



**Universidad Nacional Autónoma de México**  
Doctorado en Ciencias Biomédicas  
Instituto de Fisiología Celular

**Distribución y función de CTCF en el desarrollo temprano del  
pez cebra**

Tesis que para optar por el grado de:  
**Doctor en Ciencias**

Presenta:  
**Francisco Javier Carmona Aldana**

Director de tesis:

**Dr. Félix Recillas Targa**  
Instituto de Fisiología Celular

Comité tutor:

**Dr. Jesús Chimal Monroy**  
Instituto de Investigaciones Biomédicas

**Dr. Alfredo Varela Echavarría**  
Instituto de Neurobiología

Ciudad Universitaria, Ciudad de México. Febrero de 2019.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## **Reconocimientos académicos**

Para la realización de este trabajo, los experimentos y la discusión de los resultados se realizaron en el Instituto de Fisiología Celular, en los laboratorios del Dr. Félix Recillas Targa y del Dr. Ernesto Maldonado Olvera; y en el Instituto de Investigaciones Biomédicas, en el laboratorio del Dr. Jesús Chimal Monroy. Además, los resultados se discutieron también con la Dra. Rosa Estela Navarro González, en el Instituto de Fisiología Celular.

La evaluación, seguimiento y retroalimentación académica del Comité Tutor fue especialmente importante, por parte de la Dra. Rosa Estela Navarro González, el Dr. Jesús Chimal Monroy y el Dr. Alfredo Varela Echavarría.

Se agradece la asistencia técnica constante de la Biol. Georgina Guerrero Avendaño, el Dr. Fernando Suaste Olmos y el Biol. José Luis Ramos Balderas, en el Instituto de Fisiología Celular, y a la M. C. Marcia Bustamante Zepeda y el Dr. Víctor Daniel Garzón Cortés en el Instituto de Investigaciones Biomédicas.

El apoyo con el mantenimiento y la alimentación de los peces fue muy importante por parte de la laboratorista Gianelli Cortés González, María del Rocío Camarillo Villegas y la Dra. Rebeca López.

Se reconoce el servicio de las distintas unidades de servicio del Instituto de Fisiología Celular: la Unidad de Imagenología, la Unidad de Biología Molecular, la

Unidad de Videoconferencia, la Unidad de Cómputo, la Biblioteca y el Taller de Mantenimiento.

Se agradece la atención y el seguimiento de mi situación académica de parte de la Dra. Soledad Funes Argüello y Sara Méndez Ibáñez en la Coordinación de Enseñanza del Instituto de Fisiología Celular, y de la Lic. Gabriela Valdés Silva, en el Departamento de Genética Molecular.

Los aportes y la discusión de la tesis fueron especialmente enriquecedores por parte de los miembros del Jurado de Examen: Dra. Aurea Orozco Rivas, Dra. Mahara Valverde Ramírez, Dra. Adriana Garay Arroyo, Dra. Denhi Schnabel Peraza y Dr. Jesús Chimal Monroy.

Se agradece al Conacyt, por la beca de posgrado recibida, con número 34758. Este trabajo fue posible gracias al financiamiento otorgado al grupo del Dr. Félix Recillas (Conacyt: 42653-Q, 128464 y 220503; DGAPA-PAPIIT, UNAM: IN209403, IN203811 e IN201114; y Fronteras de la Ciencia 2015: 290) y al grupo del Dr. Jesús Chimal (Conacyt: 168642; DGAPA-PAPIIT, UNAM: IN213314 e IN211117).

## Índice

1. Abreviaturas.....	10
2. Resumen.....	12
3. Abstract.....	14
4. Introducción .....	16
5. Antecedentes.....	17
5.1 Desarrollo embrionario .....	17
5.1.1 El pez cebra como modelo de estudio del desarrollo.....	21
5.1.1.1 Búsqueda de función de genes en el pez cebra .....	23
5.2 Aspectos de cromatina en el contexto del desarrollo .....	26
5.2.1 Conceptos básicos de la conformación de la cromatina .....	27
5.2.1.1 Modificaciones postraduccionales de las histonas .....	30
5.2.1.2 Metilación del ADN .....	32
5.2.1.3 Variantes de histonas .....	35
5.2.1.4 ARNs no codificantes largos.....	35
5.2.1.5 Configuración espacial de la cromatina .....	37
5.2.2 Configuraciones epigenéticas en el desarrollo embrionario.....	44
5.2.3 CTCF: una proteína multifuncional.....	47
5.2.3.1 Funciones moleculares de CTCF.....	50

5.2.3.2 CTCF y sus funciones en el desarrollo embrionario .....	53
6. Planteamiento del problema .....	56
7. Hipótesis .....	56
8. Objetivo general .....	57
8.1 Objetivos específicos .....	57
9. Material y métodos.....	58
9.1 Cepas utilizadas, mantenimiento de los peces y método de cruzas .....	58
9.2 Generación de un anticuerpo anti-CTCF de pez cebra .....	58
9.3 Tinción por inmunofluorescencia de embriones completos .....	59
9.4 Generación de mutaciones por el sistema CRISPR-Cas9 .....	61
9.5 Inyección de morfolidos y experimentos de rescate .....	62
9.6 RT-PCR.....	62
9.7 Extracción de proteína para Western blot .....	63
10. Resultados.....	64
10.1 Distribución de CTCF durante el desarrollo embrionario del pez cebra .....	64
10.2 Fenotipo derivado de mutaciones en CTCF por el sistema CRISPR-Cas9.....	68
10.3 La reducción del nivel de la proteína CTCF por inyección de morfolidos genera malformaciones en el embrión .....	73
10.4 El ARNm del gen <i>Ctcf</i> de humano rescata el fenotipo de los morfantes .....	79

10.5 CTCF se requiere para la correcta expresión de genes proapoptóticos .....	80
11. Discusión .....	83
12. Tablas suplementarias.....	92
13. Bibliografía.....	93

## Índice de figuras

Fig. 1 Organizador de Spemann-Mangold.....	19
Fig. 2 Principales características del pez cebra como modelo experimental.....	22
Fig. 3 Composición básica de la cromatina. ....	27
Fig. 4 Esquema global de la posición de los nucleosomas en los genes. ....	28
Fig. 5 Relación entre las modificaciones postraduccionales de las histonas y el ambiente cromatínico. ....	30
Fig. 6 Efecto de la metilación del ADN sobre el reclutamiento de factores a la cromatina .....	34
Fig. 7 Metodología de captura conformacional de la cromatina (3C). ....	38
Fig. 8 Distintos niveles de configuración espacial de la cromatina. ....	40
Fig. 9 Regulación diferencial de dos promotores por su contacto con un <i>enhancer</i> .....	42
Fig. 10 Modelo de extrusión para la formación de asas de cromatina.....	44
Fig. 11 Reprogramación epigenética durante el desarrollo embrionario.....	45
Fig. 12 Esquema de la proteína CTCF. ....	49
Fig. 13 Distintos tipos de contacto mediados por CTCF.....	53
Fig. 14 Localización de la porción de la proteína que se utilizó para generar el anticuerpo anti-CTCF. ....	64

Fig. 15 Prueba de especificidad del anticuerpo anti-CTCF por Western blot. ....	65
Fig. 16 Prueba de especificidad del anticuerpo anti-CTCF por inmunofluorescencia....	66
Fig. 17 Distribución de CTCF durante el desarrollo embrionario del pez cebra.....	67
Fig. 18 Fenotipo de embriones inyectados con el sistema CRISPR-Cas9. ....	69
Fig. 19 Ensayo de endonucleasa T7. ....	71
Fig. 20 Alelos secuenciados de los embriones inyectados con el sistema CRISPR-Cas9 .....	72
Fig. 21 Sitios blanco para los morfolinis anti-CTCF y letalidad con distintas concentraciones. ....	74
Fig. 22 Validación del efecto del mo-sp por RT-PCR. ....	75
Fig. 23 Niveles de CTCF en los morfantes. ....	76
Fig. 24 Fenotipo observado en los morfantes.....	77
Fig. 25 Rescate del fenotipo del morfolino aug.....	80
Fig. 26 Expresión de genes proapoptóticos en los morfantes. ....	81

## Índice de tablas

Tabla 1 Principales características de los morfolinós y el sistema CRISPR-Cas9 .....	25
Tabla 2 Identidad de la secuencia de aminoácidos de CTCF de diferentes especies respecto a la de humano (%).....	50
Tabla suplementaria 12.1 Fenotipo de embriones inyectados con el sistema CRISPR-Cas9 a las 12 hpf (en referencia a la figura 18).....	92
Tabla suplementaria 12.2 Fenotipo de embriones inyectados con el sistema CRISPR-Cas9 a las 24 hpf (en referencia a la figura 18).....	92
Tabla suplementaria 12.3 Letalidad con distintas concentraciones de morfolino (en referencia a la Fig. 21).....	92

## 1. Abreviaturas

**ADN**, ácido desoxirribonucleico

**ARN**, ácido ribonucleico

**ARNm**, ARN mensajero

**BLAST**, herramienta básica de búsqueda local de alineamientos (*basic local alignment search tool*)

**CpG**, dinucleótido CpG

**CTCF**, factor de unión a CCCTC (*CCCTC-binding factor*)

**DAPI**, 4',6-diamino-2-fenilindol (*4', 6-diamidino-phenylindole*)

**dpf**, días postfertilización

**FT**, factor de transcripción (*transcription factor*)

**HDAC**, desacetilasa de histonas (*histone deacetylase*)

**hpf**, horas postfertilización

**ICR**, región de control de impronta (*imprinted control region*)

**kDa**, kilodaltons

**lncRNA**, ARN largo no codificante (*long non-coding RNA*)

**MBD**, proteína con dominio de unión a CpG metilado (*methyl-CpG-binding domain*)

**MeCP2**, proteína de unión a CpG metilado 2 (*methyl-CpG-binding protein 2*)

**mL**, mililitro

**mo-aug**, morfolino dirigido al sitio de inicio de la traducción, AUG

**mo-sp**, morfolino dirigido a un sitio de *splicing*

**ng**, nanogramo

**NFR**, región libre de nucleosomas (*nucleosome-free region*)

**nm**, nanómetro

**pb**, pares de bases

**PBS**, amortiguador salino de fosfatos (*phosphate-buffered saline*)

**PCR**, reacción en cadena de la polimerasa (*polymerase chain reaction*)

**pg**, picogramo

**RNAPII**, ARN polimerasa II

**RT-qPCR**, PCR cuantitativa con transcripción inversa (*reverse transcription quantitative PCR*)

**SDS-PAGE**, electroforesis en gel de poliacrilamida con dodecilsulfato de sodio (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*)

**sgRNA**, ARN guía (*single-guide RNA*)

**TAD**, dominio topológicamente asociado (*topologically associating domain*)

**Xist**, transcrito específico del cromosoma X inactivo (*X-inactive specific transcript*)

**3C**, captura conformacional de la cromatina (*chromosome conformation capture*)

## 2. Resumen

Durante el desarrollo embrionario es importante la regulación de la actividad de los genes. En esta regulación participan algunos factores que influyen en la configuración de la cromatina, como el factor de unión a CCCTC (*CCCTC-binding factor*, CTCF), una proteína con 11 dedos de zinc. Su función se relaciona con el establecimiento de contactos entre elementos funcionales del genoma y la configuración tridimensional de la cromatina. Hay trabajos que han mostrado que CTCF es necesaria para el desarrollo embrionario. En particular, en el pez cebra (*Danio rerio*), se sabe que, si se bloquea la función de CTCF, se generan fenotipos relacionados con la formación de los músculos.

El presente trabajo tiene como objetivo determinar la distribución de la proteína CTCF en las distintas etapas del embrión del pez cebra. Además, se quiere buscar si existen otras funciones, además de la formación del músculo, que pudieran explicar por qué CTCF es un factor necesario para el desarrollo embrionario.

Para ello, se llevó a cabo la generación de un anticuerpo que pudiera detectar CTCF de pez cebra. Por inmunofluorescencia, se encontró la presencia de la proteína desde el embrión de 1 célula y durante distintas etapas en las primeras 30 horas del desarrollo embrionario. CTCF se distribuye en todas las regiones del embrión, tanto en la porción anterior como en la porción posterior.

Para buscar su función, se generaron mutaciones en el gen de *ctcf* con el sistema CRISPR-Cas9. Se extrajo ADN de embriones inyectados con este sistema, y se encontró un alelo deletéreo para CTCF en el 46% (6/13) de las clonas examinadas. A las 12 horas post-fertilización (hpf), se encontró que, en estos embriones, las

estructuras características de la porción anterior y posterior tienen una forma aberrante. Además, presentan letalidad a las 24 hpf.

Para estudiar a más detalle la función de CTCF, se inyectaron dos secuencias diferentes de morfolidos, una para bloquear la traducción y una para bloquear el *splicing*. En embriones inyectados con cualquiera de estos morfolidos, se demostró que el nivel de la proteína se reduce. A las 24 hpf, se encontró que estos embriones tienen una forma aberrante en estructuras tanto anteriores como posteriores: el tamaño de la cabeza, el ojo, los límites entre distintas regiones del sistema nervioso, los somitos y la longitud y la forma de la cola.

Para determinar si la reducción de CTCF afecta genes relacionados con la proliferación y la supervivencia celular (*p53* y *bbc3/PUMA*) se probó su expresión y se vio que la misma está aumentada en los embriones inyectados con el morfolido *ctcf* que generó malformaciones.

En conclusión, se muestra que la presencia de la proteína CTCF es constante durante el desarrollo embrionario del pez cebra, en distintas regiones del embrión, y que la interrupción de la función de CTCF genera estructuras mal formadas en cualquiera de los ejes corporales del embrión. Además, se demostró que la disminución de CTCF provoca la desregulación de *p53* y *bbc3/PUMA*, que son genes proapoptóticos, y que podrían estar relacionados con la aparición del fenotipo. Finalmente, con estos resultados se contribuye al estudio de CTCF en un contexto fisiológico, lo cual puede servir como base para detallar la función de esta proteína en diferentes tejidos del pez cebra durante el desarrollo embrionario.

### 3. Abstract

Embryonic development depends greatly on the tight regulation of gene activation. Some factors important for chromatin configuration are relevant in this context. One of these is the CCCTC-binding factor (CTCF), a protein characterized by its DNA-binding domain, formed by an arrangement of 11 zinc fingers in its central region. CTCF is required for the establishment of physical contact between functional elements in the genome and the three-dimensional organization of the chromatin. Also, it has been shown that CTCF is necessary during embryonic development. In particular, in the zebrafish, the reduction of the CTCF levels is related with muscle development phenotypes.

The objective of the present work was to study the CTCF protein distribution during the different stages of the zebrafish embryonic development. Also, we looked for developmental functions, other than the muscle formation, that could explain why CTCF is required during embryonic development.

To achieve this, an antibody specific for the zebrafish CTCF was generated. By immunofluorescence, it was found that this protein is present since the 1-cell stage embryo, and through all the stages of the first 30 hours of the embryonic development. CTCF is distributed along all the regions of the embryo, both in the anterior and the posterior of the embryo.

To look for the role of the *ctcf* gene, the CRISPR-Cas9 system was injected in embryos to generate mutations in this *locus*. The target *locus* was sequenced, and deleterious alleles were found in 46% (6/13) of the clones examined. At 12 hours post-

fertilization (hpf), the anterior and posterior regions of these embryos were aberrantly formed. Further, they were lethal at 24 hpf.

To detail this, two different sequences of *ctcf* morpholinos were injected in one- or two-cell stage embryos. One morpholino was designed to block the transcript translation, and the other to interfere with its splicing. The CTCF protein level was reduced with either of these morpholinos. At 24 hpf, these embryos showed an aberrant form in structures both, in the anterior and the posterior zones: the head size, the eye, the borders delimiting the different regions of the nervous system, the somites and the length and size of the tail.

To determine if the CTCF level reduction affects some genes related with proliferation and cell survival (*p53* and *bbc3/PUMA*), their expression was tested. It was found that these genes are overexpressed in the *ctcf* morpholino-injected embryos that showed the phenotype.

In conclusion, it was found that the CTCF protein is present constantly during the embryonic development of the zebrafish, in the different regions of the embryo. A CTCF function interruption generates structures with an aberrant form in any of the embryonic body axes. A CTCF level reduction affects the regulation of the proapoptotic genes *p53* and *bbc3/PUMA*, which could be related with the phenotype. These results contribute to the understanding of the physiological function of CTCF, and might be the basis for future studies exploring in detail the role of this factor in different tissues during the zebrafish embryonic development.

## 4. Introducción

Cuando se estudia a los organismos en el contexto del desarrollo embrionario, se busca entender cómo es que, a partir de una célula original (el cigoto), se generan los demás tipos celulares que conformarán a los tejidos y órganos del individuo maduro. Además, se busca entender cómo se establece la identidad de cada una de estas células, la posición que ocuparán en el espacio y cómo se organizan en tejidos y órganos funcionales. Esto ocurre en un patrón reproducible (dentro de cierto rango) entre individuos de una misma especie (Irie y Kuratani, 2014; Levin *et al.*, 2016; White *et al.*, 2018).

El desarrollo embrionario se ha estudiado a nivel anatómico, explicando dónde y cuándo aparecen las distintas estructuras y órganos, y a nivel fisiológico, explicando cómo estas unidades anatómicas se conectan y se relacionan entre sí para servir de base a las funciones del organismo (Hamburger y Hamilton, 1951; Kimmel *et al.*, 1995; Campos-Ortega y Hartenstein, 1997). Pero también se ha estudiado desde una perspectiva celular y molecular, tratando de entender qué funciones dentro de las células guían los procesos del desarrollo (Nüsslein-Volhard y Wieschaus, 1980; Haffter *et al.*, 1996).

En este contexto se busca conocer qué ocurre en los diferentes compartimentos subcelulares y sus moléculas asociadas, y en las células, tanto individualmente como en su relación con otras células, para poder entender los patrones que les dan a los organismos sus características anatómicas y funcionales. A este nivel se estudia el

núcleo celular, tomando en cuenta la estructura de la cromatina y la regulación de la actividad genética (Borsos y Torres-Padilla, 2016; Vaquerizas y Torres-Padilla, 2016; Wagner *et al.*, 2018).

Hay varios componentes que actúan en el núcleo, entre ellos se encuentran las proteínas que regulan la configuración de la cromatina y que modulan la interacción entre elementos de regulación con un impacto directo en la actividad de los genes (Herold *et al.*, 2012; Perino y Veenstra, 2016; Arzate-Mejía *et al.*, 2018). Por esto, es interesante estudiar a estos factores y su influencia sobre el desarrollo embrionario y la fisiología de un organismo. En el caso de este trabajo, nos interesa entender cuál es la función de uno de estos factores en particular, el factor nuclear multifuncional CTCF, en el contexto del desarrollo embrionario.

## **5. Antecedentes**

### **5.1 Desarrollo embrionario**

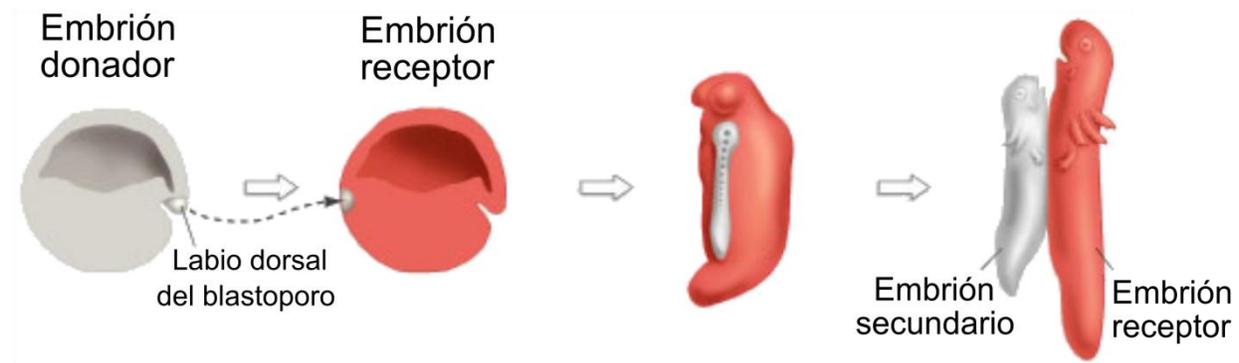
Uno de los aspectos de mayor interés en el estudio de los organismos, pretende entender la generación de nuevos individuos de una especie, preservando en cada uno de ellos los rasgos generales de la especie. La curiosidad por entender cómo se origina la vida de un organismo, llevó a generar ciertas hipótesis a través de la historia. Entre las primeras, en el siglo XVII, dos hipótesis fueron confrontadas (Gilbert, 2010):

1) **La epigénesis**, según la cual las diferentes estructuras de un embrión se forman *de novo*, al combinar un constituyente materno y uno paterno. Esta hipótesis se basaba en las ideas de William Harvey, siendo Aristóteles quien había propuesto una idea parecida.

2) **El preformismo**, basado en las ideas de Marcello Malpighi, según la cual las estructuras del embrión ya estaban formadas y funcionalmente capacitadas, y contenidas en una versión miniatura, dentro de los espermatozoides o los ovocitos. Después de la fertilización estas estructuras crecían, generando al organismo maduro.

La curiosidad y la importancia por seguir conociendo el desarrollo temprano de un organismo impulsaron más estudios y aproximaciones experimentales, que derivaron en la propuesta de novedosas teorías. Mediante observaciones detalladas se comprobó que las células germinales (los óvulos y los espermatozoides) no contenían organismos preformados sino, más bien, que la fusión de las células germinales (en especies con reproducción sexual) es necesaria para formar a una célula original, es decir al cigoto. A partir del cigoto, comienzan los programas de formación del patrón de un individuo, con sus características anatómicas y funcionales específicas. Con esto, se conciliaron algunos puntos de la Epigénesis (el organismo se forma *de novo*) con algunos del Preformismo (los principios reguladores del desarrollo están ya contenidos en el material del cual el embrión se formará). Sobre estos conocimientos básicos, enfocados en observaciones anatómicas y fisiológicas, se funda la *embriología* (Gilbert, 2010).

En los inicios de la embriología, se buscó documentar de una manera descriptiva cómo se forma anatómicamente un embrión de pollo. Esto dio lugar a descripciones muy detalladas del desarrollo del mismo (Gilbert, 2010).



**Fig. 1 Organizador de Spemann-Mangold.** Plano sagital de un embrión de tritón, en el que se representa la localización del organizador de Spemann-Mangold en el labio dorsal del blastoporo. Además, se esquematiza el experimento que realizaron Hans Spemann y Hilde Mangold en 1924. Trasplantaron células del labio dorsal del blastoporo de un embrión donador a la porción opuesta al blastoporo en un embrión receptor. Se formó un eje corporal extra, que derivó en el desarrollo de un embrión secundario. Modificado de Mani, 2013.

A la postre, los alcances de estos estudios fueron más allá de lo descriptivo, pues permitieron encontrar (entre varios otros hallazgos) centros anatómicos que guían procesos morfológicos importantes, a partir de los cuales se pudieron inferir algunos principios reguladores del desarrollo. Un ejemplo clásico es la descripción del organizador de Spemann-Mangold en anfibios, que es un grupo de células que se localiza en el labio dorsal del blastoporo, y que se requiere para la formación de los ejes corporales del embrión (Fig. 1). Al manipular experimentalmente este organizador, puede provocar la alteración de los patrones corporales, o inclusive la formación de más de un eje corporal, que puede progresar hacia un embrión secundario (Anderson y Stern, 2016). Esta evidencia sirvió como base para plantear principios conceptuales

muy relevantes acerca de los mecanismos del desarrollo, como la *inducción* (la capacidad de un grupo de células para inducir un cambio en el grado o en el destino de diferenciación de otro grupo de células), y el *establecimiento de los patrones del desarrollo* (la disposición organizada del conjunto de estructuras que conforman a un tejido, un órgano o un organismo completo) (Gurdon, 1987; Anderson y Stern, 2016).

Después, buscando las moléculas que participan en los mecanismos del desarrollo se reflexionó acerca de las teorías de la herencia de rasgos a partir de los trabajos de Gregor Mendel en el siglo XIX (Ellis *et al.*, 2011). Se buscó entender cómo se transmiten los rasgos de padres hacia hijos. Una de las hipótesis acerca de cuál componente de las células era el sustrato para la herencia de rasgos apuntaba hacia los cromosomas, lo cual fue demostrado por Thomas Hunt Morgan a principios del siglo XX, a partir de evidencias que encontró haciendo experimentos con la mosca de la fruta (*Drosophila melanogaster*). Sus trabajos establecieron las bases para formar el concepto de *gen* y décadas más tarde, usando también a la mosca, se diseñó un abordaje experimental en el que se generaron mutaciones al azar (intentando llevar a saturación estas mutaciones en el genoma) (Maienschein, 2016). En las mutantes resultantes se buscaron malformaciones durante el desarrollo embrionario, y a partir de estos cambios, se mapearon los genes mutados (Nüsslein-Volhard y Wieschaus, 1980). De esta serie de estudios surge una relación entre las características anatómicas del embrión y las moléculas que guían su desarrollo, con lo cual el concepto de *embriología* deja paso al de *biología del desarrollo* (Gilbert, 2010; Bilder, 2016).

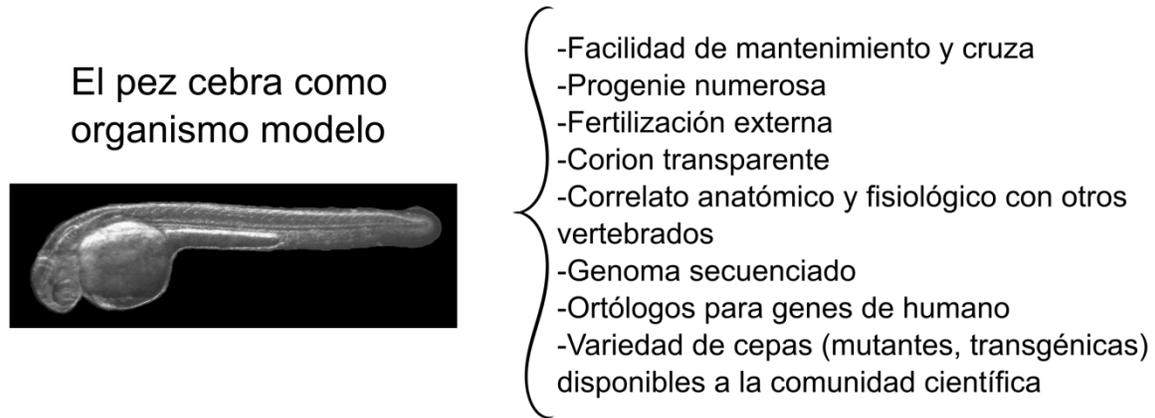
Dentro de este nuevo campo de estudio fue necesario implementar el uso de algunos organismos como referentes estándar para el estudio de la biología del

desarrollo, a los cuales se les conoce como *organismos modelo* (Gilbert, 2010). Los más comunes son la planta *Arabidopsis thaliana*, el gusano *Caenorhabditis elegans*, la mosca de la fruta (*Drosophila melanogaster*), el pez cebra (*Danio rerio*), la rana africana (*Xenopus laevis*), el pollo (*Gallus gallus*) y el ratón (*Mus musculus*), entre otros. Estos organismos han permitido describir procesos particulares, que hacen muy interesante conocer cómo la evolución ha generado distintas vías para la formación de patrones corporales y fisiológicos en el desarrollo embrionario, que derivan en la diversidad de formas y funciones entre especies. Y al mismo tiempo, se han establecido generalizaciones que permiten explicar los mecanismos que se comparten durante el desarrollo embrionario entre distintas especies, y que hacen posible que, estudiando a un organismo en particular, se pueda inferir lo que ocurre en otros organismos, en los cuales los abordajes experimentales son más complicados o inclusive imposibles (Haffter *et al.*, 1996; Johnson *et al.*, 2011; Irie y Kuratani, 2014; Bilder, 2016; Levin *et al.*, 2016).

