos obtener información acerca de cómo podría modularse el nivel de expresión de *asf1* y si CTCF tiene un papel en la regulación de genes divergentes. Además podremos obtener información acerca de cómo el factor CTCF participa en el desarrollo de este organismo.



## 4. HIPÓTESIS

El factor nuclear CTCF participa en la regulación de la expresión del gen que codifica para la chaperona de histonas ASF1 en *D. melanogaster* al unirse río arriba del sitio de inicio de la transcripción.

## 5. OBJETIVOS

### OBJETIVO GENERAL

Demostrar que el factor CTCF participa en la regulación del gen *asf1*, mediante el reconocimiento y unión a la región intergénica 5' del mismo en *D. melanogaster*.

### OBJETIVOS PARTICULARES

1. Demostrar la unión *in vitro* de CTCF en la región que se encuentra río arriba del sitio de inicio de la transcripción de *asf1*.
2. Generar mutantes donde se elimine el motivo de unión de CTCF en la región intergénica 5' de *asf1* en la línea celular S2R+.
3. Evaluar el efecto de la pérdida de función de CTCF, sobre la expresión de *asf1* y *l(3)76BDM*, en las poblaciones celulares que carezcan del motivo de unión a CTCF.

## 6. MATERIALES Y MÉTODOS

### Cultivo Celular

Las células S2R+ son similares en comportamiento y apariencia a los hemocitos de *D. melanogaster*. Esta línea celular proviene de la línea celular S2, la cual fue generada a partir de células inmortalizadas de embriones de *D. melanogaster* de 20 a 24 h de desarrollo, a las cuales se les disoció por tripsinización (Baum y Cherbas, 2008; Schneider, 1972). A diferencia de las células S2, la línea celular S2R+ expresa fz (“frizzled”) y responde a la adición extracelular de Wg (“Wingless”) (Yanagawa et al, 1998). Estas células se cultivaron a 25°C con medio Schneider (Gibco) adicionado con 10% de suero fetal bovino y 1% de penicilina/estreptomicina.

### Oligonucleótidos para los ensayos de retardo de movilidad electroforética (EMSA).

Para los ensayos con la técnica EMSA, se diseñaron pares de oligonucleótidos complementarios de entre 50 y 60 pares de bases. Estos se sintetizaron con la compañía Invitrogen y tienen las características que son desalados, purificados por PAGE y desfosforilados.

Nombre	Secuencia	Longitud en pares de bases
Sonda 1 sentido	5' -CTATTGTCGCGTTTTGACATTCCGCCACTTGTC GTGGTGAGTTTTTCGCCACCACTTGCG-3'	60
Sonda 1 antisentido	5' -CGCAAGTGGTGGCGAAAACTCACCACGACAAG TGGCGGAATGTCAAACGCGACAATAG-3'	60
Sonda 2 sentido	5' -GTCGTGGTGAGTTTTTCGCCACCACTTGCGACA CCGAACGGTCCCTCTGAAGCGCGGG-3'	58
Sonda 2 antisentido	5' -CCCGCGCTTCAGAGGGACCGTTCGGTGTGCGAA GTGGTGGCGAAAACTCACCACGAC-3'	58

Sonda eliminación sentido	5' -AACTATTGTCGCGTTTTGACATTCCGCCAGTTT TTCGCCACCACTTGCGA-3'	50
Sonda eliminación antisentido	5' -TCGCAAGTGGTGGCGAAAACTGGCGGAATGTC AAAACGCGACAATAGTT-3'	50
Sonda Fab8 sentido	5' -AAATTTCCACATTCCCGCCTTGCAGCGCCACCTG GCCTTGTAATGTAGAACTAGGAAGG-3'	60
Sonda Fab8 antisentido	5' -CCTTCCTAGTTCTACATTACCAAGGCCAGGTGGC GCTGCAAGGCGGGAATGTGGAAATTT-3'	60

### **Alineamiento de oligonucleótidos para los ensayos de retardo.**

60 ng de cada uno de los oligonucleótidos se alinearon con su complementario a 95°C por 15 min en un amortiguador de alineación compuesto por 200 mM de Tris-HCl pH 8, 100 mM de NaCl (cloruro de sodio), y 0.5 M de EDTA (ácido etilendiaminotetraacético) pH 8.0. Posteriormente se incubaron a 24°C durante 1 h.

### **Marcaje de la sonda con fósforo radioactivo (<sup>32</sup>P).**

Agua estéril	2.5 µL
Oligonucleótidos alineados	16 µL
Amortiguador 10X PNK (polinucleótido cinasa)	2.5 µL
Gamma ATP	2 µL
PNK	2 µL (30 U)
Volumen final	25 µL

Estas mezclas se realizan a 25°C y se cargan en el gel inmediatamente.

### **Purificación de la sonda marcada.**

Se preparó un gel de acrilamida desnaturante compuesto de la siguiente manera:

Urea	4 gr
Agua	30.5 mL
Acrilamida 30%	16.6 mL
Amortiguador TBE (Tris, Borato, EDTA) 5 %	2.5 mL
Persulfato de amonio	350 $\mu$ L
TEMED (tetrametiletilendiamina)	35 $\mu$ L

El gel fue pre-corrído durante 1 h a 100 V con el amortiguador de electroforesis TBE a una concentración final de 0.25%. Posteriormente, se agregaron 4  $\mu$ L de amortiguador de carga a cada reacción de marcaje, se depositó toda la reacción en los pozos y se realizó la electroforesis a 150 V durante 90 min. Una vez terminada la electroforesis, se fijó una placa autoradiográfica sobre el gel durante 15 min. Después, se procedió a revelar la placa: 30 seg en líquido revelador, 30 seg en agua, 30 seg en líquido fijador y 30 seg nuevamente en agua. Una vez que la placa estuvo seca, se recortaron recuadros correspondientes a las marcas radioactivas y tomándolos como molde, se recuperaron las bandas del gel de acrilamida. Cada recuadro de acrilamida se diluyó durante 24 h a 37°C en un amortiguador compuesto por 0.5 M de acetato de amonio, 100 mM de acetato de magnesio, 1 mM de EDTA y SDS 1% (dodecilsulfato sódico). Se centrifugaron las muestras a 11,400 rpm a 4°C por 30 min y al sobrenadante se le agregaron 165  $\mu$ L de acetato de amonio 10 M, 1  $\mu$ L de glicógeno y 800  $\mu$ L de etanol absoluto frío para precipitar las sondas. Dicha reacción fue llevada a cabo por 48 h a -20°C. Finalmente, las muestras se centrifugaron durante 1 h a 13,700 rpm a 4°C, se dejaron evaporando durante 2 h a 37°C, se resuspendieron las pastillas en 20  $\mu$ L de agua estéril y se incubaron 15 min a 37°C.

### **Cuantificación de la sonda.**

Para cuantificar los oligonucleótidos marcados radioactivamente, se hizo una dilución de 1:100, relación correspondiente a muestra: agua estéril. Posteriormente, usando dos mitades de papel circular Whatman, se agregó 1  $\mu$ L de la dilución de cada sonda en cada una de las dos mitades y se les dejó secar por 10 min. Una de las mitades de papel de cada sonda se introdujo a un vial de vidrio con 5 mL de agua, mientras que la otra mitad se lavó para determinar el porcentaje de incorporación del isótopo. Los lavados se realizaron de la siguiente manera: 3 lavados de 5 min con 0.5 M de NaHPO<sub>4</sub> (fosfato diácido de sodio), 2 lavados de 1 min con agua desionizada, 2 lavados de 1 min con etanol al 96% y se les dejó

secar por 10 min. Finalmente, se introdujeron al vial con 5 mL de agua desionizada. Estos se colocaron en canastillas del lector de centelleo líquido y se les cuantificó midiendo el isótopo  $^{32}\text{P}$ . Se compararon los valores obtenidos de cada sonda, la que no fue lavada y la que sí lo fue, para obtener la relación de incorporación de la sonda. Sólo se usaron aquellas sondas marcadas radioactivamente que obtuvieran porcentajes de incorporación mayores al 80%.

### **Extracción de proteínas nucleares.**

El protocolo para la extracción de proteínas nucleares se estandarizó específicamente para células de *Drosophila*. Se recolectaron al menos  $6 \times 10^8$  células S2R+ y se centrifugaron a 2,000 g durante 4 min a temperatura ambiente, se descartó el sobrenadante y la pastilla fue lavada con PBS frío, centrifugada a 2,000 g por 4 min a temperatura ambiente y se descartó nuevamente el sobrenadante. La pastilla fue resuspendida con 5 mL del buffer TM2+ (compuesto por 10 mM de Tris pH 7.4, 2 mM de  $\text{MgCl}_2$  (cloruro de magnesio) y 0.5 mM de PMSF (fluoruro de fenilmetilsulfonilo), y se incubó a 4°C durante 1 min. A esta, se le añadieron 360  $\mu\text{L}$  de NP-40 al 10% para lisar la membrana celular, se mezcló por inversión y se incubó a 4°C por 4 min, invirtiendo ocasionalmente el tubo. Se centrifugó la muestra a 1,000 g durante 10 min a 4°C y se descartó el sobrenadante. La pastilla consistente en únicamente núcleos se lavó con 5 mL del buffer TM2+ y se centrifugó a 1,000 g, durante 10 min a 4°C. Se desechó el sobrenadante y la pastilla se resuspendió e incubó a 4°C durante 15 min en RIPA con los inhibidores de proteasas PMSF, leupeptina y pepstatina. En cada uno de los pasos se verificó el estado de las células y núcleos respectivamente por microscopia. Finalmente se centrifugó a 6,200 g durante 7 min y se recuperó el sobrenadante con las proteínas nucleares en otro tubo.

### **Cuantificación de proteínas totales.**

Se prepararon diluciones 1:5 del reactivo Bradford (Bio-Rad) y se añadieron 200  $\mu\text{L}$  en cada uno de los pozos de una placa para ELISA. Se añadió 1  $\mu\text{L}$  de la muestra a cuantificar, así como 1  $\mu\text{L}$  de BSA a concentraciones conocidas (0.625  $\mu\text{g}/\mu\text{L}$ , 1.25  $\mu\text{g}/\mu\text{L}$ , 2.5  $\mu\text{g}/\mu\text{L}$ , 5  $\mu\text{g}/\mu\text{L}$  y 10  $\mu\text{g}/\mu\text{L}$ ) que se usaron como estándar, cada muestra se analizó por duplicado. Se midió la absorbancia a 595 nm en un lector de placas de ELISA.

## **Preparación de las muestras para los ensayos de retardo.**

Para cada una de las muestras se hicieron incubaciones con 7.5  $\mu\text{L}$  de amortiguador de unión (20 mM de HEPES pH 7.9, 0.1 mM de EDTA, 50 mM de KCl (cloruro de potasio), glicerol 5% v/v, BSA (albúmina sérica bovina) 200  $\mu\text{g}/\text{mL}$ ), 1  $\mu\text{L}$  de poli dI-dC, para disminuir las interacciones inespecíficas de sonda-proteína, y 15  $\mu\text{g}$  de proteína nuclear. En las muestras en que se usaron competidores, éstos se le añadieron en concentraciones que excedían aproximadamente 100 veces la concentración de la sonda marcada radioactivamente. Todas las muestras se llevaron a volumen final de 24  $\mu\text{L}$  con agua estéril. Las muestras, se incubaron 15 min a 4°C y después, 15 min a 25°C. Posterior a esto, se le adicionó 1  $\mu\text{L}$  de la sonda marcada radioactivamente (20,000 cpm/ $\mu\text{L}$ ) a cada muestra y se incubaron 15 min a 4°C y 15 min a 25°C. Para los experimentos de super-retardo, los anticuerpos se incubaron con los extractos proteicos, el buffer de unión y poli dI-dC por 1 hora antes de añadir la sonda marcada radioactivamente, agitando constantemente a 4°C. Se preparó un gel de acrilamida al 5% no desnaturalizante y se pre-corrió durante 1 h a 110 V a 4°C. Se cargaron las muestras y la electroforesis se realizó de manera consecutiva a 110 V por 1 h y a 150 V por 3 h a 4°C. Como buffer de electroforesis se usó TBE al 0.25%. Se secaron los geles sobre papel filtro en un secador de geles (Bio-Rad-modelo 543) durante 1 h a 90°C y se almacenaron junto a placas para revelar toda la noche. Se revelaron al siguiente día en Typhoon.

## **Anticuerpo**

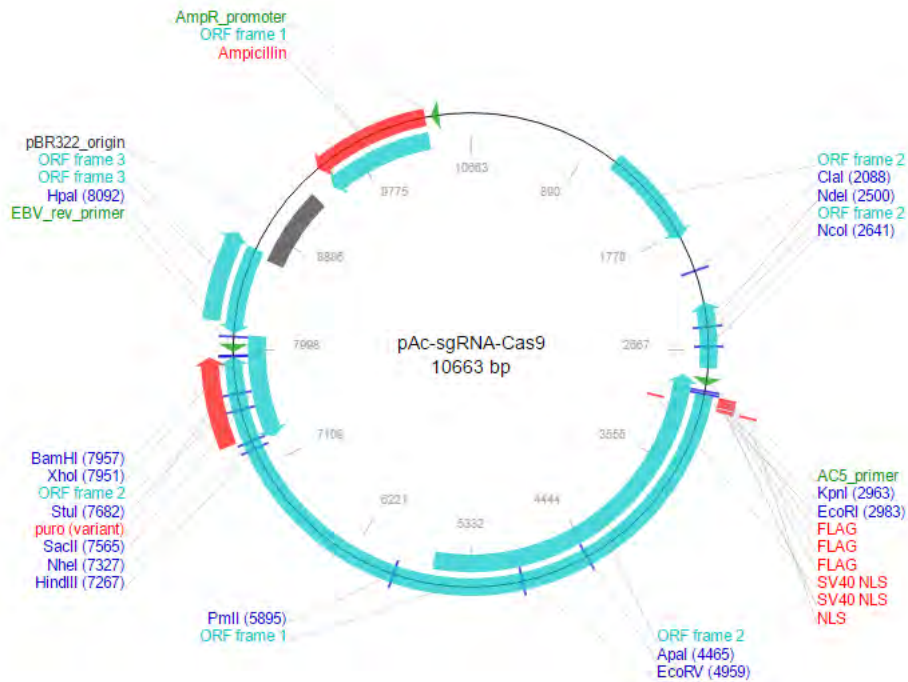
Para la producción del anticuerpo específico anti-dCTCF se tomó como referencia el anticuerpo utilizado por Gerasimova et al., (2007). El anticuerpo anti-dCTCF usado en los ensayos de retardo fue obtenido por la compañía “New England Biolabs”.

