



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Estudio de la conformación tridimensional del genoma durante el desarrollo del sistema eritroide en *Gallus gallus*

TESIS

QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias

PRESENTA:
BIOL. AURA STEPHENSON GUSSINYE

DRA. MAYRA FURLAN MAGARIL
[Instituto de Fisiología Celular](#)

MIEMBROS DEL COMITÉ TUTOR
DR. JESÚS CHIMAL MONROY
[Instituto de Investigaciones Biomédicas](#)
DRA. VIVIANA DEL CARMEN VALADEZ GRAHAM
[Instituto de Biotecnología](#)

Ciudad de México. Diciembre, 2018



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.

Dedicatoria:

A mi compañero de vida Raúl por mostrarme que la vida es una gran aventura, que el miedo es el enemigo que frena nuestros pasos, que el amor es el motor que nos impulsa a seguir adelante, que la felicidad existe y vale la pena, que nunca estaré sola pues tú estás conmigo aunque no pueda verte me seguirás recordando que las estrellas brillan aún en los cielos más oscuros.

Reconocimientos:

Esta tesis para obtener el grado de maestra se realizó bajo la dirección de la Dra. Mayra Furlan Magaril en el Departamento de Genética Molecular del Instituto de Fisiología Celular en la Universidad Nacional Autónoma de México.

El Comité Tutorial que siguió y asesoró el desarrollo de este trabajo estuvo formado por:

Dra. Mayra Furlan Magaril	Instituto de Fisiología Celular, UNAM
Dr. Jesús Chimal Monroy	Instituto de Investigaciones Biomédicas, UNAM
Dra. Viviana Valadez Graham	Instituto de Biotecnología, UNAM

Se reconoce el apoyo y la asesoría del laboratorio 122 NTE a cargo del Dr. Félix Recillas Targa donde se realizaron varios experimentos esenciales para el desarrollo de este proyecto así como la guía de todos sus integrantes y la asesoría técnica de la Biol. Georgina Guerrero Avedaño durante la realización del proyecto y de los experimentos.

Así mismo se reconoce el apoyo de la Red de Apoyo a la Investigación (RAI) y de la Unidad de Biología Molecular del Instituto de Fisiología Celular, UNAM para la realización de este proyecto de investigación.

Este proyecto fue apoyado por el donativo PAPIIT IA201817 y se reconoce también el apoyo de la beca otorgada por CONACyT durante los dos años de realización del proyecto para la presente tesis.

Adicionalmente se agradece el apoyo de la M. en C. Norma Trejo Medina, de Rosa Lilia de la Rocha Viñas y de Julio Ignacio Palacios Ordoñez de la Coordinación del Programa de Posgrado Maestría y Doctorado en Ciencias Bioquímicas de la UNAM.

El Jurado de Examen de Grado estuvo constituido por:

Presidente: Dr. Mario Enrique Zurita Ortega	Instituto de Biotecnología, UNAM
Secretario: Dra. Rosa Estela Navarro González	Instituto de Fisiología Celular, UNAM
Vocal: Dr. Víctor Julián Valdés Rodríguez	Instituto de Fisiología Celular, UNAM
Vocal: Dra. Mariana Keinrad Benítez	Instituto de Ecología, UNAM
Vocal: Dra. Ma. De la Paz Sánchez Jiménez	Instituto de Ecología, UNAM

Agradecimientos:

A mi tutora Mayra Furlan Magaril por su guía en estos dos años, por la paciencia con la que me enseñó a trabajar en el laboratorio, por responder siempre todas mis dudas e inquietudes y por mostrarme cómo ser una mejor científica y persona.

Al Dr. Félix Recillas Targa por darme un lugar en su laboratorio, permitirme aprender tanto de él como de todas las excelentes personas que se encuentran en su equipo de trabajo y por hacerme sentir como parte de su familia académica.

A Geo por tantas veces apoyarme y resolver mis dudas con tanta amabilidad y paciencia, por enseñarme y considerarme como parte de su equipo, por los cálidos saludos y abrazos.

A Esmeralda por el apoyo en el laboratorio, por las galletitas con café y por todo tu trabajo que permitió el desarrollo de este proyecto.

A mis compañeros y amigos del laboratorio por sus enseñanzas y por ser el mejor equipo en la ciencia y también en la diversión. A Roberto por emprender el primer viaje juntos, por ser mi maestro en la elaboración de buffers, por ser mi primera compañía en los retos de comenzar un posgrado, por convertirme en mi amigo y en mi equipo. A Kari por las horas de plática de nuestros proyectos, por siempre estar dispuesta a apoyarme cuando tenía dudas, por ser mi amiga y confidente en mi frustración y en mis inquietudes, por ser mi maestra en bioinformática. A Andrés por ser mi equipo pollo, mi amigo siempre cuando estaba aburrida y hacerme reír incontables horas, por apoyarme cuando estaba triste y por ayudarme a mejorar mi habilidad en la defensa personal “unagi”. A Aye por siempre tener una sonrisa para mí, por aguantarnos a Andrés y a mí hablando de cosas tontas en el laboratorio, por las pláticas en los caminos a la Tienda UNAM. A Marintia por ser un gran equipo en el laboratorio, por resolver juntas todos los problemas con las PCRs, por ser una excelente amiga. A Leo por tu gran apoyo en bioinformática. A Abraham y Amaury por seguir con entusiasmo los proyectos en el laboratorio, por ser buenos amigos dentro y fuera de la universidad, por las horas de pláticas científicas y políticas. A Martí, nuestro agregado cultural en el laboratorio, por ser un gran amigo, por siempre tener una sonrisa para mí, por darme ánimos y aguantar mis frustraciones tanto en lo académico como en la vida.

A mis compañeros y amigos del laboratorio 122 Nte por ser mis amigos y maestros en epigenética y en general en todo el proceso de crecimiento en mi posgrado. A Rodrigo por ser mi primer maestro, por enseñarme desde cómo agarrar una pipeta hasta ayudarme a crecer y madurar y por siempre estar dispuesto a escucharme aún con todas las cosas que siempre tienes que hacer, gracias, estuviste ahí cuando necesité a alguien. A Hober por ser mi maestro en todo lo que necesité, por siempre darme sus valiosas opiniones y ayudarme en mis experimentos cuando tenía dudas, por siempre estar dispuesto a interrumpir lo que sea que estuvieras haciendo y ponerme atención además de las pláticas y los buenos ratos. A Gustavo y a Josué por la ayuda con mi modelo biológico, ustedes siempre tuvieron más experiencia que yo, por los oligos y por las pláticas que siempre me daban buenas ideas para continuar. A Tania por las risas, por los buenos momentos, por tus ánimos cuando tenía dudas, por las pláticas y los cafés. A Carlos, Paco, Josué C. y Ceci por aceptarme siempre en su laboratorio y apoyarme, por las charlas y los buenos momentos. A Yane por tu amabilidad, por siempre apoyarme en todo lo que necesité con esa calidez que te caracteriza.

A mi mamá Gemma Gussinye por siempre apoyarme, por ser mi mamá cuando necesito guía, por ser mi confidente cuando tengo dudas, por ser siempre mi amiga cuando necesito reír y compartir, por ser mi hombro cuando necesito llorar. Por todo lo que siempre haces por mí, por poner mi felicidad y bienestar antes que nada y por hacerme siempre sentirme orgullosa de quien soy, en lo que soy y en lo que he logrado siempre estás tú.

A mi papá Jorge Stephenson por ser mi bastión, por jamás dejarme sola, porque siempre sé que estás ahí para apoyarme y recordarme que todo va a salir bien, por intentar siempre comprenderme aunque a veces sea complicada, porque siempre me das ánimos para levantarme y seguir adelante y porque no importa la distancia, sé que siempre está ahí mi papá.

A mi hermana Tania Stephenson por siempre ser mi ejemplo, por ser mi compañera de pláticas científicas, por mostrarme el camino y recordarme que nunca he estado sola, cuando llegué a este mundo tuve la fortuna de tener a alguien esperándome que sé que siempre estará conmigo. Por siempre hacerme sentir orgullosa de mi hermana mayor.

A Mary Carmen por adoptarme como una sobrina y porque sé que aunque no nos vemos mucho siempre has estado ahí para mí y para mi familia, por querernos tanto y acompañarnos siempre.

A Juan por no sólo ser mi cuñado, también ser mi amigo, mi familia, por ser incondicional conmigo y con aquellos que quiero, por las horas de juego, de risas, de plática, por nunca dejarme sola, por tu increíble apoyo cuando las cosas se pusieron difíciles, gracias.

A mis amigos del CAD, a Diana por ser siempre mi amiga más sincera, por estar siempre presente en mi vida sin importar nuestros caminos separados, por las risas, el crecimiento juntas desde hace tanto tiempo, por el apoyo incondicional en los momentos difíciles. A Natalia por las horas de plática y de diversión, por ser siempre mi amiga más animada, más vital, porque sé que cuento contigo siempre. A Jimena porque somos las más ñoñas, porque enfrentamos retos juntas como entrar a la UNAM siempre apoyándonos, por tu apoyo y las buenas pláticas. A Luis, aunque estés lejos siempre te recuerdo con mucho cariño. A Félix por ser de mis amigos más viejos, por conocerme tanto y quererme tal cual soy, por las pláticas marxistas.

A “los amigos de Raúl” que se convirtieron un día en mi familia. A Arturo por las horas de plática en el café, por la compañía incondicional siempre que necesito un amigo, por las palabras de aliento que me hacen sentir orgullosa, por las discusiones acerca de vampiro que enriquecen nuestras tardes y nos hacen mejores narradores. A Van por ser mi pilar cuando las cosas están mal, por escuchar todo lo que pasa por mi mente y siempre intentar comprenderme y apoyarme, por la compañía y las excursiones al centro, por ser mi narrador estrella y por las horas incontables de diversión, gracias. A Erik por las horas de plática en las banquitas, por las discusiones de homeopatía y placebos, por los mollequiles de los viernes, por ser mi amigo y compañero en el libre y por las risas, por ser mi “Team Fanatic”. A Vero la integrante más nueva en nuestra extravagante familia, por tu apoyo, por tu cariño siempre.

A mis amigos más viejos. A Ana porque eres mi amiga desde hace más de 17 años, porque sé que siempre estás ahí para mí y no importa cuánto tiempo pase podemos sentarnos aún a compartir un café en la plaza y parece que nada hubiera cambiado. A Leslie por compartir todo, lo bueno, lo malo, lo medio, porque no puedo imaginar el mundo sin una amiga como tú, por quererme como soy y apoyarme siempre. A Fran por haber pasado tantas cosas juntos, porque me conoces y me quieres tal y como soy y porque sé que estás ahí siempre sin importar lo que pase, has estado conmigo en momentos divertidos, en momentos confusos y también en momentos tristes y de verdad sé que tengo un amigo para toda la vida.

A mi compañero de vida Raúl gracias por ser mi equipo inquebrantable, mi valor cuando las cosas daban miedo, mi confidente en los más profundo y complicado de mí, mi amigo en las risas y los juegos, mi acompañante en los sueños y aventuras, mi héroe de la vida cotidiana.

Tabla de Contenido

Abreviaturas	8
Introducción	9
Eritropoyesis en Aves	9
Condensación de cromatina y enucleación en eritrocitos.....	11
Características del genoma de <i>Gallus gallus</i>	13
Estudio de la conformación tridimensional del genoma	14
Territorios Cromosomales (CT)	19
Compartimentos de la Cromatina “A” y “B”	20
Dominios Topológicamente Asociados (TADs).....	22
Proteínas que conforman la estructura del genoma	26
Importancia de la estructura del genoma en la expresión génica	27
Cambios de la estructura del genoma en procesos celulares.....	28
Antecedentes	30
Planteamiento del problema	31
Hipótesis	32
Objetivo general	32
Objetivos particulares	32
Método	33
Cultivo celular y obtención de muestras.....	33
Obtención de muestras eritroides de pollo adultos	34
Protocolo de extracción de núcleos.....	35
Diseño de oligonucleótidos (<i>primers</i>)	35
Expresión de genes.....	35
Controles de Hi-C.....	37
Extracción de RNA	40
Estandarización de <i>primers</i> para PCR cuantitativa	41
qPCR	42
Hi-C	42
Fijación	42
Lisis	43
Digestión.....	43
Rellenado y marcaje con biotina.....	44

Ligación y reversión del entrecruzamiento	44
Extracción y purificación del ADN	44
Sonicación	45
Reparación de terminaciones de fragmentos	46
Selección de tamaño	46
“Pull-down” de biotina	46
Ligación de adaptadores	47
Control final de librería	47
Amplificación final	48
Análisis de datos	48
Estandarizaciones y resolución de problemas	52
Estandarización de <i>primers</i> de expresión.	52
Estandarización de expresión de genes estructurales por qPCR	56
Estandarización de protocolo de Hi-C para fibroblastos.....	59
Resultados	62
El ARN disminuye dramáticamente en eritrocitos adultos y se compone de moléculas de tamaño pequeño	62
Existe una disminución de la expresión de genes en los eritrocitos en comparación con las células HD3 y los fibroblastos.....	66
Determinación de la estructura tridimensional del genoma durante la maduración y diferenciación eritroides	70
La estructura tridimensional del genoma cambia drásticamente entre los eritroblastos y los eritrocitos	77
Dominios Topológicos Asociados (TADs)	94
El genoma de los eritrocitos muestra un patrón altamente compartimentado.....	111
Comparación de datos obtenidos con la línea celular DT40.....	115
Discusión	121
Conclusiones	125
Perspectivas	126
Referencias	127

Abreviaturas

- ADN: Ácido Desoxirribonucleico
- ARN: Ácido Ribonucleico
- 3C: Captura Conformacional de Cromosomas
- ATAC-seq: Ensayo de Accesibilidad a la Transposasa
- NTPs: Nucleótidos Trifosfato
- ADNc: ADN complementario
- PCR: Reacción en Cadena de la Polimerasa
- RT-PCR: Reacción en Cadena de la Polimerasa con Transcriptasa Inversa
- qPCR: Reacción en Cadena de la Polimerasa Cuantitativa
- Interacción en *Cis*: dentro de un mismo cromosoma
- Interacción en *Trans*: entre distintos cromosomas
- H5: Histona 5
- 4C: 3C circular
- TAD: Dominio Topológicamente Asociado
- Hi-C: Captura Conformacional de Cromosomas del genoma completo
- CT: Territorio Cromosómico
- CTCF: Factor de unión CCCTC
- YY1: Proteína Ying Yang 1
- eRBC: eritrocito embrionario de 10 días de desarrollo
- aRBC: eritrocito adulto de gallinas
- HD3: línea celular de eritroblasto transformada por el virus aviar de eritroblastosis
- dATP: Desoxiadenosina trifosfato
- dGTP: Desoxiguanosina trifosfato
- dTTP: Timidina trifosfato
- dCTP: Desoxicitidina trifosfato
- ARNr: ARN ribosomal
- ARNm: ARN mensajero
- ARNp: ARN primario
- Mpm: Marcador de Peso Molecular
- DNAsa: Desoxirribonucleasa
- HBAD: Subunidad Alfa D de la Hemoglobina
- HBAA: Subunidad Alfa A de la Hemoglobina
- HBBA: Subunidad Beta A de la Hemoglobina
- EPB41: Proteína de Membrana Eritroide 4.1
- EPB42: Proteína de Membrana Eritroide 4.2
- RPL4: Proteína Ribosomal L4
- GAPDH: Gliceraldehído 3 Fosfato Deshidrogenasa
- U4/U6: U4/U6 Ribonucleoproteína pequeña nuclear Prp3
- GATA1: Factor de Transcripción de Globinas 1

Introducción

Eritropoyesis en Aves

La eritropoyesis es una vía de diferenciación por la cual progenitores eritroides proliferan y se diferencian hasta células llamadas reticulocitos las cuales carecen de núcleo, retículo endoplásmico y mitocondrias (1). Los estadios por los cuales atraviesan estas células se conocen como eritroblastos y se caracterizan de acuerdo a sus cualidades morfológicas (policromáticos, ortocromáticos, etc.) las cuales se alteran mientras el proceso de diferenciación avanza pudiéndose observar una disminución en el tamaño de la célula, un aumento en la heterocromatina del núcleo y por último, la expulsión del núcleo celular en el estadio anterior a los reticulocitos el cual corresponde al tipo celular inmediato anterior al eritrocito circulante (2).

La eritropoyesis en las aves es un proceso durante el cual, varios precursores eritroides pasan por distintos cambios hasta llegar a formar los eritrocitos maduros. El lugar eritropoyético inicial es el saco vitelino y éste funciona como el órgano eritropoyético principal durante la etapa anterior al nacimiento del organismo. Posteriormente, en el adulto, la eritropoyesis pasa a llevarse a cabo en la médula ósea y al bazo (3).

La eritropoyesis se divide en tres etapas, la primera es la eritropoyesis temprana la cual se caracteriza por el compromiso de las células troncales y progenitoras hacia los proeritroblastos. La segunda etapa es llamada diferenciación terminal eritroide y se identifica por el paso de varios estadios celulares ya comprometidos hacia el linaje eritroide (4), comenzando por el proeritroblasto, el eritroblasto basófilo, eritroblasto policromatófilo, reticulocito y el estadio final, el eritrocito (5) (Fig. 1A). La tercer etapa se conoce como maduración eritroide y consiste en una serie de cambios en el tamaño y la condensación de la cromatina que llevan a las células a su estado maduro en los animales adultos (3).