### **5.1.1 El pez cebra como modelo de estudio del desarrollo**

El pez cebra es uno de los organismos modelo con características atractivas para el trabajo experimental en biología del desarrollo. Este pez tiene su origen en el sureste de Asia, en el subcontinente indio (Trevarrow, 2004; Holtzman *et al.*, 2016). Entre los primeros grupos de investigación en utilizarlo para estudiar la función de los

genes durante el desarrollo embrionario fue el de George Streisinger en la década de los 80s (Patton y Zon, 2001; Holtzman *et al.*, 2016).



**Fig. 2** Principales características del pez cebra como modelo experimental.

En la actualidad, varios laboratorios en el mundo utilizan de manera rutinaria al pez cebra como modelo de estudio por sus distintas características (Fig. 2) (Kimmel *et al.*, 1995; Ingham, 1997; Patton y Zon, 2001; Trevarrow, 2004; Schier, 2013; Holtzman *et al.*, 2016; Lawrence, 2016): es un animal que puede lograr cruzarse fácilmente en cautiverio, del cual pueden obtenerse decenas de embriones; la fertilización es externa; los embriones son transparentes en las primeras etapas del desarrollo y están cubiertos por el corion, que también es transparente, lo cual facilita su observación; se pueden diluir sustancias en el agua para ensayar su efecto en el desarrollo embrionario; varias de sus características anatómicas y fisiológicas tienen un correlato con otros vertebrados; su genoma está secuenciado; el 70% de los genes humanos tienen un ortólogo en el pez; hay ceparios que se encargan de mantener líneas mutantes o transgénicas y hacerlos disponibles a la comunidad científica; dado que los peces son

animales que también se cultivan como mascotas, es fácil adquirir productos para su mantenimiento y alimentación. En resumen, el pez cebra es un organismo experimental con varias características que facilitan estudiar el desarrollo embrionario.

#### **5.1.1.1 Búsqueda de función de genes en el pez cebra**

En el pez cebra se han implementado abordajes que permiten la búsqueda experimental de la función de los genes, interrumpiendo su función y buscando el fenotipo asociado (Patton y Zon, 2001; Holtzman *et al.*, 2016). Entre las primeras metodologías que se utilizaron está la mutagénesis química, utilizando etilnitrosourea (ENU) (Haffter *et al.*, 1996; Ingham, 1997). Con este reactivo se generan mutaciones al azar y, de esta forma, se aíslan mutantes que se pueden mapear y caracterizar. Esto es útil cuando se realizan tamizajes genéticos que permiten lograr el conocimiento de un grupo de genes que funcionan en un contexto fisiológico particular. Algunas líneas mutantes que se han generado en estos ensayos, se han hecho disponibles al resto de la comunidad científica a través de ceparios especializados, los cuales se encargan de mantener distintas líneas genéticas (ya sea mutantes o transgénicas) (Fig. 2) (Holtzman *et al.*, 2016).

Sin embargo, no hay cepas mutantes generadas para todos los genes. Esto es un problema cuando se requiere hacer experimentos de genética inversa, en el cual se necesita interrumpir la función de un gen en particular (lo cual no se puede lograr con la mutagénesis química). La alternativa que se ha implementado en el pez cebra es el uso

de morfolinos (Bedell *et al.*, 2011; Holtzman *et al.*, 2016). Los morfolinos son oligonucleótidos sintéticos que se diseñan para que, por complementariedad de bases, interfieran ya sea con la traducción o con el procesamiento (*splicing*) de un transcrito.

El uso de los morfolinos se popularizó debido a que se puede dirigir el uso de los mismos hacia el gen que se desea modificar. Cabe mencionar que en los primeros trabajos en los que los usaron se vio que el fenotipo de los individuos inyectados con morfolinos (a los cuales se les denomina *morfantes*) correlacionaba con el de los mutantes para esos genes. Sin embargo, los morfolinos pueden ser inespecíficos, por lo cual se han establecido algunas prácticas de control para asegurarse de su especificidad (Eisen y Smith, 2008; Stainier *et al.*, 2017):

- 1) Usar dos morfolinos diferentes. Generalmente uno que bloquea la traducción y otro que interfiere con el procesamiento del transcrito. Y se comparan los fenotipos de uno y otro.
- 2) Realizar experimentos de rescate. Consisten en probar si el fenotipo se revierte con la inyección en conjunto del morfolino y el ARN mensajero (ARNm) del gen de prueba, sintetizado *in vitro*. Este ARNm debe tener un cambio de secuencia en el sitio de unión del morfolino.

Recientemente, la generación de cepas mutantes se ha facilitado por la implementación en el pez de la metodología de edición genética conocido como CRISPR-Cas9 (Hwang *et al.*, 2013; Jao *et al.*, 2013; Gagnon *et al.*, 2014; Holtzman *et al.*, 2016). Con este sistema se puede dirigir una mutación hacia un gen con un alto grado de especificidad y eficiencia, y esto ha permitido que se pueda comparar el

fenotipo de la interrupción de un gen por los dos sistemas: la mutación generada por el sistema CRISPR-Cas9, y el fenotipo de los morfantes (Tabla 1) (Blum *et al.*, 2015; Stainier *et al.*, 2017).

**Tabla 1 Principales características de los morfolinós y el sistema CRISPR-Cas9**

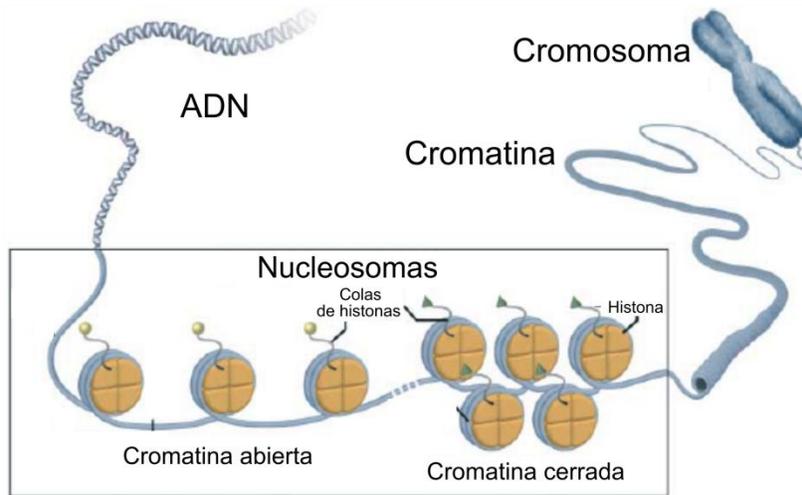
<b>Morfolinós</b>	<b>Sistema CRISPR-Cas9</b>
<ul style="list-style-type: none"> <li>• Se inyecta en embriones de 1 o 2 células.</li> <li>• La cantidad inyectada de morfolino se reparte en las sucesivas divisiones celulares. Su efecto es transitorio, afectando las primeras etapas del desarrollo embrionario.</li> <li>• Pocas compañías los producen. Se requiere un morfolino diferente para cada secuencia que se busca interferir.</li> <li>• Más costoso que el sistema CRISPR-Cas9.</li> <li>• Mayor probabilidad de inespecificidad. Requiere de varios controles.</li> <li>• Bloquea la traducción o interfiere sitios de <i>splicing</i>.</li> </ul>	<ul style="list-style-type: none"> <li>• Se inyecta en embriones de 1 o 2 células.</li> <li>• Los embriones inyectados serán mosaicos. Se necesita identificar alelos y establecer cepas.</li> <li>• Los plásmidos se adquieren una sola vez y se modifican fácilmente para usarlos después con cualquier secuencia.</li> <li>• Menos costoso que la síntesis de morfolinós.</li> <li>• Menor probabilidad de inespecificidad.</li> <li>• Más versátil. Se puede dirigir a diferentes sitios para generar mutaciones o inclusive puede introducir secuencias.</li> </ul>

Este conjunto de técnicas, cuyo uso se ha hecho rutinario en los laboratorios, ha permitido buscar la función de diferentes genes en varios contextos fisiológicos. Así se han encontrado genes importantes en el desarrollo embrionario del pez cebra, por ejemplo genes que codifican para componentes de vías de señalización, factores de transcripción e inclusive componentes de la cromatina (Haffter *et al.*, 1996; Ingham,

1997; Delgado-Olguín *et al.*, 2011; Tena *et al.*, 2011; Marsman *et al.*, 2014; Kent *et al.*, 2016; Meier *et al.*, 2018).

## 5.2 Aspectos de cromatina en el contexto del desarrollo

El ADN dentro del núcleo interactúa con otras moléculas, como proteínas y ARN, en un complejo que se conoce como *cromatina* (Fig 3). Las interacciones de los diferentes tipos de moléculas de la cromatina, entre el ADN, histonas, factores de transcripción, el complejo de la transcripción, ARNs reguladores, proteínas estructurales, etc., influyen sobre la dinámica de los distintos procesos que ocurren dentro del núcleo, como la transcripción, la replicación y la reparación del ADN (Portela y Esteller, 2010). Esto ocurre por la modulación de la cromatina a varios niveles, desde la accesibilidad de distintos elementos genéticos a proteínas reguladoras, la atracción de complejos enzimáticos que modulan la configuración de la cromatina, la interacción en el espacio de distintos elementos en el genoma, entre otros (Zhou *et al.*, 2011). Esto tiene repercusiones en procesos que fisiológicamente son finamente regulados en tiempo y espacio, como lo es el desarrollo embrionario. Recientemente se han empezado a conocer los cambios en la configuración de la cromatina que ocurren a través del tiempo y que participan en el desarrollo de un organismo (Eissenberg y Shilatifard, 2010; Zhou *et al.*, 2011; Bonn *et al.*, 2012; Burton y Torres-Padilla, 2014; Wang *et al.*, 2014; Meier y Recillas-Targa, 2017; Arzate-Mejía *et al.*, 2018).



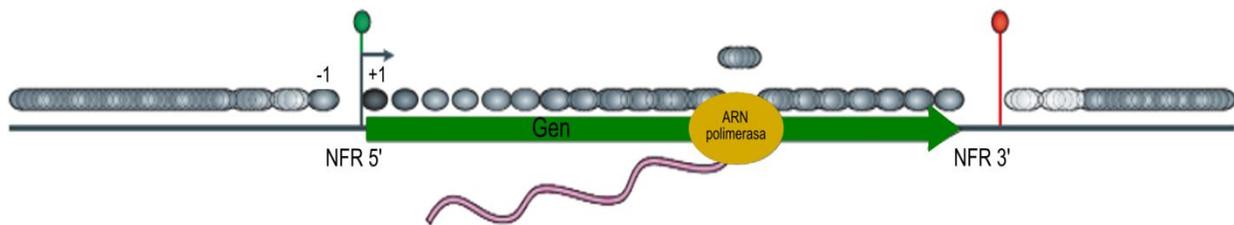
**Fig. 3 Composición básica de la cromatina.** El ADN en el núcleo se organiza en complejos formados por proteínas histonas. Estos complejos se conocen como *nucleosomas*, y constituyen la unidad más básica de organización de la cromatina. Los extremos en la región amino-terminal, que se conocen como *colas de las histonas*, son susceptibles a modificaciones químicas. Estas modificaciones provocan en la cromatina contextos abiertos (accesibles a factores nucleares) o cerrados (menos accesibles). Distintos componentes generan varios

niveles de compactación de la cromatina, encontrándose el contexto más compacto en los cromosomas mitóticos. Modificado de Pieterman *et al.*, 2014

### 5.2.1 Conceptos básicos de la conformación de la cromatina

Respecto a la conformación de la cromatina, la unidad básica de interacción entre ADN y proteínas es el *nucleosoma* (Fig. 3). Un nucleosoma se forma por un octámero de histonas (histonas H2A, H2B, H3 y H4) y un segmento de 147 pb de ADN que se enreda alrededor del complejo de histonas. El ADN que se encuentra entre dos nucleosomas se conoce como ADN de enlace (*linker DNA*) y su longitud promedio varía entre especies, siendo la más común entre 18 a 38 pb. Esta longitud puede cambiar por el desplazamiento de nucleosomas por complejos enzimáticos (los complejos remodeladores de histonas), lo cual influye sobre la densidad de nucleosomas en una sección del genoma y correlaciona con la accesibilidad de distintos elementos a factores de transcripción y complejos enzimáticos (Shilatifard, 2006; Jiang y Pugh, 2009) (Fig. 4).

Además, la histona H1 se une al ADN de enlace e influye sobre el espacio accesible de unión a factores reguladores de los elementos genéticos. Pero también provoca que la disposición de los nucleosomas como en un collar de perlas gire sobre sí misma para formar una fibra de 30 nm (también conocida como *solenoides*), que es una estructura más compacta y transcripcionalmente inactiva (Jiang y Pugh, 2009; Di Croce y Helin, 2013).



**Fig. 4 Esquema global de la posición de los nucleosomas en los genes.** Los nucleosomas se representan como óvalos grises. En los sitios que se encuentran alrededor de los genes, hay una densidad alta de nucleosomas y su posición es variable. En cambio, en los genes hay una densidad menor. Tanto en el sitio de inicio de la transcripción (NFR 5') como en el sitio de término de la transcripción (NFR 3'), existen regiones libres de nucleosomas (*nucleosome-free regions*, NFR). Alrededor de los NFR, hay nucleosomas con una posición muy fija (nucleosomas -1 y +1). Hacia el interior del cuerpo del gen, los nucleosomas tienen una posición más variable. Modificado de Jiang y Pugh, 2009.

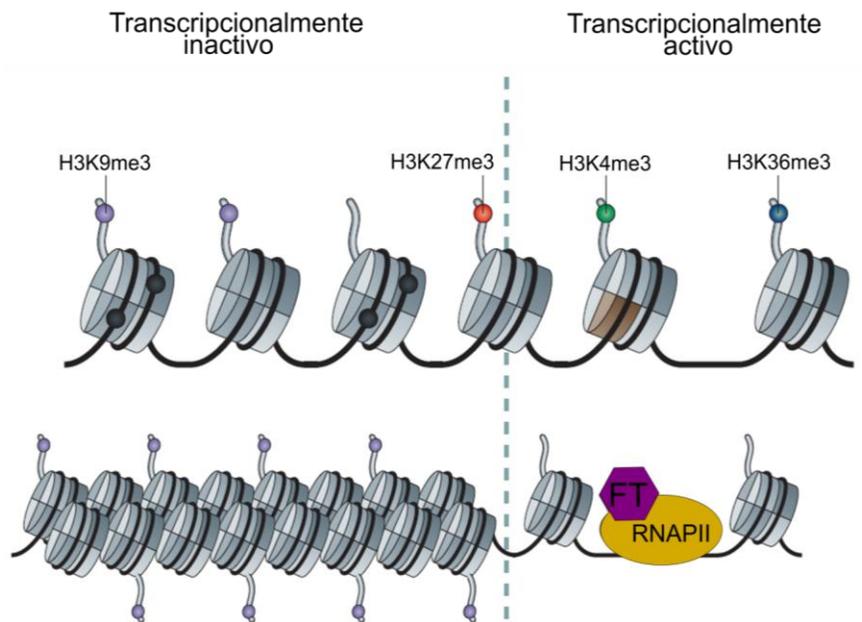
La densidad y el posicionamiento de los nucleosomas correlacionan con la accesibilidad y la actividad de los elementos genéticos (Fig. 4). Por ejemplo, alrededor de los promotores de los genes, hay una región libre de nucleosomas (*nucleosome-free region*, NFR, Fig. 4). Los nucleosomas en los bordes de esta región tienen una posición más bien fija, a comparación de aquellos que se localizan en el cuerpo del gen, que tienen cierta movilidad (Jiang y Pugh, 2009). Esto correlaciona con el posicionamiento del complejo iniciador de la transcripción en los promotores, mientras que en el cuerpo

del gen se presenta una deposición transitoria de los nucleosomas que correlaciona con la elongación por la maquinaria de la transcripción (fig. 4) (Lai y Pugh, 2017).

Cuando se estudia a los elementos genéticos, clásicamente se les define por la secuencia de nucleótidos (como los promotores, los sitios de unión a factores de transcripción y los *enhancers*). Pero en un nivel de estudio más reciente, hay que atender el contexto de cromatina en el que se encuentran estos elementos. Especialmente en las modificaciones químicas que se presentan en los distintos componentes de la cromatina y en su configuración en el espacio: modificaciones postraduccionales de las histonas, grupos químicos añadidos a los nucleótidos en el ADN (metilación, por ejemplo) o regiones puntuales de cromatina que en un contexto fisiológico determinado interaccionan entre ellas, pero en un contexto diferente interaccionan con otras regiones (Portela y Esteller, 2010; Zhou *et al.*, 2011; Cavalli y Misteli, 2013; Dekker y Mirny, 2016; García-González *et al.*, 2016; Meier y Recillas-Targa, 2017). En conjunto, se habla de *epigenética* al referirse a la regulación de los elementos genéticos y de los procesos que ocurren en el núcleo celular, por mecanismos bioquímicos que impactan en los distintos factores que constituyen a la cromatina, y que son diferentes a las propiedades reguladoras que emergen en los distintos elementos genéticos basados en la secuencia de nucleótidos (Shilatifard, 2006; Gräff y Mansuy, 2008). En el contexto de esta regulación, se han estudiado diferentes procesos epigenéticos, algunos de los cuales se describirán en las siguientes secciones.

### 5.2.1.1 Modificaciones postraduccionales de las histonas

Las histonas que componen a los nucleosomas, son proteínas con una conformación globular. Sus extremos, el amino- y el carboxilo-terminal tienen una estructura laxa, más bien lineal. Esto las hace fácilmente accesibles a los complejos enzimáticos que colocan, eliminan o cambian grupos químicos en estas regiones de las histonas (Fig. 5). Estos grupos químicos influyen en la interacción entre los mismos nucleosomas, o entre nucleosomas y otros componentes de la cromatina, lo cual, de manera muy general, se resuelve a nivel fisicoquímico y tiene como consecuencia alguno de estos dos estados (Fig. 5) (Henikoff y Shilatifard, 2011):



**Fig. 5 Relación entre las modificaciones postraduccionales de las histonas y el ambiente cromatínico.** Algunas marcas de histonas relacionadas con regiones de cromatina compacta y poco accesible a factores de transcripción (FT) son la trimetilación de la lisina 9 de la histona H3 (H3K9me3) y la trimetilación de la lisina 27 de la histona H3 (H3K27me3). Mientras que la trimetilación de la lisina 4 de la histona H3 (H3K4me3) y la trimetilación de la lisina 36 de la histona H3 (H3K36me3) se relacionan con un ambiente de cromatina abierta, accesible a FTs y a la maquinaria de la transcripción (representado aquí por la polimerasa de ARN II, RNAPII). Modificado de Zhou *et al.*, 2011.

1) Regiones en las que las propiedades fisicoquímicas de los componentes de la cromatina no ejercen una atracción estrecha. Forman contextos cromatínicos relajados, accesibles a factores reguladores. Se relacionan, por ejemplo, con regiones de transcripción activa, *enhancers* activos o regiones que (durante la fase de replicación) se replican primero.

2) Regiones en las que los componentes de la cromatina forman contactos muy estrechos. Forman contextos cromatínicos cerrados y poco accesibles a factores reguladores. Se relacionan, por ejemplo, con regiones en donde la transcripción está silenciada, *enhancers* inactivos o regiones que se replican tardíamente.

Las modificaciones que ocurren en el extremo amino-terminal de las histonas (extremos a los que comúnmente se hace referencia como *las colas de las histonas*) son metilación, acetilación, fosforilación, poli(ADP-ribosil)ación, ubiquitinación y SUMOilación, principalmente (Strahl y Allis, 2000; Shilatifard, 2006; Di Croce y Helin, 2013). El efecto de estos grupos químicos sobre los procesos nucleares depende del aminoácido específico de la histona en el que ha sido incorporado. Esto se conoce tanto por estudios de regiones particulares del genoma, como por análisis metagenómicos. La relación entre las modificaciones químicas que ocurren en aminoácidos específicos (que se conocen como *marcas de histonas*) con un proceso nuclear particular constituye el *código de histonas* (Strahl y Allis, 2000; Callinan y Feinberg, 2006; Ruthenburg *et al.*, 2007). Por ejemplo, en este código, la trimetilación de la lisina 4 de la histona H3 (H3K4me3) se relaciona con promotores de genes con transcripción activa (Ruthenburg *et al.*, 2007), y la trimetilación de la lisina 27 de la

histona H3 (H3K27me3) con regiones del genoma que son transcripcionalmente inactivas (Fig. 5) (Roh *et al.*, 2006).

Las modificaciones postraduccionales de las histonas son el resultado de la actividad de diversos factores, y la mayoría actúa formando complejos moleculares. Algunos de los componentes de estos complejos tienen una actividad catalítica, la cual es necesaria para la incorporación de las modificaciones químicas a las histonas. Los complejos más caracterizados son Trithorax y Polycomb. En Trithorax, la subunidad conocida como Leucemia de Linaje Mixto (*Mixed Lineage Leukemia*, MLL) incorpora grupos metilo a la lisina 4 de la histona H3 (Shilatifard, 2006). Por otra parte, la subunidad Homólogo de Enhancer de Zeste 2 (*Enhancer of Zeste Homolog 2*, EZH2) de Polycomb tiene la propiedad de catalizar la incorporación de grupos metilo a la lisina 27 de la histona H3 (Di Croce y Helin, 2013). Así que la función de Trithorax se relaciona con la activación de la transcripción, y la función de Polycomb con la represión (Shilatifard, 2006).

#### **5.2.1.2 Metilación del ADN**

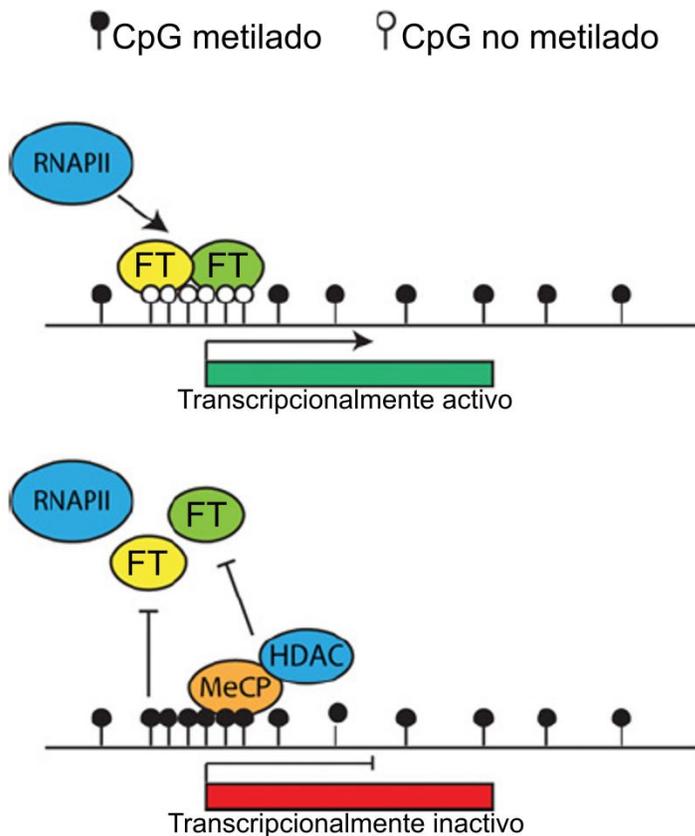
Algunas bases del ADN son susceptibles a la incorporación de un grupo metilo. La más estudiada es la incorporación de un grupo metilo en la citosina (C), cuando se encuentra unida por un enlace fosfatídico a una guanina (G). Este contexto se conoce como el *dinucleótido CpG*. La metilación ocurre con mayor frecuencia en regiones del genoma con una densidad relativamente alta de CpGs, que se conocen como *islas*

CpG, que suelen localizarse hacia los promotores de los genes (Portela y Esteller, 2010; Wu y Zhang, 2014; Meier y Recillas-Targa, 2017).

Generalmente las islas CpG no están metiladas y, cuando sí lo están, se relacionan con la represión de la transcripción. Esto es relevante en contextos fisiológicos normales, como la diferenciación hacia los diferentes tipos de tejido, o en contextos patológicos, como el cáncer (Portela y Esteller, 2010; Meier y Recillas-Targa, 2017). Además, las diferencias en los niveles de metilación que se han relacionado con cambios en los niveles de expresión son frecuentes no en el cuerpo de la isla CpG, sino en CpGs que se encuentran hacia la periferia de las islas CpG, y que se conocen como *límite de islas CpG (CpG island shore)* (Irizarry *et al.*, 2009).

Los mecanismos de represión de la transcripción por metilación de CpGs implican diferencias en la dinámica de unión de algunas proteínas con funciones reguladoras de la cromatina, y que en general, se engloban en dos contextos (Fig. 6): 1) El bloqueo de la unión de factores de transcripción o de complejos enzimáticos, o bien de proteínas que median el contacto entre elementos genéticos de regulación (Bell y Felsenfeld, 2000; Smith y Meissner, 2013; Domcke *et al.*, 2015). Esto impide que se forme el contexto propicio para la activación de la transcripción. 2) El reclutamiento de proteínas que se unen a ADN metilado, y que reclutan a complejos enzimáticos que configuran un ambiente cromatínico inaccesible para la transcripción. Este es el caso de las proteínas con dominio de unión a CpG metilado (*methyl-CpG-binding domain protein*, MBD) y de la proteína de unión a CpG metilado 2 (*methyl-CpG binding protein 2*, MeCP2), las cuales pueden formar complejos con desacetilasas de histonas, cuya

actividad genera un ambiente de cromatina cerrada (Fig. 6) (Jones *et al.*, 1998; Nan *et al.*, 1998; Nott *et al.*, 2016).