Este anticuerpo está dirigido a la porción carboxilo-terminal del gen, que fue amplificado a partir del ADNc GH14774 (“Research Genetics”) (Gerasimova et al., 2007).

## **Plásmidos.**

Para la transfección de la endonucleasa Cas9 y los RNAs guías para la técnica de edición del genoma CRISPR-Cas9 en la línea celular S2R+, se utilizó el plásmido específico para *D. melanogaster*, pAc-sg-Cas9. Este plásmido contiene a la Cas9 bajo el control del promotor de

Actina 5C y el promotor U6 para la expresión del RNA guía. Además, esta construcción tiene un cassette de resistencia a puomicina para la selección clonal (**Figura 12**).



**Figura 12.** Mapa de restricción del plásmido pAc-sg-Cas9, donado por Ji-Long Liu (Addgene plasmid # 49330).

El plásmido pGEM-T Easy Vector de Promega (**Figura 13**), se utilizó para la secuenciación por el método de Sanger, el cual se basa en la polimerización del ADN, mediante PCR, utilizando dideoxinucleótidos marcados con fluoróforos, los cuales emiten señales que son detectadas por el secuenciador. Los fragmentos de ADN amplificados por PCR se insertaron en sitios *EcoRV*, el cual no se recupera al ligarse el vector con el inserto. La cadena sentido de este plásmido está bajo el control de la ARN polimerasa T7, mientras que el de la cadena complementaria está bajo el control de la ARN polimerasa SP6.

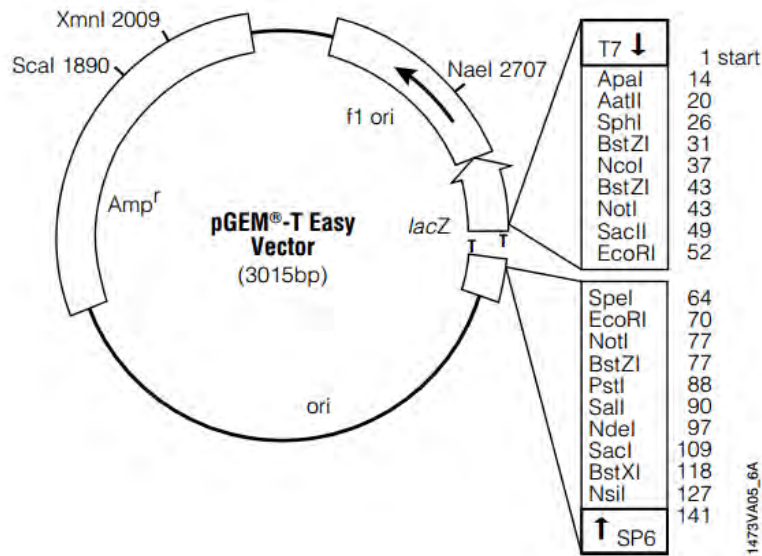


Figura 13. Mapa de restricción del plásmido pGEM-T Easy de Promega.

### Diseño de los oligonucleótidos del RNA guía para la técnica CRISPR-Cas9.

La secuencia de la región intergénica entre *l(3)76BDM* y *asfI* se ingresó a la página creada por el laboratorio del Dr. Feng Zhang, <http://crispr.mit.edu/>, la cual arrojó secuencias sugeridas para su uso como sgRNAs (RNAs guía). Esta página toma en cuenta las secuencias PAM (los oligonucleótidos NGG del genoma blanco), como la existencia de una mínima posibilidad de tener blancos inespecíficos. Para generar el sgRNA usado para este trabajo, se escogió una región que es reconocida por este RNA que dirige a la endonucleasa Cas9 sobre el motivo 1 de unión a CTCF. Para esto, se escogieron 20 nucleótidos río arriba de una secuencia PAM. Para el oligonucleótido sentido se le adicionaron las bases TTC en el extremo 5', mientras que para el antisentido se le añadieron las bases AAC en el extremo 5', para que se pudieran insertar al plásmido. Se mandaron a sintetizar a la compañía Invitrogen.

Nombre	Secuencia
Cadena Sentido	5' -TTCAAACCTCACCACGACAAGTGG-3'
Cadena Antisentido.	5' -AACCCACTTGTCGTGGTGAGTTT-3'

### Alineamiento y fosforilación de oligonucleótidos



Los oligonucleótidos sentido (500  $\mu$ M) y antisentido (500  $\mu$ M) resuspendidos en buffer EB (buffer de elución), se alinearon en un buffer al 2X compuesto por 20 mM de Tris, 2 mM de EDTA, 100 mM de NaCl pH 8.0. Se incubó la reacción a 95°C durante 5 min y posteriormente se incubó a temperatura ambiente durante 2 h. Una vez alineados, los oligonucleótidos (33  $\mu$ M) se fosforilaron con 1  $\mu$ L de PNK T4 (Roche) y 1  $\mu$ L del buffer la ligasa T4 (Roche). Se incubaron por 30 min a 37°C.

### **Digestión y desfosforilación del plásmido pAc-sgRNA-Cas9.**

Se digirieron dos microgramos del vector pAc-sgRNA-Cas9 durante 1 h a 50°C con la enzima *BspQ1* de la compañía New England Biolabs (NEB) y el buffer 3.1 (100 mM de NaCl, 50 mM de Tris-HCl, 10 mM de MgCl<sub>2</sub>, 100  $\mu$ g/mL BSA) de NEB. Posteriormente, la reacción anterior se incubó a 37°C durante 1 h con la enzima CIP (Fosfatasa alcalina de intestino de becerro) de NEB para desfosforilar el plásmido.

### **Purificación del plásmido pAc-sgRNA-Cas9.**

El vector se purificó usando el kit de purificación de banda QiaexII (QIAGEN), con las instrucciones del fabricante. Para corroborar la digestión, 5  $\mu$ L del plásmido purificado fue sometido a electroforesis por 30 min a 100 V en un gel de agarosa al 1.5%.

### **Ligación de oligonucleótidos en el vector pAc-sgRNA-Cas9.**

Un microlitro del vector purificado se incubó junto a 2 $\mu$ L de los oligonucleótidos fosforilados, 1  $\mu$ L de la ligasa de ADN T4 y 1  $\mu$ L del buffer de la ligasa T4 a 18°C durante 2 h.

### **Transformación.**

La reacción anterior se incubó con 50  $\mu$ L de bacterias competentes *E. coli* Top 10 durante 20 minutos a 4°C. Posteriormente se incubó 30 seg a 42°C y nuevamente a 4°C por 2 min. Se añadió medio LB a la mezcla y se incubó con agitación constante a 37°C durante 1 h. Se

sembraron 50  $\mu$ L de la mezcla anterior en placas Xgal-IPTG con medio LB suplementada con ampicilina y se dejaron incubando a 37°C durante 12 h.

### **Purificación de los plásmidos.**

Se tomaron algunas de las colonias y se volvieron a sembrar en placas LB con ampicilina para expandirlas y se incubaron 12 h a 37°C. Se seleccionaron aquellas colonias que habían insertado el plásmido, lo cual se verificó por PCR con oligonucleótidos específicos usando el siguiente programa en el termociclador: 1 ciclo a 95°C por 5 min, 25 ciclos: 95°C 1 min, 56°C 1 min, 72°C 1 min y un ciclo final de 72°C durante un min. Las bacterias que incorporaron el plásmido se crecieron en 5 mL de medio LB con ampicilina, se incubaron durante 12 h a 37°C con agitación constante. Por cada muestra se usaron 3 mL de las colonias bacterianas seleccionadas, se centrifugaron durante 5 min a 12,000 rpm y se descartó el sobrenadante. La pastilla se resuspendió en 200  $\mu$ L de buffer P1 (50 mM de Tris-HCl pH 8.0, 10 mM de EDTA y 100  $\mu$ g/ $\mu$ L de ARNasa A). Posteriormente, para lisar las células bacterianas, a las muestras se les añadieron 200  $\mu$ L del buffer P2 (200 mM de NaOH y SDS al 1%) y se mezclaron por inversión, dejando incubar 5 min a 25°C. Cada reacción se neutralizó adicionando 200  $\mu$ L de buffer P3 (1M de acetato de potasio pH 5.5). Se incubaron durante 10 min a 4°C y se centrifugaron a 18,000 g durante 10 min. Para precipitar el ADN se tomaron 500  $\mu$ L del sobrenadante y se les adicionaron 350  $\mu$ L de isopropanol, centrifugándose posteriormente a 18,000 g por 10 min. La pastilla de ADN se lavó con 400  $\mu$ L de etanol al 70%, volvió a centrifugarse con las mismas condiciones, se retiró el sobrenadante, se dejó secar la pastilla por 10 min, y finalmente se resuspendió en 30  $\mu$ L de buffer EB. Para corroborar que el ADN extraído contenía el plásmido más el inserto, se hicieron digestiones de las muestras. Dos  $\mu$ g de cada muestra fueron digeridas con 2 U de la enzima de restricción *Eco*RI a 37°C durante 4 h. Los productos de dichas reacciones, se corrieron en un gel de agarosa al 1.5% a 100 V durante 30 min. El ADN que correspondió a muestras que integraron el vector, fueron sometidas nuevamente a purificación de ADN por usando el Kit plasmid midi (QIAGEN catálogo 12143) y este ADN se resuspendió en 30  $\mu$ L de agua estéril. .

### **Transfección para células S2R+.**

Los plásmidos purificados fueron transfectados a células S2R<sup>+</sup> usando las siguientes condiciones, las cuales se han usado en nuestro laboratorio para la línea celular Kc162: se sembraron  $3 \times 10^5$  células y  $5 \times 10^5$  células S2R<sup>+</sup> en placas de 10mL, a las cuáles se les puso medio Schneider, se les adicionaron los plásmidos purificados a concentraciones de  $1 \mu\text{g}/\mu\text{L}$  y el reactivo Fugene (Promega) a diluciones 4:2 y 6:2 por  $\mu\text{g}$  de plásmido. Dado que los plásmidos tienen el cassette de resistencia a puomicina, a las 24 h se comenzó con la selección de colonias con dicho antibiótico durante 7 días a concentraciones de 2, 5 y 10  $\mu\text{g}/\text{mL}$ , para enriquecer las células que hubieran insertado el plásmido.

### **Extracción de ADN genómico por fenol-cloroformo.**

Se centrifugaron al menos  $3 \times 10^6$  células S2R<sup>+</sup> y las células se resuspendieron en 250  $\mu\text{L}$  de una solución compuesta por 200 mM de Tris pH 8.5 y 400 mM de NaCl. Posteriormente, se lisaron usando una solución compuesta por 10 mM de EDTA, SDS 0.4%, 50  $\mu\text{g}/\mu\text{L}$  de ARNasa A y 50  $\mu\text{g}/\mu\text{L}$  de Proteinasa K y se incubaron las soluciones a 50°C durante 4 h. Se adicionaron 500  $\mu\text{L}$  de fenol-cloroformo-álcohol isoamílico y se centrifugaron las muestras a 18,000 g. Se recuperaron los sobrenadantes por centrifugación y se limpiaron con 500  $\mu\text{L}$  de cloroformo, centrifugándose nuevamente, se recuperó el sobrenadante y se añadieron 500  $\mu\text{L}$  de isopropanol. De este último paso se obtuvo la pastilla y se lavó con etanol al 70%. Cada centrifugación se realizó a 18,000 g durante 10 min. Una vez evaporado el etanol, la pastilla se resuspendió en agua suplementada con DEPC (dietilpirocarbonato).

### **Selección y dilución clonal.**

El ADN extraído por fenol/cloroformo de las células S2R<sup>+</sup> transfectadas con el plásmido pAc-sgRNA-Cas9 se re-amplificó mediante PCR punto final y los productos se corrieron en geles de acrilamida no desnaturizante al 5% a un voltaje de 150 V durante 150 min. De las colonias de células transfectadas en las que se encontraron eliminaciones comparadas con las células control, se realizaron diluciones y se sembraron una célula por pozo en una placa de 96 pozos. Estas se continuaron seleccionando y sometiendo a más diluciones clonales.

### **Secuenciación.**

Se extrajo ADN de las clonas y se amplificaron los productos que abarcan las regiones eliminadas por PCR. Las muestras se corrieron en un gel de acrilamida no desnaturizante al 5%, a 150 V durante 150 min en buffer TBE 0.25%. De las bandas obtenidas se purificó el ADN usando el kit de purificación QIAEX II siguiendo las instrucciones del fabricante, volvieron a amplificarse por PCR y se extrajo el ADN con el kit QIAEX II. Los productos se clonaron en plásmidos pGEM-T Easy vector (Promega), y se transformaron en bacterias competentes *E. coli* Top10, se seleccionaron aquellas que presentaban el color blanco indicador de la incorporación del plásmido por la reacción de hidrólisis de galactosa. Se purificó ADN plasmídico con Mini-Preps de la misma manera descrita previamente y se enviaron a la unidad de biología molecular del Instituto de Fisiología Celular de la UNAM, donde las muestras fueron secuenciadas por Sanger.