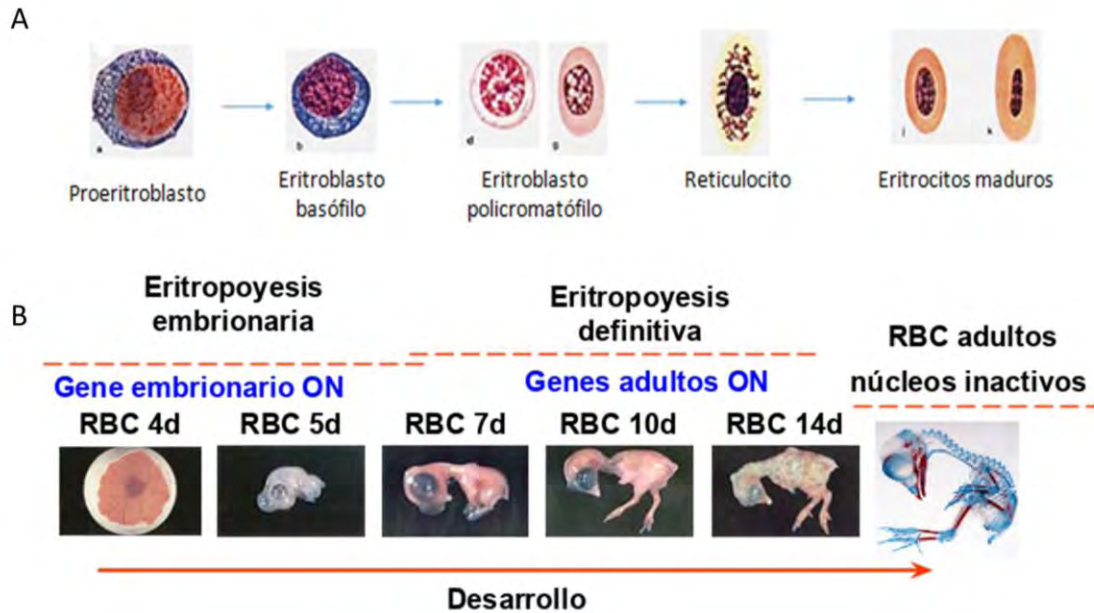


FIGURA 1 ERITROPOYESIS Y DESARROLLO EMBRIONARIO DE *GALLUS GALLUS*. A) REPRESENTACIÓN GRÁFICA DE LOS PRINCIPALES ESTADIOS DEL DESARROLLO ERITROIDE DE *GALLUS GALLUS*. (MODIFICADA DE CAMBELL, 1994). B) RELACIÓN DE LA EXPRESIÓN DE LOS GENES GLOBINA DURANTE EL DESARROLLO EMBRIONARIO DEL POLLO, PUEDE OBSERVARSE CÓMO EN EL DÍA 4 DEL DESARROLLO EL GEN EMBRIONARIO DE LA PROTEÍNA GLOBINA SE ENCUENTRA EXPRESÁNDOSE. ALREDEDOR DEL DÍA 5 COMIENZA EL CAMBIO DE LA EXPRESIÓN DE ESTOS GENES ACTIVÁNDOSE LA EXPRESIÓN DE LAS GLOBINAS ADULTAS ALFA Y BETA. ESTE CAMBIO ES COMPLETO EN DÍA 10 DONDE SÓLO SE OBSERVA LA EXPRESIÓN DE LOS GENES ADULTOS. EN EL ORGANISMO COMPLETAMENTE FORMADO Y EN EL ADULTO LOS ERITROCITOS CIRCULANTES MADUROS POSEEN UN NÚCLEO INACTIVO TRANSCRIPCIONALMENTE(6).

El proceso de maduración eritroide puede ser detectado desde pollos con 35 horas de desarrollo. En este estadio ya existen células eritroides denominadas “eritrocitos”, sin embargo, estas células poseen una forma alargada y redondeada y son llamados linaje de eritrocitos “primitivos”. Alrededor de los 5 días de desarrollo las células comienzan con un recambio y los eritrocitos “definitivos” aparecen (Fig. 1B). Estos poseen una forma más redonda y son de menor tamaño que el linaje de los eritrocitos primitivos. Pasando el día 9 de desarrollo más del 95% de los eritrocitos pertenecen al linaje de los definitivos (Fig. 1B) (7).

Los eritrocitos primitivos y definitivos atraviesan por varios cambios en el proceso de maduración, sufren una disminución de tamaño y una elevación en la condensación de la cromatina del núcleo, así mismo poseen diferencias transcripcionales (8). Las células

llamadas primitivas expresan altos niveles de genes globina embrionarios, como son el pi, rho y épsilon mientras que los eritrocitos definitivos tienen una alta expresión de los genes alfa-D, alfa-A y beta-A (5). Posteriormente los eritrocitos definitivos pasan por el proceso de maduración en el cual se observa un cambio en la forma del núcleo y una disminución del tamaño celular promedio (9).

Una característica de la maduración eritroide en mamíferos es que durante el paso del estadio de reticulocito hacia eritrocito, las células pasan un proceso donde el núcleo es expulsado de la célula (4). Esto no ocurre en el proceso de maduración eritroide de las aves incluyendo a *Gallus gallus* por lo cual los eritrocitos tanto inmaduros como maduros son un modelo donde se puede estudiar las modificaciones por las que atraviesa el núcleo celular durante el proceso de maduración y los cambios tanto a nivel estructural como transcripcional que suceden en el núcleo.

Condensación de cromatina y enucleación en eritrocitos

El proceso de enucleación es precedido por un proceso de condensación generalizada de la cromatina y por una reducción del volumen nuclear que, en mamíferos puede llegar a 1/10 de su volumen original (10). Estas características se han observado en varias especies, sin embargo, la expulsión del núcleo parece ser una característica perteneciente sólo al clado de los mamíferos. Una de las hipótesis que existen para explicar esta adaptación evolutiva reside en que el tamaño de los eritrocitos se encuentra directamente relacionado con el tamaño de los vasos sanguíneos de la especie y también con su capacidad para bombear sangre. La capacidad de bombeo presente en los mamíferos les permite irrigar vasos de diámetro muy pequeño (10).

La condensación generalizada del núcleo se ha observado en todos los vertebrados, sin embargo, son solamente los mamíferos los que atraviesan por el proceso de enucleación, otros vertebrados poseen diferentes mecanismos de condensación de la cromatina como la

presencia de la histona H5 la cual no se encuentra en los mamíferos (10). En ratón se ha observado que el proceso de condensación del núcleo se encuentra relacionado con la expresión de una desacetilasa de histonas (HDAC2) la cual promueve el paso de la cromatina accesible a heterocromatina antes de la enucleación (11). Posteriormente se ha observado que los eritroblastos de ratón extruden el núcleo formando una estructura membranal parecida a una burbuja (10).

En eritrocitos maduros aviares, los cuales retienen el núcleo, se ha observado un declive generalizado en la actividad transcripcional, así como un arresto en la replicación durante el proceso de maduración. Este proceso de “apagado” en las funciones genómicas incluye la pérdida de proteínas nucleares no histonas, la alteración de la composición de las histonas, la degeneración del nucléolo y la compactación generalizada de la cromatina acompañada de la contracción del núcleo completo (12).

Así mismo se ha observado que, acompañando a este proceso de cambios a nivel de la cromatina, la expresión del genoma muestra una reducción considerable. Se estima que más del 90% del ARN celular es degradado durante la diferenciación terminal eritroide y en un estudio realizado en el 2002 se observó que esta reducción de los niveles de RNA es apreciable desde el día 3 de incubación en embriones de pollo y parece reducirse gradualmente a lo largo del desarrollo en células eritroides en circulación (Fig. 2) (13).

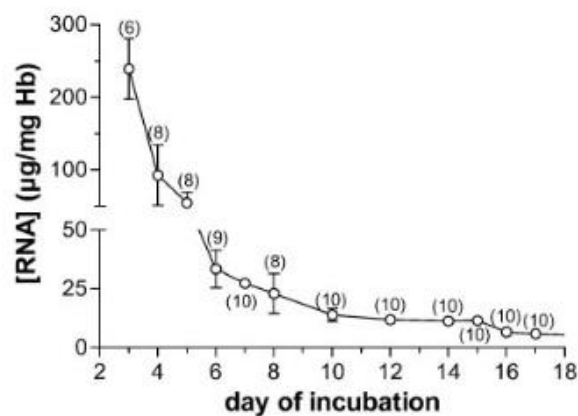


FIGURA 2 GRÁFICO REPRESENTANDO LA CANTIDAD DE ARN OBSERVADA EN ERITROCITOS DURANTE EL DESARROLLO EMBRIONARIO DE *GALLUS GALLUS*. SE PUEDE NOTAR UNA DISMINUCIÓN GRADUAL EN LA CANTIDAD DE ARN

PRESENTE EN ESTAS CÉLULAS CONFORME AVANZA EL DESARROLLO. EL NÚMERO ENTRE PARÉNTESIS REPRESENTA EL NÚMERO DE MEDICIONES PARA CADA UNO DE LOS DÍAS. (BAUMANN R, GOTZ R, DRAGON S. 2003).

El silenciamiento nuclear durante la maduración eritroide es un proceso que ocurre tanto en mamíferos como en otros vertebrados. En el caso de los mamíferos, se sabe que el proceso de hetero-cromatinización y silenciamiento génico está estrechamente relacionado con la di-metilación de la lisina 9 de la histona 3 (H3K9m) y con la des-acetilación de la lisina 12 de la histona 4 (H4K12), en contraste con las especies aviares donde este proceso se asocia con cambios en la estructura de los nucleosomas y acumulación de proteínas estructurales (14).

En el pollo los procesos por los que atraviesa el núcleo en la maduración eritroide están relacionados con los factores de condensación de la cromatina H5 (15). H5 es una proteína histona de 189 aminoácidos y 20.5 kDa que posee una localización nuclear, es depositada en eritrocitos terminalmente diferenciados formando grandes regiones de heterocromatina altamente condensada y reprimida transcripcionalmente (16). Durante la diferenciación terminal en eritrocitos aviares estos factores ocasionan la condensación de la cromatina en múltiples sitios dispersos de heterocromatina facultativa.

Características del genoma de *Gallus gallus*

El genoma de *Gallus gallus* tiene un contenido haploide de 1.1 giga bases de ADN, aproximadamente 3 veces menor que el genoma humano. Está compuesto por 38 autosomas y un par de cromosomas sexuales los cuales son denominados Z y W.

El Consorcio de Secuenciación del Genoma de Pollo (CGSC) en 2004 clasificó los 39 cromosomas de *Gallus gallus* tanto en micro como en macrocromosomas. El consorcio creó una convención de clasificación en la cual dividen los cromosomas del pollo en tres clases: cinco cromosomas (GGA 1-5) son clasificados como macrocromosomas ya que miden entre 50 y 200 Mb; cinco fueron clasificados como cromosomas intermedios (GGA 6-10)

mediendo de los 20 a los 40 Mb y los 28 restantes fueron nombrados microcromosomas (GGA 11-38) ya que poseen un tamaño promedio de 12 Mb de largo (17).

Los microcromosomas son característicos de todos los cariotipos de las especies de aves y muchos de los reptiles(18). En 1995 Newcomer pensaba que los microcromosomas eran elementos heterocromáticos especiales los cuales no poseían ni genes ni centrómeros. Sin embargo, en el 2000, Smith y colaboradores encontraron que los micro cromosomas contienen aproximadamente el 50% de todos los genes aviares y se comportan como cualquier otro de los cromosomas, se mantienen estables durante la mitosis y meiosis y tienen tanto centrómeros como telómeros funcionales (18).

La organización del genoma de *Gallus gallus* en cromosomas de variados tamaños y la diferencia en el proceso de maduración eritroide que posee con los mamíferos, brindan un modelo que permite, aplicando técnicas de análisis del genoma desarrolladas recientemente, obtener datos desconocidos acerca, tanto de la organización del genoma a nivel particular, como del proceso transcripcional y estructural que sucede en el núcleo durante el proceso de maduración eritroide. Estas técnicas han sido desarrolladas en la última década y brindan una herramienta que ha permitido obtener información acerca de la estructuración del genoma y la relación entre la estructura tridimensional del mismo y su función regulatoria en diferentes procesos celulares (19).

Estudio de la conformación tridimensional del genoma

Con base en observación y datos de microscopía, anteriormente se creía que la cromatina se encontraba establecida dispersa dentro del núcleo celular. Gracias al desarrollo de nuevas técnicas de estudio ahora se conoce que los genomas dentro del núcleo se encuentran altamente organizados y compartimentados. Una de estas técnicas es la Hibridación Fluorescente *In Situ*, con la cual se observó que los cromosomas se encuentran separados en territorios discretos que se llamaron territorios cromosómicos (20). Los territorios cromosómicos representan una característica global de esta compartimentación

y también de arquitectura nuclear (21). Cada uno de los pares de autosomas y también los cromosomas sexuales ocupan una región discreta dentro del núcleo en interfase y, aunque se desconoce lo que ocurre en todas las células, se ha observado que la preferencia de organización radial que tienen los cromosomas en este arreglo se encuentra conservada entre algunos tipos celulares (22).

Posteriormente, además de los ensayos de fluorescencia se desarrollaron métodos que permitieron alcanzar una mayor resolución en el estudio de la estructura tridimensional del genoma. La técnica de Captura de la Conformación Cromosómica (3C) fue desarrollada para estudiar y analizar la organización espacial de los cromosomas a alta resolución (23). El principio de esta técnica consiste en tratar las células a estudiar con formaldehído, lo cual genera un entrecruzamiento entre proteína-proteína y proteína-ADN de segmentos del genoma que se encuentren físicamente cerca en el momento del tratamiento. Posteriormente la cromatina es fragmentada utilizando una enzima de restricción. Después de la digestión los segmentos que se encuentran entrecruzados en proximidad son ligados.

Por último, el entrecruzamiento realizado por el formaldehído es revertido y el ADN es purificado para obtener una librería de interacciones (Fig. 3) (24). Posteriormente se determinan las frecuencias relativas de interacción entre diferentes sitios del genoma que se encuentran entrecruzados y con ello se obtiene información acerca de la organización general del genoma y la conformación tridimensional de los cromosomas (23).

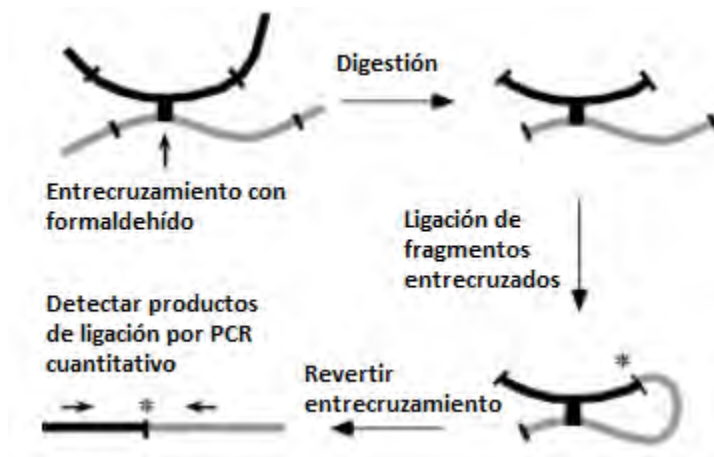


FIGURA 3 ESQUEMA DE LOS PASOS GENERALES DE LA TECNOLOGÍA DE CAPTURA DE LA CONFORMACIÓN CROMOSÓMICA. (MODIFICADA DE: DEKKER ET AL, 2002)

La tecnología 3C permitió el estudio del plegamiento de los cromosomas a una resolución que no tenía precedentes. Se utilizó para demostrar que genes podían contactar a sus secuencias reguladoras dentro del genoma aun estando a kilobases de distancia. También estas técnicas permitieron observar que los contactos regulatorios podían cambiar dramáticamente durante el desarrollo y la diferenciación y que están mediados por factores transcripcionales unidos a los sitios del genoma involucrados (25). Un ejemplo de esto fue observado por el grupo del Dr. Razin durante la caracterización del dominio alfa globina de pollo, donde utilizando la técnica de 3C pudieron observar una diferencia significativa de interacción entre un elemento regulatorio de la transcripción de los genes globina llamado *enhancer* MRE y el gen alfa-D en células eritroides en comparación a células de eritroblasto o no eritroides donde este contacto es casi nulo (26).

La tecnología clásica de 3C permite analizar contactos entre un par de secuencias seleccionadas (27) (Fig. 4). En los experimentos clásicos de 3C productos de una ligación sencilla son detectados por medio de PCR uno a la vez utilizando *primers* específicos para un locus (19). De esta forma las frecuencias de interacción son determinadas una a la vez dependiendo del candidato que se deseé estudiar. El estudio 3C clásico es usualmente utilizado para determinar frecuencias de interacción entre máximo algunas docenas de

fragmentos de restricción individuales como pueden ser genes y su interacción con promotores o algunas otras secuencias regulatorias (24).

A partir de la técnica de 3C se desarrollaron otras que permitieron el análisis de la estructura del genoma a diferentes escalas. La técnica de 4C, llamada 3C circular o 3C en Chip (19) es definida como un estudio de “uno contra todos” (Fig. 4). La tecnología 4C permite la identificación de regiones que se encuentran contactando a una secuencia de interés llamada “ancla” o “*bait*”. En contraste con 3C, no requiere un conocimiento o hipótesis a priori de las regiones candidato (27) ya que combina la tecnología 3C con una detección por PCR inversa de productos de ligación de 3C (24).

Los estudios realizados con 4C han revelado que muchos genes individuales en poblaciones celulares forman una gran cantidad de contactos con loci que se encuentran en diferentes zonas del genoma mayoritariamente en *Cis* (en su mismo cromosoma) aunque también se han observado contactos en *trans* (entre distintos cromosomas) (25).

La siguiente tecnología de estudio de la conformación de los cromosomas es llamada 5C (3C-carbon copy) y se define por el estudio de “muchos contra muchos” (25) (Fig.4), y ha brindado un análisis más profundo acerca de la forma en que la cromatina se pliega mediante medir un gran número de perfiles anclados de interacciones al mismo tiempo (19). La técnica de 5C fue la tecnología que permitió el descubrimiento de los Dominios Topológicamente Asociados o TADs al igual que estudios con la tecnología de Hi-C (25).

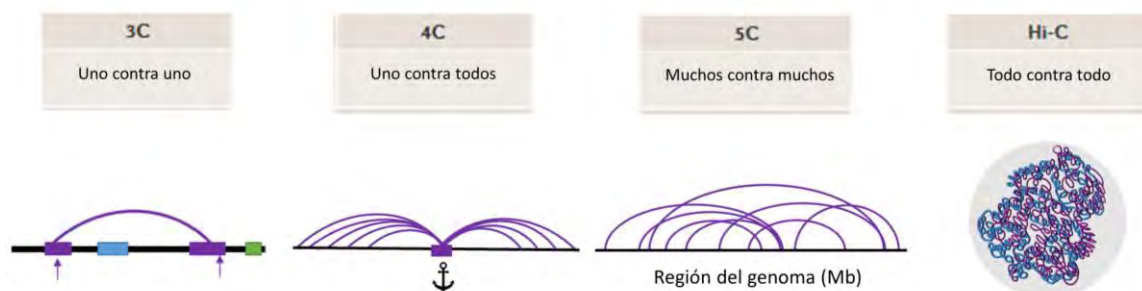


FIGURA 4 REPRESENTACIÓN GRÁFICA DE LAS TÉCNICAS DE CAPTURA DE LA CONFORMACIÓN CROMOSÓMICA. SE MUESTRA LA TÉCNICA DE 3C CON LA CUAL PUEDE ENCONTRARSE LA INTERACCIÓN DE DOS REGIONES DEL GENOMA

QUE SE ENCUENTRAN DISTALES DESDE LA PERSPECTIVA LINEAL PERO CERCANAS EN EL ESPACIO NUCLEAR (LOS CUADROS REPRESENTAN POSIBLES ELEMENTOS REGULATORIOS). LA TÉCNICA DEL 4C LA CUAL PERMITE TOMAR UNA REGIÓN LLAMADA “ANCLA” Y ENCONTRAR TODAS LAS REGIONES DENTRO DEL GENOMA QUE SE ENCUENTRAN INTERACCIONANDO CON ESTA ZONA. LA TÉCNICA DE 5C LA EN LA CUAL SE TOMA UNA REGIÓN DEL GENOMA Y BRINDA LAS INTERACCIONES DE LAS REGIONES COMPRENDIDAS DENTRO DE LA MISMA. LA TÉCNICA DE HI-C LA CUAL MUESTRA TODAS LAS INTERACCIONES DENTRO DEL GENOMA COMPLETO PERMITIENDO CALCULAR LA ESTRUCTURA TRIDIMENSIONAL DEL GENOMA EN UN TIEMPO DADO.