**Fig. 6 Efecto de la metilación del ADN sobre el reclutamiento de factores a la cromatina.**

La incorporación de grupos metilo en el ADN, principalmente en los dinucleótidos CpG, tiene efecto sobre secuencias reguladoras, como los promotores. En un contexto no metilado, los factores de transcripción (FT) y la maquinaria de la transcripción (representado aquí por la polimerasa de ARN II, RNAPII) pueden unirse al ADN. Pero la unión de algunos de estos componentes no es posible en un contexto metilado. Además, pueden incorporarse factores como la proteína de unión a CpG metilado (methyl-CpG binding protein, MeCP), la cual a su vez recluta a complejos con actividad enzimática, como las desacetilasas de histonas (HDAC) que promueven un ambiente cromatínico aún más cerrado y transcripcionalmente inactivo. Modificado de Reddington *et al.*, 2013

Además del control de la expresión génica, los mecanismos de metilación del ADN son importantes para la regulación de las secuencias repetidas del genoma, lo cual es esencial para la estabilidad estructural del genoma. En estas secuencias se crean configuraciones de heterocromatina, como es el caso de la heterocromatina pericentromérica (Smith y Meissner, 2013; Groh y Schotta, 2017).

### 5.2.1.3 Variantes de histonas

Los nucleosomas están compuestos por las histonas H3, H4, H2A y H2B. A excepción de la histona H4, hay más de una isoforma de estas histonas. Las variantes más comunes se conocen como *variantes canónicas*. Las variantes no canónicas difieren generalmente en pocos aminoácidos y pueden ser incorporadas en los procesos de ensamblaje/desensamblaje de nucleosomas (Maze *et al.*, 2014).

Estas variantes influyen en las propiedades del ambiente cromatínico y en algunas de las funciones que ocurren dentro del núcleo. Por ejemplo, la variante H3.3 suele incorporarse a nucleosomas localizados en los promotores y el cuerpo de genes que están transcripcionalmente activos, lo que relaciona a esta variante con la regulación transcripcional (Talbert y Henikoff, 2010). En otro ejemplo, la variante macroH2A se requiere en el contexto del silenciamiento del cromosoma X en las hembras de mamíferos (Talbert y Henikoff, 2010; Pasque *et al.*, 2011). Las diferencias en las variantes de histonas se asocian a cambios en las interacciones entre los componentes de los nucleosomas, y entre las histonas y otros factores asociados, como las chaperonas de histonas (Goldberg *et al.*, 2010; Maze *et al.*, 2014).

### 5.2.1.4 ARNs no codificantes largos

Aunque los ARNs no codificantes largos (lncRNAs) no se consideran como un componente epigenético por sí mismos, son factores importantes en la dinámica de la

cromatina y de los procesos epigenéticos. Se definen como transcritos de más de 200 pb que no tienen un potencial codificante significativo (Wu *et al.*, 2017; Kopp y Mendell, 2018; Uszczynska-Ratajczak *et al.*, 2018).

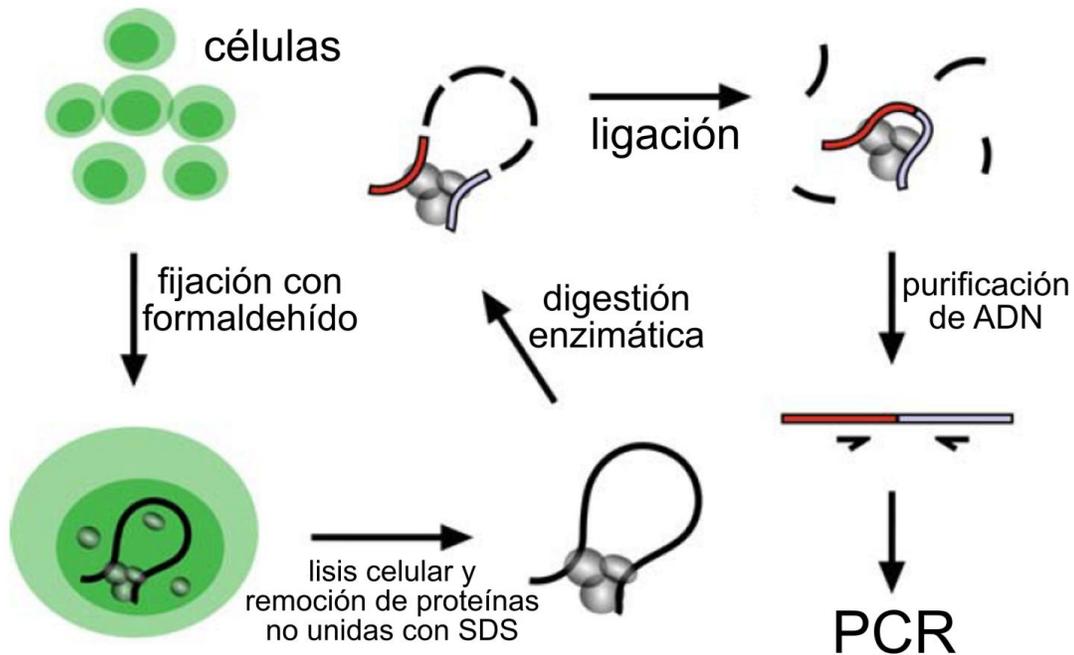
Estas moléculas se transcriben a partir de regiones intergénicas, de *enhancers* o de regiones que coinciden en sentido o antisentido con genes que codifican para proteínas (Wu *et al.*, 2017; Kopp y Mendell, 2018). Su biogénesis tiene algunas similitudes con los transcritos que codifican para proteínas (como su transcripción por la ARN polimerasa II) y algunas diferencias (para la estabilización del transcrito procesado, en algunos de ellos se incorpora 5-metilguanosa en el extremo 5', y una secuencia de poli-adenilación en el extremo 3', pero en otros ocurren mecanismos diferentes, como la incorporación de snoRP [complejo de ARN nucleolar pequeño-proteína] en ambos extremos) (Wu *et al.*, 2017).

Se han caracterizado algunos mecanismos de acción de los lncRNAs. Existe un subconjunto de lncRNAs que se requieren para el reclutamiento de factores de transcripción o de complejos modificadores de la cromatina hacia regiones específicas del genoma (Davidovich y Cech, 2015). En otros casos, algunos lncRNAs interactúan con estos componentes o con otros diferentes (como microRNAs), y en lugar de guiarlos hacia una región del genoma, evitan su unión a la cromatina (Yoon *et al.*, 2014; Morriss y Cooper, 2017). Otros lncRNAs median la interacción de una región del genoma con otra, por ejemplo, interacciones entre promotores y *enhancers* (Wang *et al.*, 2014; Zhang *et al.*, 2014).

En general, los lncRNAs son un componente importante en las funciones nucleares e impactan en contextos fisiológicos como el desarrollo embrionario (Fatica y Bozzoni, 2014). Por ejemplo, el lncRNA conocido como *transcrito específico del cromosoma X inactivo* (*X-inactive specific transcript, Xist*) es uno de los más estudiados. *Xist* se requiere para el silenciamiento del cromosoma X en las hembras de mamíferos, por lo cual es un factor importante en la compensación de dosis génica, y si se reducen los niveles de este lncRNA, se presenta un fenotipo de letalidad embrionaria (Marahrens *et al.*, 1997). *Xist* interactúa con factores de transcripción, como Yin Yang-1 (YY1), y con complejos enzimáticos como PRC2, generando un contexto de cromatina represiva. Además, en la regulación de *Xist* están implicados otros lncRNAs, como *Tsix* y *Jpx*, lo cual ofrece un panorama de la importancia y la complejidad de la regulación y las funciones fisiológicas de los lncRNAs (Fatica y Bozzoni, 2014).

#### **5.2.1.5 Configuración espacial de la cromatina**

Clásicamente, las técnicas de biología molecular permitían conocer las propiedades de distintos elementos genéticos, entendiéndolos como elementos lineales. Recientemente se han desarrollado metodologías que generan datos acerca de la posición relativa de diferentes elementos, y su configuración espacial en el interior del núcleo: cómo están localizados, unos respecto a otros, los cromosomas, *loci*, elementos genéticos, envoltura nuclear, entre otros (Fig. 8) (de Wit y de Laat, 2012).



**Fig. 7 Metodología de captura conformacional de la cromatina (3C).** Modificado de Gavrilov *et al.*, 2009.

Éstas son propiedades de los elementos nucleares que recién se están entendiendo. Al respecto, se está explorando la interacción espacial entre distintos elementos del genoma, su posición absoluta dentro de los distintos territorios al interior del núcleo, y cómo estas configuraciones se relacionan con otros mecanismos de regulación de la cromatina. Éstas constituyen un nivel superior al conocimiento de la complejidad de la regulación de los procesos nucleares (Dekker y Heard, 2015).

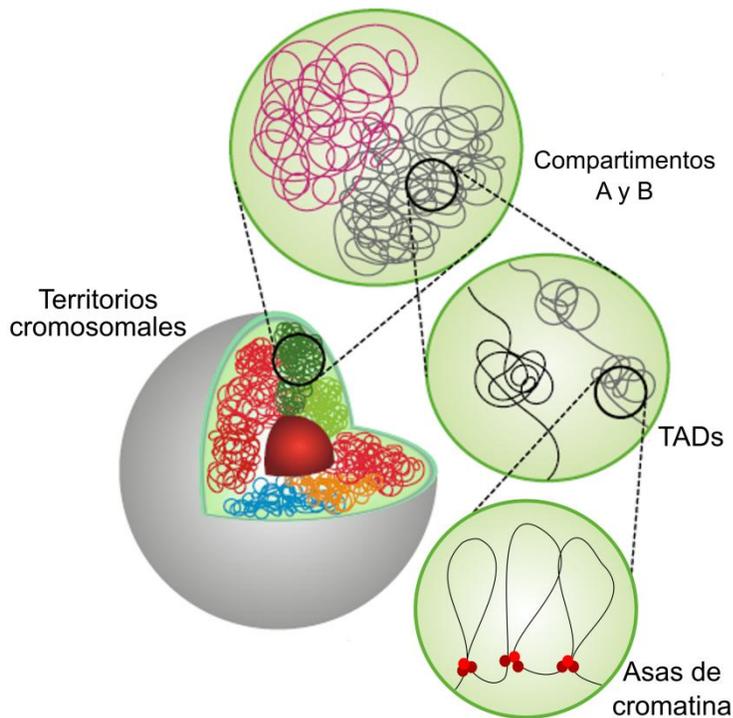
La manera de explorar estos conceptos deriva de la técnica conocida como *captura conformacional de la cromatina* (*chromosome conformation capture, 3C*; Fig. 7) (Dekker *et al.*, 2002). De manera resumida, en esta técnica, se fijan las interacciones que ocurren en la cromatina (incluyendo los elementos que linealmente se encuentran lejanos, pero que por su configuración espacial se encuentran cercanos e

interactuando) utilizando algún agente químico, como formaldehído. Enseguida, la cromatina se digiere con enzimas que tienen una amplia y constante distribución de sitios de restricción a lo largo del genoma (por ejemplo, HindIII). Esto genera extremos cohesivos en los fragmentos fijados, que posteriormente se ligan. Los productos de ligación se utilizan como templados en reacciones de PCR. Con un diseño de *primers* adecuado, se puede probar que dos regiones que linealmente son lejanas, por su configuración espacial hacen contacto una con otra, por lo que son fijadas y ligadas mediante 3C (Fig. 7) (Dekker *et al.*, 2002).

El avance conceptual que se ha derivado del uso de esta técnica ha permitido abordar experimentalmente varias hipótesis. Además, han surgido variaciones y refinamientos metodológicos, con los que se puede preguntar no sólo por la interacción entre un punto particular en el genoma con otro, si no que esto ha escalado para tener evidencia del conjunto de interacciones totales en el genoma (el paisaje conformacional completo) para diferentes tejidos, diferentes condiciones fisiológicas y diferentes tratamientos experimentales (Splinter *et al.*, 2011; Holwerda y Laat, 2012; Noordermeer y Duboule, 2013; Borsos y Torres-Padilla, 2016; Arzate-Mejía *et al.*, 2018).

A partir de este conocimiento, ahora se sabe que el genoma de varios organismos (desde plantas hasta animales, desde invertebrados hasta vertebrados) se organiza en dominios de interacción, de una manera jerárquica y no azarosa (Fig. 8) (Dekker y Heard, 2015; Doğan y Liu, 2018). En el nivel más basal están los nucleosomas en una conformación de collar de perlas, con un espesor de 10 nm (Fig. 3). Su compactación genera una hebra de cromatina más gruesa, la fibra de 30 nm. Enseguida, la interacción de elementos distales genera asas de cromatina (*chromatin*

loops; Fig. 8). Después, se encuentran regiones de decenas de kilobases con interacciones frecuentes entre sus elementos, que forman a los dominios asociados topológicamente (*topologically associating domains*, TADs). Más arriba en la jerarquía, existen los compartimentos A y B, con un tamaño que va entre centenas a decenas de miles de kilobases, y que generalmente corresponden a la eucromatina y a la heterocromatina, respectivamente. Y englobando este conjunto de estructuras se encuentran los territorios cromosomales, que representan la localización de cada cromosoma en una región discreta en el núcleo; al parecer los cromosomas con pocos genes se posicionan hacia la periferia del núcleo, y los cromosomas enriquecidos en número de genes se posicionan hacia el interior del núcleo (Fig. 8) (Dekker y Heard, 2015; Sexton y Cavalli, 2015; Bonev y Cavalli, 2016; Wang *et al.*, 2018).



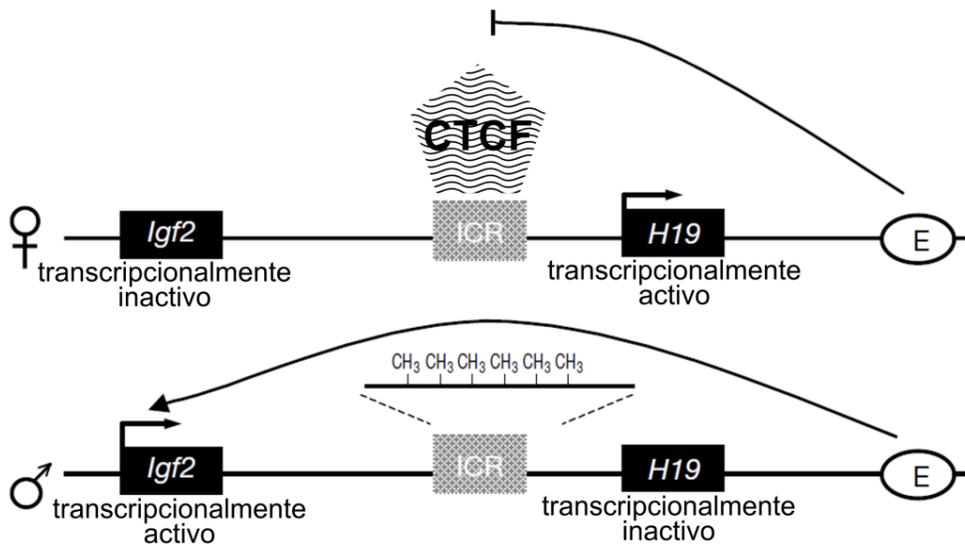
**Fig. 8 Distintos niveles de configuración espacial de la cromatina.** Dentro del núcleo celular, cada cromosoma ocupa un territorio particular, en una disposición relativa que no es azarosa. Dentro de cada territorio cromosomal hay compartimentos A y B, que se caracterizan por el grado de compactación y el estado transcripcional. Dentro de estos compartimentos se encuentran los dominios topológicamente asociados (TADs). Y dentro de los TADs, se encuentran las asas de cromatina. Modificado de Doğan y Liu, 2018.

La mayoría de las configuraciones conformacionales de los diferentes *loci* son constantes entre tejidos e individuos. Con las hipótesis planteadas actualmente se busca resolver si estos dominios constantes son necesarios para la regulación de procesos constitutivos dentro del núcleo (una hipótesis funcional), o si surgen meramente de estados energéticos, acordes a las propiedades fisicoquímicas de la cromatina (una hipótesis estructural) (Peric-Hupkes *et al.*, 2010; Finan *et al.*, 2011; Ciabrelli y Cavalli, 2015).

Pero además de los dominios que son constantes, hay otros dominios que son dinámicos, específicos y característicos de un tejido, de alguna etapa de diferenciación o de un contexto fisiológico particular. Dentro de estos dominios se han encontrado genes relacionados con su contexto, como genes específicos de la diferenciación hacia un tejido o genes del desarrollo, por ejemplo. Lo cual sugiere que los dominios conformacionales que son dinámicos, tienen una relación de regulación sobre los genes que están contenidos en ellos (Peric-Hupkes *et al.*, 2010; Chandra *et al.*, 2015; Perino y Veenstra, 2016).

Se conocen ahora varios ejemplos en los que la interacción física entre elementos genéticos regula la actividad de los genes (Holwerda y Laat, 2012). Por ejemplo, hay casos en los cuales la dinámica del contacto entre un promotor y un *enhancer* causa que la actividad de ese promotor cambie (Fig. 9) (Zhang *et al.*, 2013; Deng *et al.*, 2014). También hay casos en los que el estado de activación diferencial entre dos promotores, se relaciona con cuál de ellos establece un contacto con un *enhancer* que tienen en común (Fig. 9) (Valdes-Quezada *et al.*, 2013; Noordermeer *et al.*, 2014; Fabre *et al.*, 2017; Huang *et al.*, 2017).

En conformidad con esto, los genes que están contenidos dentro de un TAD tienen perfiles de expresión similares. Esto debido a que los elementos reguladores influyen sobre los genes que están contenidos en el mismo TAD, pero no tienen posibilidad de interactuar con genes de un TAD diferente. Si experimentalmente se elimina el borde entre dos TADs, los perfiles de expresión cambian, por la influencia de los elementos de un TAD hacia otro (Dekker y Heard, 2015; Lupiáñez *et al.*, 2015; Narendra *et al.*, 2015).



**Fig. 9 Regulación diferencial de dos promotores por su contacto con un *enhancer*.** Se muestra el ejemplo del *locus* *Igf2/H19*, el cual presenta impronta genética. En el alelo materno, *H19* se activa por un *enhancer* (E) que se encuentra en el *locus*. Río arriba del promotor de *H19* se encuentra la región de control de impronta (*imprinted control region*, ICR), en la cual se une la proteína CTCF. En cambio, en el alelo paterno, el ICR está metilado. Esto bloquea la unión de CTCF al ICR. Sin CTCF en esa posición, se establece un contacto entre el *enhancer* y el promotor de *Igf2*. En este contexto, *Igf2* se expresa y *H19* se encuentra transcripcionalmente inactivo. Modificado de Bell y Felsenfeld, 2000.

Un aspecto interesante a investigar es cómo se forman y se mantienen los TADs, de lo cual depende qué regiones en el genoma quedan organizadas dentro de cada uno

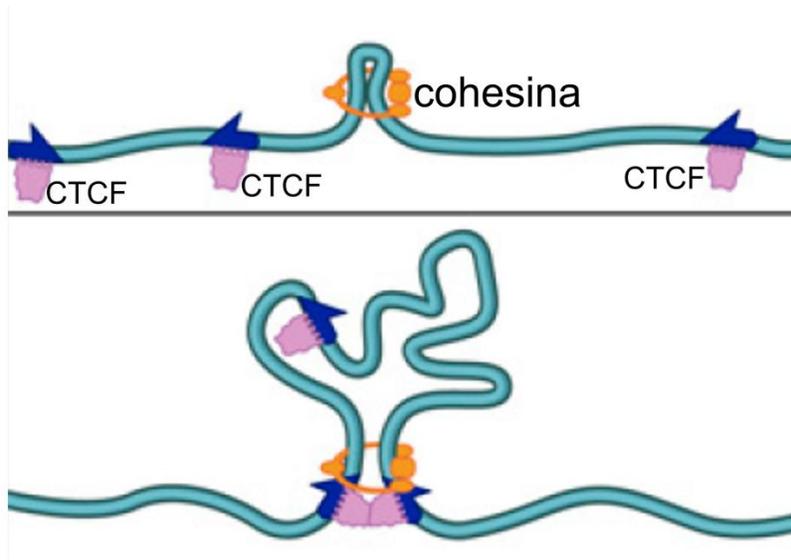
de estos dominios. Y en la delimitación de cada TAD, resultaría importante conocer cómo se establecen sus bordes (Gómez-Díaz y Corces, 2014).

Se sabe que los bordes de los TADs (*TAD borders*) están enriquecidos en genes constitutivos, genes de ARN de transferencia (tRNAs), elementos retrotransponibles intercalados cortos (*short interspersed element retrotransposons*, SINE) y el posicionamiento de algunas proteínas que, en este contexto, se han denominado *proteínas arquitectónicas* (*architectural proteins*). Las más relevantes en vertebrados son las cohesinas y el factor de unión a la secuencia CCCTC (*CCCTC-binding factor*, CTCF) (Dixon *et al.*, 2012; Holwerda y Laat, 2012; Gómez-Díaz y Corces, 2014; Bonev y Cavalli, 2016; Arzate-Mejía *et al.*, 2018).

Las cohesinas y CTCF interactúan entre sí, ya que además de encontrarse en los límites de los TADs, actúan para la formación de asas de cromatina, proceso a partir del cual se ha propuesto el modelo de extrusión (Fig. 10). Este modelo se explica de la siguiente manera: CTCF tiene varios sitios de unión en el genoma y la cohesina se ensambla en algún punto del mismo. Se sabe que forman un complejo en forma de anillo y que, por su interior, se desliza la cromatina, abriéndose un asa. El asa deja de extenderse cuando dos sitios unidos por CTCF se topan entre sí y con la cohesina, y en este punto se establecen los extremos del asa (Fig. 10). Un tema de estudio vigente es el tratar de entender cómo éstas y otras proteínas arquitectónicas funcionan para delimitar los bordes de los TADs (Bonev y Cavalli, 2016; Arzate-Mejía *et al.*, 2018).

En conjunto, los diferentes mecanismos por los cuales se establece la configuración de la cromatina, tienen un impacto sobre la regulación de la expresión de

los genes en varios contextos fisiológicos y con claras repercusiones durante el desarrollo embrionario.



**Fig. 10 Modelo de extrusión para la formación de asas de cromatina.** Las subunidades de la cohesina se ensamblan en forma de anillo, y por su interior se desliza una sección de cromatina. Cuando se encuentran dos moléculas de la proteína CTCF, unidas a la cromatina en una orientación convergente, la cohesina se deja de deslizar. En este punto se constituye el extremo de un asa de cromatina. Modificado de Sanborn *et al.*, 2015.

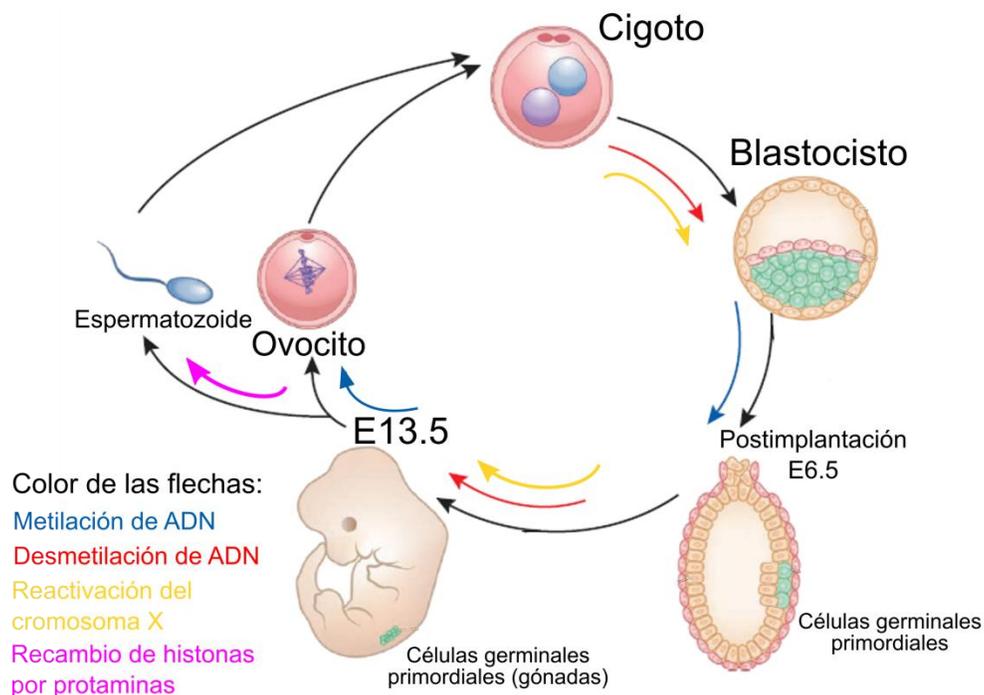
### 5.2.2 Configuraciones epigenéticas en el desarrollo embrionario

La variedad de tejidos y estructuras que se forma durante el desarrollo embrionario requiere de una coordinación fina de la expresión génica en lo espacial y lo temporal. Los mecanismos epigenéticos antes citados participan en esta coordinación (Perino y Veenstra, 2016).

Se ha investigado cuál es la relación entre la configuración epigenética de los gametos y el embrión temprano, esto es, cómo es que las configuraciones epigenéticas preestablecidas en los gametos guían los procesos del desarrollo temprano y cómo es que los estímulos ambientales a los que se enfrentan los padres influyen en la configuración epigenética de los gametos, la cual pudiera heredarse a los hijos

(herencia transgeneracional) (Daxinger y Whitelaw, 2012; Lim y Brunet, 2013; Vaquerizas y Torres-Padilla, 2016).

En este contexto, se sabe ahora que en los vertebrados ocurren dos eventos de reversión masiva de la configuración epigenética (Fig. 11). Uno sucede después de la fertilización, en el embrión temprano, y el otro tiene lugar en las células germinales primordiales, después de su especificación. En estos eventos de reversión los niveles globales de algunas modificaciones postraduccionales de las histonas cambian, y los niveles globales de metilación del ADN se reducen drásticamente (Fig. 11). En el caso de la espermatogénesis, además, las histonas se recambian parcialmente por otro tipo de proteínas, las protaminas (Chen *et al.*, 2011; Leitch *et al.*, 2013; Yoshida *et al.*, 2018).



**Fig. 11 Reprogramación epigenética durante el desarrollo embrionario.** Los eventos masivos de reprogramación ocurren después de la fertilización y después de la especificación de la línea germinal. Modificado de Cantone y Fisher, 2013.