### **Extracción de ARN total.**

Células S2R+ con 80% de confluencia se trasvasaron a tubos Falcón de 15 mL. Se centrifugaron a 1,200 g durante 5 min, se descartó el sobrenadante y se les lavó con PBS estéril frío, se les volvió a centrifugar a 1,200 g durante 5 min y se descartó el sobrenadante. Se adicionó 1 mL de TRIZOL y se homogenizó por pipeteo. Se incubaron las muestras 10 min a 25°C y se adicionaron 200 µL de cloroformo. Se agitaron durante 15 seg y se incubaron 3 min a 25°C. Después, se centrifugaron a 12,000 g durante 15 min a 4°C, la fase acuosa se transfirió a otro tubo y se añadieron 500 µL de isopropanol. Se les incubó 10 min a 25°C y se centrifugaron a 12,000 g durante 10 min a 4°C. El sobrenadante se descartó y las pastillas se lavaron dos veces con 1 mL de etanol al 75%, se centrifugaron a 7,600 g por 5 min a 4°C. Se removió el sobrenadante y se lavaron las pastillas con 1 mL de etanol absoluto. Volvieron a centrifugarse a 7,600 g por 5 min a 4°C y se removió el sobrenadante. Las pastillas se dejaron secando 10 min y se resuspendieron con 30 µL de agua DEPC.

### **Comprobación de la integridad del ARN.**

Se tomó 1µg de ARN de cada muestra y se llevó a un volumen final de 5 µL con agua estéril, se les adicionó 1 µL de buffer de electroforesis y se incubó a 65°C por 15 min, posteriormente se le agregó 1 µL de bromuro de etidio. Las muestras se cargaron en geles desnaturizantes de agarosa al 1.2% con formaldehído y la electroforesis se realizó a 80 V durante 90 min.

## Reacción de transcripción reversa

Para la generación de ADN complementario a partir de los extractos de ARN total, se tomó 1 µg de ARN de cada muestra y se les trató con DNAsa RQ1 (Promega) (1 µL por µg de ARN), se les adicionó el buffer RQ1 y se incubaron durante 30 min a 37°C. Para detener la reacción se adicionó una solución de paro y se incubó 10 min a 65°C. Este ARN tratado con DNAsa RQ1 se incubó con Retrotranscriptasa (AMV Promega), buffer para la enzima, dNTPs (0.8 µM) y oligo-dT y dicha reacción se sometió a los siguientes ciclos en un termociclador Bio-Rad; 22°C 10 min, 42°C por 1 h, 95°C por 5 min y 5°C durante 5 min.

## Reacción en cadena de la polimerasa cuantativa (qPCR).

Para cada reacción de qPCRs, se usaron 2.5 µg de ADNc, 4 picomoles de oligonucleótidos específicos para cada condición, 0.5 µM de dNTPs, polimerasa kappa, buffer de la polimerasa con 2.5 mM de magnesio y 1 µL Sybergreen (Sigma). Todos los oligonucleótidos diseñados para la PCR cuantitativa amplifican de la manera más óptima a 60°C productos de entre 100 y 200 pares de bases. Usando el termociclador OneStep, los ciclos para llevar a cabo las reacciones fueron: 1 ciclo a 94°C durante 5 min, 40 ciclos a 94°C por 30 seg, 60°C por 40 seg y 72°C por 30 seg y finalmente el ciclo final se realizó a 72°C durante 5 min. Cada reacción se realizó por triplicado usando como control negativo agua y los resultados finales se obtuvieron con 2 réplicas biológicas.

## Tabla de oligonucleótidos

Utilizando el programa PRIMER BLAST de la base de datos “National Center for Biotechnology Information” (NCBI) (Ye et al., 2012), se diseñaron los oligonucleótidos específicos para cada región génica que fueron usados para ensayos de RT-PCR.

Nombre		Secuencia	Tamaño del amplicón
ASF1	Fw	CTGAGGGGCGACACATCTTT	109 pb
	Rev	CGGTAGGAGCAGGTAAGCAG	

<b>1 (3) 76BDm</b>	Fw	CCGAGGGAGTGACGAACAAA	160 pb
	Rev	AGATCGCCCAACTTTCGTGT	
<b>RP49</b>	Fw	GCATACAGGCCCAAGATCG	113 pb
	Rev	GGGTACAATTTGGCACAATCC	
<b>Rob1</b>	Fw	AATGTCGCAGGAGGTTGAGG	149 pb
	Rev	CTGACTCATTAGGCCAGCGT	
<b>CTCF</b>	Fw	GCCAACTCCAAGATCAGTG	93 pb
	Rev	GCCATCGACCTCGTTTTTC	
<b>ASF1 CRISPR</b>	1 Fw	TCCGTAATGCCGCTGAAACT	246 pb sin eliminaciones
	2 Rev	TTCGCACAAATGGTGACAGC	

## 7. RESULTADOS

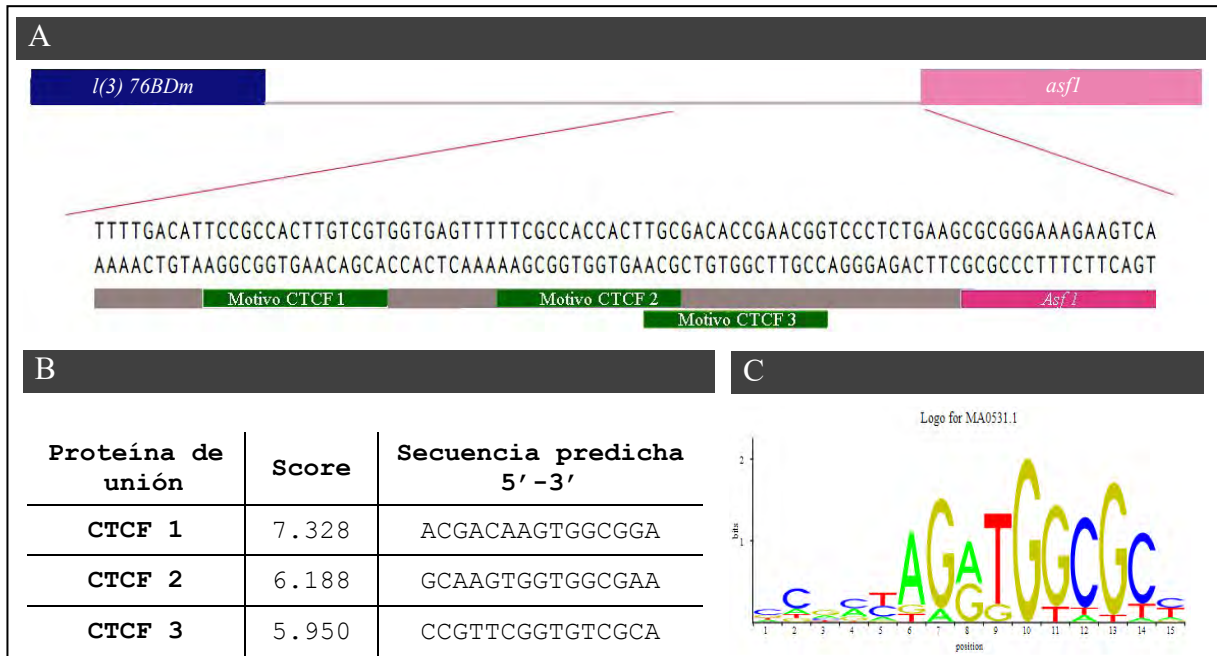
### Unión de CTCF a la región intergénica 5' de *asfl*.

A partir de los datos publicados por el grupo de Corces (Bushey et al., 2009), identificamos que la señal de CTCF se encuentra enriquecida en la región intergénica 5', cercana al sitio de inicio de la transcripción de *asfl* en *D. melanogaster*, a comparación de sitios vecinos. Sabiendo que CTCF está enriquecido en esta región decidimos comprobar si la unión de CTCF a dicha región, tiene un papel en la regulación de la transcripción de *asfl*.

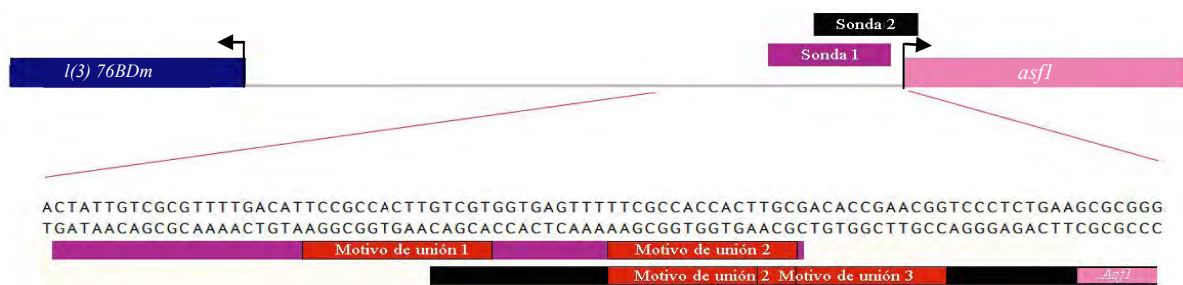
Los datos de CHIP-seq sugieren que CTCF se une a la región 5' de *asfl*, pero no indican a que motivos en el ADN lo hace. Debido a que CTCF reconoce motivos consenso en el genoma de *Drosophila* para su unión a cromatina, se utilizó la base de datos de insectos JASPAR (Mathelier, 2013), para definir la localización de estos motivos en la región 5' de *asfl*. Se introdujo la secuencia intergénica entre *l(3)76BDm* y *asfl* en JASPAR y se obtuvieron tres sitios probables de reconocimiento del factor nuclear CTCF cercanos al sitio de inicio de la transcripción de *asfl* (**Figura 14**).

Como una aproximación inicial, con ensayos *in vitro*, se decidió comprobar si CTCF era capaz de unirse a estos motivos, a través de ensayos de retardo en la movilidad electroforética (EMSA). Por medio de estos ensayos se puede demostrar la existencia de interacciones de la proteína de interés con ácidos nucleicos. Esta técnica se basa en la observación de que la movilidad electroforética de complejos formados por la unión de proteína-ADN, es menor en comparación de ácidos nucleicos libres donde no se une la proteína. Tomando como referencia la secuencia intergénica entre *l(3)76BDm* y *asfl*, se diseñaron dos pares de sondas complementarias de oligonucleótidos de ADN que contienen los motivos predichos de unión a CTCF. En cada una de estas sondas, los motivos de unión de CTCF se encuentran en la parte central de cada una de ellas, pues además de las 15 pares de bases principales para el reconocimiento por CTCF, se necesitan de aproximadamente 50 pares de bases en total, para la unión correcta de los dedos de zinc a ambos lados de la secuencia central (Nakahashi et al., 2013).

En la **Figura 15** se muestran las secuencias de ADN utilizadas para el diseño de las sondas y la ubicación de los motivos de unión de CTCF en cada una de estas. La sonda 1 contiene a los sitios 1 y 2 predichos, mientras que la sonda 2 contiene los sitios 2 y 3, los cuales se sobreponen en cuatro pares de bases.



**Figura 14. Motivos de unión de CTCF a la región intergénica 5' de *asfI*.** A) Localización de tres motivos de unión de CTCF predichos por el análisis de la base de datos JASPAR (Mathelier, 2013). Estas secuencias se encuentran cercanas al sitio de inicio de la transcripción de *asfI*. B) Secuencias de los motivos de unión de CTCF en esta región. El score se refiere a la similitud de la secuencia analizada comparada con las secuencias de nucleótidos más frecuentes a los que une un factor transcripcional. C) Logo del motivo consenso de reconocimiento por CTCF en *D. melanogaster*.



**Figura 15. Diseño de las sondas de ADN para los EMSAs.** Las secuencias que se usaron para el diseño de las sondas se encuentran río arriba del sitio de inicio de la transcripción de *asfI*. Cada una de las sondas, contiene dos motivos de unión para CTCF (cajas rojas) predichos por la base de datos JASPAR e indicados en las cajas rojas. La sonda 1 (morado) contiene los sitios 1 y 2 y consta de 60 pares de bases de ADN, mientras que la sonda 2 (negro) contiene los sitios 2 y 3 y abarca 58 pares de bases.