La técnica de Hi-C fue reportada en 2009 y es el método de Captura de la Conformación Cromosómica más global. Hi-C permite una identificación imparcial de las interacciones de la cromatina a lo largo del genoma completo (28). Este método es una adaptación del 3C que incluye un paso único en el cual, después de la digestión con enzimas de restricción, los finales de los fragmentos de ADN son rellenados con nucleótidos marcados con una molécula de biotina (Fig. 5). Esto facilita la purificación selectiva de las ligaciones de interacción que posteriormente son secuenciadas (19).

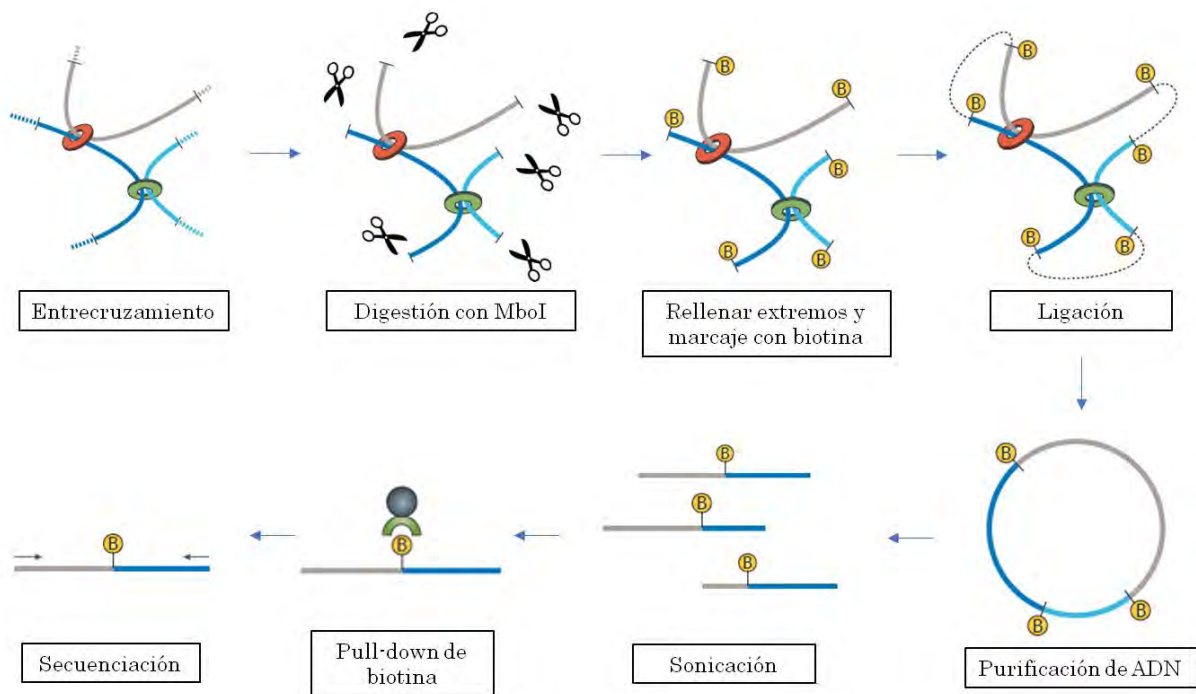


FIGURA 5 ESQUEMA REPRESENTANDO LOS PASOS DEL PROTOCOLO DE HI-C. LAS FIBRAS DE CROMATINA QUE SE ENCUENTRAN INTERACCIONANDO SE REPRESENTAN CON DISTINTOS COLORES. SE MUESTRA EL ENTRECruzamiento CON FORMALDEHÍDO REPRESENTANDO POR LOS CÍRCULOS QUE MANTIENE LAS FIBRAS DE CROMATINA JUNTAS DURANTE LOS SIGUIENTES PASOS DEL PROTOCOLO. POSTERIORMENTE SE MUESTRA UN EJEMPLO DE LOS POSIBLES ESCENARIOS QUE SE OBTIENEN DURANTE LA REALIZACIÓN DEL PROTOCOLO CON UN FRAGMENTO HÍBRIDO CIRCULAR COMPUESTO DE TRES REGIONES DEL GENOMA QUE SE ENCONTRABAN INTERACCIONANDO. LAS PERLAS DE ESTREPTAVIDINA QUE SE UTILIZAN PARA LA RECUPERACIÓN DE LOS FRAGMENTOS MARCADOS CON BIOTINA SE

ENCUENTRAN REPRESENTADAS POR LA FIGURA EN COLORES NEGRO Y VERDE. EN EL ÚLTIMO PASO SE EJEMPLIFICA LA SECUENCIACIÓN PAREADA DE LOS FRAGMENTOS POR LAS DOS FLECHAS. (MODIFICADA DE BONEV Y CAVALLI. 2016).

Esta técnica provee de mapa detallado que muestra las interacciones de “todo contra todo” (Fig. 4) a nivel del genoma completo, sin embargo, la resolución de este mapa depende de la profundidad de la secuenciación (19).

Con el desarrollo de las técnicas antes mencionadas se ha abierto un área de estudio de la organización y estructura del genoma y cómo ésta se relaciona con distintos procesos celulares. La organización del genoma se puede estudiar a distintas escalas que van desde el arreglo radial de los territorios cromosomales, pasando por los compartimentos activos e inactivos de la cromatina, la organización de dominios hasta las interacciones específicas entre secuencias reguladoras y sus genes blancos.

Territorios Cromosomales (CT)

El estudio de la conformación de la fibra de cromatina se ha llevado a cabo desde hace muchas décadas. A finales del siglo XIX y principios del XX Rabl y Boveri plantearon que los cromosomas de las células eucariotas se encontraban organizados en territorios discretos llamados “territorios cromosomales” o CT por sus siglas en inglés (revisado en Dixon *et al.* 2016). Posteriormente Stephen M. Stack, David B. Brown y William C. Dewey retomaron las ideas planteadas por Rabl y Boveri y realizaron un experimento con cultivos celulares eucariotas en interfase las cuales tiñeron con tinción Giemsa y concluyeron que los “cromosomas se mantienen en distintos dominios durante la interfase”(20). Actualmente se acepta que los territorios cromosomales son una característica importante dentro de la arquitectura global de los cromosomas en núcleos de células eucariotas (Fig.6) (29).



FIGURA 6 REPRESENTACIÓN A COLOR DE LOS TERRITORIOS CROMOSOMALES EN UN NÚCLEO DE FIBROBLASTO EN INTERFASE (MODIFICADA DE: BOLZER *ET AL.* 2005)

El posicionamiento de cromosomas de forma no aleatoria en el núcleo de las células ha presentado la posibilidad de que la funcionalidad del genoma esté relacionada con la agrupación tridimensional del mismo (24). Esto quiere decir que la vecindad de los cromosomas en el núcleo no está dada de forma aleatoria, varía entre diferentes tipos celulares y tiene consecuencias en la capacidad de los cromosomas de interactuar en *trans* con otras partes del genoma. Por esto se acepta que la posición y organización tridimensional de los cromosomas en estos CT dentro del núcleo afecta a la función, pero más particularmente a la regulación de la expresión genética (30).

Compartimentos de la Cromatina “A” y “B”

Con la técnica de Hi-C se obtuvo que el genoma de metazoarios se organiza en dos compartimentos que contienen cromatina activa e inactiva los cuales se nombraron como compartimentos A y B respectivamente (25) (Fig. 7). La cromatina tipo “A” se relaciona con la replicación temprana y se involucra con transcripción activa mientras que la cromatina tipo “B” incluye genes silenciados o inactivos y se relaciona con la replicación tardía (31). Estos compartimentos pueden medir hasta varias megabases y su existencia fue sugerida en un principio por microscopía electrónica donde se pudo apreciar como zonas

electrodensas la cromatina que se encontraba en la periferia del núcleo y en el nucléolo. En cambio, la cromatina activa se concentra cerca del centro del núcleo ya que aquí es donde se encuentran los factores transcripcionales, las RNA polimerasas y algunos otros factores (32).

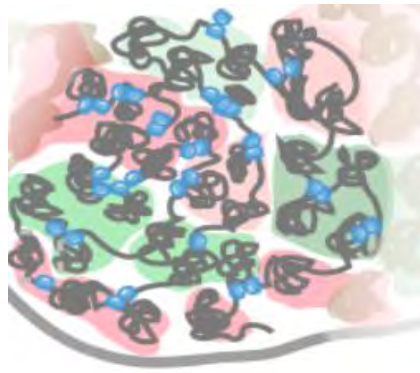


FIGURA 7 REPRESENTACIÓN DE LOS COMPARTIMENTOS DE LA CROMATINA DENTRO DEL NÚCLEO. LA FIBRA DE CROMATINA REPRESENTADA POR LA LÍNEA GRIS JUNTO CON LAS PROTEÍNAS ESTRUCTURALES, REPRESENTADAS POR LOS PUNTOS AZULES, SE ORGANIZAN DENTRO DEL GENOMA PROMOVRIENDO LA INTERACCIÓN DE GRANDES REGIONES DE LA CROMATINA. ESTO SE REPRESENTA POR LOS SOMBREADOS DE COLOR ROJO Y VERDE. ASÍ MISMO SE HA OBSERVADO QUE LAS REGIONES GRANDES QUE INTERACCIONAN ENTRE SÍ CORRESPONDEN A ZONAS CON ESTADOS TRANSCRIPCIONALES SIMILARES, YA SEA HETEROCROMATINA O EUCROMATINA, SE ENCONTRÓ QUE ZONAS ACTIVAS INTERACCIONAN PREFERENCIALMENTE CON OTRAS ZONAS ACTIVAS (COMPARTIMENTO A, SOMBREADO VERDE) Y LO MISMO OCURRE CON PARA LAS REGIONES INACTIVAS TRANSCRIPCIONALMENTE (COMPARTIMENTO B, SOMBREADO ROJO) (MODIFICADA DE NORA ET AL. 2016).

Los compartimentos A y B parecen ser específicos dependiendo del tipo celular (28). Esto se debe a que, las necesidades específicas de expresión genética cambian entre linajes celulares y algunos genes son necesarios en ciertos tipos celulares y en ciertos momentos metabólicos, lo cual genera que mediante modificaciones de la cromatina, los compartimentos puedan ser reorganizados para encontrarse cercanos a la periferia nuclear o en la parte central del núcleo (32). Un ejemplo de la variación que existe entre compartimentos se observó en un estudio realizado por Dixon y colaboradores en el cual, utilizando la técnica de Hi-C, observaron que había una gran expansión del compartimento B durante la diferenciación de células troncales humanas hacia células mesenquimales. Así mismo mencionan que consistentemente con estudios anteriores encontraron que genes que cambian del compartimento A al B tienen una tendencia a mostrar menor expresión

mientras que genes que cambian del compartimento B al A muestran una expresión más elevada (33).

La distribución de los compartimentos A y B es una característica topológica del genoma que ha sido objeto de estudio en repetidas ocasiones. En un estudio realizado en el 2015 utilizando varios linajes celulares pertenecientes a la diferenciación de células troncales embrionarias, se observó que la distribución de compartimentos A y B parecía ser específica de cada linaje. También se observó que a lo largo del proceso de diferenciación existe una expansión notoria del compartimento B que los autores relacionan con el proceso de aumento en de la zonas de heterocromatina en estos linajes durante el proceso de diferenciación (33).

Los factores que median o determinan la organización en compartimentos aún se desconocen. En un estudio del año 2017 se realizó un abatimiento inducible (KO por sus siglas en inglés) de una proteína considerada como estructural, la proteína CTCF, se observó que la falta de proteína alteraba enormemente la formación de dominios topológicamente asociados (TADs). Sin embargo, en este experimento los compartimentos no fueron afectados(34) lo que lleva a la pregunta de qué factores proteicos y de otra naturaleza se encuentran estableciendo las interacciones entre estas grandes regiones de cromatina que corresponden con marcas de eucromatina o heterocromatina.

Dominios Topológicamente Asociados (TADs)

Las técnicas de Captura de la Conformación Cromosómica y especialmente la de Hi-C permitieron el descubrimiento de una tercera escala estructura del genoma (23). Este arreglo estructural se encuentra relacionado con la regulación de la expresión genética y divide a la cromatina en dominios de cromatina que van desde algunos cientos de kilo bases hasta una o dos megabases de largo (35).

Estos dominios fueron llamados Dominios Topológicamente Asociados (TADs) y se caracterizaron por la interacción preferencial de la cromatina dentro de sí mismos, es decir las secuencias que se encuentran dentro de un mismo TAD tienen una alta frecuencia de interacción o contacto entre ellas (36). Así mismo estos TADs tienden a segregarse espacialmente unos de otros mediante regiones llamadas límites o bordes, las cuales se han observado en varias especies incluyendo *Drosophila* donde se muestran como un bandeo claro y fácilmente observable en los cromosomas politénicos (37).

Estas regiones denominadas límites tienen como característica que las interacciones entre la fibra de cromatina que se observan allí tienen una direccionalidad, es decir, la mayoría de ellas cambian de dirección justo sobre el límite. Esto se manifiesta en una reducción de los contactos entre secuencias que se encuentren en ambos lados del denominado límite(36). También se han reportado algunos análisis que indican que los grupos de genes englobados en un mismo TAD tienen una expresión correlacionada durante la diferenciación celular, aunque esto, en una perspectiva general dentro del genoma, sigue generando varias interrogantes (24).

Las formaciones de asas de cromatina en el genoma facilitan la regulación de genes mediante interacciones con elementos regulatorios que pueden ser distales. Se ha observado que aproximadamente un 70% de las formaciones de asas de cromatina ocurren dentro de los TADs. Estudios recientes ha sugerido que los TADs pueden ser definidos como dominios donde se forman asas de interacción, los cuales se mantienen unidos y delimitados por sitios convergentes de CTCF (Fig.8) (32).

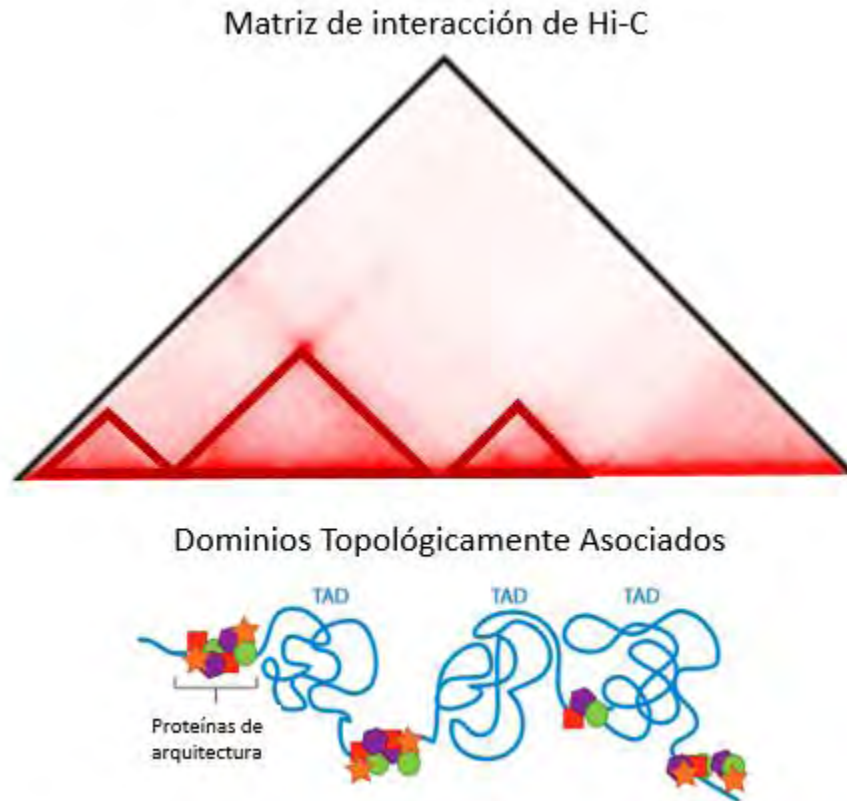


FIGURA 8 REPRESENTACIÓN DE UN MAPA DE INTERACCIONES DE HI-C. LA ESCALA DE COLOR REPRESENTA EN ROJO UN ENRIQUECIMIENTO DE INTERACCIÓN Y EN BLANCO UN DECREMENTO DE CONTACTOS. LOS TRIÁNGULOS FUE LO QUE SE OBSERVÓ EN LOS RESULTADOS DE ESTUDIOS DE 5C Y DE HI-C Y FUERON DENOMINADOS DOMINIOS TOPOLÓGICAMENTE ASOCIADOS. LA IMAGEN INFERIOR MUESTRA UNA REPRESENTACIÓN DE CÓMO PODRÍAN INTERPRETARSE LOS TADS EN LA FIBRA DE CROMATINA (MODIFICADA DE: RAO *ET AL.*, 2014).

CTCF es una proteína que posee actividad aislante y demarca dominios con firmas de cromatina distintas. Se ha planteado que, junto con las proteínas cohesinas, son uno de los principales jugadores en la arquitectura de la cromatina en mamíferos (27). CTCF no solo se asocia con las cohesinas, también ayuda a posicionar al complejo de estas proteínas, las cuales poseen una forma de anillo que ayuda a la formación de asas en la cromatina. Estudios de Hi-C mostraron que CTCF se encuentra especialmente enriquecida en los bordes de TADs y la delección de los sitios de unión de esta proteína, por ejemplo en el centro de inactivación del cromosoma X, lleva a la desorganización y fusión parcial de los TADs adyacentes promoviendo organizaciones tridimensionales a las que llamaron mega dominios (27,38).

Las primeras evidencias de los TADs en los mamíferos se obtuvieron de un experimento de 5C en el cromosoma X en ratones y también de un estudio global de Hi-C tanto en ratón como en células humanas. Esto apuntó algunas características básicas de los TADs en mamíferos y se pudo definir que los dominios eran relativamente invariables durante la diferenciación celular y estaban generalmente conservados en cuanto a su posición entre humano y ratón. Sin embargo, se pudo observar también que, aunque en la mayoría de los casos no hay gran variación entre la presencia de los TADs, hay algunos procesos en los cuales se han observado diferentes distribuciones o incluso ausencia de ciertos dominios (35). El primero de ellos es en el cromosoma inactivo X de ratón, en el cual se observó una pérdida de la estructura local de la cromatina y la formación de dos mega dominios (39), en segundo lugar los TADs no son observables en cromosomas mitóticos (35).