Estos eventos de reversión epigenética se relacionan con transiciones a estados con mayor potencial de diferenciación: desde gametos hacia las células del embrión temprano, o desde células somáticas especificadas hacia células germinales primordiales. Esto apoya la idea de que los estados de diferenciación de las células se alcanzan y mantienen por la restricción que implica el establecimiento de una configuración de cromatina particular (Skinner, 2011).

En células pluripotenciales, como las células troncales embrionarias, la cromatina se caracteriza por tener una configuración laxa, en núcleos con un tamaño más grande que el de las células diferenciadas y con una proporción relativamente menor de heterocromatina (Sathananthan *et al.*, 2002; Stancheva, 2011). Además, en estas células se encuentran sitios en el genoma que tienen una configuración de *cromatina bivalente*: marcas asociadas a activación y a represión transcripcional en el mismo *locus* (Bernstein *et al.*, 2006). La cromatina bivalente se asocia a genes que se activan temprano en el desarrollo embrionario, por lo que funcionalmente se ha interpretado como un contexto que permite que la maquinaria de la transcripción se ensamble, pero que impide que comience su actividad. En el momento adecuado se libera esta restricción y esto, aunado a múltiples señales, provoca que este grupo de genes se expresen en sincronía temporal (Boettiger y Levine, 2009; Vastenhouw y Schier, 2012).

Se ha estudiado el establecimiento de configuraciones cromatínicas durante algunos eventos de diferenciación, tanto para cada uno de los mecanismos epigenéticos por separado, como en combinación. En el nivel que se ha caracterizado más recientemente, que es la configuración espacial del genoma, el panorama conformacional en los tejidos no diferenciados ya se encuentra establecido. Conforme

el tejido se diferencia, una proporción pequeña de interacciones cambia, la cual corresponde a *loci* relacionados con la pluripotencia o con la determinación de los caracteres específicos de los tejidos maduros (Dixon *et al.*, 2015; Borsos y Torres-Padilla, 2016).

En este contexto las proteínas arquitectónicas son especialmente relevantes. Participan en la formación y el establecimiento de dominios de interacción con perfiles de expresión diferencial. La proteína arquitectónica más estudiada es el factor CTCF y ya se ha demostrado su relevancia durante el desarrollo embrionario (Herold *et al.*, 2012; Franco *et al.*, 2014; Arzate-Mejía *et al.*, 2018). La reducción de CTCF genera letalidad embrionaria en distintos modelos animales (Moon *et al.*, 2005; Gerasimova *et al.*, 2007; Mohan *et al.*, 2007; Wan *et al.*, 2008; Moore *et al.*, 2012; Gambetta y Furlong, 2018). Por lo cual, resulta interesante estudiar si su requerimiento durante el desarrollo embrionario está dado por su función como una proteína que media el establecimiento de dominios de interacción, cómo es que éstos se establecen y se mantienen, y cómo es que estas interacciones modulan la expresión de los genes durante el desarrollo embrionario.

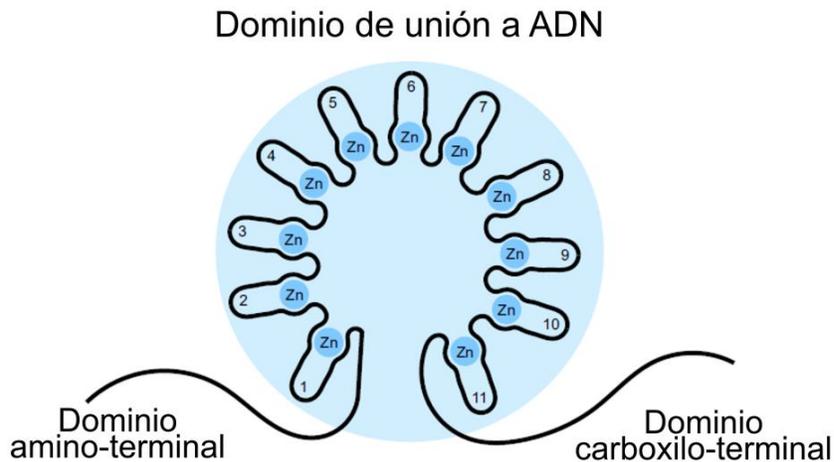
### **5.2.3 CTCF: una proteína multifuncional**

La proteína CTCF está evolutivamente conservada, y es importante para entender la dinámica de regulación de genes mediante la interacción espacial y funcional entre dominios de cromatina (Heger *et al.*, 2009; Heger *et al.*, 2012; Pauli *et al.*, 2016). Hay

ortólogos de esta proteína en insectos (*Drosophila melanogaster*), anfibios (*Xenopus laevis*), peces (pez cebra [*Danio rerio*]), aves (el pollo [*Gallus gallus*]) y mamíferos (el ratón y el humano) (Lobanenkov *et al.*, 1990; Klenova *et al.*, 1993; Filippova *et al.*, 1996; Burke *et al.*, 2002; Moon *et al.*, 2005; Pugacheva *et al.*, 2006).

En su estructura, la región funcionalmente más importante está constituida por 11 dedos de zinc, los cuales son dominios de unión a ADN y ARN, y se encuentran en la porción central de la proteína (Fig. 12) (Filippova *et al.*, 1996; Saldaña-Meyer *et al.*, 2014). Además, posee residuos en las regiones amino- y carboxilo-terminal, que son sitios de fosforilación, poli(ADP-ribosil)ación y SUMOilación, a través de los cuales se regula la actividad de este factor (Klenova *et al.*, 2001; Yu *et al.*, 2004).

CTCF tiene varios sitios de unión al genoma, que se estima van de 40 000 a 80 000 en células de mamíferos (Arzate-Mejía *et al.*, 2018). Una tercera parte de ellos se conserva entre distintos tejidos, lo cual sugiere que algunos de sus procesos reguladores actúan sobre genes cuya función es importante para cualquier tipo de célula, inclusive en especies diferentes (Martin *et al.*, 2011; Johannes *et al.*, 2013). Se ha podido determinar una secuencia consenso en estos sitios, constituida por 20 pb. En una proporción menor de estos sitios, la secuencia consenso se extiende aún más, abarcando 54 pb (Nakahashi *et al.*, 2013).



**Fig. 12 Esquema de la proteína CTCF.** Los 11 dedos de zinc en su región central constituyen el dominio de unión a ADN. En sus dominios amino- y carboxilo-terminal hay sitios blanco para algunas modificaciones postraduccionales, por las cuales se modula CTCF. Además, las regiones aledañas a los dedos de zinc son importantes para la estabilización de la unión de la proteína al ADN. Modificado de Arzate-Mejía *et al.*, 2018.

La dinámica de unión de cada uno de los dedos de zinc al ADN es diferente. Los dedos de zinc centrales (del 4 al 7) establecen con mayor fuerza la interacción de CTCF con el ADN y los dedos de zinc periféricos contribuyen a la estabilización de la unión (Nakahashi *et al.*, 2013).

Además, los dedos de zinc 10 y 11 y el extremo carboxilo-terminal, unen ARN. Por lo cual, de manera interesante, la interacción de CTCF con ARNs no codificantes podría estar implicada en algunas de sus funciones (Saldaña-Meyer *et al.*, 2014).

CTCF interactúa con varias proteínas. Por ejemplo, se asocia con las cohesinas, y con algunos factores de transcripción importantes para la pluripotencia y la diferenciación, como Oct4 y MyoD, entre otros (Donohoe *et al.*, 2009; Delgado-Olguín *et al.*, 2011; Xiao *et al.*, 2011).

Evolutivamente, la secuencia primaria de la proteína está conservada, especialmente los aminoácidos que constituyen a los dedos de zinc (Tabla 2). Lo cual permite hipotetizar que las funciones moleculares de CTCF son similares entre diferentes especies (Klenova *et al.*, 1993; Moon *et al.*, 2005; Pugacheva *et al.*, 2006). Todas estas características permiten que CTCF tenga una amplia variedad de funciones.

**Tabla 2** Identidad de la secuencia de aminoácidos de CTCF de diferentes especies respecto a la de humano (%)

	<i>Drosophila melanogaster</i>	<i>Danio rerio</i>	<i>Mus musculus</i>
Identidad de la proteína completa	27	72	98
Identidad de los dedos de zinc	46	97	100

### 5.2.3.1 Funciones moleculares de CTCF

A CTCF se le ha descrito como un factor multifuncional, por la diversidad de sus mecanismos de acción y de los contextos fisiológicos en los que actúa. Desde la perspectiva molecular, actúa como un factor de transcripción, regulando la actividad de secuencias promotoras (Klenova *et al.*, 1993). Por otra parte, es una proteína necesaria para la función de los elementos genéticos aisladores (*insulators*). Los *insulators* bloquean la influencia de un *enhancer* sobre un promotor, cuando experimentalmente se coloca en medio de éstos (Bessa *et al.*, 2009; García-González y Recillas-Targa, 2014; González-Buendía *et al.*, 2015). Además, algunos *insulators* requieren de CTCF para actuar como un elemento barrera. En este contexto, el elemento barrera se

encuentra entre dos regiones de cromatina con configuraciones distintas (por ejemplo, eucromatina y heterocromatina). Cuando el *insulator* se elimina experimentalmente, o se interfiere con la unión de CTCF al ADN, cambia la configuración de la cromatina (por ejemplo, la región de eucromatina ahora tendría configuración de heterocromatina) (Valdes-Quezada *et al.*, 2013; González-Buendía *et al.*, 2015; Arzate-Mejía *et al.*, 2018).

Un ejemplo de CTCF funcionando como una proteína *insulator* se da en algunos *loci* que funcionan bajo impronta genómica. Éste es un contexto en el que los perfiles de expresión de un gen son diferentes para el alelo paterno y el alelo materno. El *locus* IGF2/H19 funciona bajo impronta genómica: IGF2 se expresa solamente desde el alelo paterno y H19 se expresa solamente desde el alelo materno (Fig. 9) (Bartolomei *et al.*, 1991; DeChiara *et al.*, 1991). Los dos genes se regulan por el mismo *enhancer*. Hay un *insulator* (ICR en la Fig. 9) en el *locus*, el cual se encuentra metilado en el alelo paterno, y desmetilado en el alelo materno. Esto es importante para la dinámica de unión de CTCF. En el cromosoma materno, CTCF se une al *insulator*, y el *enhancer* activa a H19. Pero en el cromosoma paterno el *insulator* se encuentra metilado, lo cual impide la unión de CTCF. Entonces el *enhancer* tiene acceso al promotor de IGF2, y lo activa (Fig. 9) (Bell y Felsenfeld, 2000; Kurukuti *et al.*, 2006).

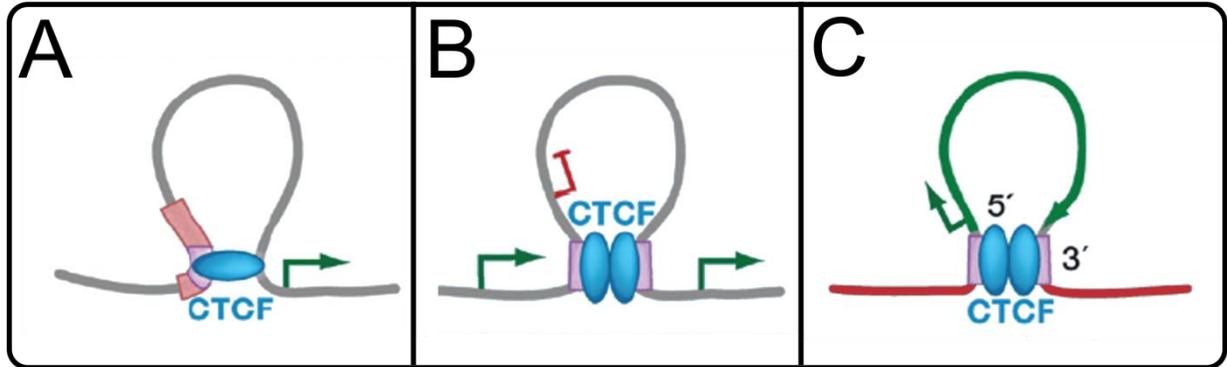
Recientemente, a partir de las técnicas que permiten estudiar la configuración espacial de la cromatina, sabemos que CTCF es una de las proteínas arquitectónicas que regula la organización del genoma (Ong y Corces, 2014; Wang *et al.*, 2018). CTCF se encuentra en la base de las asas de cromatina y es necesaria para su formación y mantenimiento (Fig. 10) (Splinter *et al.*, 2006). Un aspecto relevante es que las secuencias consenso de unión a CTCF en los extremos de las asas se caracterizan por

tener una orientación convergente. Esto es importante porque bioinformáticamente es una característica que se usa para predecir la formación de asas de cromatina (Rao *et al.*, 2014; de Wit *et al.*, 2015; Guo *et al.*, 2015).

Además, CTCF es una de las proteínas arquitectónicas que está más enriquecida en los límites de los TADs (Rao *et al.*, 2014). Actualmente se busca explorar la función de CTCF en este contexto. Algunos datos indican que dicha localización es importante para regular la influencia de elementos reguladores entre TADs diferentes, lo cual es importante para que los perfiles de expresión de los genes se mantengan (Lupiáñez *et al.*, 2015).

Combinando los datos actuales sobre la modulación de la configuración espacial del genoma por CTCF, con los datos anteriores, acerca de su actividad como una proteína *insulator*, la diversidad de funciones de este factor pueden, en parte, explicarse por la identidad de las regiones involucradas tras la formación de un contacto (Fig. 13). Por ejemplo, la formación de un contacto entre un *enhancer* y un promotor, activaría la transcripción. Si la formación del contacto deja fuera de un asa a un *enhancer*, éste no tendrá una influencia reguladora sobre los promotores en el interior del asa. En contextos de recombinación genética, como el que ocurre en el *locus* de los genes que codifican para las distintas regiones de la cadena pesada de las inmunoglobulinas en los linfocitos, CTCF acerca a las regiones que se recombinarán. Y en regiones de cromatina con una configuración particular (por ejemplo, heterocromatina), se limitará esta configuración al interior de un asa de cromatina y esta configuración no se extenderá a los territorios adyacentes (de eucromatina) (Fig. 13) (Guo *et al.*, 2011; Ong y Corces, 2014; Arzate-Mejía *et al.*, 2018; Wang *et al.*, 2018).

Por la versatilidad de sus mecanismos de acción, CTCF tiene un papel importante en varios procesos nucleares. Y esta versatilidad se extiende a los contextos fisiológicos en los que participa, como el sistema inmunológico, el cáncer y el desarrollo embrionario (Ohlsson *et al.*, 2001; Guo *et al.*, 2011; Arzate-Mejía *et al.*, 2018).



**Fig. 13 Distintos tipos de contacto mediados por CTCF.** **A.** Aproximación de un *enhancer* a un promotor, para activar su expresión. **B.** Formación de un asa, dentro de la cual queda excluido un promotor de la influencia activadora de un *enhancer*. **C.** Formación de un asa en la que se establece una configuración de cromatina distinta a las regiones aledañas. Modificado de Phillips y Corces, 2009.

### 5.2.3.2 CTCF y sus funciones en el desarrollo embrionario

CTCF está presente en una variedad amplia de tipos celulares. Esta ubicuidad permite suponer que es una proteína con una diversidad grande de funciones fisiológicas. Respecto al desarrollo embrionario, CTCF se expresa desde las etapas tempranas, lo cual sugiere que en el embrión se requiere desde el inicio del desarrollo (Moon *et al.*, 2005; Wan *et al.*, 2008; Pagano *et al.*, 2016).

En la mosca de la fruta, el ARNm de CTCF se deposita como contribución materna. Este gen se expresa desde las etapas más tempranas del desarrollo (desde el embrión de una célula hasta el final del desarrollo embrionario). El ARNm se enriquece

particularmente en el sistema nervioso del embrión (Moon *et al.*, 2005). Esta dinámica del ARNm se conserva en el pez cebra y en el ratón. En etapas tempranas, el transcrito se encuentra distribuido de manera uniforme por todas las células del embrión. Y conforme el desarrollo avanza, se enriquece particularmente en el sistema nervioso (Pugacheva *et al.*, 2006; Delgado-Olguín *et al.*, 2011).

En embriones de ratón se ha estudiado la función de CTCF en la formación del sistema nervioso, mediante la eliminación dirigida del gen en el telencéfalo y la retina anterior (*knockout* inducible). Se ha encontrado que CTCF se requiere tanto para la supervivencia de los precursores neurales, como para la regulación de su diferenciación (Watson *et al.*, 2014). En otro estudio, en el que se utilizó un *knockout* inducible en neuronas postmitóticas de la corteza y el hipocampo, se encontró que CTCF se requiere para la formación de estructuras somatosensoriales y el crecimiento de dendritas y espinas dendríticas (Hirayama *et al.*, 2012).

Además de sus funciones en el sistema nervioso durante el desarrollo embrionario, se ha descrito su participación en otros sistemas. En el ratón, se requiere en el endotelio, para la correcta formación del sistema circulatorio (Roy *et al.*, 2018). También en la diferenciación de los precursores de las extremidades (Soshnikova *et al.*, 2010). En etapas aún más tempranas, CTCF se requiere para la viabilidad de los embriones en etapas cercanas a la implantación uterina (Moore *et al.*, 2012). Si además de la expresión cigótica, se reduce la contribución materna, el efecto es aún más temprano (Wan *et al.*, 2008).

Además del ratón, CTCF se ha estudiado en otros modelos experimentales. En la mosca, regula la expresión de los genes *Hox*. Así que las moscas mutantes para *ctcf*, presentan transformaciones homeóticas (cambio de la identidad de segmentos metaméricos en el abdomen) y no son viables, mueren poco después de que eclosionan de la pupa (Moon *et al.*, 2005; Gerasimova *et al.*, 2007; Mohan *et al.*, 2007; Gambetta y Furlong, 2018).

En el pez cebra, CTCF regula la expresión de factores maestros de la diferenciación del músculo, *myod* y *miogenina*. Como resultado de una reducción en los niveles de CTCF, se produce menos Actina y Miosina, y la cola de las larvas tiene una forma aberrante (Delgado-Olguín *et al.*, 2011).

Los fenotipos para CTCF en la mosca de la fruta y el pez cebra se han descrito en etapas avanzadas del desarrollo (Gerasimova *et al.*, 2007; Mohan *et al.*, 2007; Delgado-Olguín *et al.*, 2011). Sin embargo, en estos dos organismos, se ha demostrado la expresión de CTCF desde las etapas iniciales del embrión, e inclusive, que existe contribución materna para este factor (Moon *et al.*, 2005; Pagano *et al.*, 2016). Esto sugiere que posiblemente existan más funciones para CTCF que las que se han descrito.

En este trabajo, tenemos interés en inferir la función de CTCF durante las etapas tempranas del desarrollo embrionario del pez cebra.

## **6. Planteamiento del problema**

El factor CTCF está implicado funcionalmente en distintos procesos en el núcleo celular, formando parte de mecanismos por los cuales se regula la actividad de varios genes, la configuración de la cromatina y la organización espacial del genoma. Es importante dentro de diversos contextos fisiológicos en diferentes especies animales, como el desarrollo embrionario.

En el pez cebra, se ha descrito que CTCF se requiere para la formación de los músculos en el embrión, mediante la regulación de la expresión de algunos factores de diferenciación. Además, se ha demostrado que se expresa desde las etapas iniciales del desarrollo embrionario, y que existe contribución materna. Por lo tanto, la pregunta que planteamos en este trabajo es: ¿CTCF tiene funciones en etapas más tempranas del desarrollo del pez cebra, que impactan en la formación de otros tejidos, además de los músculos?

## **7. Hipótesis**

La proteína CTCF está presente de manera continua en las distintas etapas del desarrollo embrionario del pez cebra, y contribuye a la formación de varios tejidos.

## **8. Objetivo general**

Determinar las funciones y la distribución de la proteína CTCF en el desarrollo embrionario del pez cebra.

### **8.1 Objetivos específicos**

**8.1.1** Determinar la dinámica de la distribución de la proteína CTCF durante el desarrollo embrionario del pez cebra.

**8.1.2** Determinar las funciones de CTCF en la formación del embrión de pez cebra.

**8.1.3** Buscar los mecanismos moleculares por los cuales CTCF guía el desarrollo embrionario del pez cebra.

## 9. Material y métodos

### 9.1 Cepas utilizadas, mantenimiento de los peces y método de cruza

La cepa silvestre de peces cebras (*Danio rerio*) TAB-WIK, resultó de la cruce entre las cepas TAB-14 y WIK, que se obtuvieron del Centro Internacional de Recursos de Pez Cebras (*Zebrafish International Resource Center*, ZIRC). Las condiciones de mantenimiento de los peces fueron las siguientes: 26-28°C de temperatura, 14h:10h de ciclo de luz: oscuridad y 6.8-7.2 de pH. Se alimentaron con nauplios de artemia de 2 días post-eclosión y con alimento comercial en hojuelas (TetraMin). Los embriones que se utilizaron para las inyecciones se obtuvieron de cruces individuales (1 hembra x 1 macho). Los embriones que se utilizaron para las tinciones con inmunofluorescencia se obtuvieron de cruces masivas. La clasificación de las etapas de los embriones se basa en la de Kimmel y colaboradores (Kimmel *et al.*, 1995).

### 9.2 Generación de un anticuerpo anti-CTCF de pez cebras

Se generó cDNA a partir de ARN extraído de embriones, con el sistema Reverse Transcription System (Promega) de acuerdo al protocolo de Peterson y Freeman (Peterson y Freeman, 2009). El fragmento del gen *ctcf* que codifica para los aminoácidos del 723 al 798 se amplificó por PCR con los *primers* CTCFzf Fw1 5'-

GAATTCCGGCTGCTGCCATTATCCAGGCT-3' (*forward*) y CTCFzf Rv 5'-AAGCTTTCACCGGTCCATCATGCTAAG-3' (*reverse*). Este fragmento se clonó en el plásmido pGEM-T Easy (Promega) y después se subclonó en el vector de expresión pET28b (EMD Millipore), con los sitios de restricción *EcoRI* y *HindIII*. Como sistema de expresión se utilizó la cepa bacteriana *E. coli* BL21 (DE3). Después de incubar por toda una noche se tomaron 100 mL de este cultivo, se inocularon en medio de cultivo nuevo y se dejó hasta fase exponencial. Para la inducción se agregó 1 mM de IPTG, y se incubó por 4 h a 37°C a 250 rpm. Se recuperó la fracción celular y se suspendió en *buffer* salino de fosfatos (*phosphate-buffered saline*, PBS). Las células se lisaron por sonicación con 10 ciclos de 25 s cada uno. Los restos celulares se removieron por centrifugación a 14 000 rpm por 10 min. El sobrenadante se pasó por una columna de Ni-NTA (Qiagen). La columna se lavó 15 veces con 10 volúmenes de PBS+200 mM de imidazol. El péptido purificado se utilizó para inmunizar conejos, y el antisuero se probó por Western blot con extractos de proteínas de embriones. El anticuerpo se purificó con el sistema AminoLink Plus Immobilization Kit (Thermo Scientific).

### **9.3 Tinción por inmunofluorescencia de embriones completos**

El protocolo para teñir embriones en etapas desde 1 célula hasta 75% de epibolia está basado en el que siguieron Sun y colaboradores (Sun *et al.*, 2008): Los embriones se fijaron durante una noche en PBS con paraformaldehído al 4%. Todos los lavados se hicieron con PBS+Tween-20 al 0.1% (PBT). El corion se removió con pinzas de punta fina. Para la deshidratación y la rehidratación se utilizaron mezclas de metanol y PBT,

con cambios graduales de la proporción de uno y otro. Para la permeabilización se usó PBT+Triton X-100 al 0.5% durante 15 min. El bloqueo y las incubaciones con los anticuerpos se hicieron en PBT+BSA al 1%. El anticuerpo primario fue nuestro propio anti-Ctcf. Como anticuerpo secundario se usó un anticuerpo anti-conejo hecho en cabra con el fluoróforo Alexa 488 (Molecular Probes) en una dilución 1:1000. Las imágenes se tomaron en un microscopio confocal Olympus FV10.

Para los embriones desde la etapa de *bud* y posteriores, el corion se removió con pinzas de punta fina. La fijación se hizo con paraformaldehído al 4% en PBS, durante una noche a 4°C. Para los lavados se utilizó una solución de bloqueo (BSA al 0.1%, Triton X-100 al 1%, diluidos en PBS). El bloqueo se hizo con la solución de bloqueo durante una noche a 4°C. Para la permeabilización se utilizó acetona fría, durante 10 min a -20°C. Después de lavar, los embriones se incubaron con el anticuerpo primario diluido en solución de bloqueo durante una noche a 4°C. El anticuerpo primario se removió mediante lavados. El anticuerpo secundario se agregó diluido en solución de bloqueo y se incubó por 2h a temperatura ambiente.

La señal de autofluorescencia en embriones tempranos (embriones de tapas anteriores a la de *bud*) se corrigió con el complemento Image Calculator de ImageJ (Billinton y Knight, 2001; Schindelin *et al.*, 2012). A la señal obtenida por el filtro para el fluoróforo del anticuerpo, se le sustrajo la señal de autofluorescencia, obtenida a través del filtro para DAPI.

#### 9.4 Generación de mutaciones por el sistema CRISPR-Cas9

El ensayo se basó en el protocolo publicado por Gagnon y colaboradores (Gagnon *et al.*, 2014). Se diseñó un sgRNA contra el primer exón codificante de *ctcf* utilizando la herramienta bioinformática CHOPCHOP (Montague *et al.*, 2014). La secuencia del oligonucleótido específico (con el promotor SP6) fue 5'-ATTTAGGTGACACTATAGGGACCGACTGAGGCCGTGG-GTTTTAGAGCTAGAAATAGCAAG-3' y la secuencia del oligonucleótido constante fue 5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTAACTTGCTATTTCTAGCTCTAAAC-3'. La transcripción del sgRNA se hizo con el sistema MEGAscript SP6 (Ambion). Para generar el ARNm de Cas9, el plásmido pCS2-Cas9 (Addgene) se transcribió *in vitro* con el sistema mMACHINE mMACHINE (Ambion). Cada embrión se inyectó con 133 pg del sgRNA y 200 pg de ARNm de Cas9, diluidos en *buffer* Danieau (NaCl 0.58 mM, KCl 7 mM, MgSO<sub>4</sub> 4 mM, Ca(NO<sub>3</sub>)<sub>2</sub> 6 mM, HEPES 50 mM). Los embriones inyectados se dejaron desarrollar en medio E3 (NaCl 5 mM, KCl 0.17 mM, CaCl<sub>2</sub> 0.33 mM, MgSO<sub>4</sub> 0.33 mM, azul de metileno al 0.1%) a una temperatura de 28.5°C. Se extrajo ADN de grupos de embriones de 24 horas postfertilización (hpf). Se utilizaron los *primers* CTCFseqF 5'-CCAACCTTTACCTCAATTGCACA-3' (*forward*) y CTCFseqR 5'-TTGGTATGTCTTCAC TCCTTG-3' (*reverse*) en ensayos de PCR, y los productos de amplificación se utilizaron en ensayos con la endonucleasa T7 (NEB), o bien se clonaron y se prepararon para secuenciación.