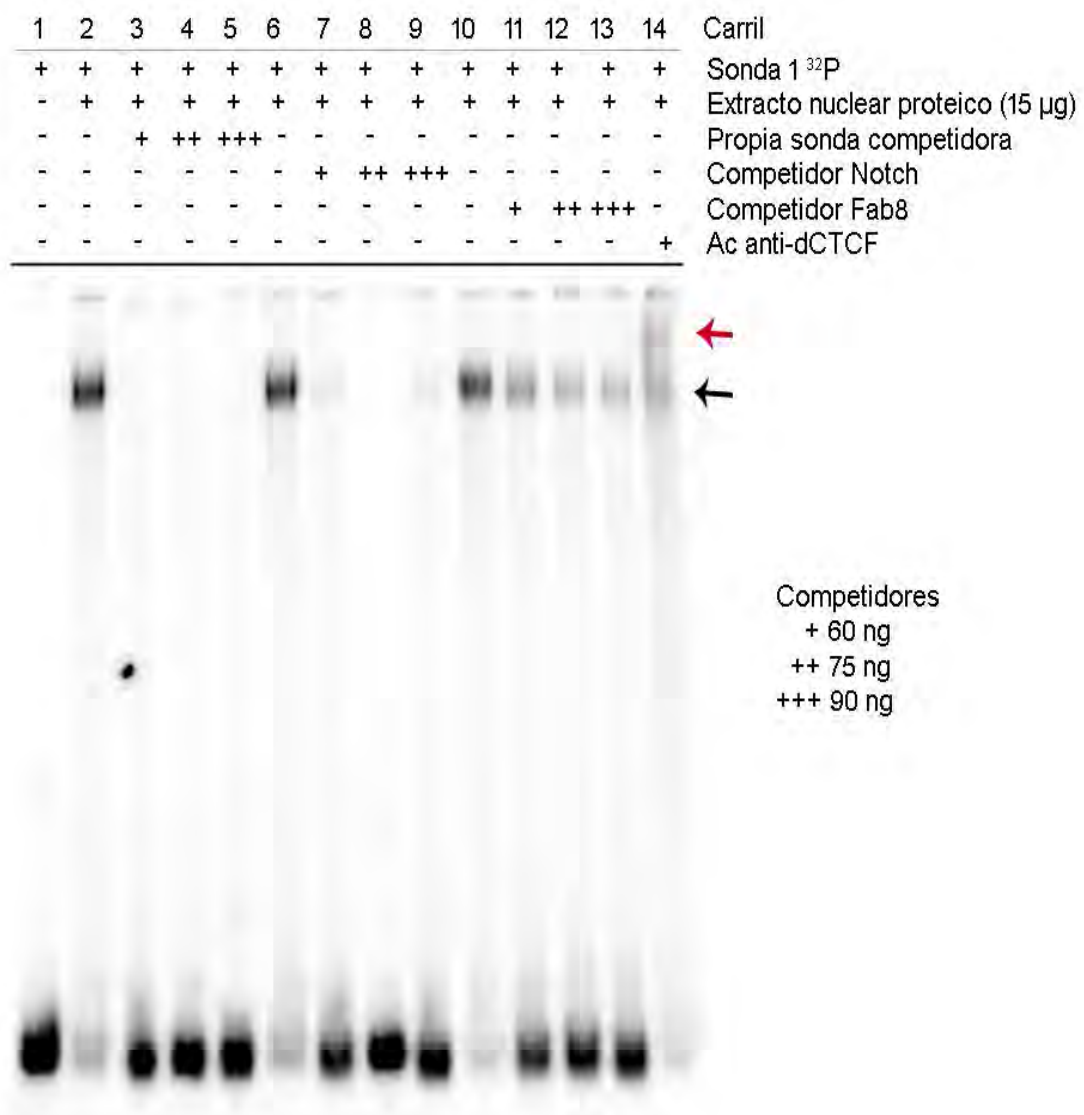


Las sondas de oligonucleótidos se marcaron con  $^{32}\text{P}$  y se realizaron los ensayos de retardo. Se usaron dos aproximaciones para demostrar la especificidad de la unión de CTCF con estas sondas: 1) Se usaron sondas de oligonucleótidos competidoras y 2) Se hicieron super-retardos con un anticuerpo específico contra CTCF de *D. melanogaster*. Como competidores se usaron sondas frías: las propias sondas 1 y 2, una sonda de la región *Fab8*, que une a CTCF con gran eficacia (Moon et al., 2005) y una sonda de una región intergénica cercana al sitio de inicio de la transcripción de *Notch*. Por ensayos en nuestro laboratorio, hemos demostrado que esta última sonda también une a CTCF. Todas estas sondas competidoras se usaron en concentraciones crecientes que exceden a las sondas marcadas radioactivamente, a concentraciones de 60 ng (100 veces más), a 75 ng y 90 ng. Finalmente, para los super-retardos se utilizaron un anticuerpo específico contra CTCF (anti-dCTCF) y como control negativo el anticuerpo IgG anti-conejo, para comprobar que CTCF es parte de los complejos de ADN-proteína formados.

En la **Figura 16** se muestra el resultado del ensayo de retardo utilizando a la sonda 1. En el carril 1 se puede observar como migra la sonda 1 marcada con  $^{32}\text{P}$  cuando no interacciona con proteína alguna. En los carriles 2, 6 y 10, se incubó la sonda 1 marcada radioactivamente con 15  $\mu\text{g}$  de extracto de proteínas nucleares de la línea celular S2R+ y puede observarse que alguna proteína nuclear interacciona con la sonda y forma un complejo proteína-sonda, ya que la sonda migra más lentamente que cuando la sonda está libre. La señal de la sonda se pierde cuando se incorporan los distintos competidores (la propia sonda, sonda Notch y sonda Fab8). La sonda que mejor compite es la sonda 1 pues la señal del complejo desaparece desde los 60 ng de competidor (carriles 3, 4 y 5). La siguiente mejor competidora es la sonda Notch (carriles 7 y 9) y por último la sonda Fab8 (carriles 11, 12 y 13). Que la sonda Fab8 compita con la menor eficiencia, podría deberse a que CTCF puede reconocer ambas sondas con la interacción de distintos dedos de zinc en cada una de ellas. En el carril 14 se muestra el super-retardo, donde se demuestra que en dicho complejo de proteína-sonda participa CTCF. En esta muestra se utilizó el anticuerpo anti-dCTCF, que ocasiona que se disminuya la intensidad de la señal radioactiva del complejo y aparece una banda, que migra más lentamente. Es en esta banda que el anticuerpo está uniéndose al complejo, lo cual retarda su movilidad electroforética.

Estos resultados indican que CTCF puede unirse al motivo 1 de la sonda 1. El motivo 2 se encuentra en la porción final de la sonda y resulta menos probable que CTCF pueda

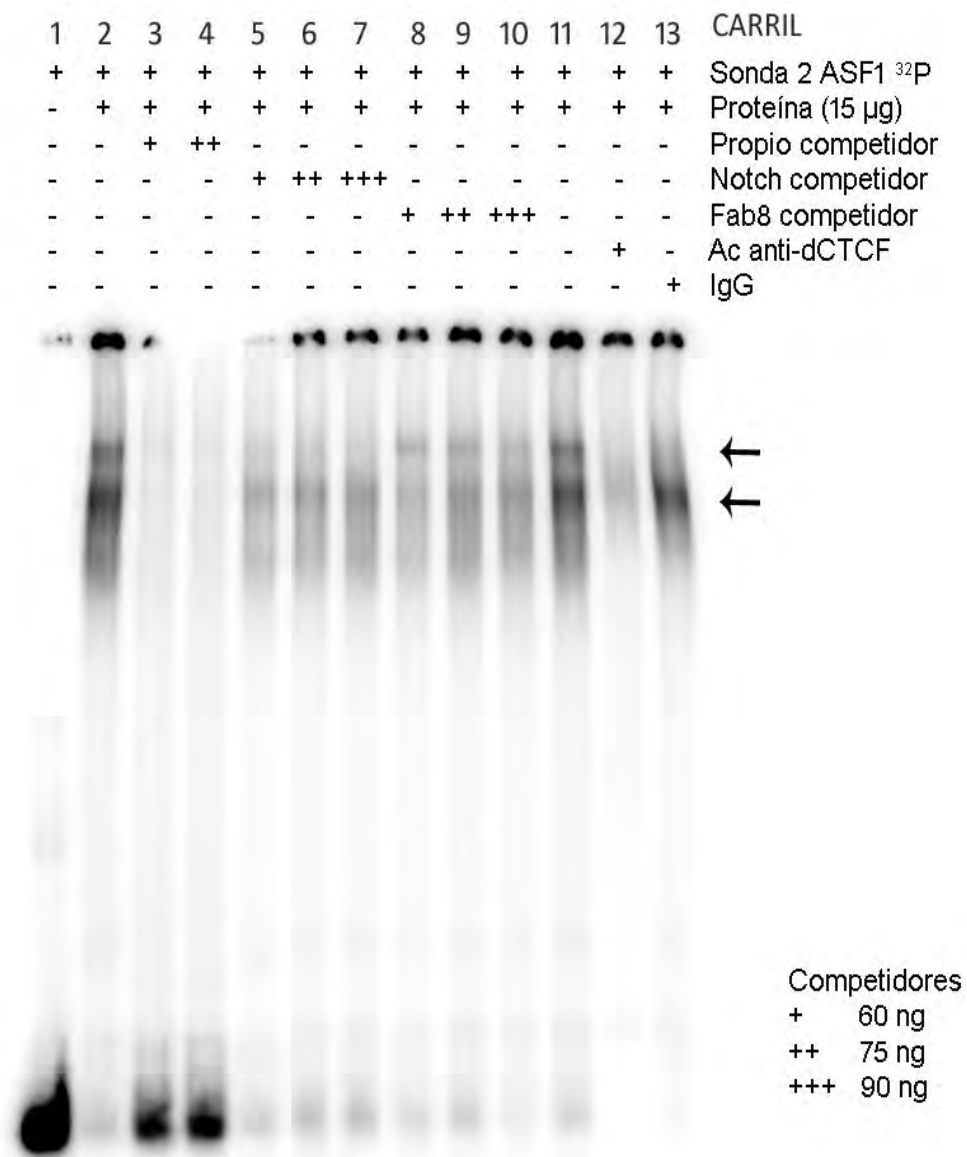
unirse correctamente a esta secuencia, debido a que son necesarias más bases a los lados de la secuencia central.



**FIGURA 16. Ensayo de retardo de movilidad electroforética de la sonda 1.** En los carriles 2, 6 y 10 se puede observar que se forma un complejo de la sonda 1 radioactiva con proteínas del extracto nuclear de la línea celular S2R+ (flecha negra). En los carriles 3, 4 y 5 se usa como competidor a la propia sonda en exceso, en los carriles 7, 8 y 9 se usa la sonda competidora Notch, mientras que en los carriles 11, 12 y 13 la sonda competidora Fab8, en el carril 14 es el super-retardo usando un anticuerpo anti-dCTCF (flecha roja).

En la **Figura 17** se muestra el resultado del ensayo de la sonda 2. En el carril 1 se muestra la migración de la sonda 2 libre. En los carriles 2 y 11 se puede observar que algunas proteínas del extracto nuclear pueden reconocer esta sonda, lo que permite la formación de dos complejos. En los carriles 3 y 4 se usa como competidor la propia sonda en exceso y se

observa que la señal desaparece desde los 60 ng. Cuando se usan los competidores Notch (carriles 5, 6 y 7) y Fab8 (carriles 8, 9 y 10) se observa que la intensidad de las marcas radioactivas de ambos complejos disminuye, en ambas condiciones. Finalmente en el carril 11, se muestra el super-retardo, donde el anticuerpo anti-dCTCF secuestra a dichos complejos, disminuyendo así la señal de las bandas de menor movilidad. Estos datos sugieren que CTCF se une a secuencias en el ADN de ambas sondas o al menos forma parte de los complejos de ADN-proteínas observados en los ensayos de movilidad electroforética. El que CTCF sea parte de las proteínas de estos complejos se corrobora con los super-retardos



**FIGURA 17. EMSA de la sonda 2.** Se forman dos complejos entre la sonda 2 radioactiva y proteínas del extracto nuclear de la línea celular S2R+ (carriles 2 y 11. Flechas negras). Competidores: la propia sonda en exceso (carriles 3 y 4), la sonda Notch (carriles 5, 6 y 7) y la sonda Fab8 (carriles 8, 9 y 10). Se muestra el super-retardo con el anticuerpo anti-dCTCF (carril 12) y con el anticuerpo IgG, como control negativo (carril 13).

## Generación de la mutación del motivo de unión 1 a CTCF.

CTCF puede reconocer y unirse a las sondas 1 y 2. Para evaluar el papel que tiene este factor en la regulación de *asfl* se mutó el motivo 1 de unión a CTCF en la línea celular S2R+. Además, se analizó el efecto sobre la transcripción de *l(3)76BDm*, ya que la distancia entre ambos es de 379 pares de bases y se ha sugerido que CTCF podría tener un papel en la regulación diferencial de genes divergentes (Smith et al., 2009). No se mutaron los otros 2 motivos, ya que se encuentran muy cercanos al sitio de inicio de la transcripción de *asfl* y la unión de otros factores necesarios para el inicio de la transcripción podría ser afectada. Para lograr dicho objetivo, procedimos a mutar el motivo de unión 1, por medio de la técnica CRISPR-Cas9.