La formación y el mantenimiento de TADs son procesos altamente estudiados en el área de la estructuración tridimensional del genoma, pero el cómo se aíslan las interacciones entre dominios sigue siendo una incógnita. En un estudio realizado en el 2017, se observó que la degradación inducible de la proteína CTCF generaba una alteración en la capacidad aislante de las fronteras de los TADs alterando la formación de los mismos (34). Así mismo en otro estudio realizado en mayo del mismo año observaron que, al reducir la proteína SCC4 encargada de colocar a las cohesinas sobre el ADN, las fronteras de los TADs se tornaban difusas y su capacidad aislante se redujo (40). Estas evidencias muestran que los factores proteicos tienen una importancia fundamental en el establecimiento de los Dominios Topológicamente Asociados y parecen contribuir con la capacidad del genoma de aislar ciertas regiones de poder interactuar con otras generando complejos mecanismos de regulación.

Aunque se ha observado que los TADs son relativamente invariables entre las células, existe una diferencia entre la frecuencia de interacción entre secuencias presentes dentro de ellos que se observan en algunos tipos celulares en diferentes etapas de desarrollo. Esto se debe

a que hay distintos elementos regulatorios que entran en contacto en los distintos estadios del desarrollo celular (p ej. potenciadores, promotores y silenciadores) y esto denota la complejidad de la regulación de los genes. Sin embargo, la situación actual sugiere que los TADs influyen y delimitan el panorama regulatorio de los genes y que tanto la formación de asas, como el acercamiento de los promotores de los genes con sus potenciadores o silenciadores se dan, en su mayoría, dentro de estos dominios (35).

Proteínas que conforman la estructura del genoma

CTCF es una proteína que contiene un dominio de unión a DNA compuesto por once dedos de zinc que muestra una homología cercana al 100% entre el ratón, pollo y humano (41).

CTCF se encuentra involucrada en diversos roles de regulación génica, se ha visto involucrada en activación y represión de promotores, posee actividad de bloqueo de *enhancers*, barrera y más recientemente se ha relacionado con el aislamiento transcripcional e interacciones dentro de los cromosomas como son formaciones de asas (43, 36).

Estudios de locus específicos han apuntado a que una de las funciones de CTCF es mediar de forma directa el plegamiento de los cromosomas, así mismo, gracias a matrices de alta resolución de Hi-C, se ha observado que existe un enriquecimiento de sitios de unión a CTCF en los bordes de los TADs (34). Estas observaciones han llevado a llamar a CTCF como una proteína estructural del genoma. Esto se pudo observar claramente en el estudio antes mencionado realizado en 2017 en el cual, mediante una estrategia de degradación condicional en células troncales de ratón, se demostró que CTCF es un factor determinante para el plegamiento de los cromosomas en mamíferos, ya que se observó una pérdida tanto de TADs como de asas en ausencia de la proteína por lo tanto se concluyó que CTCF está principalmente involucrada en la organización de la arquitectura genómica al nivel de las sub-megabases, lo cual engloba tanto a los asas como a las TADs. De forma interesante este estudio mostró que, al reducir los niveles de CTCF hasta que resultaran indetectables,

los compartimentos A y B de la cromatina no se vieron afectados, por lo que se ha comenzado a mencionar que deben existir varios mecanismos moleculares que influyan la arquitectura genómica a distintos niveles (34).

Otra proteína considerada estructural en el genoma es la proteína cohesina la cual está formada por tres subunidades (SMC1, SMC3 y RAD21). Esta proteína posee una estructura en forma de anillo con la cual puede sostener a las hebras de ADN (42). Así mismo se ha observado que se localiza en el mismo sitio donde se encuentra la proteína CTCF y juntas parecen tener un rol muy importante en la formación y el mantenimiento de las asas de cromatina así como de creación de fronteras entre los TADs (40).

Otra proteína que actualmente ha ingresado a la categoría de arquitectural es Ying Yang 1 (YY1) la cual posee características similares a las de CTCF (presencia de dedos de zinc, formación de homo-dímeros, expresión constitutiva), es un factor transcripcional que forma dímeros y se ha encontrado formando asas de cromatina preferentemente entre promotores de genes y enhancers (43).

Importancia de la estructura del genoma en la expresión génica

La estructuración del genoma, especialmente en el nivel de los TADs tiene una importancia regulatoria para la expresión de los genes. Se ha observado en varias ocasiones que genes dentro de un mismo TAD tienen un patrón de regulación conjunta que permite a las regiones regulatorias contactar a los genes dentro del dominio y evitar contactar promotores fuera de él(29). De hecho, cuando las fronteras de los TADs son alteradas se ha observado activación ectópica de promotores génicos fuera del dominio que generan alteraciones en el desarrollo (ej. Polidactilia) (44) y también extensión de zonas activas donde normalmente debería haber heterocromatina, en el loci de los genes Hox (45).

Existen variados ejemplos de la función que se ha encontrado a esta organización tridimensional del genoma, sin embargo, aún no queda clara la relación que existe entre la organización que se observa con las técnicas de conformación de cromosomas, las proteínas estructurales que se han descrito y las alteraciones en la expresión que se han observado.

Cambios de la estructura del genoma en procesos celulares

Como ya se ha mencionado el estudio de la estructura tridimensional del genoma ha brindado datos nuevos acerca de posibles escalas de organización y regulación de la expresión genética. Sin embargo, existen limitadas evidencias que muestren si la arquitectura del genoma se ve alterada en cambios globales del genoma.

Uno de los estudios que han mostrado la relación entre la estructura del genoma y un proceso transcripcional ocurren en la inactivación del cromosoma X en mamíferos. La inactivación de este cromosoma es iniciada por un RNA largo no codificante que es un transcrito específico del cromosoma X inactivo llamado *Xist* (Fig.9). Este transcrito detona una cascada de eventos que llevan al empaquetamiento de uno de los cromosomas X que será inactivado en forma de heterocromatina facultativa mientras que el otro cromosoma, llamado Xa no es afectado (46).

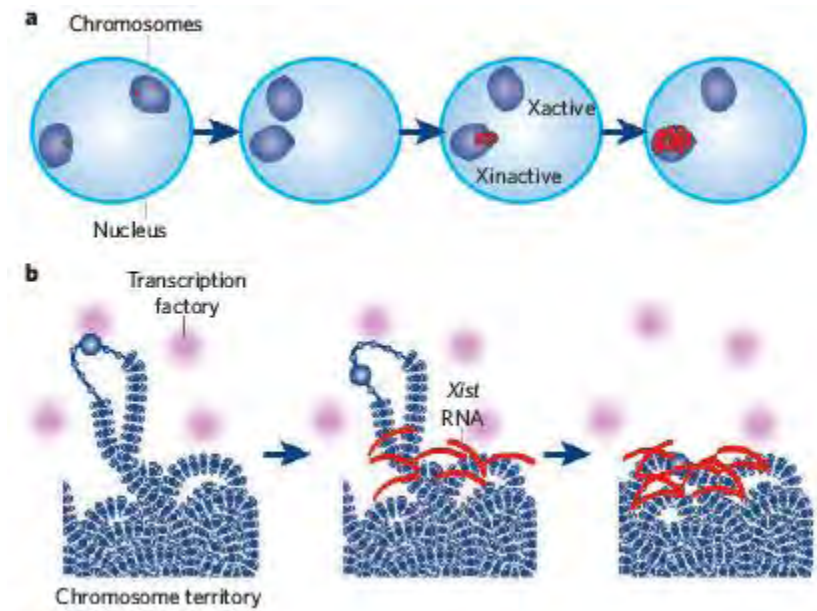


FIGURA 9. ESQUEMA DEL SILENCIAMIENTO DEL CROMOSOMA X EN RATÓN. A. MOVIMIENTO DE LOS 2 CROMOSOMAS X DURANTE EL PROCESO DE INACTIVACIÓN, SEPARACIÓN ENTRE Xi (INACTIVO) Y Xa (ACTIVO). LOS PUNTOS ROJOS EJEMPLIFICAN LA TRANSCRIPCIÓN DEL ARN LARGO NO CODIFICANTE QUE SE PRODUCE SÓLO EN Xi. B. LA COBERTURA DEL CROMOSOMA X INACTIVO POR EL ARN XIST EXCLUYE A POL II Y A LA MAQUINARIA DE TRANSCRIPCIÓN DEL TERRITORIO CROMOSÓMICO DE Xi. LOS GENES QUE SE ENCUENTRAN FUERA DE LA ZONA QUE XIST CUBRE SE VAN RETRAYENDO LENTAMENTE HASTA LOGRAR LA INACTIVACIÓN COMPLETA. (MODIFICADA DE: FRASER P Y BICKMORE W. 2007)

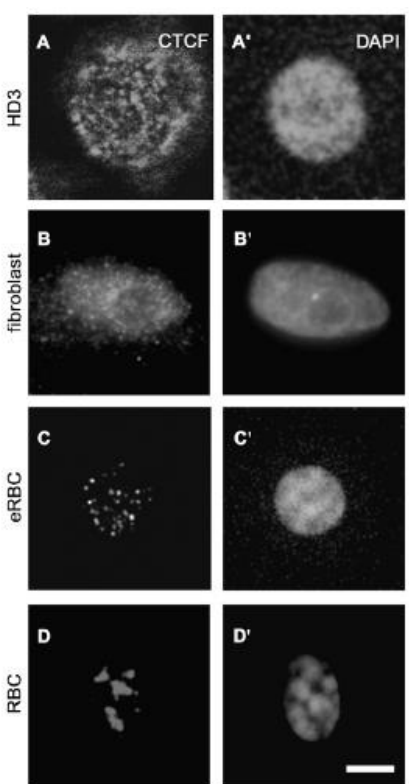
Giorgetti y colaboradores estudiaron la estructura del cromosoma X en ratón. Ellos realizaron un experimento de Hi-C en los dos cromosomas X de una hembra. Observaron que los compartimentos y los TADs del cromosoma X activo (Xa) eran similares a los detectados en los demás autosomas, sin embargo el cromosoma X inactivo (Xi) no mostraba ningún tipo de compartimentos y se encontraba particionado en dos dominios de interacción masivos que englobaban por completo a todo el cromosoma (39). Esto mostró que el cromosoma Xi, además de sufrir un silenciamiento transcripcional generalizado y una compactación de cromatina, también sufre una pérdida de compartimentación y TADs.

Por otra parte, el desarrollo eritroide es un proceso en el cual se sabe que la cromatina también pasa por un proceso de silenciamiento y compactación. No se sabe qué mecanismos median este silenciamiento a diferencia de lo que ocurre en el cromosoma X, sin embargo, el cambio transcripcional global es una característica que se mantiene en los

dos casos. Por esta razón y debido a que el proceso de maduración eritroide del pollo no atraviesa por el proceso de enucleación, consideramos que es un modelo adecuado para el estudio de la estructura tridimensional del genoma durante un proceso de silenciamiento genómico globalizado.

Antecedentes

No se conoce a profundidad el proceso por el que atraviesa el núcleo de las células durante la maduración de eritrocitos en *Gallus gallus*, sin embargo, se tienen algunas evidencias de que el genoma sufre silenciamiento masivo en los eritrocitos maduros. En este proceso de silenciamiento se encontraron evidencias de algunos cambios en la distribución de proteínas que podrían estar indicando una reorganización a nivel de la arquitectura de la cromatina. El estudio fue realizado en células eritroides de pollo tanto embrionarias como



terminalmente diferenciadas. En estas células se encontró que durante el curso de diferenciación la proteína CTCF, la cual está altamente relacionada con la estructura y mantenimiento de los TADs (27), era progresivamente excluida de la cromatina hacia áreas libres de ADN, las cuales llamaron dominios intercromatínicos (ICD por su siglas en inglés) y a también hacia las áreas de la periferia del núcleo (Fig.10) (47).

FIGURA 10 DISTRIBUCIÓN NUCLEAR DE CTCF EN CÉLULAS PROLIFERATIVAS Y NO PROLIFERATIVAS. HD3 DE POLLO (A Y A') ERITROBLASTO HD3 (B Y B') FIBROBLASTOS DE POLLO (C Y C') ERITROCITOS EMBRIONARIOS (10 DÍAS) (D Y D') ERITROCITOS ADULTOS TEÑIDOS CON ANTICUERPO POLICLONAL DE CONEJO CONTRA CTCF. SE PUEDE OBSERVAR EL CAMBIO EN LA ORGANIZACIÓN DE LA DISTRIBUCIÓN DE LA PROTEÍNA (KANTIDZE ET AL. 2007).

En el estudio que se realizó en células HD3, las cuales son eritroblastos de pollo transformados con el virus AEV que pueden proliferar activamente pero no expresan los

genes globina, se observó la distribución de CTCF. Se encontró que era más cercana a las células no eritroides (Fig.10) en comparación con los eritrocitos embrionarios (Fig.11 A y B). En estas células se pudo observar una variación drástica, CTCF estaba concentrada en un número limitado de compartimentos relativamente grandes y se encontraba localizada hacia la periferia nuclear en zonas que llamaron “libres de ADN” (Fig.11A'' y B'')(47). Estos patrones de exclusión de proteínas hacia el espacio inter-ADN se observaron en algunos candidatos más como la proteína Kaiso, la cual posee una actividad represora de la expresión génica mediante la des-acetilación de histonas, y la Polimerasa II, lo cual es una característica del proceso general de silenciamiento que sucede en los eritrocitos cuando llegan a la madurez (47).

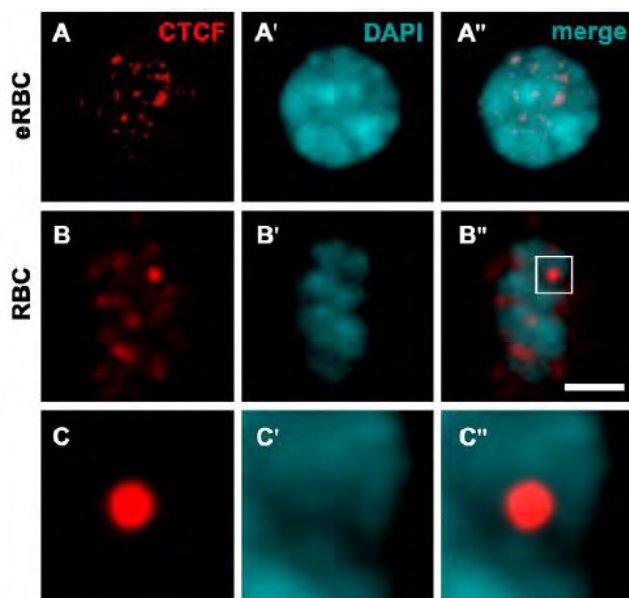


FIGURA 11 LOCALIZACIÓN POR INMUNOTINCIÓN DE CTCF EN NÚCLEOS DE CÉLULAS DEL SISTEMA ERITROIDE DE POLLO TEÑIDAS CON TINCIÓN DAPI (KANTIDZE ET AL. 2007).

Planteamiento del problema

La regulación de la expresión genética es un proceso complejo que aún requiere de mucha investigación para ser comprendido. Uno de los procesos que parecen tener una relación importante con los cambios en la expresión del genoma es la arquitectura del mismo, la cual se ha visto, en casos particulares, atraviesa por cambios en respuesta o ayuda a los

procesos por los que pasan las células al diferenciarse. El desarrollo del sistema eritroide ha sido estudiado y caracterizado ampliamente desde hace mucho tiempo debido a la importancia biológica y a la alta especialización de las células que lo comprenden. Por todo esto corresponde a un modelo interesante para el estudio de la regulación genética por la que atraviesan las células durante el proceso de diferenciación y maduración. *Gallus gallus*, al ser un organismo con un genoma bien caracterizado y una importancia económica para la sociedad humana, brinda un modelo de estudio novedoso y adecuado para conocer un proceso global como el silenciamiento génico general y la relación entre un cambio drástico en la expresión y la organización de la arquitectura de la cromatina.

Hipótesis

La estructura tridimensional del genoma es un factor de regulación de la expresión genética por lo cual se podrá observar un cambio en la organización de la arquitectura del genoma denotado por la pérdida de estructuras, especialmente de dominios topológicamente asociados (TADs) relacionada con el silenciamiento de la transcripción durante la maduración de eritrocitos además de una ganancia de heterocromatina específicamente en el estadio de los eritrocitos adultos de *Gallus gallus*.

Objetivo general

Comparar la estructura tridimensional de la cromatina en los diferentes estadios de desarrollo durante el proceso de maduración de eritrocitos en *Gallus gallus*.

Objetivos particulares

1. Comparar la expresión genética en eritrocitos inmaduros y maduros de *Gallus gallus* así como en la línea celular HD3.

2. Analizar las interacciones del genoma utilizando la técnica de Hi-C en la línea celular HD3 y en los eritrocitos de *Gallus gallus*.
3. Construir una matriz de interacciones de Hi-C para la observación de TADs y compartimentos de la cromatina con los datos obtenidos.
4. Explorar y determinar las características topológicas particulares del genoma del pollo.
5. Comparar la estructura tridimensional de la cromatina en los tipos celulares analizados (HD3, eRBC y RBC) y observar si existe una diferencia en la arquitectura de la misma.

Método

Cultivo celular y obtención de muestras

Se utilizaron células de la línea HD3 de eritroblastos inmortalizados de pollo. Se cultivaron en medio HD3 (DMEM, 10% SFB, 2% Suero de pollo) a 37°C. Se calculó el número de células por conteo en cámara de Neubauer y se dividieron en pellets de 10 millones, los cuales se almacenaron a -80°C para su posterior uso en las técnicas de Hi-C y extracción de ARN.

Para la obtención de eritrocitos embrionarios se obtuvieron huevos fecundados de *Gallus gallus* de diez días de desarrollo, se realizó un protocolo de extracción de células el cual consiste en remover la parte superior del cascarón en el lado donde se encuentra la cámara de aire. Posteriormente se procede a retirar la parte exterior de la membrana que permite el intercambio de gases en el huevo, sin romper la parte interior de la misma sólo para dejar expuestas las arterias extraembrionarias de esa zona. Con una aguja se realiza una punción en las arterias de tal forma que la sangre salga hacia la parte superior de la membrana y se acumule allí donde es recolectada con una pipeta y colocada en solución de PBS con 10% de EDTA. Esta forma de extracción permite obtener la sangre con un bajo riesgo de contaminación con las sustancias nutritivas que rodean al embrión.

Posteriormente se realizó un protocolo de conteo indirecto de las células sanguíneas para la separación de la muestra en pellets celulares de 5 millones destinados tanto para el protocolo de Hi-C como para la extracción de RNA los cuales se almacenaron a -80°C .