## 9.5 Inyección de morfolinos y experimentos de rescate

Las secuencias de los morfolinos *ctcf* fueron: 1) Morfolino para bloquear la traducción (mo-aug): 5'-GCCTCAGTCGGTCCCCCTTCCATG-3'; 2) Morfolino para bloquear el *splicing* (mo-sp): 5'-CAAAACAGATCACAAACCTGAAAGC-3'. Para las inyecciones, los morfolinos se diluyeron en *buffer* Danieau. Se utilizó un inyector PV380 Pneumatic PicoPump (World Precision Instruments). Los embriones inyectados se dejaron desarrollar en medio E3 a una temperatura de 28.5°C. A las 24 hpf, se fotografiaron para documentar el fenotipo con un estereomicroscopio Nikon SMZ1500, con una cámara AxioCam MRc (Zeiss). Para los experimentos de rescate, el ARNm de *CTCF* se transcribió *in vitro* con el sistema mMESSAGE mMACHINE (Ambion), a partir del plásmido pCI-7.1, el cual contiene el cDNA completo de *CTCF* de humano (el cual fue donado por la Dra. Elena Klenova, de la Universidad de Essex, Reino Unido). Los embriones se inyectaron con una mezcla de 4 ng de mo-aug, y 50 o 100 pg de ARNm de *CTCF*.

## 9.6 RT-PCR

Se extrajo ARN de grupos de 50 a 70 embriones en la etapa de 90% de epibolia (para comprobar el efecto del mo-sp) o embriones de 24 hpf (para los ensayos de expresión de *p53* y *bbc3/PUMA*). La extracción de ARN se llevó a cabo según el protocolo de Peterson y Freeman (Peterson y Freeman, 2009). Se generó cDNA con el

sistema Reverse Transcription System (Promega). Se utilizaron *primers* para amplificar una región de los exones 2 y 3 de *ctcf* (para comprobar el efecto del mo-sp), los cuales tuvieron las secuencias: FCF (*forward*) 5'-ACAGAAGTCATGGAGGCTGC-3' y FCR (*reverse*) 5'-CTGTGTTGTAGCGCAGCTTG-3'. Las secuencias de los *primers* para *p53*,  *$\beta$ -actina* y *bbc3/PUMA* se tomaron de trabajos ya publicados (Rhodes *et al.*, 2010; Toruno *et al.*, 2014). Los ensayos de RT-qPCR se llevaron a cabo con el sistema StepOne Real-Time PCR System (Applied Biosystem).

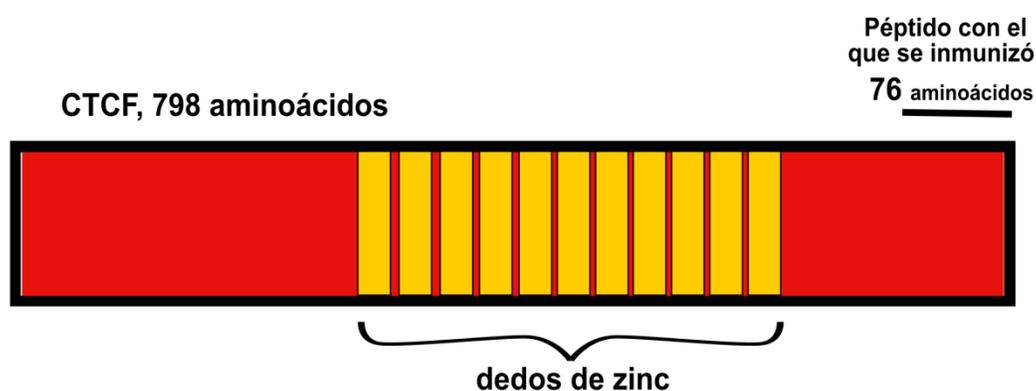
## 9.7 Extracción de proteína para Western blot

El protocolo de extracción de proteína se basó en el de Link y colaboradores (Link *et al.*, 2006). De manera breve, el corion de embriones en la etapa de 90% de epibolia se removió con pinzas de punta fina. Los embriones se transfirieron a un tubo de 1.5 mL con *buffer* de devitelización (Tris-HCl 625 mM pH 6.8, KCl 1.8 mM, NaHCO<sub>3</sub> 1.25 mM). Se pasaron a través de una punta de pipeta para romper el vitelo, y se agitaron en un *vortex*. Se centrifugaron a 300 g durante 30 s y el sobrenadante se desechó. Se agregó *buffer* de carga (Tris-HCl 625 mM pH 6.8, glicerol al 50%, SDS al 15%,  $\beta$ -mercaptoetanol al 12.5%, azul de bromofenol) según un volumen de 0.5  $\mu$ l/embrión. La mezcla se calentó a hervor durante 5 min. 5  $\mu$ l del extracto de proteína se utilizó en cada carril en electroforesis por SDS-PAGE.

## 10. Resultados

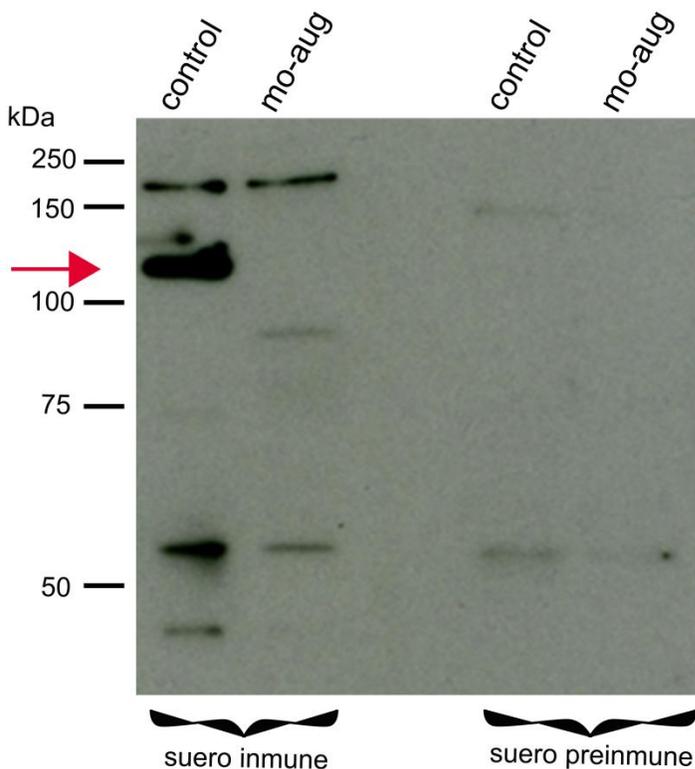
### 10.1 Distribución de CTCF durante el desarrollo embrionario del pez cebra

Para investigar cómo se distribuye la proteína CTCF en distintas etapas del desarrollo embrionario del pez cebra, se generó un anticuerpo policlonal. Se inmunizaron conejos con un péptido constituido de los últimos 76 aminoácidos de la proteína (la proteína completa tiene 798 aminoácidos, Fig. 14). La región de la proteína que se utilizó para las inmunizaciones se encuentra después de los dedos de zinc. En un análisis de este fragmento por BLAST no se recuperó ninguna secuencia diferente a CTCF, no tiene estructura secundaria y las predicciones computacionales le dan una puntuación alta de antigenicidad (Predicted Antigenic Peptides, Immunomedicine group, Universidad Complutense de Madrid).

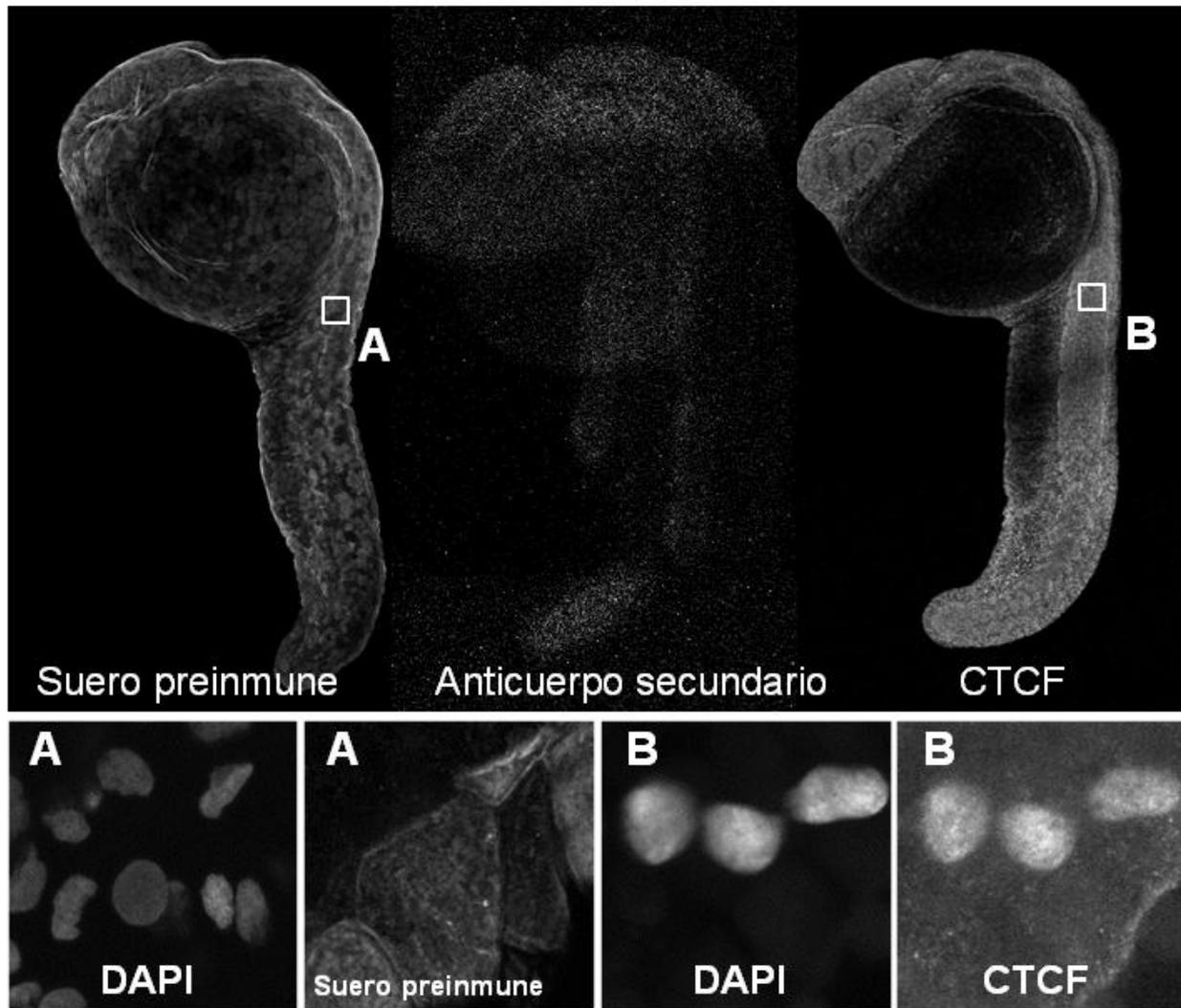


**Fig. 14** Localización de la porción de la proteína que se utilizó para generar el anticuerpo anti-CTCF.

Para demostrar que las inmunizaciones fueron exitosas, se probaron los antisueros en ensayos de Western blot. Se observó una banda de 130 kDa, que no se revela con el suero preinmune (Fig. 15). Además, esta banda tampoco se revela cuando se usan extractos de proteínas de embriones *knockdown* para CTCF (embriones inyectados con morfolidos, los cuales se detallan más adelante). En inmunotinciones de embrión completo, el anticuerpo purificado genera una señal distribuida por todo el cuerpo del embrión (Fig. 16). En contraste, cuando se utiliza el suero preinmune, se obtiene una señal débil e inespecífica en la superficie del embrión. Además, con el anticuerpo purificado, la señal de CTCF es nuclear, pues coincide con la señal de DAPI, la cual es la localización que se espera para esta proteína (Fig. 16A). En cambio, la señal del suero preinmune, además de ser débil, no coincide con la de DAPI. En conjunto, estas pruebas permitieron validar el anticuerpo y determinar que es específico para CTCF de pez cebra.

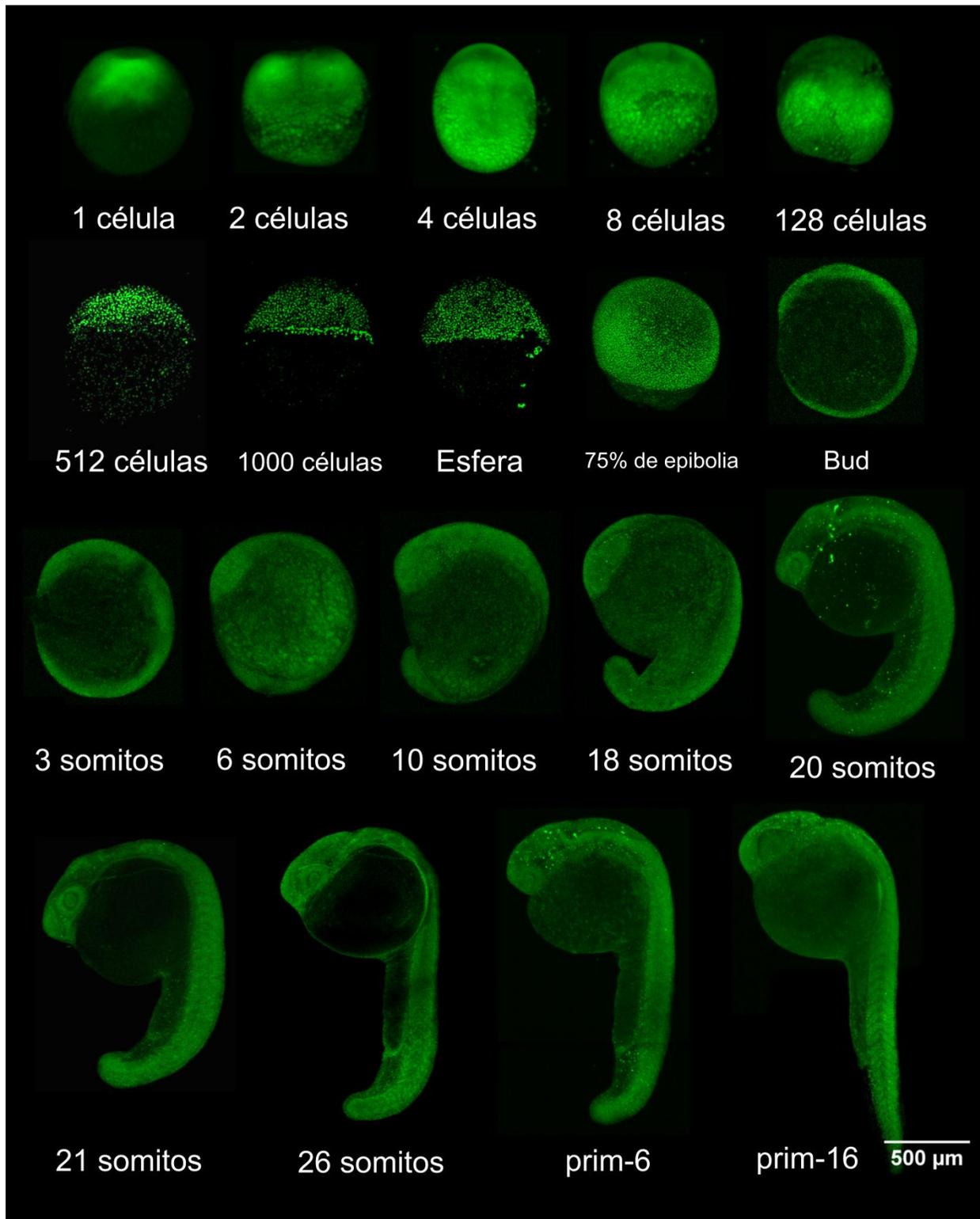


**Fig. 15 Prueba de especificidad del anticuerpo anti-CTCF por Western blot.** Se utilizaron extractos de proteína de embriones inyectados con vehículo (control) o con el mo-aug. Se comparó el suero inmune con el suero preinmune. Con el suero inmune, se detectó una banda diferencial de 130 kDa.



**Fig. 16 Prueba de especificidad del anticuerpo anti-CTCF por inmunofluorescencia.** Ensayos de inmunofluorescencia de embrión completo. Se muestran embriones de una etapa del desarrollo que corresponde a la de 26 somitos. En estas tinciones se utilizó el anticuerpo purificado. En las regiones amplificadas en A y B se compara la tinción nuclear de DAPI, con la tinción del suero preimmune y con el anticuerpo purificado.

El anticuerpo purificado se utilizó para determinar la distribución de CTCF en el desarrollo del pez cebra con ensayos de inmunofluorescencia de embriones completos (Fig. 17). Se usaron embriones que van de la etapa de 1 célula (justo después de la fertilización) hasta la etapa de prim-16, que abarcan las primeras 30 horas del desarrollo.



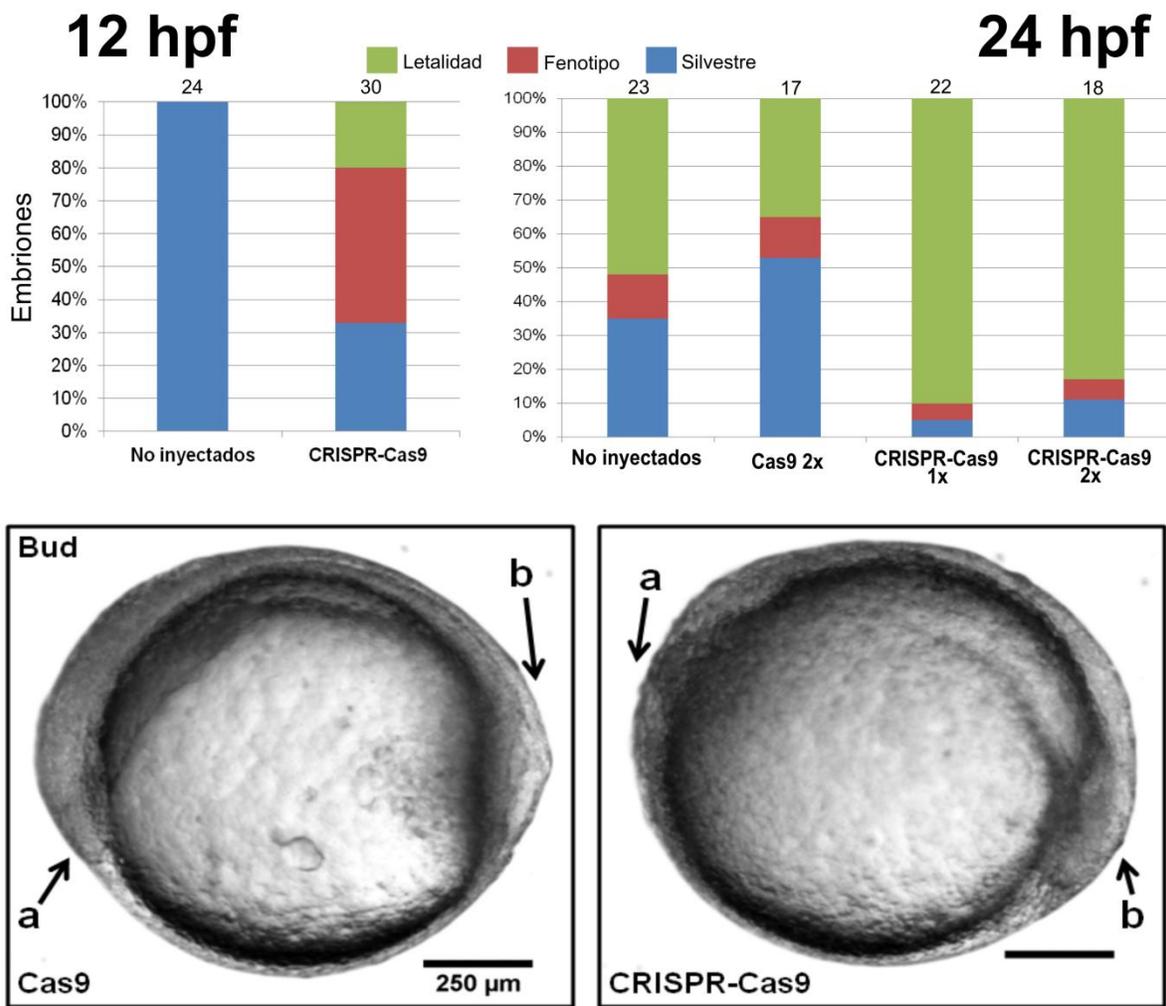
**Fig. 17 Distribución de CTCF durante el desarrollo embrionario del pez cebra.** Inmunotinciones de embrión completo, utilizando el anticuerpo purificado anti-CTCF.

llo embrionario (determinación de las etapas con base a la nomenclatura de Kimmel (Kimmel *et al.*, 1995). Con este anticuerpo se determinó que CTCF está presente desde el embrión de 1 célula, y se mantiene presente durante todas las etapas examinadas (Fig. 17). Además, se observó que CTCF se distribuye en todas las células de los embriones en etapas tempranas (desde el embrión de 1 célula hasta 75% de epibolia), y a lo largo de todo el cuerpo en embriones de las etapas siguientes (desde la etapa de *bud* hasta prim-16). En ninguna de estas etapas encontramos un enriquecimiento particular en alguna región del embrión. Con estos datos concluimos que CTCF es una proteína nuclear, ubicua, y que está presente en todas las etapas del desarrollo del pez cebra.

## 10.2 Fenotipo derivado de mutaciones en CTCF por el sistema CRISPR-Cas9

Buscando la función de CTCF durante el desarrollo mediante una estrategia de pérdida de función, se generaron mutaciones en este *locus* usando el sistema de edición genética CRISPR-Cas9. Se diseñó un RNA guía (sgRNA) dirigido hacia el primer exón codificante del gen *ctcf*, en una región muy próxima al sitio de inicio del marco de lectura abierto (Fig. 20). Se llevaron a cabo inyecciones del sistema CRISPR-Cas9 y se compararon con embriones inyectados únicamente con la Cas9 (Fig. 18 y Tabla suplementaria 12.2). Se dejaron desarrollar, y se revisaron a las 24 horas post-fertilización (hpf). Se encontró que la inyección fue letal en la mayoría de los embriones inyectados, modificados genéticamente con el sistema CRISPR-Cas9 (Fig. 18).

Comparando con embriones inyectados únicamente con la Cas9 (sin el sgRNA) o con embriones que se dejaron desarrollar sin haberlos inyectado, consecuentemente se determinó que la letalidad de la inyección con el sistema CRISPR-Cas9 es entre 30 a 40% mayor a la letalidad que se encuentra en las condiciones control (Cas9 sola, o embriones sin inyección; Fig. 18).

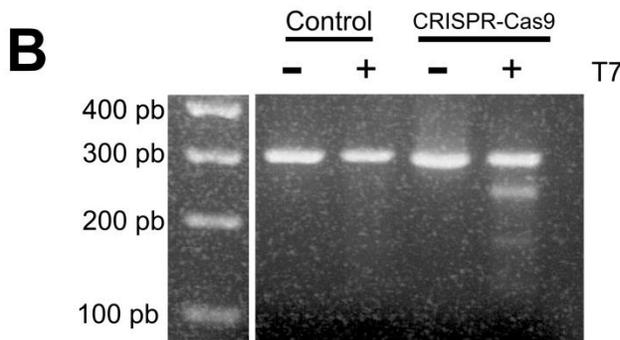
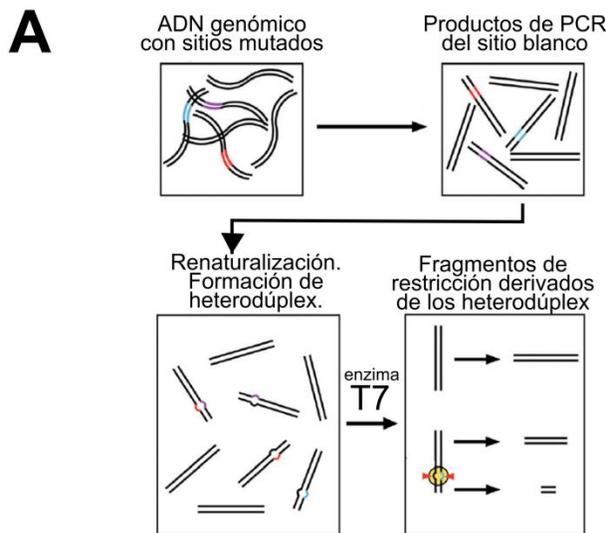


**Fig. 18 Fenotipo de embriones inyectados con el sistema CRISPR-Cas9.** Se muestra la proporción de letalidad, embriones con fenotipo de malformaciones (Fenotipo) y embriones con fenotipo silvestre (Silvestre). Las cifras en las gráficas indican el número de embriones analizados. Además, se muestran fotografías representativas de embriones en la etapa de *bud* (10 hpf), inyectados con la Cas9 (que tienen fenotipo silvestre), y embriones inyectados con el sistema CRISPR-Cas9 (que tienen fenotipo de malformaciones). Para comparación se indican con flechas la cabeza (a) y la parte más caudal del tronco (b). Las gráficas se basan en los datos mostrados en las Tablas suplementarias 12.1 y 12.2

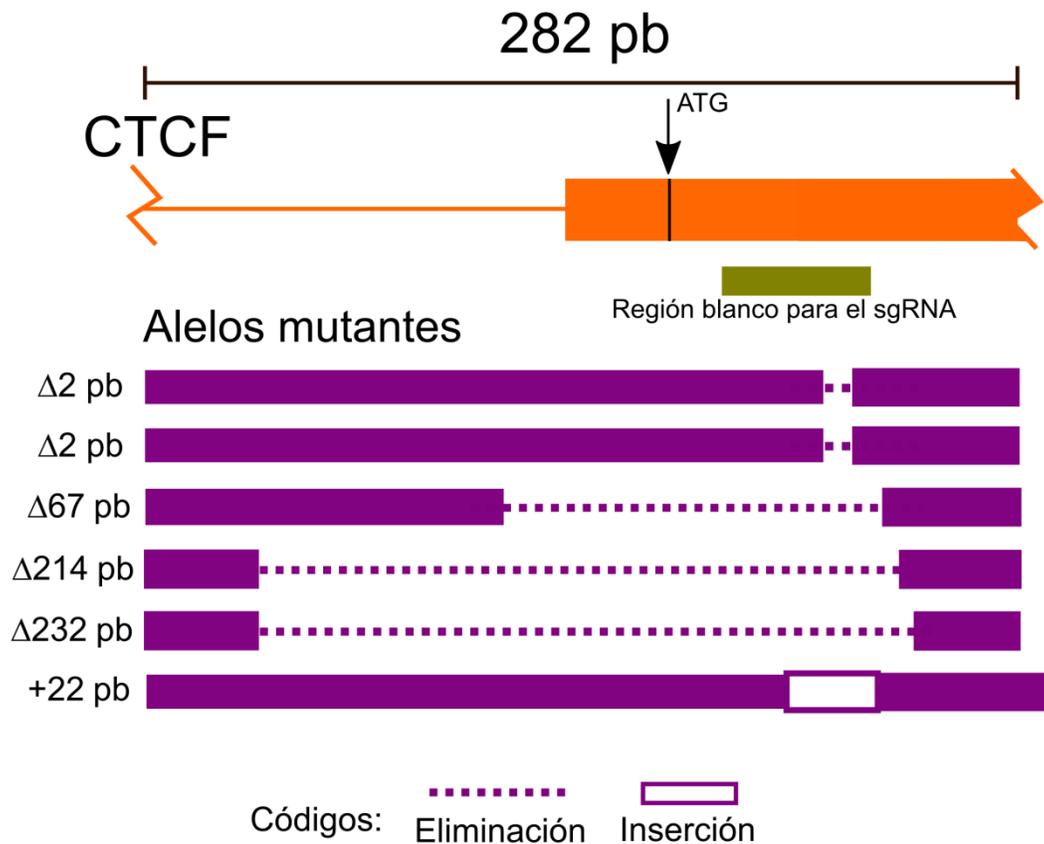
Para determinar cuál era el fenotipo de los embriones inyectados con el sistema CRISPR-Cas9 antes de presentar letalidad, se realizó una nueva serie de inyecciones, y las observaciones se hicieron en una etapa más temprana, a las 12 hpf (Fig. 18 y Tabla 12.1). A este tiempo, también se encontró una letalidad mayor en comparación con embriones no inyectados. Además, en los embriones inyectados con el sistema CRISPR-Cas9 se observó que, en la porción anterior, donde ya se empieza a formar la cabeza, ésta aparece aplanada, en lugar de la forma de bulbo que ya se alcanza a observar en la cabeza en formación de los embriones no inyectados (Fig. 18). La porción posterior, donde se empieza a formar la cola, también se observa aplanada, en vez del brote abultado que termina en una punta aguda, que es como se observa en los embriones no inyectados (Fig. 18). El fenotipo de la porción posterior es similar a lo que en nuestro grupo reportamos previamente, en el que se disminuyeron los niveles de CTCF utilizando morfolinós (Delgado-Olguín *et al.*, 2011). A partir de estas observaciones concluimos que la inyección del sistema CRISPR-Cas9 dirigido al *locus* de CTCF, genera malformaciones en el embrión. Este fenotipo es muy evidente a los 3 días postfertilización (dpf; Fig. 18).