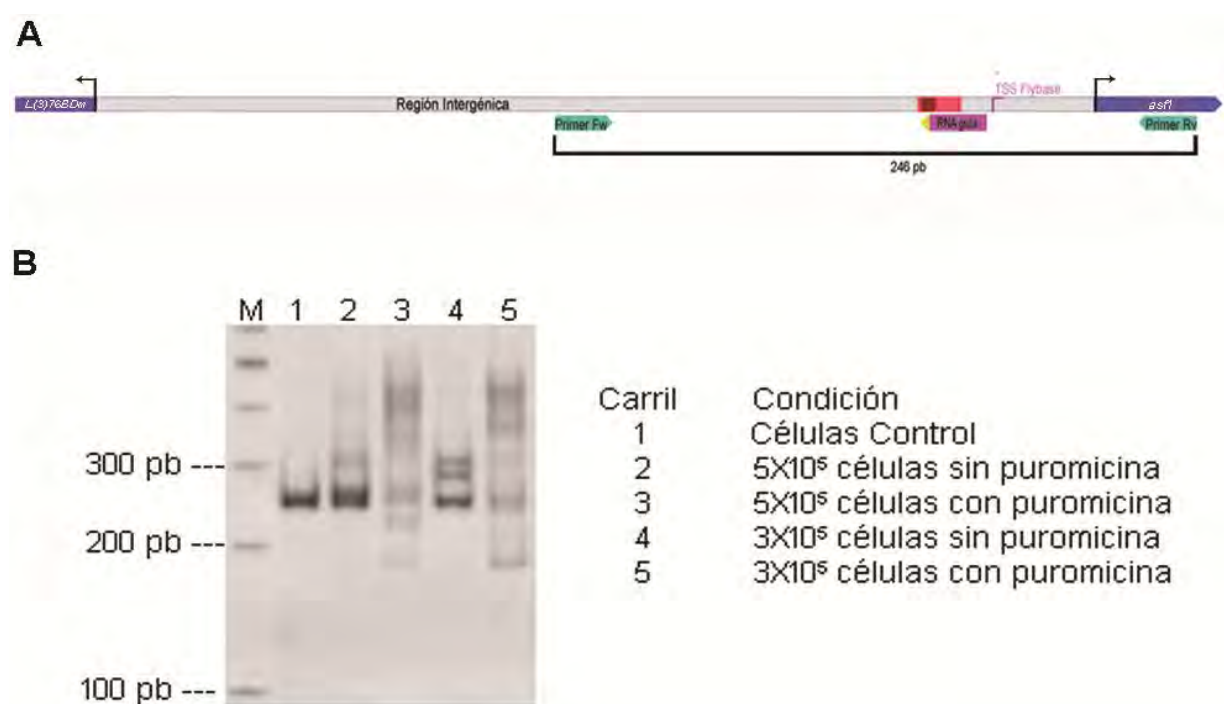
Para este trabajo utilizamos un plásmido (pAc-sgRNA-Cas9) que codifica para la Cas9 y un sgRNA específico para el reconocimiento de parte de motivo 1 de CTCF (ver Materiales y Métodos). Para el diseño del sgRNA, se usó la página <http://crispr.mit.edu/>, la cual muestra secuencias que pueden ser utilizadas para la técnica CRISPR-Cas9, ya que toma en cuenta tanto las secuencias PAM, como que haya una baja probabilidad de blancos inespecíficos en el genoma.

Con este análisis, se obtuvo como posible secuencia blanco, una región de 20 nucleótidos que abarca la mayor parte del motivo 1 de CTCF (**Figura 19**). Esta secuencia se clonó dentro del plásmido pAc-sgRNA-Cas9, que se transformó en bacterias *E. coli* Top 10 competentes y posteriormente se transfectó en la línea celular S2R+ (línea celular embrionaria de *D. melanogaster*). Esta línea celular tiene un nivel de expresión alto de *asfl*, lo que nos permitiría evaluar si las mutaciones en la región de estudio afectan la expresión de dicho gen.

Las transfecciones se realizaron usando diferentes condiciones que influyen sobre la eficiencia de la transfección: se transfectó 1  $\mu\text{g}/\mu\text{L}$  de plásmido en  $3 \times 10^5$  células o  $5 \times 10^5$  células S2R+ y a estas se les adicionaron 2, 5 o 10  $\mu\text{g}/\text{mL}$  de puomicina, para seleccionar positivamente la integración del vector. Para evaluar las mutaciones resultantes, se llevaron a cabo ensayos de PCR a partir de ADN genómico obtenido de colonias celulares transfectadas con el plásmido y de las correspondientes células control no transfectadas. Para ello se usaron un par de oligonucleótidos que en células control amplifican 246 pares de bases que contienen el motivo de unión 1 a CTCF y abarcan parte la región codificante de *asfl* y parte de la región intergénica (**Figura 18**). El cambio en el tamaño de los amplicones en las células

transfectadas refleja la naturaleza de las mutaciones resultantes (inserciones y eliminaciones). La ventaja de esta técnica, es que pueden generarse eliminaciones de 1 hasta aproximadamente 50 o más pares de bases.

Para darnos una idea general del tamaño de las mutaciones obtenidas, se usaron geles de acrilamida al 5% y se realizó la electroforesis durante 2 h a 150 V. Esta metodología permitió apreciar diferentes bandas con diferencias de tamaño de aproximadamente 10 pb. Las condiciones que tuvieron las mejores eficiencias de transfección fueron aquellas donde se usaron  $3 \times 10^5$  y  $5 \times 10^5$  células con  $5 \mu\text{g/mL}$  de puromicina. En la **Figura 18** muestro que en los eventos de mutación hubo tanto eliminaciones de nucleótidos como inserciones, a comparación de condiciones control, donde se evaluaron células S2R+ que no fueron transfectadas con el plásmido que contiene al sistema CRISPR-Cas9.



**Figura 18. Productos derivados del sistema CRISPR-Cas9 en la línea celular S2R+.** Se evaluó la naturaleza de las mutaciones resultantes de la transfección del sistema CRISPR-Cas9 mediante análisis por PCR. **A)** Para la amplificación de la región que contiene al motivo 1 de CTCF (caja roja), se usaron dos oligonucleótidos: Primer Fw y Primer Rev (azul claro), que en células control amplifican 246 pares de bases. **B)** Productos que se obtuvieron de la transfección en la línea celular S2R+ con diferentes condiciones de transfección y selección. En los carriles 2 y 3 se usaron  $5 \times 10^5$  células y en los carriles 4 y 5,  $3 \times 10^5$  células, tanto con  $5 \mu\text{g/ml}$  de puromicina como sin ella. Puede observarse que a diferencia de las células control (carril 1), que solo dan un amplicón de 246 pares de bases (flecha roja), en los carriles 3 y 5 existen poblaciones celulares con mutaciones de diferente naturaleza, ya que tienen productos de amplificación tanto de menor como de mayor tamaño, indicando inserciones y eliminaciones respectivamente.

Con el fin de comprobar la naturaleza de las mutaciones a nivel de secuencia, se recuperó el ADN de las bandas correspondientes a la amplificación de esta región, tanto de células control, como de las bandas de menor tamaño de las células transfectadas. El ADN recuperado se clonó en plásmidos pGEM T-Easy (Promega) y se transformaron en bacterias *E. coli* Top10 competentes. Los plásmidos purificados se secuenciaron en la unidad de Biología Molecular del Instituto de Fisiología Celular de la UNAM. Las eliminaciones resultantes (**Figura 19**), remueven parte del motivo de unión 1 a CTCF y van desde 13 hasta 65 pares de bases. Además, las secuencias a ambos lados de las eliminaciones son iguales a las de la condición control y corresponden perfectamente a la de esta región anotada en el “Genome Browser” de la Universidad de California Santa Cruz. Esto indica que, no se generaron mutaciones en las secuencias aledañas, corroborando la especificidad de los sgRNAs diseñados.



**Figura 19. Eliminaciones en el motivo 1 de CTCF generadas con la técnica CRISPR-Cas9.** El RNA guía de la Cas9 utilizado, es complementario a 20 nucleótidos en el genoma de *D. melanogaster*, de los cuales 11 abarcan parte del motivo 1 de CTCF, incluyendo la secuencia “core” (nucleótidos principales para el reconocimiento de CTCF). En amarillo se muestra la secuencia PAM y en la línea rosa está el sitio de inicio de la transcripción de *asfI* (azul), predicho por la base de datos Flybase. Los productos de PCR obtenidos de las bandas de menor tamaño a la control, resultaron en eliminaciones que van de 13 a 65 pares de bases. Todas ellas remueven ocho nucleótidos del motivo 1 de CTCF. Para la amplificación por PCR de estos productos, se usaron dos oligonucleótidos: Primer Fw y Primer Rev (azul claro), que en células control amplifican 246 pares de bases.

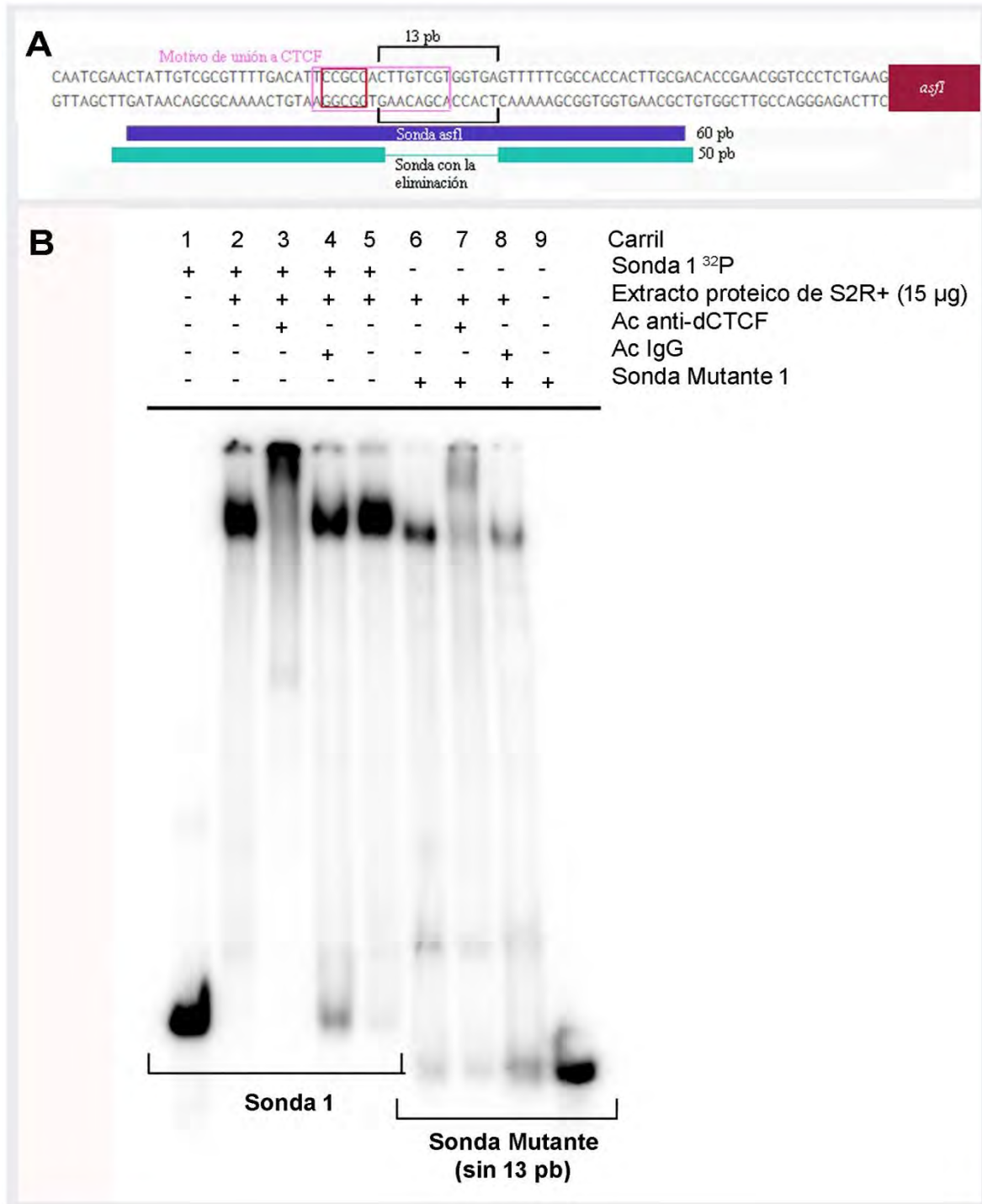
Las secuencias mutadas obtenidas muestran que la técnica CRISPR-Cas9, usando el sgRNA descrito, ha sido eficaz en la eliminación de parte del motivo 1 de CTCF y ha generado mutaciones de distintos tamaños. Estas secuencias, al ser analizadas en la base de datos JASPAR, ya no contienen el sitio de unión 1 a CTCF en esta región. Esto nos servirá

para saber si la unión de CTCF a este motivo tiene un efecto en la regulación de la transcripción de *asf1* en un contexto *in vivo*. Además, a diferencia del uso de un RNA de interferencia que afectaría la expresión de CTCF de manera global, los resultados obtenidos con estas mutaciones serán específicos de la unión de CTCF a esta región del genoma.

### **La eliminación de parte del motivo 1 de CTCF impide la unión de este factor.**

Con el fin de validar que la mutación del motivo 1 de CTCF afecta su unión a esta región, realizamos ensayos de retardo de movilidad electroforética con una sonda que contiene una de las mutaciones obtenidas. Para evitar la probabilidad de que se esté afectando la unión de otros factores transcripcionales al ADN, se eligió probar la mutación que corresponde a la eliminación del segmento de menor tamaño del motivo 1 de unión a CTCF. Se diseñó una sonda de 50 pares de bases, tomando como molde la sonda 1 usada anteriormente, pero en este caso, llevando una remoción de 13 pares de bases, para demostrar que con tan sólo esta pequeña eliminación se pierde la unión de CTCF en el motivo 1. El diseño de la sonda y los resultados obtenidos se muestran en la **Figura 20**.

Para el EMSA se usaron, tanto la sonda 1 como la sonda mutante marcadas radioactivamente. Ninguna de las sondas forma complejo alguno cuando no se incuban con proteínas (carriles 1 y 9). Ambas, al incubarse con los extractos nucleares proteicos de células S2R<sup>+</sup> forman un complejo que migra a una altura similar, sugiriendo que las sondas reconocen a las mismas proteínas. Sin embargo, la intensidad de la señal del complejo en la sonda mutante (carril 6) es mucho menor comparando la intensidad del complejo que se forma con el de la sonda 1 (carriles 2 y 5). Además, al usar el anticuerpo que reconoce específicamente a CTCF, en ambas condiciones se observa un super-retardo (carriles 3 y 7), confirmando que la proteína CTCF es parte de dichos complejos y estos retardos no se generan al usar el anticuerpo IgG de conejo (carriles 4 y 8). Estos ensayos se realizaron por duplicado.