Para la obtención de cultivos primarios de fibroblastos se obtuvieron huevos fecundados de diez días de desarrollo al igual que para la obtención de los eritrocitos inmaduros. Se procedió a extraer el embrión del huevo y sacrificarlo. Posteriormente se extrajo la mayor cantidad de tejido muscular posible del pecho del embrión, siempre intentando evitar contaminación de órganos y tejido adiposo, y con ese tejido se realizó un protocolo para obtener cultivos celulares primarios de fibroblastos embrionarios. El cultivo se expandió para congelación en nitrógeno y también se obtuvieron muestras de 10 millones de células para realizar los protocolos de Hi-C y extracción de RNA las cuales se almacenaron a -80°C .

Los eritrocitos adultos se obtuvieron de gallinas de 34 semanas de edad las cuales se sacrificaron para posteriormente recolectar la sangre en buffer anticoagulante. Se sacrificaron tres gallinas para obtener tres réplicas biológicas. Se realizó el protocolo de conteo indirecto al igual que con las células embrionarias y se calculó el número de células presentes por mililitro de solución. Posteriormente la sangre se dividió en tubos de 1.5 y 15 ml y se centrifugó para generar pellets de 800 millones de células (presentes en 1ml de la solución de sangre) algunos de los cuales se prepararon para el protocolo de Hi-C mientras que otros se congelaron y se almacenaron a -80°C para su posterior uso.

Obtención de muestras eritroides de pollo adultos

Para la obtención de muestras de eritrocitos adultos se adquirieron tres gallinas hembras de 38 semanas de edad como réplicas biológicas. Se procedió a su sacrificio y a la recolección de la sangre la cual representó aproximadamente entre 50 y 70ml por cada ejemplar. La sangre fue recolectada en una solución para evitar su coagulación y posteriormente se realizó el protocolo de conteo indirecto para células eritroides. Se calculó el número de células presente en la mezcla de sangre y solución anticoagulante y se

realizaron alícuotas de 1ml con un aproximado de 8×10^8 células. Estas alícuotas se centrifugaron y se congelaron por *Flash freeze* para su posterior almacenamiento a -80°C .

También se separaron 6ml de cada réplica biológica a los cuales se les realizó el paso de fijación con formaldehído del protocolo de Hi-C. Estas células fueron también almacenadas a -80°C .

Protocolo de extracción de núcleos

Para la purificación de núcleos se utilizó un buffer al 1% de NP-40. Las células se lavaron en PBS frío estéril y posteriormente se centrifugaron a 1,200 rpm por cinco minutos a 4°C . Se eliminó el sobrenadante y se re suspendieron las células en buffer NI (15mM Tris-HCl pH 7.5, 300mM de Sacarosa, 10mM HEPES, 60mM KCl, 4mM MgCl, 5mM NaCl). Se agregaron 2.5ml del mismo buffer NI adicionado con 1% de NP-40 (Igepal) y se mezcló por pipeteo por 1 minuto. Se centrifugó a 1,400 rpm por cinco minutos a 4°C . Se eliminó el sobrenadante y se realizaron dos lavados con 10 y 5 ml de buffer NI. De esta forma se obtuvieron los núcleos de las células eritroides.

Para la purificación de núcleos de fibroblastos se utilizó el mismo protocolo añadiendo un paso adicional en el momento de la lisis. Al agregar el buffer NI-NP-40 se utilizó un homogeneizador para añadir una presión mecánica y facilitar la extracción de los núcleos celulares.

Diseño de oligonucleótidos (*primers*)

Expresión de genes

Para la medición de expresión de genes se diseñaron *primers* que amplificaran una región no mayor a 150 pb ya que ese tamaño es el recomendado para reacciones de PCR cuantitativas.

Se eligieron tres grupos de genes, un grupo de expresión constitutiva, otro de expresión eritroide específica y uno más de expresión de las proteínas globinas. Posteriormente se utilizó el buscador de la Universidad de Santa Cruz para buscar las regiones del genoma

pertinentes en la última anotación hasta el momento que corresponde a galGal 5.0. Los *primers* fueron diseñados utilizando la plataforma Benchling y el programa Primer3. Se diseñaron tres *primers* por gen, uno de los cuales es sentido y los otros dos anti-sentido abarcando una región que amplifica exón-exón y otra que amplifica exón-intrón (Fig. 12). Esto es debido a que se quería evaluar tanto los ARNs que ya habían pasado por el proceso de *splicing* como los ARN nacientes.



FIGURA 12 DISEÑO DE *PRIMERS* DE EXPRESIÓN EXÓN-EXÓN Y EXÓN-INTRÓN. SE DISEÑARON PARA AMPLIFICAR REGIONES MENORES DE 150PB.

Los genes elegidos y las secuencias diseñadas se muestran en la siguiente tabla.

TABLA 1. *PRIMERS* DISEÑADOS PARA LA MEDICIÓN DE LA EXPRESIÓN DE GENES.

Genes	Secuencia	Largo
Gen de alfa actina de pollo	ACAGCACGTTGTTGGCATACAG	22
	GACAGCACGTTGTTGGCATAACA	22
	ACAGCACGTTGTTGGCATACAG	22
	GACAGCACGTTGTTGGCATAACA	22
	<u>TTAGCTCACGCCTGCTCTTCTT</u>	22
	<u>TTAGCTCACGCCTGCTCTTCTT</u>	22
	GAAACCCTCTCCAGCCTTCCT	22
	AACCCTCTCCAGCCTTCCTTC	22
Gen de beta actina de pollo	TGTAGTGCCTCCAGAAAGCACA	22
	TGTAGTGCCTCCAGAAAGCACA	22
	<u>TGCTTCCTCAGTGTCTGCTGT</u>	22
	CTGCCAGAGACTCTCTCCAG	22
Gen ribosomal de pollo	TCGAGGAATACGAGCAACAGCT	22
	TCGAGGAATACGAGCAACAGCT	22
	<u>GGGGAACAGCAGTGTGTCAC</u>	22
	CATTCGCCCTGATGTGGTGAAC	22
Gen de glucuronidasa de pollo	CTGCGGATCAGTTCCTCCATCA	22
	CTGCGGATCAGTTCCTCCATCA	22
	<u>ATCCAATCTCAACCTCCCCTGG</u>	22

	ATGGACCTGTGTGACACCTACG	22
Gen globina alfa A	GCTGCTGACAAGAACAACGTCA	22
	GCTGCTGACAAGAACAACGTCA	22
	<u>GGGAAGAGGATCAGGGGAGAGA</u>	22
	GTGACAGATCGAAGTGGGGGAA	22
Gen globina alfa D	CAACCTGCATGCCTACAACCTG	22
	CAACCTGCATGCCTACAACCTG	22
	<u>CCTGTAACCCCAAGATCCCCTG</u>	22
	AGACAGGAACTTGTCGAAGGCA	22
Gen globina beta A	TGCATTGTGACAAGCTGCATGT	22
	TGCATTGTGACAAGCTGCATGT	22
	<u>CAGCATGGGGAGATCTATGGGG</u>	22
	GTGAAGTCCTTGCTGAAGTGGG	22
Gen factor nuclear eritroide 2	GTAAAGGGTTTCAGCGGTCGTG	22
	GTAAAGGGTTTCAGCGGTCGTG	22
	<u>GTGGTGAAGGGCAGTGAAAAGG</u>	22
	GTCTCCTCTCCTCCCAGAGACA	22
Gen proteína de membrana eritroide, banda 4.1	TTTTATCCTCCGGATCCCGCTC	22
	TTTTATCCTCCGGATCCCGCTC	22
	<u>AATCCGATTCCAGCAGCTTCTC</u>	22
	AGCTCAGACTGCACTGTGTAGG	22
Gen proteína de membrana eritroide, banda 4.2	GGGCAGCCTTTCGTTATCACAG	22
	GGGCAGCCTTTCGTTATCACAG	22
	<u>GGCTGTTGGACGAGAGCTCTAA</u>	22
	CCTGGGTTCCGTCTGCTTTAGA	22

*secuencias subrayadas corresponden a intrones.

Controles de Hi-C

A lo largo del protocolo de Hi-C se requieren realizar varios pasos de verificación de que la técnica está funcionando correctamente. Para esto es necesario diseñar oligonucleótidos específicos para el genoma y la enzima de restricción que se va a utilizar.

Se diseñaron *primers* para verificar el paso de digestión, ligación y también la presencia de alguna interacción de regiones del genoma que no se encontraran aledañas la una a la otra (Tabla 2).

TABLA 2 PRIMERS DISEÑADOS PARA CONTROLES DE HI-C UTILIZANDO LA ENZIMA DE RESTRICCIÓN *Mbol*.

Tipo de control	Secuencia	Tamaño	Descripción
Control de interacción a corta distancia	GCTGTGAGTGAAGCAGCATTGT	22	Enhancer alfa MRE del gen alfa globina de pollo
	CTTCACTCACAGCATTGCCAGG	22	Enhancer alfa MRE del gen alfa globina de pollo
	CCCGGCCTTATCTTATCCGCA	22	Gen Globina Alfa-D de pollo
	AGAACACATCCCTTCTGGCCTC	22	Gen Globina alfa-A de pollo
Control sin sitio de restricción (control de digestión)	ACTTCCCCACTTCGACCTTTC	22	
	CAGGTTGTAGGCATGCAGGTTG	22	
Control de ligación	CACCGCAGAGGCTGTAACAAAA	22	Gen de histona 1 de pollo digerido con Mbol. Primer sitio de restricción.
	CTACATTAGGCCGTCCACACCT	22	Gen de histona 1 de pollo digerido con Mbol. Segundo sitio de restricción.
	AAAATACGAGCATCTGACCCGC	22	Gen de histona 1 de pollo digerido con Mbol. Tercer sitio de restricción.
Control de ligación	CAAACCCCGAAATTGGCTGGAA		Promotor del gen Sim1
	TGTCCTTTCTCCCCCTCAGAGA		Enhancer específico de aves del gen Sim1

Control de interacción

Para el control de interacción se eligió la región de los genes Alfa globina donde en el 2008 mediante el protocolo de 3C se observaron las interacciones de los elementos que

conforman este locus. Se encontró que los promotores de los genes Alfa A y Alfa D interactuaban con una región hacia el extremo 5' llamada MRE la cual se ha descrito como un enhancer (26).

El diseño de los *primers* se comenzó eligiendo la región reportada y posteriormente se realiza una digestión virtual de dicha región para observar los sitios de restricción de la enzima que se utilizará. En este caso se localizaron los sitios *Mbol* y, como se muestra en la figura 13, se diseñaron *primers* cercanos a los sitios de restricción con el fin de encontrar una señal de amplificación mediante PCR cuando la digestión, rellenado y ligación se realizarán de acuerdo al protocolo de Hi-C.



FIGURA 13 LOS *PRIMERS* SE DISEÑAN EN UNA ZONA CERCANA AL SITIO DE RESTRICCIÓN PREDICHO. EN LA LITERATURA SE ENCUENTRAN REPORTADAS ALGUNAS INTERACCIONES DEL GENOMA DE *GALLUS GALLUS* QUE FUERON ENCONTRADAS MEDIANTE TÉCNICAS COMO EL 3C O EL 4C. LA B REPRESENTA LA MOLÉCULA DE BIOTINA QUE ES AÑADIDA EN UNO DE LOS PASOS DEL PROTOCOLO DE HI-C. LOS *PRIMERS* DISEÑADOS EN REGIONES LEJANAS DEL GENOMA, AL ESTAR EN CERCANÍA ESPACIAL Y DESPUÉS DEL PROCESO DE LIGACIÓN BRINDAN UNA SEÑAL RECONOCIBLE MEDIANTE PCR.

Control de digestión

Debido a que los pasos de digestión y ligación en el protocolo de Hi-C marcan en gran medida la calidad del resultado del procedimiento, se diseñaron *primers* para corroborar que la enzima de restricción, el rellenado de los extremos cohesivos y la ligación habían funcionado correctamente.

Estos *primers* se diseñaron utilizando la plataforma de *benchling* y el programa *Primer 3*. El diseño consistió en elegir dos sitios de restricción de la enzima que se esté utilizando, en este caso *Mbol*, que sean contiguos. Se diseñaron dos *primers* sentido cerca de los dos sitios de restricción como se muestra en a figura 14. La base teórica de este diseño consiste

en que, al momento de la digestión y la religación de la fibra de ADN existe una probabilidad de que el segundo fragmento gire y se ligue frente al anterior permitiendo la amplificación por PCR de esa zona con los *primers* diseñados la cual no sería posible en el genoma sin digerir del organismo.



FIGURA 14 DISEÑO DE *PRIMERS* DE CONTROL DE LIGACIÓN. SE DISEÑAN DE FORMA SENTIDO-SENTIDO EN DOS SITIOS DE RESTRICCIÓN CONTIGUOS ASUMIENDO QUE DESPUÉS DEL PROCESO DE DIGESTIÓN Y LIGADO EXISTIRÁ UNA PROPORCIÓN DE FRAGMENTOS QUE HABRÁN GIRADO Y PERMITIRÁN UNA AMPLIFICACIÓN QUE EN LA PORCIÓN ORIGINAL NO SERÍA POSIBLE.

Además de los controles de interacción los *primers* diseñados (como se muestra en la figura 14) permiten observar la eficiencia de ligación que se obtuvo en el protocolo. Esto se debe a que posterior al paso de la digestión con *MboI* se realiza un rellenado con la polimerasa Klenow de los extremos cohesivos que deja la enzima de restricción y se introduce un nucleótido marcado con biotina. La enzima *MboI* tiene la característica de que en el proceso de rellenado se recupera el sitio de restricción anterior además de crearse ahora un nuevo sitio de restricción para la enzima *ClaI* (Fig. 15). De este modo, mediante la realización de una reacción de PCR y posteriormente la digestión con las enzimas antes mencionadas se pudo observar cualitativamente si el proceso fue realizado correctamente.



FIGURA 15 CUANDO SE REALIZA LA DIGESTIÓN Y LIGACIÓN CON LA ENZIMA DE RESTRICCIÓN *MBOI* SE CREA UN NUEVO SITIO DE RESTRICCIÓN DE LA ENZIMA *CLAI* QUE PERMITE OBSERVAR LA EFICIENCIA DE DIGESTIÓN EN EL PROTOCOLO.

Extracción de RNA

La extracción de RNA se realizó por el método e Trizol. Se utilizó 1 ml de Trizol por muestra (de todos los tipos celulares a analizar) y se re suspendieron utilizando una pipeta o un homogeneizador en las muestras de sangre de pollos adultos. Posteriormente se añadió 0.2 ml de cloroformo a las muestras y se agitaron vigorosamente. Se centrifugaron a 12,000 g por 15 minutos a 4°C y se recuperó la sección acuosa o el sobrenadante de cada una de las muestras. Se les añadió 0.5 ml por cada mililitro de Trizol a las muestras y se dejaron precipitando a -20°C por 12 horas aproximadamente. Pasado este tiempo se centrifugaron a 12,000 g por 10 minutos y se realizaron dos lavados con 1 ml de etanol al 75% y un último lavado con 1ml de etanol absoluto. Las muestras se re suspendieron en agua inyectable o tratada con DEPC y se cuantificaron utilizando en NanoDrop.

Estandarización de *primers* para PCR cuantitativa

La estandarización de las condiciones para la amplificación de los genes y regiones elegidas para medición de expresión se realizó probando distintas condiciones de PCR.

Se variaron tanto la concentración final de magnesio como la temperatura de extensión. En el caso del magnesio se probaron las condiciones de 2.5mM y 3mM de magnesio ya que el kit utilizado para realizar la PCR cuantitativa (KAPA SYBR Fast KR0389-v9.13) contiene una concentración final de 2.5mM de MgCl.

Las temperaturas probadas variaron entre 60 y 62°C para eliminar las bandas que amplificaran regiones inespecíficas en las muestras.

Los *primers* de expresión que funcionaron correctamente bajo estas condiciones se trasladaron a las condiciones de PCR cuantitativa y se procedió a realizar la estandarización pertinente.

qPCR

Se utilizó el Kit KAPA SYBR Fast KR0389-v9.13 para la realización de las reacciones de PCR cuantitativa. Los *primers* se diseñaron en regiones que reconocían amplificaciones exón-exón e intrón-exón para la medición de ARN mensajeros maduros y pre-ARNs.

Para la estandarización de los pares de *primers* se procedió a probar distintas condiciones de temperatura en las reacciones, así como cantidad de cDNA y concentración de *primers*.

Hi-C

Fijación

Para la fijación se utilizaron células recolectadas en el momento ya sea de cultivo celular, en el caso de HD3 y fibroblastos o de sangre de animales. La fijación se realizó momentos después de la recolección de la sangre.

La muestra se re suspendió en 35ml de DMEM 10%SFB, posteriormente se añadieron 5ml de formaldehído al 16% para una concentración final del 2% y se incubó por 10 minutos a temperatura ambiente mezclando por inversión 10 veces cada dos minutos de incubación.

Inmediatamente después de los 10 minutos se añadieron 6 ml de glicina 1M fría para detener la reacción de entrecruzamiento y se incubó por 5 minutos a temperatura ambiente, de la misma forma que el paso anterior, mezclando por inversión cada dos minutos. Posteriormente se mantuvo 15 minutos en hielo sin agitación.

Se centrifugó a 1500 rpm por 10 minutos a 4°C. Posteriormente se descartó el sobrenadante y el pellet se re suspendió en PBS1X frío hasta un volumen final de 50ml.

En ese paso se separaron las muestras (en caso de ser necesario) en alícuotas de 10 millones de células. Paso seguido se volvieron a centrifugar con las mismas condiciones utilizadas en el paso anterior y se descartó el sobrenadante. Los pellets celulares se congelaron en nitrógeno líquido para ser almacenados a -80°C y posteriormente utilizados.

Lisis

Después de descongelar los pellets celulares en hielo de los cuatro tipos que se utilizaron se procedió a realizar el protocolo de lisis celular para extraer y trabajar únicamente con los núcleos celulares. Para esto se preparó un buffer de lisis celular (10mM Tris-HCl pH8, 0.2% NP 40/Igepal, 10mM NaCl y una pastilla de inhibidor de proteasas) y las células fueron resuspendidas en 1ml de este buffer frío. Se incubaron por 30 minutos en hielo mezclando por inversión cada 5 minutos. Posteriormente se centrifugaron a 300g por 5 minutos a 4°C y se descartó el sobrenadante. Se realizó un lavado con buffer de lisis y uno con buffer NE2 1.25X frío. Posteriormente se resuspendieron en 360µl de buffer NE2 1.25X.

Para permeabilizar los núcleos se añadieron 11 µl de SDS al 10% (0.3% final) y se incubaron a 37°C por 45 minutos a 950 rpm.

Posteriormente se añadieron 75 µl de Tritón X-100 al 10% (1.6% final) y se volvieron a incubar en las mismas condiciones que en el paso de SDS.

Digestión

Para la digestión de la cromatina se utilizó la enzima *Mbol*. Se añadieron 250 U de enzima divididas en dos sesiones. Las primeras 200 U se añadieron a los núcleos y después de mezclar gentilmente con la pipeta se colocaron en incubación a 37°C a 950 rpm por toda la noche.