Para buscar cuáles fueron las lesiones moleculares causadas en el *locus* de CTCF por la inyección del sistema CRISPR-Cas9, se extrajo ADN de embriones de 12 hpf que se inyectaron con este sistema. Se desecharon los embriones con letalidad, y el resto de los embriones se agruparon y se utilizaron para la extracción de ADN. La región blanco del sgRNA se amplificó por PCR, para posteriormente realizar un ensayo de endonucleasa T7 (Fig. 19). En este ensayo se amplifica por PCR una región que contiene la secuencia blanco del sistema CRISPR-Cas9 (Fig. 19A). Enseguida, se

desnaturalizan los amplicones por temperatura. Se reduce la temperatura para que se vuelvan a aparear las hebras. Si hay un apareamiento de hebras incompleto, debido a que en algunas de las hebras hay inserciones o deleciones (*indels*), derivados del sistema CRISPR-Cas9, este mal apareamiento genera un sitio de corte para la endonucleasa T7 (Fig. 19A). Se amplificó una región de 282 pb, que contiene a la secuencia blanco del sgRNA. Se incubó con la enzima T7. Al correr esta reacción en un gel se revelaron fragmentos de restricción en los embriones que fueron inyectados con el sistema CRISPR-Cas9, y que no aparecieron en embriones no inyectados (Fig.19B). Este resultado demostró que el sistema CRISPR-Cas9 generó *indels* en el locus.



**Fig. 19 Ensayo de endonucleasa T7. A.** Esquema que representa los pasos básicos de este ensayo. Para probar si con la inyección del sistema CRISPR-Cas9 se generaron mutaciones, la región a la que se dirigió el sistema se amplifica por PCR. Enseguida, los productos de PCR se renaturalizan, para que se formen heterodúplex (fragmento de ADN en el que una hebra contiene un alelo silvestre, y la otra hebra contiene un alelo mutante). La enzima T7 detecta los heterodúplex y los corta. Al correr en un gel de agarosa, se detectan los fragmentos de restricción. Modificado de (Wyvekens *et al.*, 2015). **B.** Se extrajo ADN de embriones no inyectados (Control) y embriones inyectados con el sistema CRISPR-Cas9 para generar mutaciones en el gen de *ctcf*. Se llevaron a cabo PCRs, que generaron un amplicón de 282 pb (Fig. 20). Con el producto de PCR se llevó a cabo un ensayo con la endonucleasa T7.



**Fig. 20 Alelos secuenciados de los embriones inyectados con el sistema CRISPR-Cas9.** Se muestra un esquema con una porción del gen *ctcf*. El sgRNA está dirigido a un exón, justo río abajo del comienzo del marco de lectura abierto (ATG). Además, se esquematizan los alelos que se obtuvieron de la secuenciación de fragmentos de PCR ligados en vectores de clonación, a partir de ADN de embriones inyectados con el sistema CRISPR-Cas9.

Para conocer en detalle las alteraciones que se produjeron en el *locus*, se clonaron amplicones derivados de la PCR, los cuales fueron secuenciados. De 13 clonas que se secuenciaron, en 6 de ellas se encontraron *indels* (6/13=46%, Fig. 20). Uno de estos alelos corresponde a una eliminación de 2 pb ( $\Delta 2$ , Fig. 20) y se encontró en dos clonas independientes. Este alelo genera un cambio en el marco de lectura, el cual puede formar una proteína con una secuencia de aminoácidos diferente a partir del

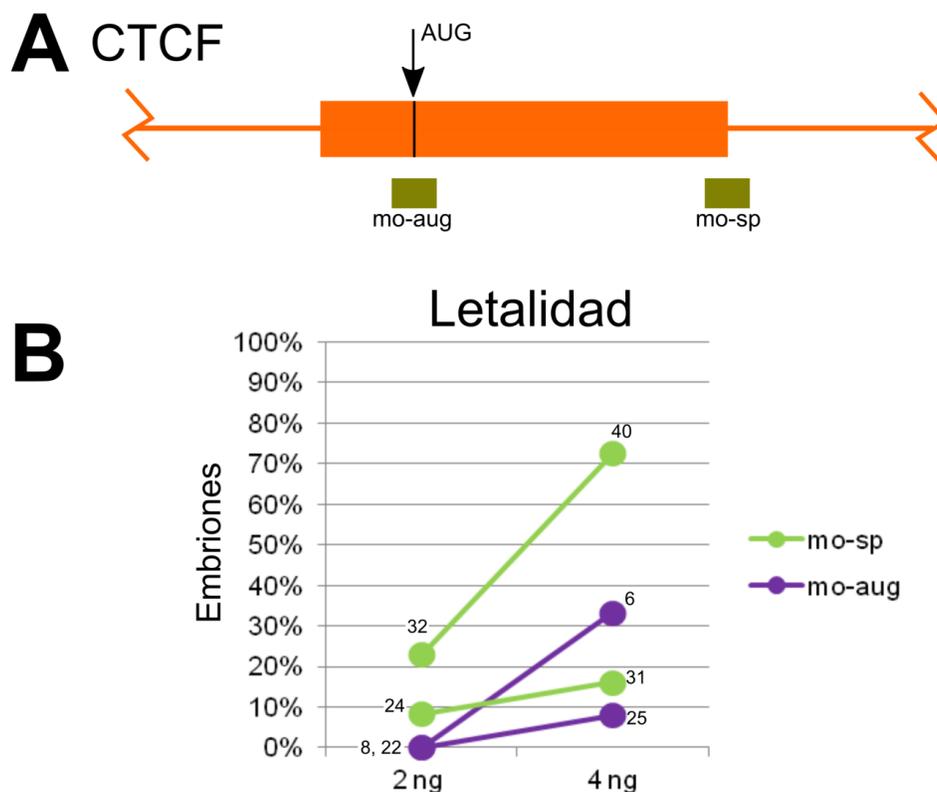
noveno aminoácido. En tres clonas se encontraron eliminaciones mayores: de 67, 214 y 232 pb ( $\Delta 67$ ,  $\Delta 214$ ,  $\Delta 232$ , Fig. 20). Estas eliminaciones comienzan desde un intrón y se extienden hasta el exón contiguo, en una región río abajo del sitio de inicio de la traducción, de lo cual se espera un *splicing* aberrante. Además, en una de las clonas se encontró una inserción de 22 pb (+22, Fig. 20) río abajo del sitio de inicio de la traducción, que recorrería el marco de lectura abierto después del sexto aminoácido. En conjunto, se demostró que el sistema CRISPR-Cas9 generó alteraciones en el *locus* del gen *ctcf*, que derivaron en malformaciones durante el desarrollo embrionario.

### **10.3 La reducción del nivel de la proteína CTCF por inyección de morfolinó genera malformaciones en el embrión**

Para continuar el estudio de la función de CTCF en el desarrollo embrionario del pez cebra, se generaron embriones *knockdown*, inyectando morfolinó (Fig. 21). Los morfolinó son oligonucleótidos sintéticos que tienen un grupo químico que se llama *anillo morfolino*, en lugar de la ribosa (por extensión, al oligonucleótido completo, se le llama *morfolino*). Además, las bases en los morfolinó se unen por enlaces fosforodiamidato, en lugar de enlaces fosfodiéster. Esto hace que sean menos polares que los ácidos nucleicos de las células, y que las nucleasas no actúen sobre ellos. Para su uso, los morfolinó se diseñan con una secuencia que sea complementaria a un transcrito al cual se busca afectar. Se unen a ARN, ya sea pre-ARNm o ARNm, e

interfieren ya sea con el *splicing* o con la traducción de ese transcrito, y como consecuencia abate el nivel de la proteína (Nasevicius y Ekker, 2000).

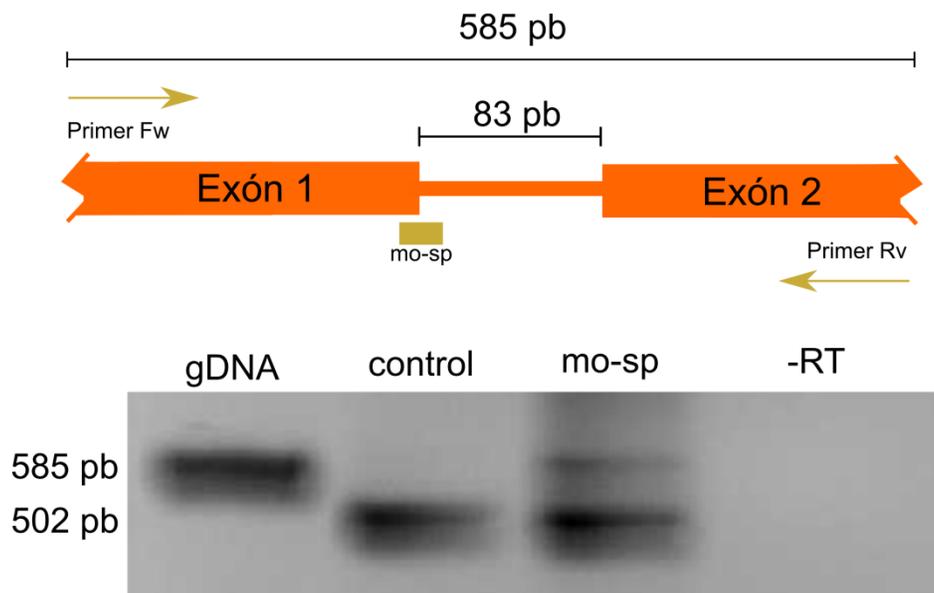
Se utilizaron dos morfolininos *ctcf* que se diseñaron en el contexto de un trabajo previo de nuestro grupo de investigación (Delgado-Olguín *et al.*, 2011). El morfolino aug (mo-aug) tiene como blanco el sitio de inicio de la traducción (Fig. 21A). El morfolino sp (mo-sp) tiene como blanco un sitio de *splicing*, al final del primer intrón (Fig. 21A). El efecto que se espera del mo-sp es que durante el *splicing* ese intrón no sea removido. Esto generaría un sitio prematuro de término de la traducción en el ARNm.



**Fig. 21 Sitios blanco para los morfolininos anti-CTCF y letalidad con distintas concentraciones. A.** Esquema de una porción del gen *ctcf* con la anotación de los sitios a los que están dirigidos los morfolininos. El mo-aug está dirigido al sitio de inicio de la traducción (AUG), y el mo-sp está dirigido hacia el sitio donador de *splicing* que se encuentra río abajo del AUG. **B.** Porcentaje de letalidad en embriones inyectados con cada uno de los morfolininos con dos concentraciones diferentes. Las cifras en la gráfica, indican el número de embriones analizados.

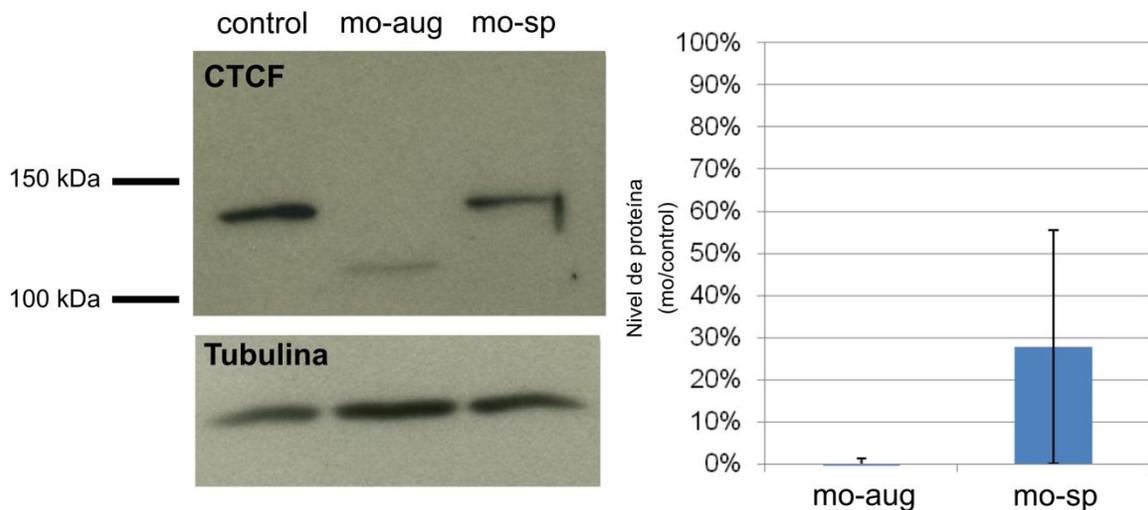
Se inyectaron el mo-aug o el mo-sp (por separado), en embriones en la etapa de una a dos células (dicho de otro modo, se generaron *morfantes*). Se inyectaron dos cantidades diferentes, para ensayar cuál utilizar en los experimentos posteriores: 2 o 4 ng por embrión (Fig. 21B y Tabla suplementaria 12.3). Posteriormente se evaluó la letalidad de los morfantes a las 48 y 72 hpf. Con la inyección de 2 ng de morfolino no hubo diferencia en la letalidad en comparación a embriones no inyectados o embriones inyectados con un morfolino control. Pero con 4 ng, sí hubo un aumento en la letalidad (Fig. 21B). Por lo tanto, los siguientes experimentos se hicieron inyectando 4 ng de morfolino.

Una vez establecidas dichas condiciones se realizaron ensayos de RT-PCR para determinar el efecto del mo-sp (Fig. 22). Se realizaron nuevas inyecciones y se extrajo RNA total de embriones a las 9 hpf. Por RT-PCR, se amplificó un fragmento con *primers*



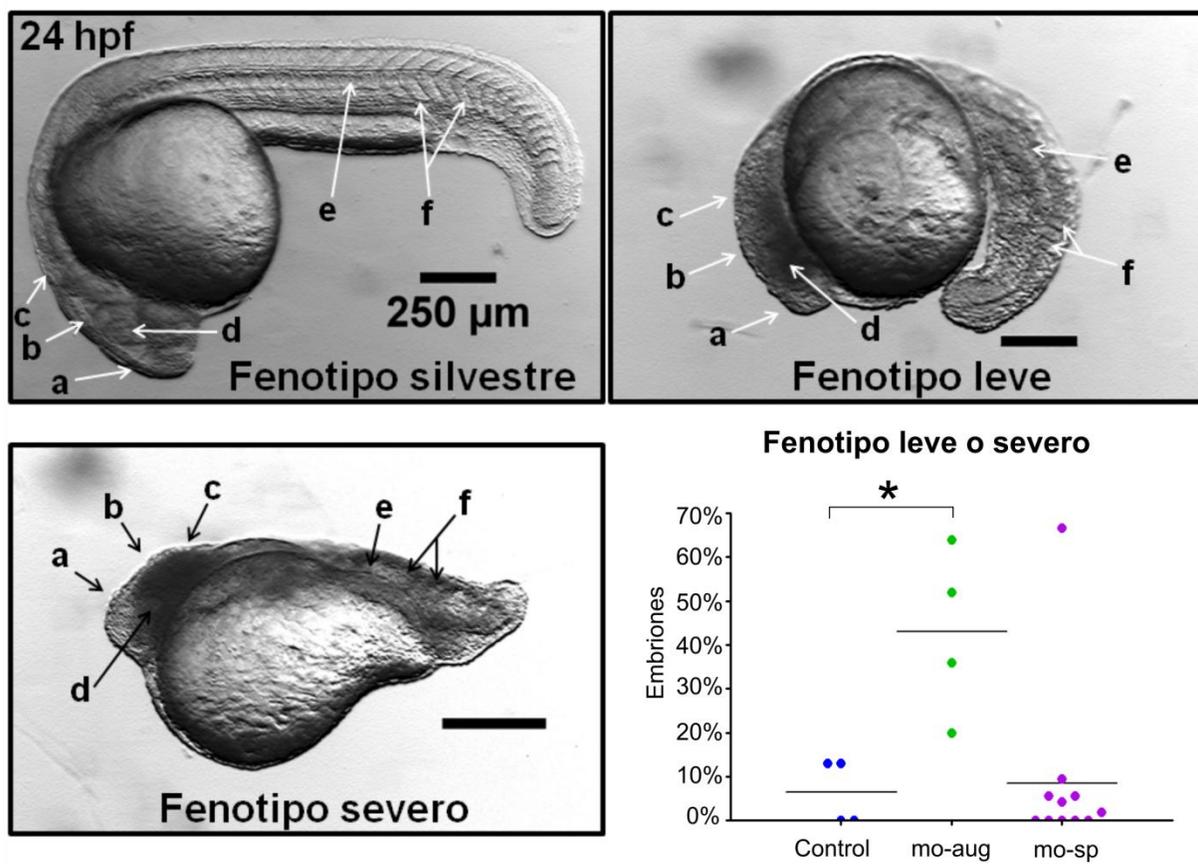
**Fig. 22 Validación del efecto del mo-sp por RT-PCR.** Se inyectaron embriones con el mo-sp y se hicieron extracciones de RNA. Por RT-PCR se amplificó una región que contiene al intrón (de 83 pb) que en embriones control es cortado por *splicing*. Para comparación, se muestra una PCR a partir de ADN genómico (gADN).

que enmarcan a un intrón (Fig. 22), el cual normalmente debería procesarse por *splicing*, y que con el morfolino se espera que no se procese. El tamaño de la región enmarcada por los *primers* en el DNA genómico (gADN) es de 585 pb, y el del intrón es de 83 pb, por lo cual se espera que un transcrito procesado tenga un tamaño de 502 pb (Fig. 22). En embriones inyectados con el vehículo, con esos *primers* se obtiene un producto de amplificación de 502 pb (Fig. 22). En los embriones inyectados con el mo-sp se genera, además del producto de amplificación de 502 pb, uno de 585 pb (Fig. 22). La diferencia de tamaño de estos productos de amplificación está dada por el tamaño del intrón. Como un control adicional, se usaron los mismos *primers* para amplificar por PCR a partir de ADN genómico (en el cual se amplificaría el intrón). Resulta un producto de amplificación de 585 pb (Fig. 22). Por lo tanto, la inyección del mo-sp generó un ARNm aberrante.



**Fig. 23 Niveles de CTCF en los morfantes.** Se hicieron comparaciones del nivel de CTCF por Western Blot, entre los dos tipos de morfantes y embriones inyectados con vehículo (control). Por densitometría, se determinaron los niveles de CTCF normalizados con Tubulina, y se compararon con el nivel del control (media +/- desviación estándar; n=5).

Para comprobar el efecto en los niveles de la proteína CTCF, se realizaron extracciones de proteínas de embriones de 8 a 9 hpf (etapas entre 75 y 90% de epibolia, según Kimmel *et al.*, 1995). Por Western blot, se demostró que CTCF se reduce a niveles indetectables con el mo-aug (Fig. 23). Con el mo-sp se reduce a 25% del nivel que se detecta en embriones inyectados con el vehículo (Fig. 23). Estos datos demuestran que tanto el mo-aug como el mo-sp provocan la reducción de los niveles de CTCF.



**Fig. 24 Fenotipo observado en los morfantes.** Se inyectaron embriones con 4 ng de mo-aug o mo-sp. A las 24 hpf, se observaron y clasificaron en fenotipo silvestre, o fenotipo con malformaciones, leve o severo. Con las flechas se señalan: el prosencéfalo (a), el mesencéfalo (b), el rombencéfalo (c), el ojo (d), la notocorda (e), y los somitos (f). En la gráfica se muestra el porcentaje de embriones inyectados con vehículo (Control) o con morfolino, y que resultaron en un fenotipo leve o severo (\* $p=0.006$ , prueba  $t$  de una cola, con datos no apareados).

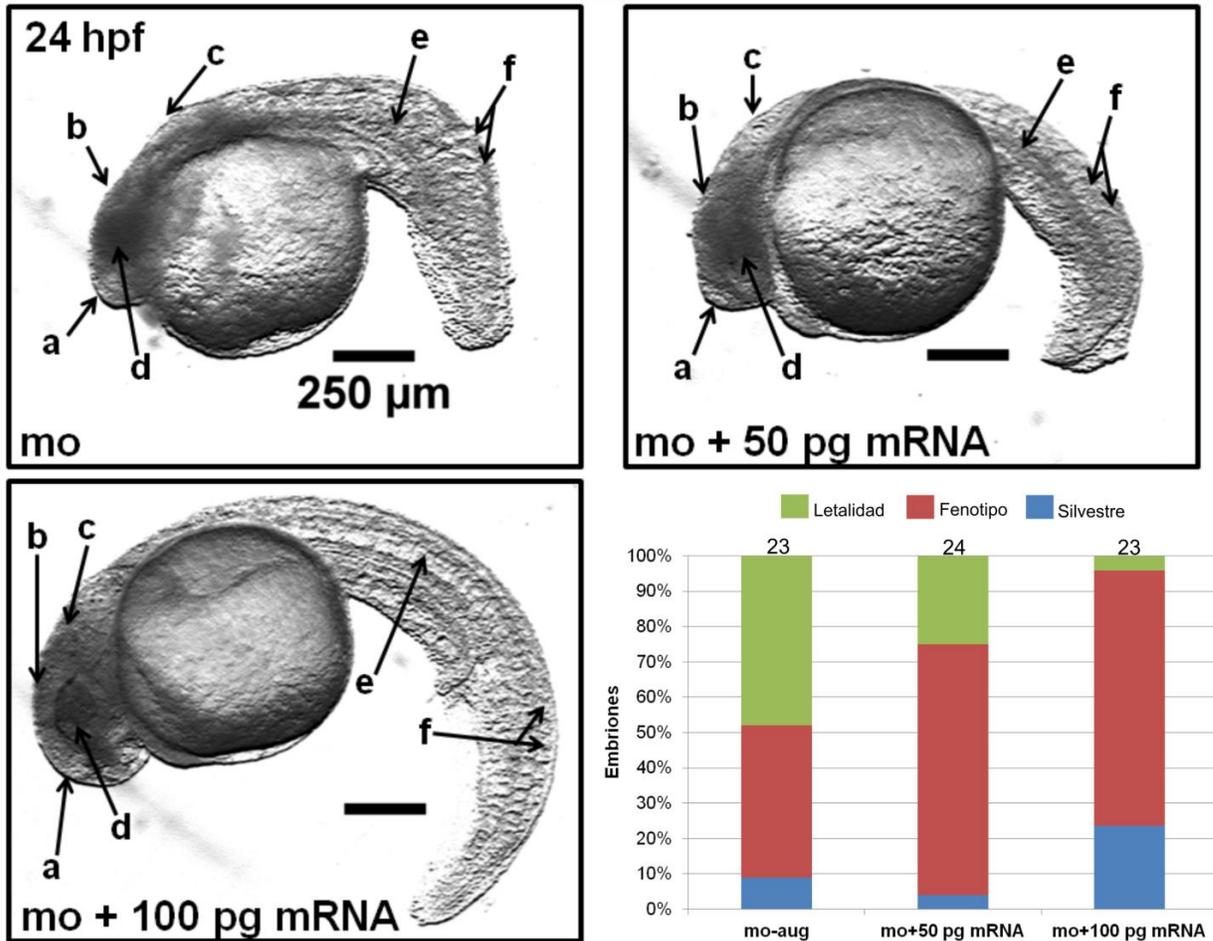
Una vez que el efecto de los morfolinós fue validado, se analizó el fenotipo que generaron. A las 24 hpf, en los embriones inyectados con el mo-aug o con el mo-sp se observan malformaciones en distintas regiones del cuerpo (Fig. 24): La longitud del cuerpo de los morfantes está disminuida. Además, en la porción anterior, los ojos presentan un tamaño más pequeño. Las distintas regiones del cerebro (el prosencéfalo, el mesencéfalo y el rombencéfalo), no tienen los límites definidos entre una y otra región. En la parte media, la notocorda no se nota recta, con la turgencia característica de esta estructura. Además, en la parte media y posterior, los somitos, que normalmente tienen un surco recto y definido entre uno y otro, y una forma característica de punta de flecha, no se forman adecuadamente. Peor aún, la cola se forma con una longitud menor (Fig. 24). Este dato es consistente con observaciones previas realizadas por nuestro grupo de investigación (Delgado-Olguín *et al.*, 2011). La severidad de estas malformaciones en los morfantes fue variable, pues algunos embriones presentaron un fenotipo moderado, otros un fenotipo severo, e inclusive hubo embriones inyectados con el morfolino que presentaron un fenotipo silvestre (Fig. 24). Esta variabilidad puede deberse a factores técnicos (por la variabilidad en la concentración del morfolino entre inyecciones), o a factores biológicos (diferencias en la expresividad del fenotipo). Aunque tanto el fenotipo moderado como el fenotipo severo se encontraron en embriones inyectados con cualquiera de los dos morfolinós, la proporción y la variabilidad del fenotipo fue estadísticamente significativo sólo con el mo-aug (Fig. 24). En conjunto, estas observaciones indican que la reducción de CTCF en etapas tempranas del desarrollo, genera malformaciones que afectan distintas regiones del cuerpo del embrión.