**Figura 20. Ensayo de movilidad electroforética de la sonda mutante.** A) Respecto a la sonda 1, la sonda mutante carece de 13 pares de bases, que corresponden a la eliminación más pequeña obtenida por CRISPR-Cas9 en la línea celular S2R+. Se elimina parte del motivo de unión 1 a CTCF. B) Migración de la sonda 1 libre (carril 1). Se incubó la sonda 1 con extractos proteicos nucleares de células S2R+(carriles 2 y 5). Super-retardo de la sonda 1 con el anticuerpo anti-dCTCF (carril 3). Anticuerpo inespecífico IgG de conejo como control del super retardo (carril 4). Sonda mutante con extractos proteicos nucleares (carril 6), nótese que se observa la formación de un complejo, el cual cambia la velocidad de migración al hacerse el super-retardo con el anticuerpo anti-dCTCF (carril 7). Pero la migración del complejo no cambia cuando se usa el control IgG (carril 8) y la sonda libre no forma complejo alguno (carril 9). Con esto se demuestra que en la sonda mutante, CTCF se une menos eficazmente que cuando el motivo 1 se encuentra intacto.

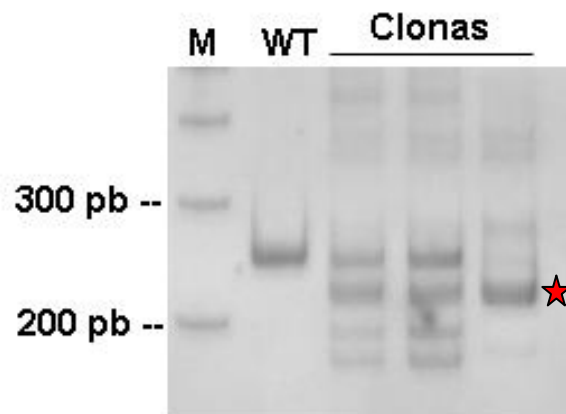


Con estos resultados se valida que CTCF es capaz de reconocer el motivo 1 predicho por la base de datos JASPAR y que la mutación de este, reduce la unión de dicho factor a esta región. Cabe mencionar que algunas de las razones por las que la unión de CTCF no se elimina totalmente en la sonda mutante podría deberse a que: 1) aún se mantienen algunos nucleótidos clave (parte del “core”) para el reconocimiento de dicho factor, 2) como sucede en humanos (Filippova et al., 1996), algunos de los dedos de zinc pueden reconocer nucleótidos río arriba de la región “core” o 3) parte del motivo 2 está presente en la sonda mutante. Sin embargo la unión de CTCF se reduce y en un contexto de cromatina (células) esta unión podría no estarse dando o se estaría viendo afectada en gran medida.

Si observamos cambios en la expresión de *asf1* en poblaciones celulares que contienen las eliminaciones del motivo 1 de unión a CTCF, los resultados obtenidos en este último ensayo EMSA, nos darán la certeza de que CTCF si participa en la regulación de *asf1* por su unión específica a esta región del genoma.

### **Regulación de la expresión de *asf1* por el factor CTCF.**

Para evaluar si la expresión de *asf1* y de *l(3)76B<sup>Dm</sup>* se afecta con la eliminación mínima (13 pb) del motivo 1 de unión de CTCF, se obtuvieron poblaciones celulares enriquecidas de manera individual con las mutantes. Esto se logró haciendo diluciones clonales de las células con la condición de transfección de  $3 \times 10^5$  células con 5  $\mu\text{g/ml}$  de puromicina para la selección. De las diluciones que se hicieron, una de ella se encuentra enriquecida en casi su totalidad, con células que tienen una eliminación de aproximadamente 20 pares de bases. A pesar que la población celular utilizada aun tiene células con mutaciones de diferente tamaño, consideramos que la mutante de nuestro interés se encuentra lo suficientemente enriquecida para obtener datos confiables del papel de CTCF en esta región. Cabe aclarar que se utilizó esta población, debido a que las clonas celulares con regiones eliminadas de menor tamaño (13 pares de bases) se encontraban muy poco enriquecidas. La obtención de estas aun se encuentra en proceso y cuando las obtengamos se realizarán los mismos ensayos que con la eliminación de 20 bases. El análisis de los fragmentos que contienen la eliminación de la región de CTCF de las poblaciones celulares S2R<sup>+</sup> enriquecidas con diferentes eliminaciones se muestra puede ver en la **Figura 21**.



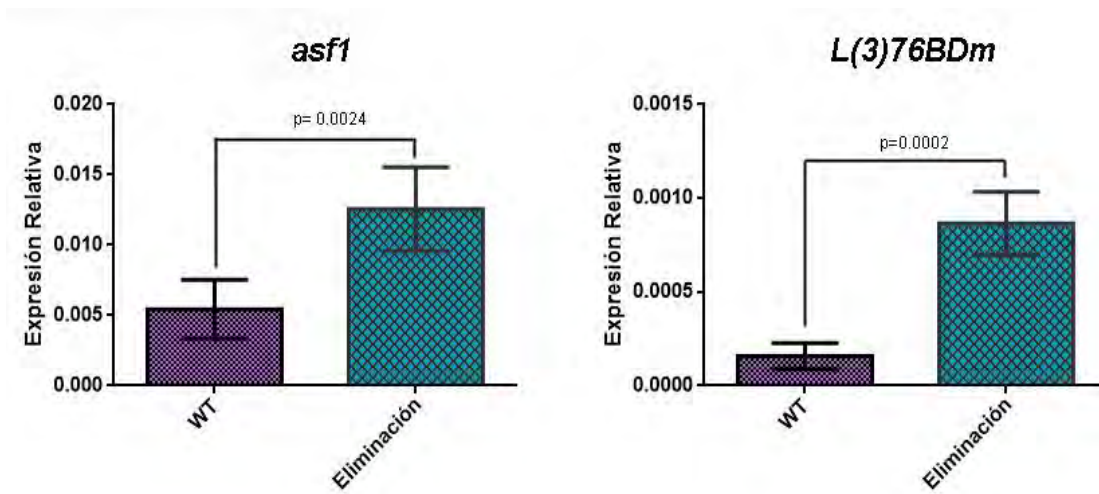
**Figura 21. Clonas celulares utilizadas para los ensayos de PCR cuantitativa.** En el último carril (estrella roja), se muestra la colonia celular de S2R<sup>+</sup> utilizada para evaluar el efecto de la eliminación del motivo de unión 1 de CTCF. Esta población celular está enriquecida con una mutante que carece de aproximadamente 20 pares de bases a comparación con las células control (WT). En los carriles centrales se muestran clonas que contienen diferentes mutaciones obtenidas por la técnica CRISPR-Cas9.

Para evaluar la expresión de *asf1* y *l(3)76BDm*, se extrajo ARN total de la clona celular seleccionada y de células S2R<sup>+</sup> control para hacer transcripción reversa y finalmente PCR cuantitativa. Como control de expresión constitutiva se utilizó el gen *rp49* que produce la proteína ribosomal 49. Estos ensayos se realizaron con dos muestras biológicas independientes, cada una por triplicado.

En la **Figura 22** se muestran los niveles de expresión de la población celular con la eliminación comparada con células control. Cuando se elimina parte del motivo de unión a CTCF, los niveles de expresión de ambos genes aumentan. La expresión de *asf1* aumenta poco más del doble, mientras que la de *l(3)76BDm* aumenta aproximadamente cinco veces más. Cabe mencionar que los datos muestran la expresión relativa, es decir los valores se normalizaron utilizando la expresión de *rp49*.

En resumen, con estos resultados se confirma que la unión de CTCF al motivo 1 tiene un papel en el control de la transcripción del gen de la chaperona de histonas ASF1, así como en el gen que se encuentra río arriba, *l(3)76BDm*. La ausencia de este motivo permite que ambos genes incrementen su expresión. Esto indica que CTCF podría ser importante para la regulación diferencial de genes divergentes en *D. melanogaster*, además de que por sus funciones ya descritas, CTCF estaría ejerciendo dicha regulación, permitiendo que la región

promotora de cada uno de estos genes se contacte con elementos regulatorios, como potenciadores, que permitan este aumento en su expresión. Sin embargo para confirmar esto, aun deben realizarse otros ensayos.



**Figura 22.** Cambio en la expresión de *asf1* y *l(3)76BDm* con la eliminación del motivo 1 de unión al factor CTCF. En la figura de la izquierda se muestra la expresión de *asf1* en condiciones normales (WT), en células S2R+ y se observa como este gen se sobre-expresa en la población celular que contiene la mutante al motivo de unión. Lo mismo sucede con *l(3)76BDm* (figura derecha). Para el análisis estadístico se usó una prueba t-student.

## 8. DISCUSIÓN

La regulación transcripcional debe ser finamente controlada durante el desarrollo de un organismo, ya que la habilidad de las células para crecer, diferenciarse y responder ante diversos estímulos requiere de la expresión espacio-temporal coordinada de cientos de genes (Schertel et al., 2015). Un mecanismo efectivo del control de la transcripción génica involucra una combinatoria de factores transcripcionales, la ARN polimerasa, la modulación de la estructura de la cromatina y otros reguladores que deben unirse a elementos en el ADN (Cave, 2011). Entre los factores transcripcionales, el factor nuclear CTCF contribuye de manera indispensable a la regulación de la expresión de genes y la organización tridimensional del genoma.

A pesar de que CTCF tiene un papel como organizador global de la cromatina, ha sido asociado también a la regulación de genes de manera individual. Los resultados mostrados en este trabajo son uno de los pocos ejemplos en la literatura donde se demuestra que CTCF está involucrado directamente en la regulación de un gen indispensable para el desarrollo de *Drosophila*. Esto es relevante pues la mayoría de los estudios han analizado los cambios en el genoma completo ante la disminución de CTCF, usando ARNs de interferencia o aboliendo totalmente su expresión (Schwartz et al., 2012; Van Bortle et al., 2012). Sin embargo, existen otros ejemplos de cómo CTCF regula genes específicos en vertebrados (Bell et al., 1999; Guo et al., 2015; Narendra et al., 2015).

Las proteínas de unión a “insulators” están involucradas en la regulación de la expresión de genes a través de la formación de asas y es posible que el establecimiento de dominios por medio de estos elementos, se requiera durante el desarrollo para crear perfiles de transcripción durante la especificación celular (Bushey et al., 2009). La proteína de unión a secuencias “insulator” CTCF se une de manera ubicua en sitios conservados en el genoma, sin embargo, muchos otros sitios a los que se une varían de célula a célula, lo que indica que juega un papel importante en el mantenimiento de la diversidad fenotípica (Torrano et al., 2005). Las mutaciones nulas de CTCF en *D. melanogaster* son letales, mientras que las hipomórficas resultan en fenotipos homeóticos (Mohan et al., 2007). Uno de los genes que podría estar involucrado en los problemas de desarrollo ante mutaciones de CTCF, es el gen *asfl*, dado su papel multifuncional como chaperona de histonas. Por estas razones quisimos

dilucidar si la unión de CTCF a la región intergénica 5' de *asfl* es necesaria para regular la expresión de este gen.

Los datos de ChIP-seq (Bushey et al., 2009) indican que CTCF está enriquecido en la región intergénica 5' de *asfl*, pero no nos dicen a que región específica del genoma se une. Para definir la región a la que une CTCF buscamos los motivos de unión consenso que podría reconocer esta proteína, usando la base de datos JASPAR (Mathelier, 2013; **Figura 14**). Con esta aproximación definimos que existen tres sitios de unión probables cercanos al sitio de inicio de la transcripción de *asfl*, a los que podría unirse CTCF. Mediante ensayos de retardo en la movilidad electroforética (EMSAs) demostramos que CTCF es capaz de reconocer y unirse a la sonda 1 (contiene al motivo 1 y a motivo 2 de unión a CTCF) y a la sonda 2 (contiene a los motivos 2 y 3) (**Figura 15**). Ya que la sonda 2 contiene motivos de unión a CTCF que están dentro del sitio de inicio de la transcripción de *asfl*, predichos por la base de datos flybase, decidimos no mutar esta región, pues podríamos afectar la unión de factores transcripcionales necesarios para la unión de la ARN polimerasa II.

Una ventaja de haber realizado los EMSAs es que nos permitió delimitar los motivos de unión que reconoce CTCF y al usar poblaciones celulares S2R+ que contienen las mutaciones del motivo 1 de unión a CTCF, sabemos que los cambios en la expresión de *asfl* se ven afectados por la reducción de la unión de CTCF al motivo 1 de unión. Podemos observar también que se reduce la unión de CTCF en la sonda mutante en comparación con la sonda 1 (**Figura 20**). El hecho de que no se elimine por completo esta unión puede deberse a tres factores: 1) a pesar de que se elimina parte del motivo 1, aún quedan los nucleótidos principales (“core”) para el reconocimiento de CTCF. 2) CTCF podría unirse con menor afinidad, ya que se ha comprobado que algunos de los dedos de zinc de este factor pueden unirse a regiones vecinas del motivo principal de reconocimiento (Nakahashi et al., 2013). 3) en la porción 5' de la sonda 1, parte del motivo 2 de unión a CTCF permanece y CTCF podría estarse uniendo a este, a pesar de que sean necesarias aproximadamente 50 pares de bases para que la unión de CTCF sea estable (Nakahashi et al., 2013). Además, para estar seguros de la pérdida de la unión de CTCF al motivo previamente descrito, aún es necesario realizar ensayos de inmunoprecipitación de la cromatina (ChIP), que confirmarán que dicha unión se pierde o reduce en un contexto cromatínico, es decir *in vivo*.