La mañana siguiente se añadieron 50 U de *Mbol* adicionales y se volvieron a incubar bajo las mismas condiciones por dos horas más.

Habiendo transcurrido las dos horas los tubos se colocaron por 20 minutos a 65°C con la finalidad de inactivar la enzima de restricción y proceder con los pasos subsecuentes.

Rellenado y marcaje con biotina

Trabajando los tubos en hielo se realizó una mezcla de dinucleótidos (1.5µl de dGTP 10 mM; 1.5µl de dCTP 10 mM; 1.5µl de dTTP 10mM y 37.5µl de dATP-biotina-14 0.4mM (Invitrogen 19524-016). Se añadió también 50 U del fragmento largo de la polimerasa de ADN I (NEB M0210L) a cada tubo junto con la mezcla preparada anteriormente. Se mezclaron con la pipeta y se colocaron en incubación por 75 minutos a 37°C en un programa de agitación de 700 rpm cada 30 segundos por 10 segundos.

Ligación y reversión del entrecruzamiento

Se preparó una mezcla de ligación de 100µl de *buffer* de ligación 10X, 10µl de BSA 10mg/ml, 15 U de ligasa de ADN T4 1U/µl y 425µl de agua. Se añadió la mezcla a cada uno de los tubos y posteriormente se incubaron a 16°C por toda la noche para promover la eficiencia máxima de ligación.

Al día siguiente se añadieron 50µl de Proteinasa K 10mg/ml y se incubaron las muestras a 37°C por dos horas con la finalidad de degradar la mayor cantidad de proteínas posible para posteriormente dejar en incubación a 65°C toda la noche.

Después de la reversión del entrecruzamiento se procedió a realizar la purificación del ADN, sin embargo, antes de realizarla las muestras se trataron con 10µl de RNAsa A 10mg/ml incubándose a 37°C por una hora.

Extracción y purificación del ADN

Para la extracción del ADN se realizó el método fenol-cloroformo en el cual se añadieron 2µl de glucógeno a cada muestra y se procedió a añadir 1 volumen de fenol-cloroformo pH 8. Se mezclaron con ayuda del vortex y se centrifugaron a máxima velocidad por 15 minutos. La fase acuosa se transfirió a un nuevo tubo y se realizó una segunda extracción con 100µl de TLE para recuperar la mayor cantidad de ADN posible de cada muestra.

Se precipitó el ADN con 2 volúmenes de etanol al 100% y 0.1 volumen de acetato de sodio 3M, se dejaron a -20°C por toda la noche.

Transcurrido ese tiempo se precipitaron a máxima velocidad (14000 rpm) a 4°C por 30 minutos. Los pellets se lavaron dos veces con etanol al 70% y se resuspendieron en TLE.

El ADN se cuantificó con Qubit en diluciones 1:50 para una medición precisa.

Sonicación

Para el proceso de sonicación de las muestras, el cual nos permitió homogenizar el tamaño de los fragmentos de ADN, se utilizó un COVARIS perteneciente a la Unidad de Biología Molecular del El Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán en la Ciudad de México. Se tomaron 10mg de cada una de las muestras de Hi-C y se colocaron el 130ml de agua inyectable.

Los parámetros del COVARIS se ajustaron para llevar el tamaño de los fragmentos de ADN a un aproximado de 400pb y fueron los siguientes:

- Nivel de agua: entre 10 y 15
- Factor "Duty": 10%
- Pico de potencia incidente: 140w
- Ciclos por ráfaga: 200
- Tiempo: 55s

Reparación de terminaciones de fragmentos

Las muestras se transfirieron a tubos de 1.5ml y posteriormente se les añadió 16µl de Buffer de ligación al 10X, 2µl de dATP 10mM, 5µl de Polimerasa de ADN T4 (NEB M0203S) y 7µl. Después se incubaron las muestras por 30 minutos a 20°C.

Posteriormente se añadió 5µl de dNTPs, 4µl de Buffer de ligación 10X y 5µl de Cinasa de polinucleótidos T4 o PNK (10 u/µl), 1µl de Polimerasa Klenow y 25µl de agua inyectable para obtener un volumen total de 200µl.

Las muestras se incubaron por otros 30 minutos a 20°C.

Selección de tamaño

Para seleccionar el rango de tamaño adecuado para el proceso de secuenciación se realizó una purificación secuencial con perlas SPRI utilizando 0.55 volúmenes de perlas para la primera selección y 0.95 para la segunda.

Posteriormente el ADN se diluyó en 100µl de TLE y las muestras se llevaron a un volumen final de 300µl.

“Pull-down” de biotina

Para el “pull down” de las moléculas de biotina se utilizaron 150 perlas de Invitrogen “Dynabeads MyOne Streptavidin C1” por cada una de las librerías.

Se lavaron las perlas con un buffer con 0.05% de Tween, posteriormente se lavaron con un buffer libre de Tween (EDTA 1mM, Tris-HCL pH8 5mM, NaCl 2M)

Se mezclaron las perlas con las muestras y se dejó que se incubaran por 30 minutos en rotación para que la selección de las moléculas de biotina fuera eficiente.

Con ayuda de un imán se aislaron las perlas y se removió el sobrenadante con todas las moléculas no marcadas con biotina. Posteriormente se realizaron una serie de lavados a las perlas con buffer con Tween y buffer NEB2.

Posteriormente las muestras se resuspendieron en una mezcla de buffer NEB2, Polimerasa Klenow sin actividad de exonucleasa y dATP, esto se realizó para colocar una adenina en cada uno de los extremos de los fragmentos y de esa forma se pudieran ligar los adaptadores de true seq posteriormente.

Después de la incubación a 37°C con los dATPs se procedió a realizar varios lavados, primero con buffer con Tween, después con buffer sin Tween y por último con buffer de ligación 1x en el cual finalmente se colocaron las perlas con la librería de Hi-C.

Ligación de adaptadores

Se utilizaron 4µl de adaptadores True Seq para secuenciación (15mM) y se ligaron utilizando 800 unidades de ligasa T4. Se incubaron por 2 horas.

Posteriormente se recapturaron las perlas en el imán y se lavaron nuevamente con los buffers antes utilizados para re suspenderlas finalmente en buffer NEB2 1X.

Control final de librería

Primero se realizaron pruebas amplificando por PCR de punto final con 10 y 15 ciclos para evitar una sobre amplificación de la librería, pero poder contar con el material suficiente para la secuenciación.

Posteriormente se realizó una PCR con 10 ciclos para cada una de las muestras por duplicado. Se tomó la réplica de PCR de cada una de las muestras y se digirió con la enzima *ClaI* para observar la eficiencia de formación del nuevo sitio en la librería completa.

Amplificación final

Las muestras se amplificaron 10 ciclos mediante PCR punto final, posteriormente se purificó el ADN con ayuda de *SPRI beads*. Se recolectó el sobrenadante en el cual se encuentra la librería amplificada de Hi-C se mandó a secuenciar.

Análisis de datos

El análisis de los datos obtenidos del proceso de secuenciación masiva requirió la utilización de varias herramientas bioinformáticas. Para el primer paso del análisis el cual corresponde a la depuración de los datos provenientes de la secuenciación se utilizó el programa HiCUP (48), posteriormente para la construcción de las matrices de interacción de utilizaron dos programas HiC-Pro (49) y Juicer (50). La visualización de las matrices de interacción y la construcción de mapas de regiones específicas se realizó con ayuda de los programas HiCPlotter (51) y JuiceBox(52). Por último, para la identificación de TADs y de compartimentos de la cromatina se utilizaron, una función específica del programa Juicer llamada eigenvector y el programa TADtool (53).

HiCUP

HiCUP es una herramienta informática diseñada para analizar datos de organización tridimensional del genoma generados por experimentos como Hi-C. HiCUP toma los archivos pareados FASTQ los cuales se obtienen posteriormente al proceso de secuenciación(48). A partir de estos archivos HiCUP realiza un mapeo tomando como

referencia el genoma del organismo, en este caso el Galgal5, el cual es última versión publicada del genoma de *Gallus gallus*.

El segundo paso que realiza el programa HiCUP consiste en la depuración de datos, la cual consiste en eliminar del set todas las lecturas que representen artefactos del experimento de Hi-C. Ejemplos de estas lecturas son: secuencias re-ligadas consigo mismas, secuencias no ligadas, secuencias continuas linealmente dentro del genoma, duplicas de PCR, entre otras. El programa toma todas estas lecturas y las agrupa en una categoría llamada “Di-tags inválidas” y agrupa todas las lecturas informativas de interacciones de Hi-C en la categoría de “Di-tags válidas”.

Por último, HiCUP brinda un reporte de control de calidad (QC report) por sus siglas en inglés en el cual grafica la información del experimento e informa cuál es el porcentaje de lecturas pareadas válidas que encontró en el experimento, cuántas de estas lecturas corresponden a una interacción en *cis* (Intra cromosómica) o en *trans* (Inter cromosómica), cuáles fueron los artefactos que encontró en el experimento y en qué porcentaje. También brinda gráficas del tamaño promedio de los fragmentos secuenciados entre otros controles de calidad que ofrecen una visión general acerca de la ejecución del experimento.

Los archivos después del análisis en HiCUP poseen el formato BAM/SAM y son los que se utilizan para el posterior análisis con otros algoritmos.

HiCUP fue instalado desde la dirección <https://www.bioinformatics.babraham.ac.uk/projects/hicup/> utilizando la versión 0.6.1 del programa.

HiC-Pro

HiC-Pro es un programa diseñado para analizar datos crudos de Hi-C y normalizarlos. Puede procesar datos digeridos con una enzima de restricción o también utilizando una enzima DNAsa (49). El programa brinda una visión general de la calidad del experimento con

gráficos similares a los construidos por HiCUP y también procesa los datos para la construcción de las matrices. Después del análisis construye una carpeta con los datos pareados en formato “matrix” y brinda tanto los datos crudos como normalizados por el método de Corrección Iterativa y descomposición en eigenvectores (ICE por sus siglas en inglés), de esta forma se pueden construir los dos tipos de matrices con un visualizador.

ICE es un algoritmo que está diseñado para procesar lecturas de secuenciación pareadas como es el caso de datos de 3C o de Hi-C y aplica un método de normalización basado en el principio de que todos los loci dentro de un mapa de Hi-C deberían tener la misma visibilidad y de esta forma descompone los mapas de frecuencia de contactos en un grupo de sesgos y construye un mapa para normalizar las lecturas utilizando la probabilidad relativa de contacto entre dos loci genómicos logrando igualar la visibilidad entre todos los loci (54).

HiC-Pro fue instalado desde la dirección <https://github.com/nservant/HiC-Pro.git> utilizando la versión 2.7.0.

Juicer

Juicer es un programa de análisis de datos de Hi-C desarrollado por Lieberman Aiden y colaboradores que permite la visualización de matrices de interacción de forma interactiva con la ayuda de su programa complementario Juicebox. Así mismo Juicer contiene varias herramientas que permiten visualizar y anotar distintas características de los datos como son la función de eigenvector que permite obtener el Componente Principal de la distribución de los datos (50).

Juicer puede ser instalado desde la dirección <https://github.com/theaidenlab/juicer.git> en la versión 1.5.

TADtool

TADtool es una herramienta bioinformática que permite llamar Dominios Topológicamente Asociados utilizando también una ventana visual que permite observar los cambios en los TADs de acuerdo al movimiento de los parámetros del programa.

TADtool requiere una matriz normalizada de Hi-C y un archivo donde se especifiquen las coordenadas de todos los “bins” correspondientes a esa matriz (53). El programa contiene dos algoritmos que permiten el reconocimiento de las estructuras llamadas TADs. Uno de ellos es el Índice de Aislamiento el cual calcula la medida por cada segmento obteniendo el promedio de las interacciones que ocurren a través de ellos. Valles o mínimos que se encuentren a lo largo de los segmentos representan regiones de baja frecuencia de interacción dentro del Hi-C y por lo tanto son interpretados como fronteras de un TAD (55). De esta forma el programa puede reconocer los TADs en las distintas regiones de las matrices. La calibración de los parámetros se realizó con base en el experimento de Hi-C de HD3 para la comparación con los otros dos tipos celulares.

TADtool fue instalado desde la dirección <https://github.com/vaquerizaslab/tadtool.git> utilizando la versión 0.76 del programa.

HiCPlotter

HiCPlotter es una herramienta de visualización de matrices de Hi-C y características de las mismas. Requiere de una matriz de interacción de Hi-C, como ejemplo una calculada por el programa HiCPro y posteriormente permite construir una imagen para visualizarla. Adicionalmente el programa permite añadir características o datos a las mismas en forma de carriles o “tracks” que pueden mapearse en la parte inferior de los mapas(51), algunos ejemplos de estos pueden ser: datos de Inmunoprecipitación de cromatina, coordenadas de genes, TADs, compartimentos entre otros.

HiCPlotter fue instalado desde la dirección <https://github.com/kcakdemir/HiCPlotter.git> utilizando la versión 0.7.3.

Estandarizaciones y resolución de problemas

Estandarización de *primers* de expresión.

Para poder comenzar con el proceso de estandarización de los *primers* se realizó el protocolo de extracción de ARN para cada uno de los tipos celulares (Fibroblastos, HD3, eritrocitos embrionarios y eritrocitos adultos). Posteriormente se realizaron RT-PCRs para cada muestra utilizando *primers* al azar para obtener representada toda la población de ARNs presente en las células incluyendo aquellos que no hubieran atravesado por el proceso de *splicing*.

Se probaron dos temperaturas de hibridación (60°C y 62°C).

A continuación, se muestran los geles de agarosa al 1% de las pruebas realizadas para los pares de *primers* que amplifican desde un exón hasta otro exón y permitirán reconocer ARNm que han sido procesados por medio del corte y empalme (Fig. 16).

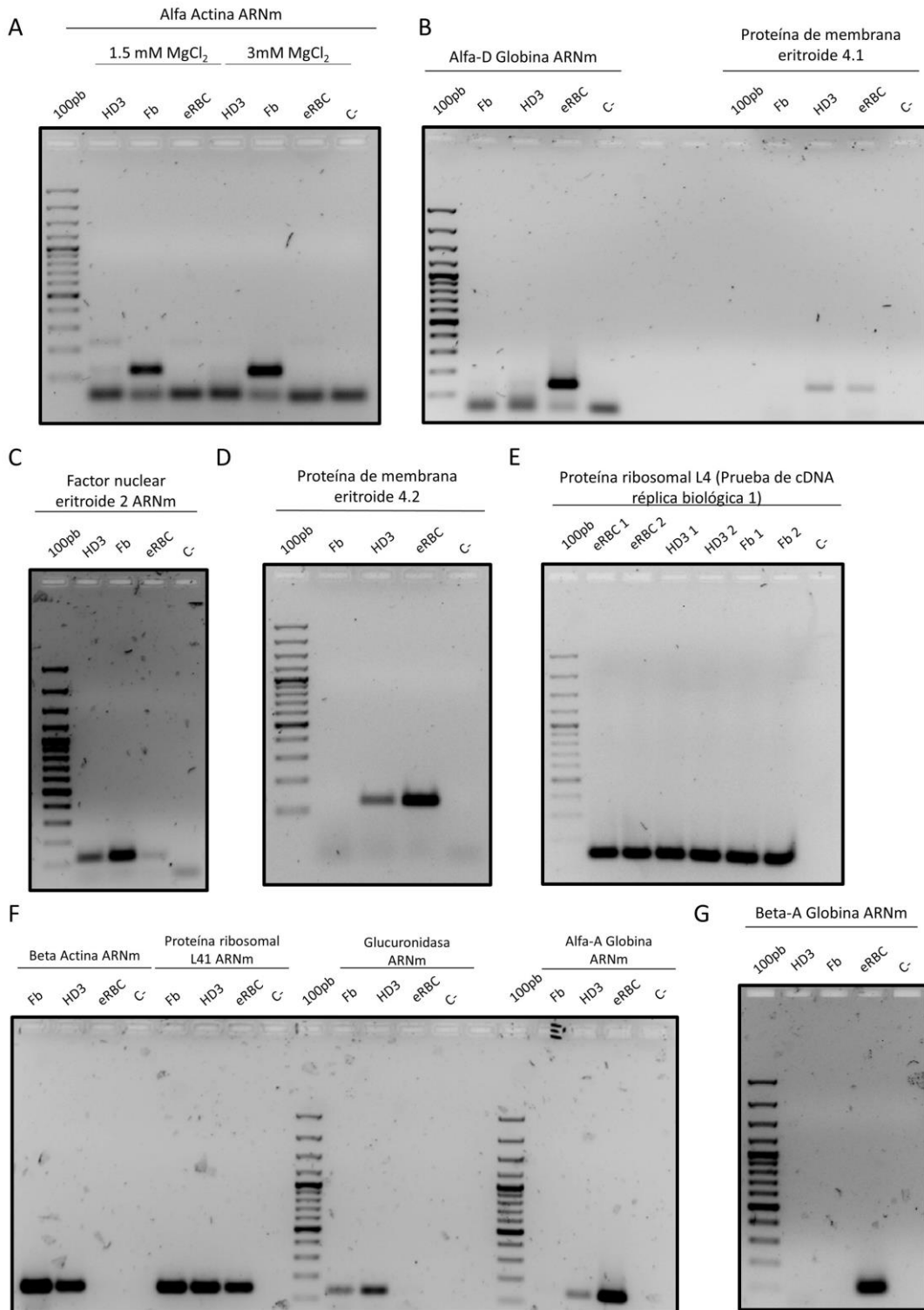


FIGURA 16 GELES 1% DE AGAROSA DE PCRs PARA MEDICIÓN DE EXPRESIÓN DE GENES CANDIDATO. A) PCR DE ALFA ACTINA (ACTA) ARNm UTILIZANDO DOS CONCENTRACIONES DE MAGNESIO. B) PCR DE ALFA GLOBINA D (HBAD) CON UNA SOLA CONDICIÓN DE MAGNESIO (1.5MM) Y PCR DE EPB41 MOSTRÓ SER ESPECÍFICA PARA LAS CÉLULAS

ERITROIDES. C) PCR FACTOR NUCLEAR ERITROIDE. EL GEN MOSTRÓ NO TENER UNA EXPRESIÓN ESPECÍFICA EN LAS CÉLULAS ERITROIDES ASÍ QUE FUE DESCARTADO PARA LOS POSTERIORES ESTUDIOS. D) PROTEÍNA DE MEMBRANA 4.2 (EPB42) MOSTRÓ UNA EXPRESIÓN ESPECÍFICA PARA LAS CÉLULAS DEL LINAJE ERITROIDE. E) LA EXPRESIÓN DEL TRANSCRITO DE LA PROTEÍNA RIBOSOMAL L4 (RPL4) SE DISEÑÓ COMO UN GEN CONSTITUTIVO DE EXPRESIÓN ESTABLE POR LO QUE SE UTILIZÓ COMO CONTROL PARA VERIFICAR QUE LAS RÉPLICAS TÉCNICAS EN LA RT-PCR FUNCIONABAN DE FORMA EFICIENTE Y SIMILAR. F) LA EXPRESIÓN DE BETA ACTINA NO ES HOMOGÉNEA ENTRE LAS MUESTRAS POR LO QUE SE DESCARTÓ COMO PARTE DEL SET PARA MEDIR LA EXPRESIÓN. EN EL CASO DE RPL4 SU EXPRESIÓN PARECE SER SEMI CONSTITUTIVA POR LO QUE SE INCLUYÓ EN LOS SETS PARA LAS PRUEBAS POR QPCR AL IGUAL QUE ALFA GLOBINA A QUE MOSTRÓ EXPRESIÓN CASI EXCLUSIVAMENTE EN LAS ERBC. EN EL CASO DE LA EXPRESIÓN DE LA ENZIMA GLUCURONIDASA NO SE OBSERVÓ AMPLIFICACIÓN HOMOGÉNEA DEL TRANSCRITO NI PRESENCIA DEL MISMO EN LAS ERBC POR LO QUE TAMBIÉN SE EXCLUYÓ DE LAS MUESTRAS A CUANTIFICAR. G) EXPRESIÓN DE BETA GLOBINA A MOSTRÓ SER ESPECÍFICA PARA LAS ERBC.