#### 10.4 El ARNm del gen *Ctcf* de humano rescata el fenotipo de los morfantes

Para demostrar la especificidad del fenotipo que genera el mo-aug, se llevaron a cabo experimentos de rescate, en los que se inyectó ARNm de *Ctcf* junto con el morfolino. El ARNm se generó por transcripción *in vitro*, a partir de una clona con la secuencia del gen *Ctcf* de humano. Se usó la secuencia de humano porque no se vería afectada por el morfolino, pues no contiene la secuencia blanco de éste.

Se inyectaron dos concentraciones diferentes del ARNm junto con el mo-aug, y se hicieron comparaciones con el fenotipo de embriones inyectados únicamente con el mo-aug (Fig. 25). Las diferencias que se encontraron fueron: 1) La proporción de letalidad se redujo respecto a los morfantes (Fig. 25); 2) Como eso generó que más embriones sobrevivieran, la proporción de embriones con fenotipo aumentó en el ensayo de rescate respecto a los morfantes (Fig. 25).

Sin embargo, la severidad del fenotipo fue visiblemente menor en comparación a los morfantes (Fig 25), pues fueron embriones con la cabeza menos reducida y con menos tejido necrótico, ojos con una malformación menos pronunciada, la notocorda con una longitud más grande y cercana hacia la porción más posterior, los somitos con bordes mejor delimitados, la cola menos reducida y en general, una forma más cercana a la de embriones normales (Fig. 25). Por lo tanto, el ARNm del gen *Ctcf* de humano rescató el fenotipo del mo-aug, demostrando la especificidad del fenotipo.

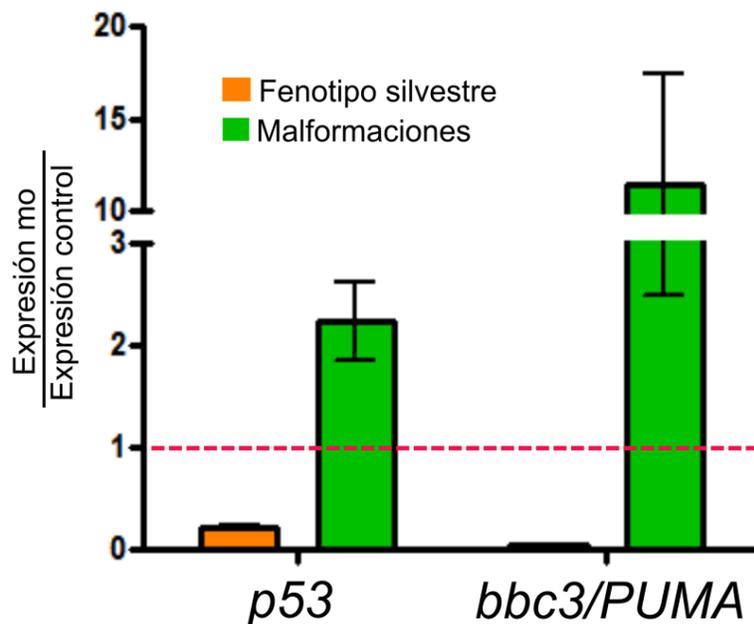


**Fig. 25 Rescate del fenotipo del morfolino aug.** Se inyectaron embriones con morfolino (mo) o con una combinación de mo y ARNm de *Ctcf* de humano (mo+ARNm). Se probaron dos cantidades diferentes de ARNm. Se graficó la proporción de letalidad, embriones con fenotipo de malformaciones (Fenotipo) o con fenotipo silvestre (Silvestre). Las fotografías son representativas de los embriones en la categoría de Fenotipo, y para comparación se indican: el prosencéfalo (a), el mesencéfalo (b), el telencéfalo (c), el ojo (d), la notocorda (e) y los somitos (f). Las cifras en la gráfica indican el número de embriones analizados.

### 10.5 CTCF se requiere para la correcta expresión de genes proapoptóticos

Como es sabido, CTCF es un regulador de la actividad de diversos genes, por lo que se decidió determinar si la expresión de algunos genes está desregulada en los morfantes. Como encontramos tejido necrótico en los morfantes (Fig. 24), además de

un tamaño reducido de varias zonas del cuerpo, bajo el razonamiento de que en los morfantes podría estar ocurriendo muerte celular exacerbada, decidimos analizar la expresión de dos genes proapoptóticos: *p53* y *PUMA* (que en el pez cebra tiene el nombre de *bbc3*). Se extrajo ARN total de embriones inyectados con el mo-aug en la etapa de epibolia al 90% y de embriones inyectados con el vehículo. Se evaluó la expresión de *p53* y *bbc3* por RT-qPCR y no fue posible detectar diferencias (datos no mostrados). Debido a que en embriones en la etapa de 90% de epibolia no se detecta alguna alteración morfológica, se decidió repetir el experimento con embriones en una etapa más avanzada, embriones de 24 hpf. Se removieron los embriones muertos, y los embriones que se inyectaron con el mo-aug y que sobrevivieron se separaron en dos grupos, los que presentaron el fenotipo mostrado en la Figura 24, y los que no



**Fig. 26 Expresión de genes proapoptóticos en los morfantes.** Se inyectaron embriones con el mo-aug. A las 24 hpf, se separaron en embriones con fenotipo silvestre y embriones con el fenotipo de malformaciones mostrado en la Fig. 24. Se extrajo ARN, y se comparó el nivel de expresión de los genes *p53* y *bbc3/PUMA*, por RT-qPCR, entre embriones inyectados con el vehículo (Control) y los morfantes (mo). Se graficó la media +/- desviación estándar; n=4.

presentaron fenotipo (“Malformaciones” y “Fenotipo silvestre”, Fig. 26). Por RT-qPCR se determinó que la expresión tanto de *p53* como de *bbc3* es elevada en los morfantes con malformaciones, en comparación con embriones inyectados con vehículo (Fig. 26). En cambio, en los morfantes con fenotipo silvestre, la expresión de ambos genes, *p53* y *bbc3*, está disminuida en comparación con embriones inyectados con el vehículo. Con estos datos se concluye que CTCF se requiere para el mantenimiento de la correcta expresión de los genes proapoptóticos *p53* y *bbc3*.

## 11. Discusión

A partir de los estudios que han arrojado datos acerca de la función de la proteína CTCF se ha logrado conocer que actúa a través de diferentes mecanismos moleculares de manera diferencial en distintos tejidos. Esta versatilidad se extiende además a sus funciones en varias especies animales, por lo que cual se ha descrito como un factor multifuncional (Ohlsson *et al.*, 2001; Heger *et al.*, 2012; Arzate-Mejía *et al.*, 2018). En este trabajo describimos la distribución de esta proteína y su requerimiento en etapas tempranas del desarrollo embrionario del pez cebra.

En diferentes especies, *ctcf* se expresa desde las etapas más tempranas del desarrollo, como el embrión de una célula, el blastómero y la gástrula. En estas etapas el ARNm tiene virtualmente los mismos niveles (evaluado por hibridaciones *in situ*) en diferentes tejidos. Pero más tarde, durante la organogénesis, se enriquece en algunas estructuras, especialmente en el sistema nervioso, si bien su expresión ocurre también en los demás tejidos (Burke *et al.*, 2002; Moon *et al.*, 2005; Delgado-Olguín *et al.*, 2011).

A partir de tinciones por inmunofluorescencia de embrión completo, mostramos la distribución de CTCF durante las primeras 30 horas del desarrollo embrionario del pez cebra (Fig. 17). Hasta donde sabemos, es la primera vez que se muestra la presencia de esta proteína en un organismo completo. Por lo cual, ahora podemos hacer comparaciones con los estudios anteriores que mostraron el patrón espacial del ARNm por hibridación *in situ*. Además, la presencia de la proteína se había

documentado en líneas celulares de origen diferente. Esto habría permitido inferir que CTCF está en cualquier tipo celular de un organismo (Roadmap Epigenomics Consortium *et al.*, 2015). Ahora, con nuestros resultados, demostramos que esto es cierto en el contexto del desarrollo embrionario.

Describimos su distribución desde el embrión de una célula hasta etapas en las que ya está avanzada la organogénesis, incluyendo la blástula, la gástrula y las primeras etapas de la organogénesis (Fig. 17). La presencia de la proteína es constante a través de diferentes etapas, y también su nivel es constante entre diferentes tejidos. Nosotros no encontramos niveles diferenciales de CTCF a nivel de proteína entre distintos tejidos en el pez cebra, lo cual es distinto a lo que se había documentado antes respecto a la expresión del ARNm (Pugacheva *et al.*, 2006; Delgado-Olguín *et al.*, 2011). Para realizar este análisis fue útil la generación de un anticuerpo específico para la proteína del pez (Fig. 16). Este anticuerpo podría utilizarse en un ensayo con una aproximación más cuantitativa: diseccionar diferentes tejidos embrionarios, para comparar los niveles de CTCF por Western blot. Al menos, con los datos que presentamos, encontramos que CTCF es una proteína ubicua, tanto en su patrón de distribución temporal como en el espacial.

Si bien ha sido poco estudiada la regulación transcripcional de CTCF, nuestros datos llevan a pensar que su regulación postranscripcional también podría ser un factor importante. Esto debido a que las dinámicas de cambio de la distribución espacial del transcrito en diferentes tejidos, no se reflejan en cambios en la distribución espacial de la proteína. En el contexto del desarrollo embrionario se ha demostrado previamente, a partir de análisis masivos de transcritos y las proteínas

para las que codifican, que no hay necesariamente una correlación entre el cambio del nivel de un ARNm y el de la proteína para la que codifica (Peshkin *et al.*, 2015). Es decir, una duplicación en los niveles del ARNm, no necesariamente se refleja en la duplicación de los niveles de la proteína. Habría que explorar varios factores que podrían regular postranscripcionalmente a CTCF, como la dinámica de traducción del ARNm, la posible agregación en gránulos de ARN (en los cuales, diferentes niveles del ARNm podrían detectarse, pero su traducción estaría bloqueada) o bien la dinámica de degradación de la proteína. Además, cabe la posibilidad de que estos procesos tengan una dinámica diferente entre diversos tejidos.

La presencia de la proteína en todos los tejidos y etapas del desarrollo embrionario sugiere fuertemente su requerimiento funcional constante. Nuestros resultados y los de otros grupos, demuestran que esto es cierto en el pez cebra (Fig. 18 y 24). CTCF es necesario para la viabilidad del embrión temprano, como lo encontramos en nuestro trabajo, y que ha sido confirmado en un trabajo reciente (Meier *et al.*, 2018) y se requiere, entre otros procesos, para la diferenciación hacia músculo (Delgado-Olguín *et al.*, 2011; Marsman *et al.*, 2014).

En embriones de pez cebra con niveles reducidos de CTCF cambia la expresión de un número importante de genes (Delgado-Olguín *et al.*, 2011; Meier *et al.*, 2018). Tal como se espera, varios procesos parecen estar desregulados. Pues en el fenotipo que encontramos, tanto en los embriones inyectados con el sistema CRISPR-Cas9 como en los morfantes, varias estructuras tienen malformaciones en los distintos ejes corporales del embrión (Fig. 18 y 24). Esto hace sentido con la distribución de la proteína, la cual se extiende por todo el embrión, por lo cual varios

tejidos sufren una malformación (Fig. 17). Esto es diferente a lo que suele ocurrir con proteínas que se distribuyen sólo en algunas regiones particulares del embrión, las cuales se espera que contribuyan a la formación de estructuras más restringidas.

Si bien nuestros resultados demuestran que la reducción de CTCF es una condición que no es viable para el desarrollo embrionario, es difícil inferir funciones específicas en los diferentes tejidos. Sería muy interesante conocer qué sucede al reducir los niveles de este factor en un tejido particular, por ejemplo, en el sistema nervioso. Esto es viable debido a que el sistema experimental ya se ha diseñado para el pez cebra, el cual consta de un plásmido con dos elementos: la secuencia para la Cas9 regulada por un promotor que se expresa sólo en un tejido particular, y la secuencia de un sgRNA regulada por un promotor ubicuo (Ablain *et al.*, 2015). Este sistema permitiría generar una mutación de manera restringida y dirigida sólo a un tejido, que sería aquél en el cual se expresa la Cas9. Llama la atención que, aunque el sistema ya se ha diseñado, aún no hay muchos trabajos en los que lo hayan utilizado, tal vez por limitantes experimentales, como la disponibilidad de construcciones con promotores específicos con los cuales se logre una inducción de la expresión de las Cas9 lo suficientemente fuerte.

El análisis de la función de CTCF podría ser complejo, aun dirigiéndolo a un tejido particular, como se ha demostrado en el sistema nervioso del embrión de ratón (Watson *et al.*, 2014). Se encontró que, en el telencéfalo y la retina, CTCF lleva a cabo una doble función: es un regulador de la diferenciación, y un factor de mantenimiento de las células precursoras neurales.

Algo que podría guiar el estudio de la función de CTCF en tejidos particulares, es el análisis de la serie de genes cuya expresión cambia en los morfantes *ctcf* de pez cebra. En un trabajo previo de nuestro grupo se detectaron 187 genes por análisis de microarreglos (Delgado-Olguín *et al.*, 2011). Otro grupo encontró recientemente cambios en la expresión de 269 genes por RNA-seq (Meier *et al.*, 2018). Entre los genes desregulados se encuentran *miogenina* y *wnt11a*, los cuales se requieren para la diferenciación del músculo, una función que ya ha sido demostrada en el pez. Un candidato interesante a explorar es *notch1a*, el cual tiene sitios consenso de unión a CTCF río arriba de su secuencia promotora (datos no mostrados). Este gen tiene funciones en la somitogénesis y la neurogénesis en el pez cebra. Se podría explorar la regulación de este gen a través de CTCF, y conocer si su función se relaciona con el fenotipo de diferenciación muscular, y el de la reducción del tamaño de la cabeza y el ojo que encontramos en este trabajo (Fig. 18 y 24).

Nos pareció importante mostrar el fenotipo de falta de función de CTCF mediante dos técnicas diferentes: el sistema CRISPR-Cas9 y la inyección de morfolidos (Fig. 18 y 24). El fenotipo es el mismo con los dos sistemas: embriones que no son viables y con una formación aberrante en los distintos ejes corporales. Cabe mencionar que ha habido controversia al comparar el resultado generado por inyección de morfolidos con los mutantes de algunos genes en varios trabajos: hay casos en los que el fenotipo del morfante no se parece al del mutante (Kok *et al.*, 2015). Se ha comentado que esto puede deberse a que, en los mutantes, la eliminación del gen es completa. Y que en esta condición, otros genes que tienen la misma función pueden activarse, por lo cual el fenotipo puede rescatarse

parcialmente (Blum *et al.*, 2015). En los morfolidos la estructura de los genes no se altera, por lo que estos mecanismos de rescate no se echan a andar.

En el caso de la inactivación de CTCF por una mutación, no parece factible el rescate de su función por algún otro gen, pues no se ha encontrado otra proteína con funciones similares a CTCF en vertebrados. Por lo cual, es posible que no exista otro gen que pudiera rescatar el fenotipo al generar mutaciones con el sistema CRISPR-Cas9. Si se toma en cuenta la evidencia que existe en el ratón, la eliminación del gen *CTCF* genera embriones que no son viables, lo cual es similar a nuestros resultados con el pez cebra (Moore *et al.*, 2012).

En invertebrados, la inactivación de CTCF tiene resultados diferentes. En la mosca, la formación de contactos de regulación y la organización del genoma dependen no sólo de CTCF, sino además de un grupo de proteínas con funciones similares (Nègre *et al.*, 2010). Recientemente se demostró que CTCF es dispensable durante el desarrollo embrionario en la mosca, posiblemente porque la eliminación de CTCF puede rescatarse por alguna de estas otras proteínas (Gambetta y Furlong, 2018). De aquí se pueden desprender preguntas interesantes acerca de la razón por la cual un sistema como el de la mosca, con más de una proteína arquitectónica, evolucionó en vertebrados hacia un sistema que depende aparentemente de una sola proteína para estas funciones.

Otro punto a considerar al comparar los fenotipos que encontramos con el sistema CRISPR-Cas9 y la inyección de morfolidos, es que los embriones inyectados con el sistema CRISPR-Cas9 son embriones mosaico (Fig. 20). Sería

interesante caracterizar alelos nulos y analizar su fenotipo. Además, se podrían combinar los dos sistemas: inyectar morfolidos en embriones mutantes para un alelo nulo, con el fin de interrumpir la contribución materna en embriones que no expresan el componente cigótico, y poder analizar el efecto de la ausencia completa de CTCF.

Aunque en nuestros resultados los embriones con niveles reducidos de CTCF tienen una forma aberrante, es posible distinguir los ejes corporales (por ejemplo, dónde se encuentra la región anterior y la posterior; Fig. 24). Este resultado se parece a lo que se encontró al reducir esta misma proteína en las regiones precursoras de las extremidades en el ratón (Soshnikova *et al.*, 2010). Se formó una extremidad con una forma aberrante, aunque los ejes se distinguen. En este contexto se demostró que CTCF se requiere como un factor de supervivencia celular. En nuestros datos encontramos algo similar: con la reducción de CTCF se sobreactivan los genes proapoptóticos *p53* y *bbc3/PUMA* (Fig. 26). Algo interesante por averiguar es el patrón temporal y espacial de la muerte celular que deriva de la reducción de CTCF.

Con los experimentos de rescate se demostró que el fenotipo de los morfolidos es específico (Fig. 25). Pero además, al rescatar con la secuencia de *CTCF* de humano, se demostró que la conservación de la proteína no está dada únicamente a nivel de identidad de aminoácidos, sino también a nivel de la función molecular. Esto ya se había demostrado para la función en la regulación de factores de diferenciación del músculo, y nuestros datos demuestran que esto también es cierto para el desarrollo embrionario temprano (Delgado-Olguín *et al.*, 2011).

Así como es interesante demostrar la conservación de la función de CTCF entre diferentes especies, resultaría atractivo demostrar también las diferencias. En un trabajo de nuestro grupo se encontró que CTCF de humano une ARN mediante los dedos de zinc 10 y 11 y el extremo carboxilo-terminal (Saldaña-Meyer *et al.*, 2014). Si bien en el pez cebra los dedos de zinc tienen identidad de secuencia con el humano (Tabla 2), el extremo carboxilo-terminal no la tiene (Pugacheva *et al.*, 2006). Esto podría tener impacto en las funciones reguladoras de CTCF. Por ejemplo, en el humano, CTCF es un regulador positivo de la expresión de *P53* (Soto-Reyes y Recillas-Targa, 2010; Saldaña-Meyer *et al.*, 2014). En esta regulación es importante el dominio de unión a ARN. Con los resultados de este trabajo encontramos que en el pez, CTCF actúa como un regulador negativo de *p53* (Fig. 26). Un aspecto interesante a explorar es si esta diferencia entre regulación positiva o negativa, está dada por el dominio de unión a ARN.

Habría que explorar también los mecanismos de regulación directa de CTCF en diferentes *loci* en el pez cebra. Por ejemplo, se han detectado sitios consenso para la unión de CTCF en el *locus* de *p53* (Rhodes *et al.*, 2010). Sería interesante demostrar que CTCF se posiciona en estos sitios y buscar si se establecen contactos entre regiones reguladoras de este gen. Tal vez esto podría explicar los cambios diferenciales de expresión de *p53* en los morfantes (Fig. 26). Según estos datos, cuando separamos embriones inyectados con morfolino entre embriones con fenotipo y embriones sin fenotipo, en los primeros hay una sobreactivación de *p53* (Fig. 26). Pero en los segundos hay una expresión reducida de *p53*. Podría explorarse la posibilidad de que aunque en estos dos grupos de embriones

inyectados con morfolino el nivel global de CTCF se reduce, el posicionamiento de las moléculas residuales de CTCF es diferente. Y que esto pudiera influir de manera distinta sobre las regiones reguladoras, en un contexto resultando en la sobreactivación de *p53*, y en el otro en la expresión reducida de *p53*.

En conclusión, en este trabajo de investigación demostramos que CTCF es una proteína esencial para el desarrollo embrionario del pez cebra. Que es una proteína ubicua, tanto en lo temporal como en lo espacial. Se requiere para la formación de estructuras en los diferentes ejes corporales del embrión. Y que regula genes necesarios para la supervivencia celular.

Con nuestros datos establecemos un panorama desde el cual podría explorarse los mecanismos por los cuales CTCF regula diferentes genes en contextos fisiológicos diversos. En una aproximación dirigida hacia el conocimiento de su papel específico en los diferentes tejidos, para tener una idea más fina de sus funciones durante el desarrollo embrionario.

## 12. Tablas suplementarias

**Tabla suplementaria 12.1** Fenotipo de embriones inyectados con el sistema CRISPR-Cas9 a las 12 hpf (en referencia a la figura 18).

	Letalidad		Fenotipo		Normal	
	n/x	%	n/x	%	n/x	%
No inyectados	0/24	0	0/24	0	24/24	100
CRISPR-Cas9	6/30	20	14/30	47	10/30	33

**Tabla suplementaria 12.2** Fenotipo de embriones inyectados con el sistema CRISPR-Cas9 a las 24 hpf (en referencia a la figura 18).

Grupo experimental	Letalidad	
	x/n	%
No inyectados	12/23	52
ARNm Cas9 (2x)	6/17	35
ARNm Cas9 + sgRNA-CTCF	20/22	91
ARNm Cas9 + sgRNA-CTCF (2x)	15/18	83

**Tabla suplementaria 12.3** Letalidad con distintas concentraciones de morfolino (en referencia a la Fig. 21).

	Experimento 1.		Experimento 1.		Experimento 2.	
	Revisión a las 48 hpf.		Revisión a las 72 hpf.		Revisión a las 48 hpf.	
	x/n	%	x/n	%	x/n	%
No inyectados	5/19	26.3	6/19	31.6	1/14	7.1
mo-control	2/2	100	2/2	100	0/34	0
mo-sp 2ng	9/39	23	9/39	23	2/24	8.3
mo-sp 4ng	29/40	72.5	31/40	77.5	5/31	16.1
mo-aug 2ng	0/8	0	0/8	0	0/22	0
mo-aug 4ng	2/6	33.3	5/6	83.3	2/25	8

### 13. Bibliografía

- Ablain, J., Durand, E. M., Yang, S., Zhou, Y. y Zon, L. I. (2015) A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish, *Developmental Cell*, 32:756–764
- Anderson, C. y Stern, C. D. (2016) *Organizers in Development*. 1a ed, *Current Topics in Developmental Biology*. Elsevier Inc. doi: 0.1016/bs.ctdb.2015.11.023.
- Arzate-Mejía, R. G., Recillas-Targa, F. y Corces, V. G. (2018) Developing in 3D: the role of CTCF in cell differentiation, *Development*, 145:137729
- Bartolomei, M. S., Zemel, S. y Tilghman, S. M. (1991) Parental imprinting of the mouse H19 gene, *Nature*, 351:153–155
- Bedell, V. M., Westcot, S. E. y Ekker, S. C. (2011) Lessons from morpholino-based screening in zebrafish, *Briefings in Functional Genomics*, 10:181–188
- Bell, A. C. y Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene, *Nature*, 405:482–485
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., *et al.* (2006) A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells, *Cell*, 125:315–326
- Bessa, J., Tena, J. J., De La Calle-Mustienes, E., Fernández-Miñán, A., Naranjo, S., *et al.* (2009) Zebrafish Enhancer Detection (ZED) vector: A new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish, *Developmental Dynamics*, 238:2409–2417
- Bilder, D. (2016) The Maturation of Development, *Developmental Cell*, 38:569–570
- Billinton, N. y Knight, A. W. (2001) Seeing the wood through the trees: A review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence, *Analytical Biochemistry*, 291:175–197
- Blum, M., De Robertis, E. M., Wallingford, J. B. y Niehrs, C. (2015) Morpholinos: Antisense and Sensibility, *Developmental Cell*, 35:145–149
- Boettiger, A. N. y Levine, M. (2009) Synchronous and Stochastic *Drosophila* Embryo, *Science*, 325:23–25
- Bonev, B. y Cavalli, G. (2016) Organization and function of the 3D genome, *Nature Reviews Genetics*, 17:661–678
- Bonn, S., Zinzen, R. P., Girardot, C., Gustafson, E. H., Perez-Gonzalez, A., *et al.* (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development, *Nature Genetics*, 44:148–156
- Borsos, M. y Torres-Padilla, M. E. (2016) Building up the nucleus: Nuclear organization in the establishment of totipotency and pluripotency during mammalian development, *Genes and Development*, 30:611–621