Por medio de la técnica CRISPR-Cas9 logramos la eliminación de una parte de uno de los tres motivos de unión a CTCF, que se encuentran río arriba del sitio de inicio de la

transcripción de *asfl*. Con la eliminación específica del motivo de unión 1 a CTCF, la diferencia de expresión de *asfl* corresponde a cambios en la unión de CTCF en esta región específica del genoma, lo cual no sucedería con el uso de ARNs de interferencia contra CTCF, ya que la unión de dicho factor a numerosos sitios en el genoma se vería alterada.

Cabe mencionar que estos resultados nos dan una primera aproximación para dilucidar la manera en que CTCF participa en esta región regulatoria, pues demostramos que su unión es necesaria para regular los niveles de ARNm de *asfl* y *L(3)76BDm* que se expresan. Sin embargo, es indispensable complementar con otros ensayos que confirmen el mecanismo por el cual actúa CTCF. Para los ensayos futuros es importante usar poblaciones celulares que contengan una sola de las mutaciones. Asimismo, deben realizarse ensayos que indiquen con que otros factores se asocia CTCF para lograr dicha regulación y las regiones regulatorias génicas que interaccionan.

Los efectos observados sobre el aumento en la expresión de *asfl* podrían explicarse por las diferentes propiedades de CTCF dentro de los mecanismos propuestos para sus funciones.

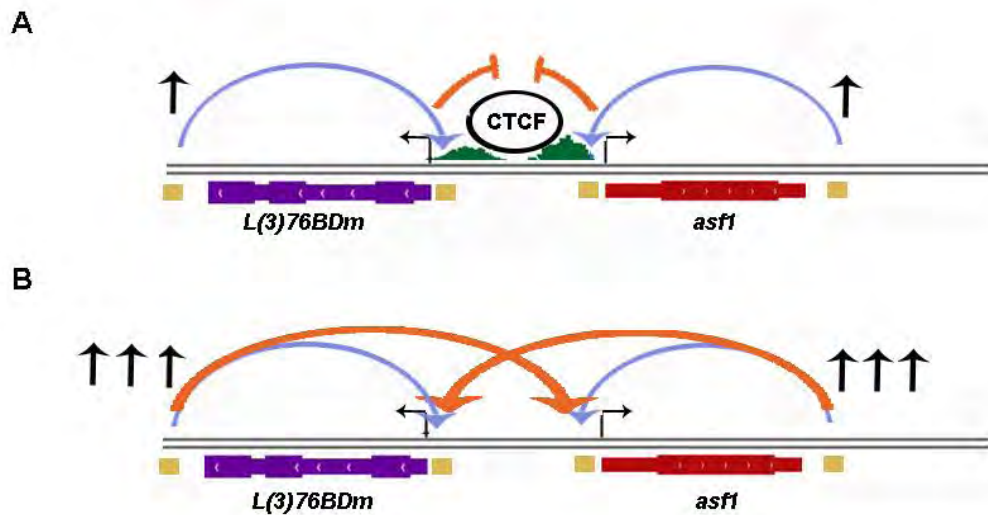
Ha sido ampliamente descrito que este factor nuclear está involucrado en la organización del genoma, gracias a su capacidad de mediar interacciones tridimensionales entre regiones de cromatina, a través de la formación de asas, en los diferentes organismos en los que se encuentra presente. CTCF puede permitir o evitar la interacción entre promotores y regiones regulatorias, concretamente con potenciadores y silenciadores, dependiendo de la naturaleza de los sitios que entran en contacto gracias a la acción de este factor. (Ohlsson et al., 2001; Ong y Corces, 2014). De hecho, se ha sugerido que las asas mediadas por CTCF demarcan dominios regulatorios de genes regulados por potenciadores (Oti et al., 2016).

Diversos estudios han demostrado que mediante la inversión o eliminación de los motivos de unión a CTCF, por medio de la técnica CRISPR-Cas9, ocurren cambios en las interacciones de promotores y potenciadores e incluso la organización tridimensional del genoma se modifica (Guo et al, 2015; Narendra et al, 2015). En la línea celular Kc162 de *D. melanogaster*, alrededor de un 70% de las regiones ocupadas por CTCF pertenecen a promotores y potenciadores y se ha propuesto además, que la presencia de CTCF afecta las interacciones y la arquitectura del genoma por medio de estos elementos regulatorios (Cubéñas-Potts et al, 2016). La ocupación de CTCF en el motivo 1 en la región intergénica 5'

de *asfl* podría estar promoviendo la interacción entre el promotor y un silenciador o podría prevenir la interacción del promotor con un potenciador. Entonces cuando CTCF no se encuentra en esta región génica, las interacciones entre regiones regulatorias se ven afectadas y permiten que aumente la transcripción del gen (**Figura 23**).

Asimismo, no solo se afecta la expresión de *asfl*. *l(3)76BDm*, que está divergente, a 379 pares de bases y río arriba de *asfl*, también incrementa su expresión ante la pérdida del motivo 1 de unión a CTCF. Este aumento en la expresión de ambos genes podría suceder debido que las interacciones entre elementos regulatorios se ve afectada para estos genes (**Figura 23**). Cabe mencionar, que los patrones de expresión de *asfl* y *l(3)76BDm* a lo largo del desarrollo son distintos. Mientras que la expresión de *asfl* varía a lo largo del desarrollo, la de *l(3)76BDm* se mantiene constante a relativamente bajos niveles (Emmert, 2014), lo que indica que deben ser diferencialmente regulados. Esto es importante, ya que CTCF tendría un papel en la regulación diferencial de genes divergentes, participando en el establecimiento de dominios de expresión distintos para cada uno de los genes, papel que en *D. melanogaster* aun no queda totalmente dilucidado (Smith et al., 2009; Xie et al., 2007). Esto no excluye la posibilidad de que el aumento en la expresión de *asfl* influya también sobre la transcripción de otros genes, ya que está involucrada en procesos de ensamblaje y desensamblaje de nucleosomas y en el reclutamiento a genes específicos, por lo que aún debe realizarse otros ensayos para evaluar si los niveles de la proteína de la chaperona de histonas varían, así como el efecto global sobre otros genes.

Es importante mencionar que este no es el único mecanismo por el que CTCF podría actuar en esta región génica. Otra opción es que la ocupación de la región intergénica 5' de *asfl* por CTCF interfiera con la transcripción, debido a que podría bloquear a la ARN polimerasa II y por lo tanto la transcripción. En vertebrados ha sido demostrado que CTCF puede interaccionar físicamente con la subunidad grande de la ARN polimerasa II y co-localiza con ella en diversos sitios en el genoma. La mayoría de las secuencias donde se encuentran ambas no están siendo transcritas, sugiriendo que CTCF colabora con el estancamiento de la ARN polimerasa II (Chernukhin et al., 2007). Además, en *Drosophila* CTCF co-localiza con el exosoma, maquinaria de degradación de ARN, principalmente en promotores de genes transcritos de manera divergente (Lim et al., 2013), lo cual podría estar sucediendo en este caso.



**Figura 23. Modelo de la regulación que ejerce CTCF sobre los genes *asf1* y *l(3)76BDm*.** Ya que CTCF participa en la formación de asas cromatínicas permitiendo o evitando el contacto entre distintos elementos regulatorios en el ADN, lo cual tiene consecuencias en la transcripción génica, se propone lo siguiente. **A)** Cuando CTCF se une río arriba del sitio de inicio de la transcripción de *asf1* (rojo), permite que tanto este gen, como el que se encuentra de manera divergente, *l(3)76BDm* (morado) se regulen de una manera diferencial. **B)** Cuando CTCF no se une a esta región, se formarían otras asas cromatínicas y por lo tanto, los contactos de los promotores de cada uno de estos genes, podrían estar contactando otros elementos regulatorios (cajas amarillas), como potenciadores, lo que promueve el incremento en la expresión tanto de *asf1* como de *l(3)76BDm* (flechas negras).

Finalmente, CTCF puede influir en el reclutamiento de otros complejos proteicos, lo que permite el mantenimiento de los dominios transcripcionales. Dentro de las proteínas que podrían estar regulando y manteniendo el dominio transcripcional de *asf1* están los miembros de los complejos Polycomb y Trithorax, que son responsables del mantenimiento de la represión o la activación transcripcional respectivamente. CTCF es necesaria para el mantenimiento de la heterocromatina de dominios silenciados por Polycomb (Pc) en *D. melanogaster*. Cuando disminuyen los niveles de CTCF la marca represiva H3K27me3 también se reduce en estos dominios (Van Bortle et al., 2012). Una opción es que CTCF sea requerida para mantener marcas de silenciamiento en la región regulatoria de *asf1* y cuando no se une, permita el aumento en la transcripción por la disminución de estas marcas o porque pueda reclutar a un complejo activador. Por esto, es importante definir con que otros factores transcripcionales y proteínas de unión a “insulators” interacciona la proteína CTCF que se encuentra en el motivo 1 estudiado en este proyecto. Definir estas interacciones nos dará más indicios de los mecanismos de regulación que está ejerciendo CTCF sobre *asf1*.



## 9. CONCLUSIONES

- En la región intergénica 5' de *asfl* hay tres motivos de unión a los que CTCF puede unirse *in vitro*.
- La eliminación de parte del motivo 1 de unión a CTCF, en la región intergénica 5' de *asfl*, reduce el reconocimiento y unión de CTCF a esta región.
- En células S2R+ con mutaciones que eliminan parte del motivo 1 de unión a CTCF se incrementa la expresión de *asfl* y *l(3)76BDM* a comparación de células silvestres. Este aumento en la expresión sugiere que la unión de CTCF en esta región es necesaria para la regulación transcripcional de ambos genes.

## 10. PERSPECTIVAS

Para dejar claros los mecanismos regulatorios de CTCF sobre *asfl* serán necesarios futuros ensayos tanto en líneas celulares como en individuos de *D. melanogaster* en diferentes etapas del desarrollo.

En primer lugar, deberá aislarse una clona celular S2R+, transfectada con el sistema CRISPR-Cas9 que contenga únicamente la eliminación de menor tamaño para confirmar los resultados de expresión obtenidos en este trabajo. También deben realizarse ChIPs, para asegurar que CTCF ha perdido su capacidad de unirse al motivo 1 en un contexto cromatínico, utilizando células control y la clona que contenga la eliminación de menor tamaño.

Con el fin de identificar las regiones regulatorias que están en contacto con CTCF, deberán hacerse ensayos de captura conformacional de cromosomas, tanto en condiciones control (células con el motivo 1 de unión a CTCF intacto), como con células que tengan eliminado parte del motivo 1 de unión a CTCF, para observar como su ausencia tiene un

efecto en la organización tridimensional de la cromatina en esta región y encontrar así elementos regulatorios que afectan la expresión de *asfl*.

Deben realizarse ChIPs para identificar la presencia de otros componentes de los complejos Polycomb y Trithorax que puedan estar involucrados en el establecimiento de este dominio transcripcional.

Un aspecto muy interesante a evaluar es el efecto de la unión de CTCF a este motivo en individuos en diferentes etapas del desarrollo de *D. melanogaster*. Con lo cual se conocerá si dicha unión es dependiente de la etapa del desarrollo, si la expresión de *asfl* se ve modificada y qué procesos celulares se ven afectados.

## 11. REFERENCIAS

- Artavanis-Tsakonas, S., Rand, M. D. y Lake, R. J. (1999).** Notch signaling: cell fate control and signal integration in development. *Science* **284**: 770-776.
- Avvakumov, N., Nourani, A. y Cote, J. (2011).** Histone chaperones: modulators of chromatin marks. *Molecular Cell* **41**: 502-514.
- Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Muller, M., Shanower, G., Schedl, P., Gyurkovics, H. y Karch, F. (2000).** The *Fab-8* boundary defines the distal limit of the bithorax complex *iab-7* domain and insulates *iab-7* from initiation elements and a PRE in the adjacent *iab-8* domain. *Development* **127**: 779-790.
- Bassett, A. R., Tibbit, C., Ponting, C. P. y Liu, J. L. (2013).** Mutagenesis and homologous recombination in *Drosophila* cell lines using CRISPR/Cas9. *Biology Open* **3**: 42-49.
- Baum, B. y Cherbas, L. (2008).** *Drosophila* cell lines as model systems and as an experimental tool. *Drosophila: Methods and Protocols*. C. Dahmann. Totowa, NJ, Humana Press: 391-424.
- Belozеров, V. E., Majumder, P., Shen, P. y Cai, H. N. (2003).** A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of *Drosophila*. *The EMBO Journal* **22**: 3113-3121.
- Bell, A. C., West, A. G. y Felsenfeld, G. (1999).** The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**: 387-396.
- Berger, S. L. (2007).** The complex language of chromatin regulation during transcription. *Nature* **447**: 407-412.
- Bonchuk, A., Maksimenko, O., Kyrchanova, O., Ivlieva, T., Mogila, V., Deshpande, G., Wolle, D., Schedl, P. y Georgiev, P. (2015).** Functional role of dimerization and CP190 interacting domains of CTCF protein in *Drosophila melanogaster*. *BMC Biology* **13**: 63.
- Brahmachari, V. y Jain, S. (2013).** Heterochromatin. In *Encyclopedia of Systems Biology*, pp 880-880. *Springer*.
- Burgess, R. J. y Zhang, Z. (2013).** Histone chaperones in nucleosome assembly and human disease. *Nature Structural & Molecular Biology* **20**: 14-22.