En el caso de los *primers* que amplifican para ADNc correspondiente a ARNm se observó una expresión adecuada por PCR de punto final para las tres globinas (HBAD, HBAA y HBBA), también para el gen ribosomal L4 (RPL4) y para dos de los genes eritroide específicos (EPB41 y EPB42). Por esta razón se decidió continuar con la estandarización de algunos de los pares de *primers* y, en el caso del par de Glucuronidasa y el factor nuclear eritroide dos, se decidió buscar otros candidatos que pudieran funcionar como ARN constitutivos.

Posteriormente se realizaron pruebas similares en los pares de *primers* que fueron diseñados para amplificar regiones desde un exón del gen hasta un intrón reconociendo a ARNm que aún no han pasado por el procesamiento post-transcripcional. El interés en este tipo de transcritos se relacionó con la hipótesis de que el silenciamiento transcripcional que ocurre en los eritrocitos adultos debe estar relacionado con varios procesos regulatorios que pueden comenzar desde el paso de la transcripción o quizá regular el proceso de *splicing* posterior. Debido a que no se sabe mucho acerca de lo que ocurre en estas células se esperaba comparar los dos tipos de transcrito y observar si existía alguna diferencia que brindara información acerca del proceso regulatorio de la expresión del genoma.

A continuación, se muestran las pruebas realizadas para los *primers* exón-intrón (Fig. 17):

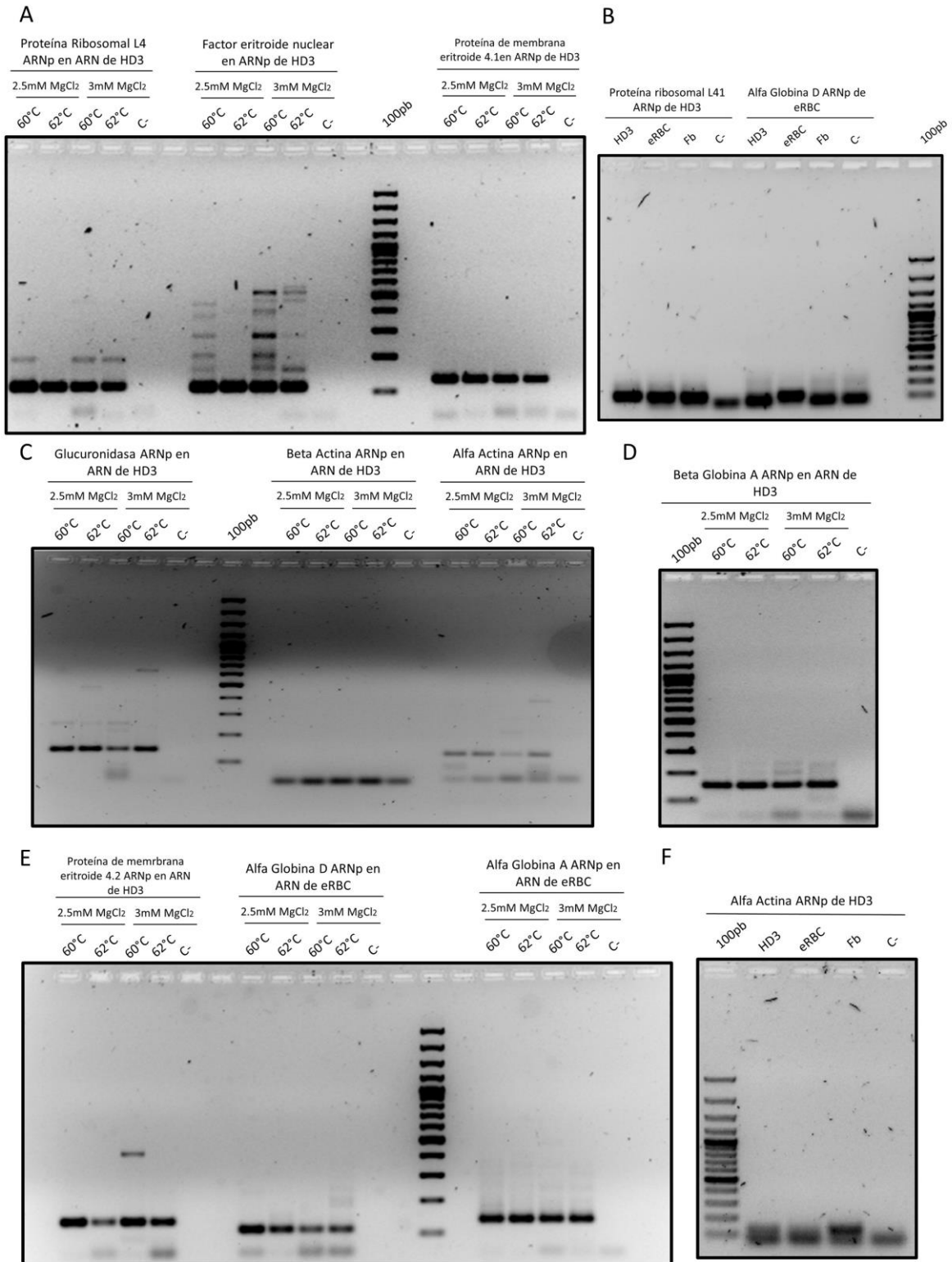


FIGURA 17 GELES DE AGAROSA 1% MOSTRANDO LOS RESULTADOS DE PRUEBAS DE ESTANDARIZACIÓN DE *PRIMERS* DISEÑADOS PARA MEDIR LA EXPRESIÓN DE GENES ANTERIOR AL PROCESO DE *SPLICING*. A) RPL4 Y EPB41 MOSTRARON RECONOCER ESPECÍFICAMENTE A SU TRANSCRITO EN ALGUNA DE LAS CONDICIONES PROBADAS, SIN EMBARGO, EL

FACTOR ERITROIDE NUCLEAR TENÍA DEMASIADOS BLANCOS INESPECÍFICOS POR LO QUE SE DECIDIÓ DESCARTAR ESE TRANSCRITO. B) PRUEBA DE RPL41 HBAD EN TODOS LOS TIPOS CELULARES ARN PRIMARIO. C) LOS *PRIMERS* DE GLUCURONIDASA Y LAS DOS ACTINAS (ALFA Y BETA) MOSTRARON TENER AMPLIFICACIÓN INESPECÍFICA. D) HBBA ARN PRIMARIO MOSTRÓ SER ESPECÍFICO BAJO CIERTAS CONDICIONES, EN ESPECIAL CON LA CONCENTRACIÓN DE 2.5MM DE MAGNESIO. E) EN EL CASO DE LAS DOS ALFAS GLOBINAS Y EPB42 LAS CONDICIONES DE 2.5MM Y 60°C RESULTARON ADECUADAS PARA SU USO EN PCR Y POSTERIORES PRUEBAS EN qPCR. F) LOS *PRIMERS* DE ALFA ACTINA PRESENTARON UNA EXPRESIÓN DIFERENTE EN CADA UNO DE LOS TIPOS CELULARES POR LO QUE SE DESCARTAR ESTE GEN COMO PARTE DEL GRUPO CONSTITUTIVO.

Se puede observar que, en general, los *primers* diseñados para la amplificación de regiones correspondientes a ARN no procesado presentan una mayor dificultad de estandarización. Eso puede deberse a varias razones, la primera es que hay probablemente una menor cantidad de transcrito primario dentro de las células debido a que el proceso de transcripción y procesamiento es casi simultáneo. En segundo lugar, es importante asegurarse que, durante la conversión de ARN a ADN complementario (ADNc) el tratamiento con DNAsa sea completamente efectivo ya que, si no, es posible que queden porciones de ADN genómico que los *primers* que reconocen exón-intrón amplifiquen.

Estandarización de expresión de genes estructurales por qPCR

Adicionalmente a la expresión que se midió de los genes candidato elegidos durante el desarrollo del proyecto se añadieron dos genes adicionales de interés biológico para complementar la información obtenida. El primer candidato que se añadió a la medición fue el factor de transcripción GATA1 el cual se encuentra estrechamente relacionado con el correcto desarrollo de los eritrocitos y los megacariocitos(56) por lo cual podría ser considerado dentro del grupo de los genes eritroides. La razón por la cual no se eligió este gen dentro de los candidatos primarios fue que, dentro del genoma del pollo, este gen no se encuentra completamente anotado y existen inconsistencias con el mapeo de las secuencias correspondientes a los intrones en diferentes buscadores de genomas. Sin embargo, se decidió utilizar los exones que tuvieran un tamaño mediano para diseñar los oligonucleótidos considerando que los de mayor tamaño podrían incluir con mayor

probabilidad un número variado de exones e intrones que aún no se encuentran correctamente anotados(Fig. 18).

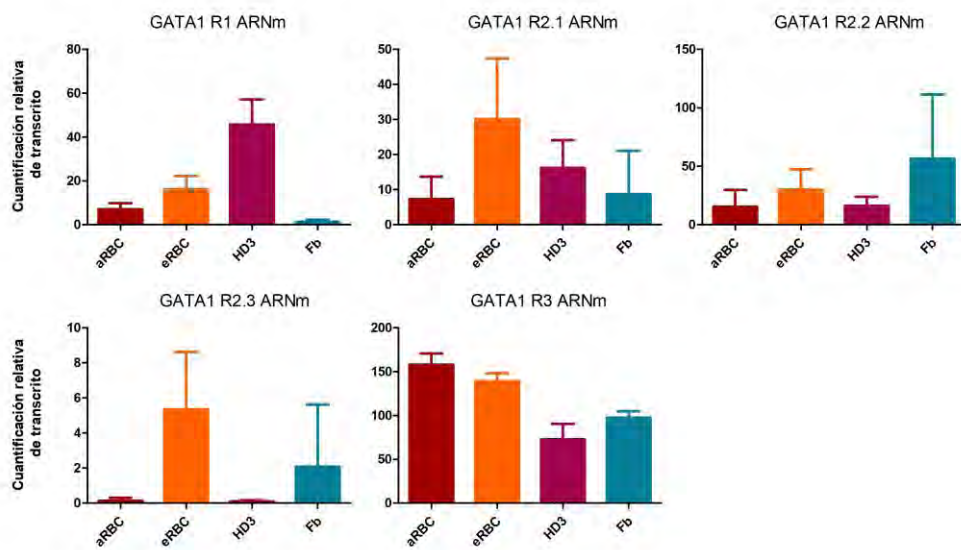


FIGURA 18 GRÁFICOS DE EXPRESIÓN DEL ARN MENSAJERO DEL FACTOR GATA1. SE MUESTRAN LAS RÉPLICAS 1, 2 Y 3 LAS CUALES SE REALIZARON CON DISTINTAS MUESTRAS BIOLÓGICAS. LA RÉPLICA 2 TIENE TRES RÉPLICAS TÉCNICAS UTILIZANDO DIFERENTES CANTIDADES DE ADNC EN LAS CURVAS ESTÁNDAR (COMENZANDO EN 50NG, 100NG Y 150NG COMO CANTIDAD INICIAL). SE PUEDE OBSERVAR QUE DENTRO DE LAS RÉPLICAS NO HAY UNA OBSERVACIÓN ESTABLE DE LA EXPRESIÓN DE ESTE GEN.

Los resultados de la medición de este factor se realizaron utilizando tres muestras biológicas distintas, sin embargo, no se encontró un patrón estable en la cuantificación de la expresión. Así mismo, dado que este gen se encuentra relacionado con la eritropoyesis, la observación de cantidades similares del mismo presentes en los fibroblastos indica que probablemente existió un problema durante el experimento ya sea con un blanco inespecífico de los oligonucleótidos o bien con la capacidad de los mismos de reconocer el blanco sobre el ruido de los dímeros que pudieran estar formando durante la reacción. Esto ocurrió de manera similar con el gen YY1 el cual también se eligió por estar reportado como productor de una proteína estructural en el genoma importante en el mantenimiento de asas de cromatina (43) (Fig. 19).

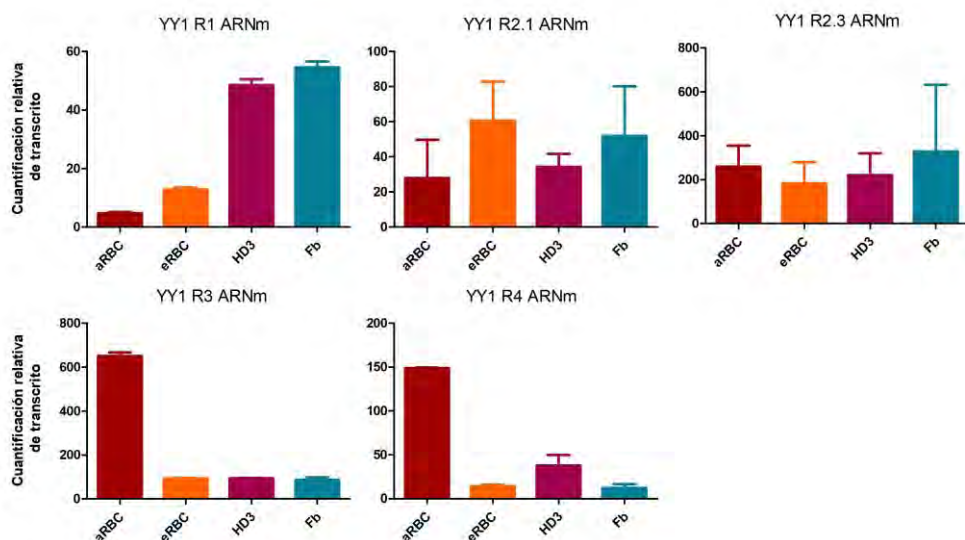


FIGURA 19 GRÁFICOS DE EXPRESIÓN DEL GEN YING YANG 1. SE MUESTRAN 4 RÉPLICAS BIOLÓGICAS DEL EXPERIMENTO. LA RÉPLICA 2 POSEE DOS RÉPLICAS TÉCNICAS UTILIZANDO DISTINTAS CANTIDADES DE ADN PARA LA CURVA ESTÁNDAR (100NG Y 150NG). SE PUEDE OBSERVAR QUE EL PATRÓN DE EXPRESIÓN DENTRO DE LAS RÉPLICAS ES INCONSISTENTE.

En el caso de YY1 se observó que el experimento 2 presentaba variaciones dentro de las réplicas técnicas que ponían en duda la calidad de la medición. En el caso de la réplica 1 el patrón observado parecía relacionarse con los genes medidos con anterioridad, sin embargo, en las réplicas 3 y 4 se observó un patrón completamente contrario. Se cree que posiblemente pudo haber una menor efectividad en la eliminación de todo el material genético correspondiente a ADN en las células adultas durante el tratamiento con la DNAsa. Esto podría ocurrir si la eficiencia de la enzima ha disminuido y especialmente en el caso de los eritrocitos adultos debido al gran número celular (800 millones de células) que se utiliza para extraer cantidades significativas de ARN.

Los dos factores medidos podrían brindar información valiosa acerca de los procesos biológicos por los que están atravesando los distintos tipos celulares por lo que se propone diseñar nuevos oligonucleótidos y realizar nuevas extracciones de ARN que permitan corroborar estos datos.

Estandarización de protocolo de Hi-C para fibroblastos

El primer experimento de Hi-C realizado en fibroblastos de cultivo primario de embriones de pollo de 10 días resultó con una baja eficiencia de acuerdo a los controles realizado por lo que se tomó la decisión de no continuar de procesar esa librería. Por lo tanto, se comenzó un proceso de estandarización de la técnica utilizando el mismo procedimiento que se realizó para la extracción de núcleos. Se tomaron dos muestras fijadas con formaldehído al 2% y almacenadas a -80 °C y se procesaron de acuerdo al protocolo de Hi-C con la diferencia de que a una de ellas se le sometió a un proceso de homogenización con ayuda de un homogeneizador de 2ml diez veces.

Posteriormente las muestras se procesaron de la misma forma hasta llegar a la purificación del material genético donde se realizaron nuevamente los controles para corroborar la eficiencia del experimento.

El primer control que se realizó consistió en tomar una alícuota de las dos muestras de Hi-C después del paso de digestión con la enzima de restricción Mbol y observarlas en un gel de agarosa al 1% en comparación con una muestra no digerida (Fig. 20)

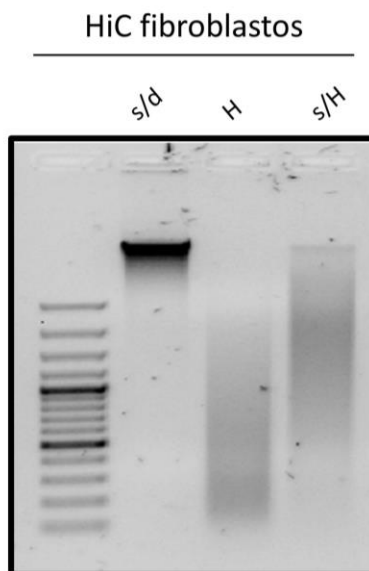


FIGURA 20 GEL DE AGAROSA 1% DE CONTROL DE DIGESTIÓN DE HI-C EN FIBROBLASTOS DE POLLO. EL PRIMER CARRIL MUESTRA EL MARCADOR DE PESO MOLECULAR 100PB, EL SEGUNDO ADN DE UN EXPERIMENTO DE HI-C EN FIBROBLASTOS QUE AÚN NO HA SIDO DIGERIDO. EL TERCER CARRIL MUESTRA LA DIGESTIÓN EN LA MUESTRA HOMOGENIZADA (H) 10 VECES DURANTE LA LISIS Y EL CUARTO CARRIL MUESTRA LA DIGESTIÓN DE LA MUESTRA QUE NO FUE HOMOGENIZADA (S/H).