- Burke, L. J., Hollemann, T., Pieler, T. y Renkawitz, R. (2002) Molecular cloning and expression of the chromatin insulator protein CTCF in *Xenopus laevis*, *Mechanisms of Development*, 113:95–98
- Burton, A. y Torres-Padilla, M.-E. (2014) Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis, *Nature Reviews Molecular Cell Biology*, 15:723–735
- Callinan, P. A. y Feinberg, A. P. (2006) The emerging science of epigenomics, *Human molecular genetics*, 15 Spec:95–101
- Campos-Ortega, J. A. y Hartenstein, V. (1997) *The Embryonic Development of Drosophila melanogaster*. 2a ed. Editado por J. A. Campos-Ortega y V. Hartenstein. Springer-Verlag Berlin Heidelberg.
- Cantone, I. y Fisher, A. G. (2013) Epigenetic programming and reprogramming during development, *Nature Structural and Molecular Biology*, 20:282–289
- Cavalli, G. y Misteli, T. (2013) Functional implications of genome topology, *Nature Structural and Molecular Biology*, 20:290–299
- Chandra, T., Ewels, P. A., Schoenfelder, S., Furlan-Magaril, M., Wingett, S. W., et al. (2015) Global reorganization of the nuclear landscape in senescent cells, *Cell Reports*, 10:471–484
- Chen, X., Lu, C., Prado, J. R. M., Eun, S. H. y Fuller, M. T. (2011) Sequential changes at differentiation gene promoters as they become active in a stem cell lineage, *Development*, 138:2441–2450
- Ciabrelli, F. y Cavalli, G. (2015) Chromatin-driven behavior of topologically associating domains, *Journal of Molecular Biology*, 427:608–625
- Di Croce, L. y Helin, K. (2013) Transcriptional regulation by Polycomb group proteins, *Nature Structural & Molecular Biology*, 20:1147–1155
- Davidovich, C. y Cech, T. R. (2015) The recruitment of chromatin modifiers by long noncoding RNAs: Lessons from PRC2, *RNA*, 21:2007–2022
- Daxinger, L. y Whitelaw, E. (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals, *Nature Reviews Genetics*, 13:153–162
- DeChiara, T. M., Robertson, E. J. y Efstratiadis, A. (1991) Parental imprinting of the mouse insulin-like growth factor II gene, *Cell*, 64:849–859
- Dekker, J., Rippe, K., Dekker, M. y Kleckner, N. (2002) Capturing Chromosome Conformation, *Science*, 295:1306–1311
- Dekker, J. y Heard, E. (2015) Structural and functional diversity of Topologically Associating Domains, *FEBS Letters*, 589:2877–2884
- Dekker, J. y Mirny, L. (2016) The 3D Genome as Moderator of Chromosomal Communication, *Cell*, 164:1110–1121
- Delgado-Olguín, P., Brand-Arzamendi, K., Scott, I. C., Jungblut, B., Stainier, D. Y., et al. (2011) CTCF promotes muscle differentiation by modulating the activity of myogenic regulatory factors, *Journal of Biological Chemistry*, 286:12483–12494

- Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., *et al.* (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping, *Cell*, 158:849–860
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., *et al.* (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature*, 485:376–380
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., *et al.* (2015) Chromatin architecture reorganization during stem cell differentiation, *Nature*, 518:331–336
- Doğan, E. S. y Liu, C. (2018) Three-dimensional chromatin packing and positioning of plant genomes, *Nature Plants*, 4:521–529
- Domcke, S., Bardet, A. F., Adrian Ginno, P., Hartl, D., Burger, L., *et al.* (2015) Competition between DNA methylation and transcription factors determines binding of NRF1, *Nature*, 528:575–579
- Donohoe, M. E., Silva, S. S., Pinter, S. F., Xu, N. y Lee, J. T. (2009) The pluripotency factor Oct4 interacts with CTCF and also controls X-chromosome pairing and counting, *Nature*, 460:128–132
- Eisen, J. S. y Smith, J. C. (2008) Controlling morpholino experiments: don't stop making antisense, *Development*, 135:1735–1743
- Eissenberg, J. C. y Shilatifard, A. (2010) Histone H3 lysine 4 (H3K4) methylation in development and differentiation, *Developmental Biology*, 339:240–249
- Ellis, T. H. N., Hofer, J. M. I., Timmerman-Vaughan, G. M., Coyne, C. J. y Hellens, R. P. (2011) Mendel, 150 years on, *Trends in Plant Science*, 16:590–596
- Fabre, P. J., Leleu, M., Mormann, B. H., Lopez-Delisle, L., Noordermeer, D., *et al.* (2017) Large scale genomic reorganization of topological domains at the HoxD locus, *Genome Biology*, 18:1–15
- Fatica, A. y Bozzoni, I. (2014) Long non-coding RNAs: New players in cell differentiation and development, *Nature Reviews Genetics*, 15:7–21
- Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., *et al.* (1996) An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes, *Molecular and Cellular Biology*, 16:2802–2813
- Finan, K., Cook, P. R. y Marenduzzo, D. (2011) Non-specific (entropic) forces as major determinants of the structure of mammalian chromosomes, *Chromosome Research*, 19:53–61
- Franco, M. M., Prickett, A. R. y Oakey, R. J. (2014) The Role of CCCTC-Binding Factor (CTCF) in Genomic Imprinting, Development, and Reproduction, *Biology of Reproduction*, 91:1–9
- Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., *et al.* (2014) Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs, *PLoS ONE*, 9:5–12

- Gambetta, M. C. y Furlong, E. E. M. (2018) The Insulator Protein CTCF Is Required for Correct Hox Gene Expression, but Not for Embryonic Development in *Drosophila*, *Genetics*, 301350.2018
- García-González, E., Escamilla-Del-Arenal, M., Arzate-Mejía, R. y Recillas-Targa, F. (2016) Chromatin remodeling effects on enhancer activity, *Cellular and Molecular Life Sciences*, 73:2897–2910
- García-González, E. y Recillas-Targa, F. (2014) A regulatory element affects the activity and chromatin structure of the chicken  $\alpha$ -globin 3' enhancer, *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1233–1241
- Gavrilov, A., Eivazova, E., Pirozhkova, I., Lipinski, M., Razin, S., *et al.* (2009) Chromosome Conformation Capture (from 3C to 5C) and Its ChIP-Based Modification, en Collas, P. (ed.) *Chromatin Immunoprecipitation Assays. Methods in Molecular Biology (Methods and Protocols)*. Humana Press, pp. 171–188
- 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
- Gilbert, S. F. (2010) Developmental Anatomy, en *Developmental Biology*. 9th ed. Sinauer Associates Inc., pp. 5–30.
- Goldberg, A. D., Banaszynski, L. A., Noh, K. M., Lewis, P. W., Elsaesser, S. J., *et al.* (2010) Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions, *Cell*, 140:678–691
- Gómez-Díaz, E. y Corces, V. G. (2014) Architectural proteins: Regulators of 3D genome organization in cell fate, *Trends in Cell Biology*, 24:703–711
- González-Buendía, E., Escamilla-Del-Arenal, M., Pérez-Molina, R., Tena, J. J., Guerrero, G., *et al.* (2015) A novel chromatin insulator regulates the chicken folate receptor gene from the influence of nearby constitutive heterochromatin and the  $\beta$ -globin locus, *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1849:955–965
- Gräff, J. y Mansuy, I. M. (2008) Epigenetic codes in cognition and behaviour, *Behavioural Brain Research*, 192:70–87
- Groh, S. y Schotta, G. (2017) Silencing of endogenous retroviruses by heterochromatin, *Cellular and Molecular Life Sciences*, 74:2055–2065
- Guo, C., Yoon, H. S., Franklin, A., Jain, S., Ebert, A., *et al.* (2011) CTCF-binding elements mediate control of V(D)J recombination, *Nature*, 477:424–430
- Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., *et al.* (2015) CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function, *Cell*, 162:900–910
- Gurdon, J. B. (1987) Embryonic induction--molecular prospects, *Development*, 99:285–306
- Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., *et al.*

- (1996) The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*, *Development*, 123:1–36
- Hamburger, V. y Hamilton, H. L. (1951) A series of normal stages in the development of the chick embryo, *Journal of Morphology*, 88:49–92
- Heger, P., Marin, B., Bartkuhn, M., Schierenberg, E. y Wiehe, T. (2012) The chromatin insulator CTCF and the emergence of metazoan diversity, *Proceedings of the National Academy of Sciences*, 109:17507–17512
- Heger, P., Marin, B. y Schierenberg, E. (2009) Loss of the insulator protein CTCF during nematode evolution, *BMC Molecular Biology*, 10:84
- Henikoff, S. y Shilatifard, A. (2011) Histone modification: Cause or cog?, *Trends in Genetics*, 27:389–396
- Herold, M., Bartkuhn, M. y Renkawitz, R. (2012) CTCF: insights into insulator function during development, *Development*, 139:1045–1057
- Hirayama, T., Tarusawa, E., Yoshimura, Y., Galjart, N. y Yagi, T. (2012) CTCF Is Required for Neural Development and Stochastic Expression of Clustered Pcdh Genes in Neurons, *Cell Reports*, 2:345–357
- Holtzman, N. G., Kathryn Iovine, M., Liang, J. O. y Morris, J. (2016) Learning to fish with genetics: A primer on the vertebrate model *Danio rerio*, *Genetics*, 203:1069–1089
- Holwerda, S. y Laats, W. De (2012) Chromatin loops , gene positioning , and gene expression, *Frontiers in Genetics*, 3:1–13
- Huang, P., Keller, C. A., Giardine, B., Grevet, J. D., Davies, J. O. J., *et al.* (2017) Comparative analysis of three-dimensional chromosomal architecture identifies a novel fetal hemoglobin regulatory element, *Genes and Development*, 31:1704–1713
- Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., *et al.* (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system, *Nature Biotechnology*, 31:227–229
- Ingham, P. W. (1997) Zebrafish genetics and its implications for understanding vertebrate development, *Human Molecular Genetics*, 6:1755–1760
- Irie, N. y Kuratani, S. (2014) The developmental hourglass model: a predictor of the basic body plan?, *Development*, 141:4649–4655
- Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., *et al.* (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores, *Nature Genetics*, 41:178–186
- Jao, L.-E., Wentz, S. R. y Chen, W. (2013) Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system, *Proceedings of the National Academy of Sciences*, 110:13904–13909
- Jiang, C. y Pugh, B. F. (2009) Nucleosome positioning and gene regulation: Advances through genomics, *Nature Reviews Genetics*, 10:161–172

- Johannes, S., Holwerda, B. y Laat, W. De (2013) CTCF\_ the protein, the binding partners, the binding sites and their chromatin loops, *Philosophical Transactions of the Royal Society B*, 368:20120369
- Johnson, A. D., Richardson, E., Bachvarova, R. F. y Crother, B. I. (2011) Evolution of the germ line-soma relationship in vertebrate embryos, *Reproduction*, 141:291–300
- Jones, P. L., Veenstra, G. J. C., Wade, P. A., Vermaak, D., Kass, S. U., *et al.* (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, *Nature Genetics*, 19:187–191
- Kent, B., Magnani, E., Walsh, M. J. y Sadler, K. C. (2016) UHRF1 regulation of Dnmt1 is required for pre-gastrula zebrafish development, *Developmental Biology*, 412:99–113
- Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. y Schilling, T. F. (1995) Stages of embryonic development of the zebrafish, *Developmental dynamics*, 203:253–310
- Klenova, E. M., Nicolas, R. H., Paterson, H. F., Carne, A. F., Heath, C. M., *et al.* (1993) CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms, *Molecular and cellular biology*, 13:7612–24
- Klenova, E. M., Chernukhin, I. V, El-kady, A., Lee, R. E., Pugacheva, E. M., *et al.* (2001) Functional Phosphorylation Sites in the C-Terminal Region of the Multivalent Multifunctional Transcriptional Factor CTCF, 21:2221–2234
- Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., *et al.* (2015) Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish, *Developmental Cell*, 32:97–108
- Kopp, F. y Mendell, J. T. (2018) Functional Classification and Experimental Dissection of Long Noncoding RNAs, *Cell*, 172:393–407
- Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., *et al.* (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2, *Proceedings of the National Academy of Sciences*, 103:10684–10689
- Lai, W. K. M. y Pugh, B. F. (2017) Understanding nucleosome dynamics and their links to gene expression and DNA replication, *Nature Reviews Molecular Cell Biology*, 18:548–562
- Lawrence, C. (2016) *New frontiers for zebrafish management*, *Methods in Cell Biology*, 135:483-508
- Leitch, H. G., Tang, W. W. C. y Surani, M. A. (2013) *Primordial Germ-Cell Development and Epigenetic Reprogramming in Mammals*. 1a ed, *Current Topics in Developmental Biology*
- Levin, M., Anavy, L., Cole, A. G., Winter, E., Mostov, N., *et al.* (2016) The mid-developmental transition and the evolution of animal body plans, *Nature*,

531:637–641

- Lim, J. P. y Brunet, A. (2013) Bridging the transgenerational gap with epigenetic memory, *Trends in Genetics*, 29:176–186
- Link, V., Shevchenko, A. y Heisenberg, C. P. (2006) Proteomics of early zebrafish embryos, *BMC Developmental Biology*, 13:1–9
- Lobanenkov, V. V, Nicolas, R. H., Adler, V. V, Paterson, H., Klenova, E. M., *et al.* (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene, *Oncogene*, 5:1743–53
- Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., *et al.* (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions, *Cell*, 161:1012–1025
- Maienschein, J. (2016) Garland Allen, Thomas Hunt Morgan, and Development, *Journal of the History of Biology*, 49:587–601
- Mani, M. (2013) *Spemann-Mangold organizers, What's new - to me...* Disponible en: <https://madhavmani.wordpress.com/> (Consultado: el 22 de octubre de 2018).
- Marahrens, Y., Panning, B., Dausman, J., Strauss, W. y Jaenisch, R. (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis, *Genes and Development*, 11:156–166
- Marsman, J., O'Neill, A. C., Kao, B. R. Y., Rhodes, J. M., Meier, M., *et al.* (2014) Cohesin and CTCF differentially regulate spatiotemporal runx1 expression during zebrafish development, *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1839:50–61
- Martin, D., Pantoja, C., Fernández-Miñán, A., Valdes-Quezada, C., Moltó, E., *et al.* (2011) Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes, *Nature Structural and Molecular Biology*, 18:708–714
- Maze, I., Noh, K. M., Soshnev, A. A. y Allis, C. D. (2014) Every amino acid matters: Essential contributions of histone variants to mammalian development and disease, *Nature Reviews Genetics*, 15:259–271
- Meier, K. y Recillas-Targa, F. (2017) New insights on the role of DNA methylation from a global view, *Frontiers in bioscience*, 22:644–668
- Meier, M., Grant, J., Dowdle, A., Thomas, A., Gerton, J., *et al.* (2018) Cohesin facilitates zygotic genome activation in zebrafish, *Development*, 145:156521
- Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heintz, N., *et al.* (2007) The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning, *EMBO Journal*, 26:4203–4214
- Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. y Valen, E. (2014) CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing, *Nucleic Acids Research*, 42:401–407

- Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., *et al.* (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., *et al.* (2012) Loss of maternal CTCF is associated with peri-implantation lethality of CTCF null embryos, *PLoS ONE*, 7:10.1371
- Morriss, G. R. y Cooper, T. A. (2017) Protein sequestration as a normal function of long noncoding RNAs and a pathogenic mechanism of RNAs containing nucleotide repeat expansions, *Human Genetics*, 136:1247–1263
- Nakahashi, H., Kwon, K. R. K., Resch, W., Vian, L., Dose, M., *et al.* (2013) A Genome-wide Map of CTCF Multivalency Redefines the CTCF Code, *Cell Reports*, 3:1678–1689
- Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., *et al.* (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, *Nature*, 393:386–389
- Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., *et al.* (2015) CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation, *Science*, 347:1017–1021
- Nègre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., *et al.* (2010) A comprehensive map of insulator elements for the *Drosophila* genome, *PLoS Genetics*, 6:10.1371
- Noordermeer, D., Leleu, M., Schorderet, P., Joye, E., Chabaud, F., *et al.* (2014) Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci, *eLife*, 2014:1–21
- Noordermeer, D. y Duboule, D. (2013) Chromatin looping and organization at developmentally regulated gene loci, *Wiley Interdisciplinary Reviews: Developmental Biology*, 2:615–630
- Nott, A., Cheng, J., Gao, F., Lin, Y. T., Gjoneska, E., *et al.* (2016) Histone deacetylase 3 associates with MeCP2 to regulate FOXO and social behavior, *Nature Neuroscience*, 19:1497–1505
- Nüsslein-volhard, C. y Wieschaus, E. (1980) Mutations affecting segment number and polarity in *Drosophila*, *Nature*, 287:795-801
- 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
- Ong, C. T. y Corces, V. G. (2014) CTCF: An architectural protein bridging genome topology and function, *Nature Reviews Genetics*, 15:234–246
- Pagano, J. F. B., Rauwerda, H., de Leeuw, W. C., Wackers, P., de Jong, M., *et al.* (2016) Transcriptome data on maternal RNA of 24 individual zebrafish eggs from five sibling mothers, *Data in Brief*, 8:69–72
- Pasque, V., Gillich, A., Garrett, N. y Gurdon, J. B. (2011) Histone variant macroH2A

- confers resistance to nuclear reprogramming, *EMBO Journal*, 30:2373–2387
- Patton, E. E. y Zon, L. I. (2001) The art and design of genetic screens: zebrafish, *Nature Reviews Genetics*, 2:956–966
- Pauli, T., Vedder, L., Dowling, D., Petersen, M., Meusemann, K., *et al.* (2016) Transcriptomic data from panarthropods shed new light on the evolution of insulator binding proteins in insects, *BMC Genomics*, 17:1–10
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., *et al.* (2010) Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation, *Molecular Cell*, 38:603–613
- Perino, M. y Veenstra, G. J. C. (2016) Chromatin Control of Developmental Dynamics and Plasticity, *Developmental Cell*, 38:610–620
- Peshkin, L., Wühr, M., Pearl, E., Haas, W., Freeman, R. M., *et al.* (2015) On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development, *Developmental Cell*, 35:383–394
- Peterson, S. M. y Freeman, J. L. (2009) RNA Isolation from Embryonic Zebrafish and cDNA Synthesis for Gene Expression Analysis, *Journal of Visualized Experiments*, 1470.
- Phillips, J. E. y Corces, V. G. (2009) CTCF: Master Weaver of the Genome, *Cell*, 137:1194–1211
- Pieterman, C. R. C., Conemans, E. B., Dreijerink, K. M. A., Laat, J. M. De y Timmers, H. T. M. (2014) Thoracic and duodenopancreatic neuroendocrine tumors in multiple endocrine neoplasia type 1: natural history and function of *menin* in tumorigenesis, *Endocrine-Related Cancer*, 21:121-142
- Portela, A. y Esteller, M. (2010) Epigenetic modifications and human disease, *Nature Biotechnology*, 28:1057–1068
- Pugacheva, E. M., Kwon, Y. W., Hukriede, N. A., Pack, S., Flanagan, P. T., *et al.* (2006) Cloning and characterization of zebrafish CTCF: Developmental expression patterns, regulation of the promoter region, and evolutionary aspects of gene organization, *Gene*, 375:26–36
- Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, *Cell*, 159:1665–1680
- Reddington, J. P., Pennings, S. y Meehan, R. R. (2013) Non-canonical functions of the DNA methylome in gene regulation, *Biochemical Journal*, 451:13–23
- Rhodes, J. M., Bentley, F. K., Print, C. G., Dorsett, D., Misulovin, Z., *et al.* (2010) Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved, *Developmental Biology*, 344:637–649
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., *et al.* (2015) Integrative analysis of 111 reference human epigenomes, *Nature*, 518:317–329

- Roh, T.-Y., Cuddapah, S., Cui, K. y Zhao, K. (2006) The genomic landscape of histone modifications in human T cells, *Proceedings of the National Academy of Sciences of the United States of America*, 103:15782–7
- Roy, A. R., Ahmed, A., DiStefano, P. V., Chi, L., Khyzha, N., *et al.* (2018) The transcriptional regulator CCCTC-binding factor limits oxidative stress in endothelial cells, *Journal of Biological Chemistry*, 293:8449–8461
- Ruthenburg, A. J., Allis, C. D. y Wysocka, J. (2007) Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark, *Molecular Cell*, 25:15–30
- Saldaña-Meyer, R., Gonzalez-Buendia, E., Guerrero, G., Narendra, V., Bonasio, R., *et al.* (2014) CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, *Wrap53*, *Genes & Development*, 28:723–734
- Sanborn, A. L., Rao, S. S. P., Huang, S.-C., Durand, N. C., Huntley, M. H., *et al.* (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes, *Proceedings of the National Academy of Sciences*, 112:6456–6465
- Sathananthan, H., Pera, M. y Trounson, A. (2002) The fine structure of human embryonic stem cells, *Reproductive biomedicine online*, 4:56–61
- Schier, A. F. (2013) Zebrafish earns its stripes, *Nature*, 496:443–444.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., *et al.* (2012) Fiji: An open-source platform for biological-image analysis, *Nature Methods*, 9:676–682
- Sexton, T. y Cavalli, G. (2015) The role of chromosome domains in shaping the functional genome, *Cell*, 160:1049–1059
- Shilatifard, A. (2006) Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression, *Annual Review of Biochemistry*, 75:243–269
- Skinner, M. K. (2011) Role of epigenetics in developmental biology and transgenerational inheritance, *Birth Defects Research Part C - Embryo Today: Reviews*, 93:51–55
- Smith, Z. D. y Meissner, A. (2013) DNA methylation: Roles in mammalian development, *Nature Reviews Genetics*, 14:204–220
- Soshnikova, N., Montavon, T., Leleu, M., Galjart, N. y Duboule, D. (2010) Functional Analysis of CTCF During Mammalian Limb Development, *Developmental Cell*, 19:819–830
- Soto-Reyes, E. y Recillas-Targa, F. (2010) Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines, *Oncogene*, 29:2217–2227
- Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., *et al.* (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus, *Genes and Development*, 20:2349–2354

- Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J. G., *et al.* (2011) The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA, *Genes and Development*, 25:1371–1383
- Stainier, D. Y. R., Raz, E., Lawson, N. D., Ekker, S. C., Burdine, D., *et al.* (2017) Guidelines for morpholino use in zebrafish, *PLoS Genetics*, 13:1007000
- Stancheva, I. (2011) Revisiting heterochromatin in embryonic stem cells, *PLoS Genetics*, 7:7–9
- Strahl, B. D. y Allis, C. D. (2000) The language of covalent histone modifications, *Nature*, 403:41–45
- Sun, X. J., Xu, P. F., Zhou, T., Hu, M., Fu, C. T., *et al.* (2008) Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes, *PLoS ONE*, 3:10.1371
- Talbert, P. B. y Henikoff, S. (2010) Histone variants ancient wrap artists of the epigenome, *Nature Reviews Molecular Cell Biology*, 11:264–275
- Tena, J. J., Alonso, M. E., De La Calle-Mustienes, E., Splinter, E., De Laat, W., *et al.* (2011) An evolutionarily conserved three-dimensional structure in the vertebrate lrx clusters facilitates enhancer sharing and coregulation, *Nature Communications*, 2:1–9
- Toruno, C., Carbonneau, S., Stewart, R. A. y Jette, C. (2014) Interdependence of bad and puma during ionizing-radiation-induced apoptosis, *PLoS ONE*, 9:10.1371
- Trevarrow, B. (2004) *Zebrafish Facilities for Small and Large Laboratories*, *Methods In Cell Biology*. 2a ed. Editado por H. W. I. Detrich, M. Westerfield, y L. I. Zon. Elsevier Academic Press
- Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigó, R. y Johnson, R. (2018) Towards a complete map of the human long non-coding RNA transcriptome, *Nature Reviews Genetics*, 19:535-548
- Valdes-Quezada, C., Arriaga-Canon, C., Fonseca-Guzmán, Y., Guerrero, G. y Recillas-Targa, F. (2013) CTCF demarcates chicken embryonic  $\alpha$ -globin gene autonomous silencing and contributes to adult stage-specific gene expression, *Epigenetics*, 8:827–838
- Vaquerizas, J. M. y Torres-Padilla, M. E. (2016) Developmental biology: Panoramic views of the early epigenome, *Nature*, 537:494–496
- Vastenhouw, N. L. y Schier, A. F. (2012) Bivalent histone modifications in early embryogenesis, *Current Opinion in Cell Biology*, 24:374–386
- Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., *et al.* (2018) Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo, *Science*, 360:981–987
- Wan, L.-B., Pan, H., Hannonhalli, S., Cheng, Y., Ma, J., *et al.* (2008) Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development, *Development*, 135:2729–2738

- Wang, D. C., Wang, W., Zhang, L. y Wang, X. (2018) A tour of 3D genome with a focus on CTCF, *Seminars in Cell and Developmental Biology*, 9521:30160-5
- Wang, H., Li, W., Guo, R., Sun, J., Cui, J., *et al.* (2014) An intragenic long noncoding RNA interacts epigenetically with the RUNX1 promoter and enhancer chromatin DNA in hematopoietic malignancies, *International Journal of Cancer*, 135:2783–2794
- Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., *et al.* (2014) Programming and inheritance of parental DNA methylomes in mammals, *Cell*, 157:979–991
- Watson, L. A., Wang, X., Elbert, A., Kernohan, K. D., Galjart, N., *et al.* (2014) Dual Effect of CTCF Loss on Neuroprogenitor Differentiation and Survival, *The Journal of Neuroscience*, 34:2860–2870
- White, M. D., Zenker, J., Bissiere, S. y Plachta, N. (2018) Instructions for Assembling the Early Mammalian Embryo, *Developmental Cell*, 45:667–679
- de Wit, E., Vos, E. S. M., Holwerda, S. J. B., Valdes-Quezada, C., Verstegen, M. J. A. M., *et al.* (2015) CTCF Binding Polarity Determines Chromatin Looping, *Molecular Cell*, 60:676–684
- de Wit, E. y de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization, *Genes & development*, 26:11–24
- Wu, H., Yang, L. y Chen, L. L. (2017) The Diversity of Long Noncoding RNAs and Their Generation, *Trends in Genetics*, 33:540–552
- Wu, H. y Zhang, Y. (2014) Reversing DNA methylation: Mechanisms, genomics, and biological functions, *Cell*, 156:45–68
- Wyvekens, N., Tsai, S. Q. y Joung, J. K. (2015) Genome editing in human cells using CRISPR/Cas nucleases, *Current Protocols in Molecular Biology*, 112:31.3.1-31.3.18
- Xiao, T., Wallace, J. y Felsenfeld, G. (2011) Specific Sites in the C Terminus of CTCF Interact with the SA2 Subunit of the Cohesin Complex and Are Required for Cohesin-Dependent Insulation Activity, *Molecular and Cellular Biology*, 31:2174–2183
- Yoon, J. H., Abdelmohsen, K. y Gorospe, M. (2014) Functional interactions among microRNAs and long noncoding RNAs, *Seminars in Cell and Developmental Biology*, 34:9–14
- Yoshida, K., Muratani, M., Araki, H., Miura, F., Suzuki, T., *et al.* (2018) Mapping of histone-binding sites in histone replacement-completed spermatozoa, *Nature Communications*, 9:3885
- Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., *et al.* (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation, *Nature Genetics*, 36:1105–1110
- Zhang, H., Jiao, W., Sun, L., Fan, J., Chen, M., *et al.* (2013) Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming, *Cell Stem Cell*, 13:30–35

Zhang, H., Zeitz, M. J., Wang, H., Niu, B., Ge, S., *et al.* (2014) Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the *Kcnq1* locus, *Journal of Cell Biology*, 204:61–75

Zhou, V. W., Goren, A. y Bernstein, B. E. (2011) Charting histone modifications and the functional organization of mammalian genomes, *Nature Reviews Genetics*, 12:7–18