**Bushey, A. M., Ramos, E. y Corces, V. G. (2009).** Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes & Development* **23**: 1338-1350.

**Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W. H., Seepany, H., Gao, Z., Day, L. A. y Greenblatt, J. F. (2010).** The program for processing newly synthesized histones H3. 1 and H4. *Nature Structural & Molecular Biology* **17**: 1343-1351.

**Campos, E. I., Smits, A. H., Kang, Y. H., Landry, S., Escobar, T. M., Nayak, S., Ueberheide, B. M., Durocher, D., Vermeulen, M. y Hurwitz, J. (2015).** Analysis of the histone H3. 1 interactome: a suitable chaperone for the right event. *Molecular Cell* **60**: 697-709.

**Cave, J. W. (2011).** Selective repression of Notch pathway target gene transcription. *Developmental Biology* **360**: 123-131.

**Chernukhin, I., Shamsuddin, S., Kang, S. Y., Bergström, R., Kwon, Y. W., Yu, W., Whitehead, J., Mukhopadhyay, R., Docquier, F. y Farrar, D. (2007).** CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Molecular and Cellular Biology* **27**: 1631-1648.

**Cuartero, S., Fresán, U., Reina, O., Planet, E. y Espinàs M. L. (2014).** Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function. *The EMBO Journal* **33**: 637-647.

**Cubeñas-Potts, C., Rowley, M. J., Lyu, X., Li, G., Lei, E. P. y Corces, V. G. (2016).** Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Research*: gkw1114.

**Curators, F. (2008).** Assigning Gene Ontology (GO) terms by sequence similarity in FlyBase. [www.flybase.org](http://www.flybase.org).

**De Koning, L., Corpet, A., Haber, J. E. y Almouzni, G. (2007).** Histone chaperones: an escort network regulating histone traffic. *Nature Structural & Molecular Biology* **14**: 997-1007.

**De Wit, E., Vos, E. S., Holwerda, S. J., Valdes-Quezada, C., Versteegen, M. J., Teunissen, H., Splinter, E., Wijchers, P. J., Krijger, P. H. y De Laat, W. (2015).** CTCF binding polarity determines chromatin looping. *Molecular Cell* **60**: 676-684.

**Doudna, J. A. y Charpentier, E. (2014).** The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**: 1258096.

**Eitoku, M., Sato, L., Senda, T. y Horikoshi, M.** (2008). Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cellular and Molecular Life Sciences* **65**: 414-444.

**Fan, H. F., Liu, Z. N., Chow, S. Y., Lu, Y. H. y Li, H.** (2015). Histone chaperone-mediated nucleosome assembly process. *PloS One* **10**: e0115007.

**Felsenfeld, G. y Groudine, M.** (2003). Controlling the double helix. *Nature* **421**: 448-453.

**Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P. E., Collins, S. J. y Lobanekov, V. V.** (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology* **16**: 2802-2813.

**Gerasimova, T. I., Lei, E. P., Bushey, A. M. y Corces, V. G.** (2007). Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Molecular Cell* **28**: 761-772.

**Ghirlando, R. y Felsenfeld, G.** (2016). CTCF: making the right connections. *Genes & Development* **30**: 881-891.

**Gonzalez-Muñoz, E., Arboleda-Estudillo, Y., Out, H. H. y Cibelli, J. B.** (2014). Histone chaperone ASF1a is required for maintenance of pluripotency and cellular reprogramming. *Science* **345**: 822-825.

**Goodfellow, H., Krejčí, A., Moshkin, Y., Verrijzer, C. P., Karch, F. y Bray, S. J.** (2007). Gene-specific targeting of the histone chaperone Asf1 to mediate silencing. *Developmental Cell* **13**: 593-600.

**Green, E. M., Antczak, A. J., Bailey, A. O., Franco, A. A., Wu, K. J., Yates, J. R. y Kaufman, P. D.** (2005). Replication-independent histone deposition by the HIRA complex and Asf1. *Current Biology* **15**: 2044-2049.

**Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D. U., Jung, I., Wu, H., Zhai, Y. y Tang, Y.** (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**: 900-910.

**Herold, M., Bartkuhn, M. y Renkawitz, R.** (2012). CTCF: insights into insulator function during development. *Development* **139**: 1045-1057.

**Holohan, E. E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S. y White, R.** (2007). CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. *PLoS Genetics* **3**: e112.

**Huang, H., Strømme, C. B., Saredi, G., Hödl, M., Strandsby, A., González-Aguilera, C., Chen, S., Groth, A. y Patel, D. J.** (2015). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nature Structural & Molecular Biology* **22**: 618-626.

**Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P. y Siksnys, V.** (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biology* **10**: 841-851.

**Kellum, R. y Schedl, P.** (1992). A group of *scs* elements function as domain boundaries in an enhancer-blocking assay. *Molecular and Cellular Biology* **12**: 2424-2431.

**Kyrchanova, O. y Georgiev, P.** (2014). Chromatin insulators and long-distance interactions in *Drosophila*. *FEBS Letters* **588**: 8-14.

**Lim, S. J., Boyle, P. J., Chinen, M., Dale, R. K. y Lei, E. P.** (2013). Genome-wide localization of exosome components to active promoters and chromatin insulators in *Drosophila*. *Nucleic Acids Research*: gkt037.

**Liyanage, V. R., Jarmasz, J. S., Murugesan, N., Del Bigio, M. R., Rastegar, M. y Davie, J. R.** (2014). DNA modifications: function and applications in normal and disease states. *Biology* **3**: 670-723.

**Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M. y Laxova, R.** (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**: 1012-1025.

**Maksimenko, O., Bartkuhn, M., Stakhov, V., Herold, M., Zolotarev, N., Jox, T., Buxa, M. K., Kirsch, R., Bonchuk, A. y Fedotova, A.** (2015). Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin. *Genome Research* **25**: 89-99.

**Messiaen, S., Guiard, J., Aigueperse, C., Fliniaux, I., Tourpin, S., Barroca, V., Allemand, I., Fouchet, P., Livera, G. y Vernet, M.** (2016). Loss of the histone chaperone ASF1B reduces female reproductive capacity in mice. *Reproduction* **151**: 477-489.

**Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heintz, N., Bardenhagen, I., Leers, J., White, R. A., Renkawitz, R. y Saumweber, H.** (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *The EMBO Journal* **26**: 4203-4214.

**Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith S. T., Munhall, A., Grewe, B., Bartkuhn, M. y Arnold, R. (2005).** CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO reports* **6**: 165-170.

**Moore, J. M., Rabaia, N. A., Smith, L. E., Fagerlie, S., Gurley, K., Loukinov, D., Disteche, C. M., Collins, S. J., Kemp, C. J. y Lobanenko, V. V. (2012).** Loss of maternal CTCF is associated with peri-implantation lethality of *CTCF* null embryos. *PLoS One* **7**: e34915.

**Moshkin, Y. M., Armstrong, J. A., Maeda, R. K., Tamkun, J. W., Verrijzer, P., Kennison, J. A. y Karch, F. (2002).** Histone chaperone ASF1 cooperates with the Brahma chromatin-remodelling machinery. *Genes & Development* **16**: 2621-2626.

**Moshkin, Y. M., Kan, T. W., Goodfellow, H., Bezstarosti, K., Maeda, R. K., Pilyugin, M., Karch, F., Bray, S. J., Demmers, J. A. y Verrijzer, C. P. (2009).** Histone chaperones ASF1 and NAP1 differentially modulate removal of active histone marks by LID-RPD3 complexes during NOTCH silencing. *Molecular Cell* **35**: 782-793.

**Nakahashi, H., Kwon, K. R. K., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S. y Yamane, A. (2013).** A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Reports* **3**: 1678-1689.

**Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., Mazzoni, E. O. y Reinberg, D. (2015).** CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**: 1017-1021.

**Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M. y Senda, T. (2007).** Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* **446**: 338-341.

**Negre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., Henikoff, J. G., Feng, X., Ahmad, K., Russell, S., White, R. A., Stein, L., Henikoff, S., Kellis, M. y White, K. P. (2010).** A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genetics* **6**: e1000814.

**Ni, X., Zhang, Y. E., Nègre, N., Chen, S., Long, M. y White, K. P. (2012).** Adaptive evolution and the birth of CTCF binding sites in the *Drosophila* genome. *PLoS Biology* **10**: e1001420.

**Nikolaev, L., Akopov, S. B., Didych, D. y Sverdlov, E. (2009).** Vertebrate protein CTCF and its multiple roles in a large-scale regulation of genome activity. *Current Genomics* **10**: 294-302.

**Ohlsson, R., Renkawitz, R. y Lobanenkov, V. (2001).** CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in Genetics* **17**: 520-527.

**Ohtani, K. y Dimmeler, S. (2011).** Epigenetic regulation of cardiovascular differentiation. *Cardiovascular Research* **90**: 404-412.

**Ong, C. T. y Corces, V. G. (2014).** CTCF: an architectural protein bridging genome topology and function. *Nature Reviews Genetics* **15**: 234.

**Oti, M., Falck, J., Huynen, M. A. y Zhou, H. (2016).** CTCF-mediated chromatin loops enclose inducible gene regulatory domains. *BMC Genomics* **17**: 1.

**Parseghian, M. H. (2015).** What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery. *AIMS Biophysics* **2**: 724–772.

**Ramachandran, S. y Henikoff, S. (2015).** Replicating nucleosomes. *Science Advances* **1**: e1500587

**Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. y Zhang, F. (2013).** Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**: 2281-2308.

**Ransom, M., Dennehey, B. K. y Tyler, J. K. (2010).** Chaperoning histones during DNA replication and repair. *Cell* **140**: 183-195

**Recillas-Targa, F., Valadez-Graham, V. y Farrell, C. M. (2004).** Prospects and implications of using chromatin insulators in gene therapy and transgenesis. *Bioessays* **26**: 796-807.

**Saldaña-Meyer, R., González-Buendía, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F. y Reinberg, D. (2014).** CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes & Development* **28**: 723-734.

**San Juan, B. P., Andrade-Zapata, I. y Baonza, A. (2012).** The bHLH factors Dpn and members of the E(spl) complex mediate the function of Notch signalling regulating cell proliferation during wing disc development. *Biology Open* **1**: 667-676.

**Sanematsu, F., Takami, Y., Barman, H. K., Fukagawa, T., Ono, T., Shibahara, Kei-ichi. y Nakayama, T. (2006).** Asf1 is required for viability and chromatin assembly during DNA replication in vertebrate cells. *Journal of Biological Chemistry* **281**: 13817-13827.



**Schertel, C., Albarca, M., Rockel-Bauer, C., Kelley, N. W., Bischof, J., Hens, K., Van Nimwegen, E., Basler, K. y Deplancke, B.** (2015). A large-scale, in vivo transcription factor screen defines bivalent chromatin as a key property of regulatory factors mediating *Drosophila* wing development. *Genome Research* **25**: 514-523.

**Schneider, I.** (1972). Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *Development* **27**: 353-365.

**Schwartz, Y. B. y Cavalli, G.** (2017). Three-Dimensional Genome Organization and Function in *Drosophila*. *Genetics* **205**: 5-24.

**Schwartz, Y. B., Linder-Basso, D., Kharchenko, P. V., Tolstorukov, M. Y., Kim, M., Li, H. B., Gorchakov, A. A., Minoda, A., Shanower, G. y Alekseyenko, A. A.** (2012). Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Research* **22**: 2188-2198.

**Smith, S. T., Wickramasinghe, P., Olson, A., Loukinov, D., Lin, L., Deng, J., Xiong, Y., Rux, J., Sachidanandam, R., Sun, H.** (2009). Genome wide ChIP-chip analyses reveal important roles for CTCF in *Drosophila* genome organization. *Developmental Biology* **328**: 518-528.

**Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverly, T., Mozden, N., Misra, S. y Rubin, G. M.** (1999). The Berkeley *Drosophila* Genome Project gene disruption project: Single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153**: 135-177.

**Swaminathan, V., Kishore, A. H., Febitha, K. y Kundu, T. K.** (2005). Human histone chaperone nucleophosmin enhances acetylation-dependent chromatin transcription. *Molecular and Cellular Biology* **25**: 7534-7545.

**Torrano, V., Chernukhin, I., Docquier, F., D'Arcy, V., León, J., Klenova, E. y Delgado, M. D.** (2005). CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *Journal of Biological Chemistry* **280**: 28152-28161.

**Van Bortle, K., Ramos, E., Takenaka, N., Yang, J., Wahi, J. E. y Corces, V. G.** (2012). *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Research* **22**: 2176-2187.

**Vazquez, J. y Schedl, P.** (2000). Deletion of an insulator element by the mutation facet-strawberry in *Drosophila melanogaster*. *Genetics* **155**: 1297-1311.

**Weber, C. M. y Henikoff, S.** (2014). Histone variants: dynamic punctuation in transcription. *Genes & Development* **28**: 672-682.

**Xie, X., Mikkelsen, T. S., Gnirke, A., Lindblad-Toh, K., Kellis, M. y Lander, E. S.** (2007). Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proceedings of the National Academy of Sciences* **104**: 7145-7150.

**Yanagawa, S, Lee, J. S. e Ishimoto, A.** (1998). Identification and Characterization of a Novel Line of *Drosophila* Schneider S2 Cells That Respond to Wingless Signaling. *Journal of Biological Chemistry* **273**: 32353-32359.

**Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. y Madden, T. L.** (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**: 134.

**Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavoosidana, G., Mukhopadhyay, R. y Kanduri, C.** (2004). Poly (ADP-ribosyl) ation regulates CTCF-dependent chromatin insulation. *Nature Genetics* **36**: 1105-1110.