En este gel se observó que las dos muestras fueron eficientemente digeridas con la enzima sin embargo hubo una mayor eficiencia de digestión en la muestra homogenizada.

Posteriormente se realizaron los controles internos y de interacción mediante el uso de oligonucleótidos diseñados para esto. Para este control se utilizaron los dos pares de oligonucleótidos probados anteriormente en los experimentos con las células eritroides y un par más que reconoce una posible interacción entre el promotor del gen *Sim1*, involucrado en el desarrollo de las plumas(57) y se observó que las bandas amplificadas por la PCR correspondían a los tamaños esperados y observados en los experimentos anteriores de Hi-C (Fig. 21). Sin embargo, al realizar el control de digestión con la enzima *Clal* se observaron irregularidades en el gel (Fig.22) al presentar bandas que no fueron digeridas ni con la enzima *Mbol* ni con la enzima *Clal*. Esto llevó a pensar que los oligonucleótidos se encontraban comprometidos por lo que se realizaron pruebas utilizando distintos reactivos, pero no se logró obtener un buen resultado.

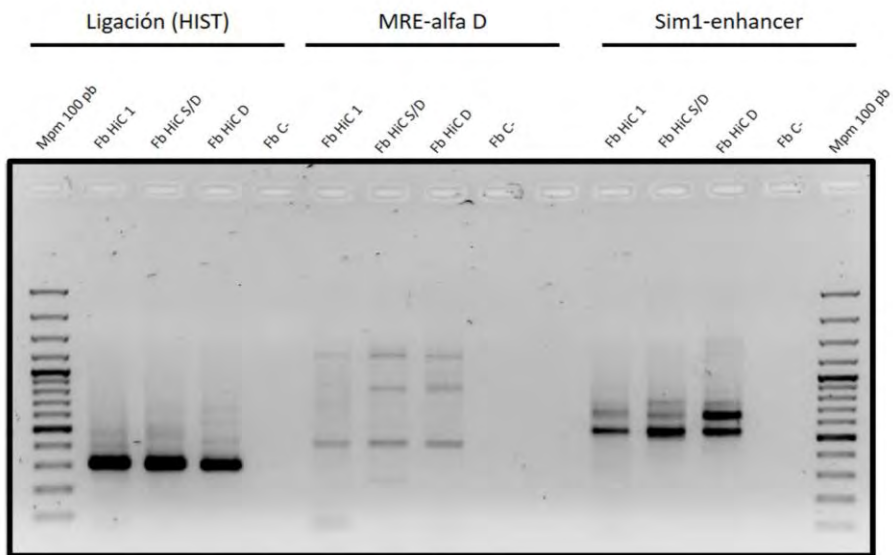


FIGURA 21 GEL 1% DE AGAROSA DE CONTROLES DE EXPERIMENTO DE HI-C EN FIBROBLASTOS DE POLLO. FB HI-C 1: LIBRERÍA DE EXPERIMENTO ANTERIOR DE HI-C CON BAJA EFICIENCIA REALIZADA EN FIBROBLASTOS. FB HI-C S/D: HI-C REALIZADO EN FIBROBLASTOS SIN EL PROCESO DE HOMOGENIZACIÓN. FB HI-C D: HI-C REALIZADO EN FIBROBLASTOS CON DIEZ HOMOGENIZACIONES EN EL PASO DE LISIS.

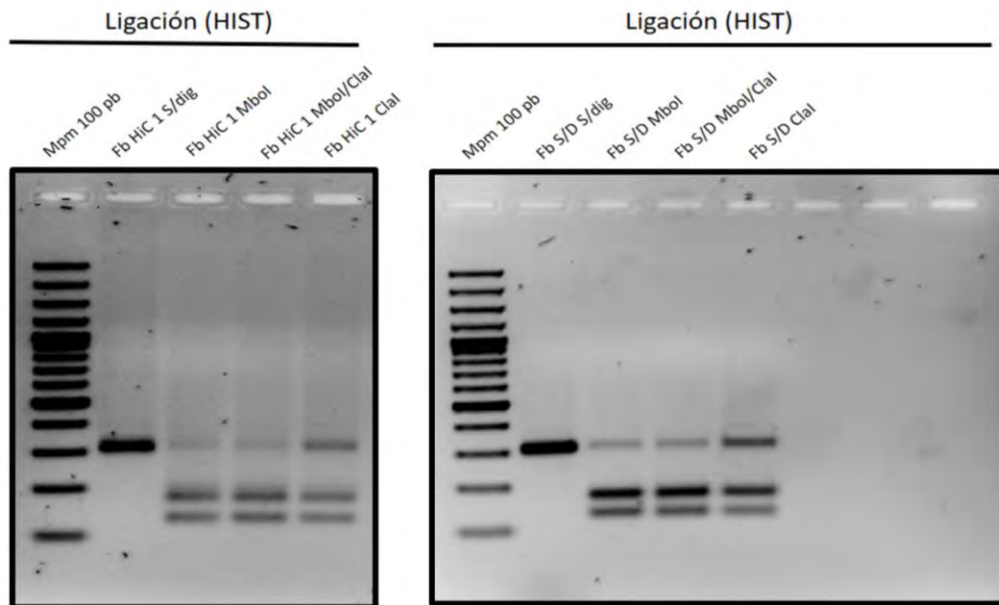


FIGURA 22 GEL DE AGAROSA 1.5% CON DIGESTIONES UTILIZANDO MBOI Y CLA1 DE LAS MUESTRAS OBSERVADAS EN LA FIGURA 31. SE OBSERVA UNA AMPLIFICACIÓN QUE DENOTA INEFICIENCIA EN EL EXPERIMENTO.

Una de las hipótesis acerca de la razón por la cual el protocolo no pudo ser estandarizado reside en la permeabilidad de la membrana plasmática de estas células, la cual puede ser menor que la observada en los eritrocitos y por lo tanto será necesario realizar varias pruebas con distintos detergentes que permitan a las enzimas introducirse hasta el núcleo de forma eficiente y realizar el corte, relleno con nucleótidos y ligación. Este protocolo continuará en proceso de estandarización alterando los tiempos de permeabilización y también se diseñarán más oligonucleótidos que permitan acceder a una mayor cantidad de controles durante el desarrollo del experimento.

Resultados

El ARN disminuye dramáticamente en eritrocitos adultos y se compone de moléculas de tamaño pequeño

Para determinar el contenido total de RNA así como la expresión específica de genes se utilizaron células que corresponden a tres estadios de la diferenciación eritroide de *Gallus gallus*, como se muestra en la figura 23. Las HD3 corresponden a un estadio de eritroblasto, mientras que los eRBC corresponden a eritrocitos circulantes de 10 días de desarrollo embrionario y las aRBC corresponden a eritrocitos adultos. También se utilizó a fibroblastos embrionarios de 10 días de desarrollo como un control no eritroide.



FIGURA 23 REPRESENTACIÓN DE LA DIFERENCIACIÓN ERITROIDE EN *GALLUS GALLUS*. LAS CÉLULAS UTILIZADAS COMO MODELO CORRESPONDEN A UN ESTADIO CERCANO A UN PROERITROBLASTO (HD3) Y A LOS ESTADIOS TERMINALES, EL ERITROCITO INMADURO (ERBC) Y EL ERITROCITO MADURO (ARBC).

Se realizó el protocolo de extracción para los cuatro tipos celulares, sin embargo, en el caso de las células eritroides adultas, la cantidad de ARN obtenida no fue suficiente por lo cual

se utilizó una mayor cantidad de células congeladas para llevar a cabo la extracción (8×10^8 células). Con esta cantidad celular se obtuvo suficiente ARN para corroborar su calidad en un gel desnaturizante donde se cargaron 500ng de material (Fig. 24A). En los aRBC se observó un patrón distinto al observado en las otras muestras donde podían observarse las dos bandas características de las subunidades 28S y 18S del ARNr. Así mismo, en los eRBC ya se puede observar un cambio en la razón entre las moléculas ribosomales 28S y 18S y una acumulación de moléculas de menor tamaño.

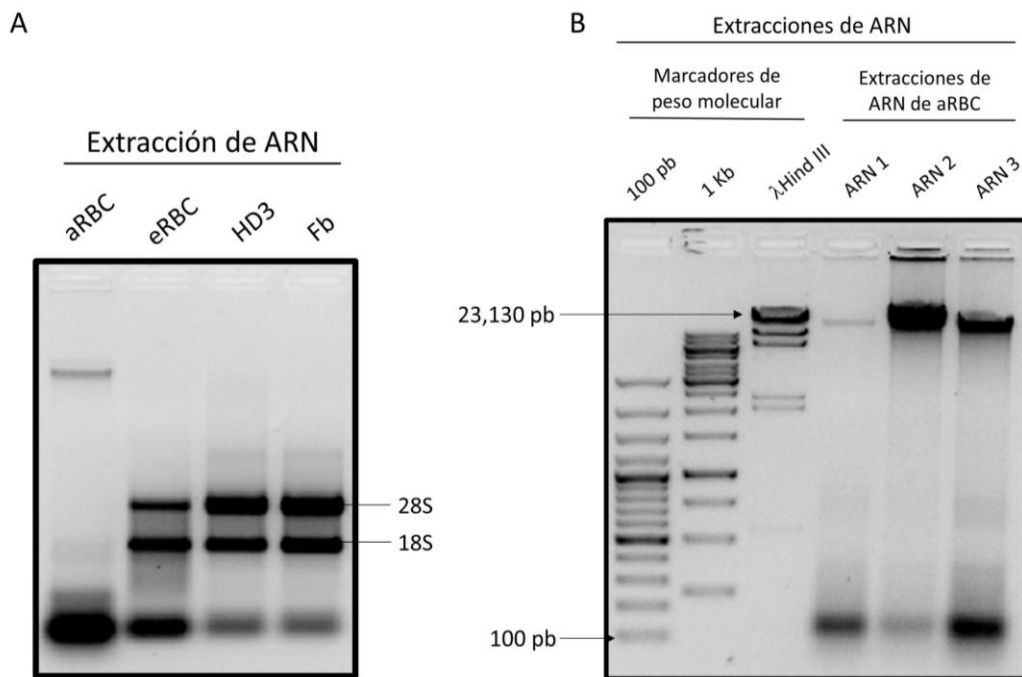


FIGURA 24 GELES DESNATURALIZANTES DE ARN OBTENIDO DE LAS CUATRO MUESTRAS CELULARES, SE MUESTRAN LOS CUATRO TIPOS CELULARES, ERITROCITOS ADULTOS (aRBC) ERITROCITOS EMBRIONARIOS (eRBC), LA LÍNEA CELULAR HD3 Y LOS FIBROBLASTOS (Fb). A) PRIMERA EXTRACCIÓN DE ARN, SE OBSERVÓ UN PATRÓN ATÍPICO EN LAS aRBC DONDE LOS ARN RIBOSOMALES NO PUDIERON APRECIARSE. B) MUESTRAS EXCLUSIVAMENTE DE aRBC COMPARADAS CON MARCADORES DE PESO MOLECULAR PARA EL CÁLCULO APROXIMADO DEL TAMAÑO DE LAS BANDAS OBSERVADAS, SE PUEDEN APRECIAR TRES EXTRACCIONES DE ARN.

Para corroborar la observación previa y descartar la degradación del material durante la extracción, se realizaron tres extracciones de ARN adicionales (Fig. 24B). En los tres casos,

el patrón observado en la primera extracción de ARN se mantuvo. También corroboramos que las bandas de alto peso molecular en el gel corresponden a material genómico, ya que las mismas disminuyen con el tratamiento con DNasa I (Fig. 25 A y B)

El perfil de ARN en los aRBC muestra la presencia de moléculas pequeñas que pueden ser resultado ya sea de la degradación de las especies de ARN durante la maduración de los eritrocitos (13) o como parte de la maquinaria regulatoria que quizá esté involucrada en el proceso de compactación y silenciamiento de la cromatina.

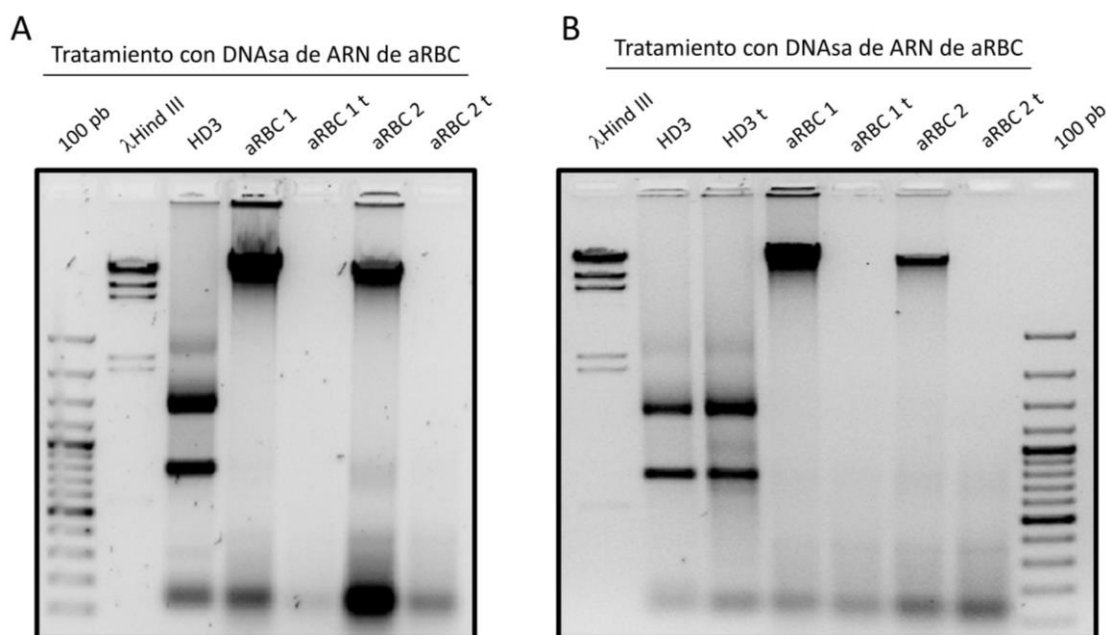


FIGURA 25 TRATAMIENTO CON DNASA DE LAS MUESTRAS DE ARN OBTENIDAS DE ERITROCITOS ADULTOS DE POLLO COMPARADAS CON UNA MUESTRA DE ARN DE HD3 ARBC (ARN SIN TRATAR) ARBC T (ARN POSTERIOR A TRATAMIENTO CON DNASA. A) PRIMER TRATAMIENTO, SE UTILIZARON 500NG DE ARN PARA CADA TRATAMIENTO. B) SEGUNDO TRATAMIENTO CON DNASA, EN ESTE CASO LOS 500NG DE ARN DE HD3 TAMBIÉN FUERON SOMETIDOS AL PROTOCOLO.

Una vez corroborado el patrón de ARN en los geles se procedió a realizar la comparación directa de la cantidad de ARN presente en cada una de las muestras. Se tomaron 10 millones de células de cada pellet y se realizó el protocolo de extracción con 1ml de Trizol para cada muestra. En la tabla 3 se muestran las cantidades obtenidas por cada una de las muestras.

TABLA 3 CANTIDAD DE MICROGRAMOS DE ARN OBTENIDOS DE EXTRACCIÓN CON EL MÉTODO DE TRIZOL DE 10 MILLONES DE CÉLULAS.

Tipo celular	Número de células	RNA total (µg)
HD3	10 millones	98.8
eRBC	10 millones	19.6
aRBC	10 millones	0.052
Fibroblastos	10 millones	119

Se pudo observar que la cantidad presente en las células que ya se encontraban diferenciadas en eritrocitos era mucho menor, en comparación con el ARN encontrado en los eritroblastos HD3 y en los fibroblastos. (Fig. 26) Así mismo la diferencia entre los niveles de ARN encontrados en eRBC y aRBC eran notorias también. Esta observación concuerda con lo descrito por Baumann, *et al.* Acerca de la disminución de ARN observada en la sangre de embriones de pollo durante el desarrollo eritroide (13). En su investigación ellos proponen la hipótesis principal de la degradación del ARN en los eritrocitos comprometidos reside en la necesidad de los NTPs como reguladores alostéricos de la hemoglobina (13).

Extracción de ARN

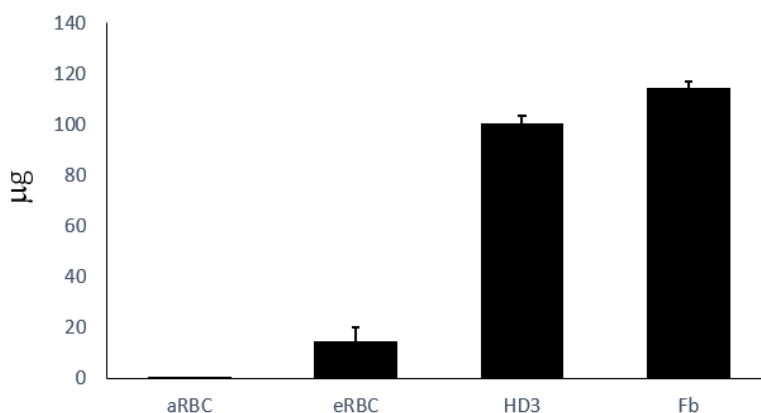


FIGURA 26 CANTIDAD DE ARN OBTENIDA DE TRES EXTRACCIONES POR EL MÉTODO DE TRIZOL UTILIZANDO 10 MILLONES DE CÉLULAS.

Existe una disminución de la expresión de genes en los eritrocitos en comparación con las células HD3 y los fibroblastos

Después del proceso de estandarización de los *primers* se comenzaron las pruebas por PCR cuantitativa utilizando el kit *KAPA SYBR Fast qPCR* (KR0389-V9.13) en un termociclador de 48 pozos StepOne.

Los genes candidato se dividieron en tres grupos:

- Constitutivos:
 - RPL4: Gen ribosomal L4
 - U4/U6: Proteína perteneciente a la maquinaria de splicing
 - GAPDH: Gen de la Gliceraldehído 3 Fosfato Deshidrogenasa
- Eritroides específicos:
 - EPB41: Gen de la proteína de membrana eritroide 4.1
 - EPB42: Gen de la proteína de membrana eritroide 4.2
- Genes globina:
 - HBAD: Gen de la proteína Alfa Globina D
 - HBBA: Gen de la proteína Beta Globina A

Dentro del grupo de los genes eritroides específicos se pudo observar una disminución de la expresión significativa tanto en los fibroblastos como en las aRBC (Fig. 27). EPB41 mostró tener expresión tanto en HD3 como en eRBC y aRBC sin embargo, su expresión se encuentra significativamente reducida en aRBC en relación con eRBC y HD3. EPB42 se encuentra enriquecida únicamente en eRBC y no se encuentra expresada en HD3, probablemente debido a que estas células aún no se encuentran terminalmente diferenciadas.

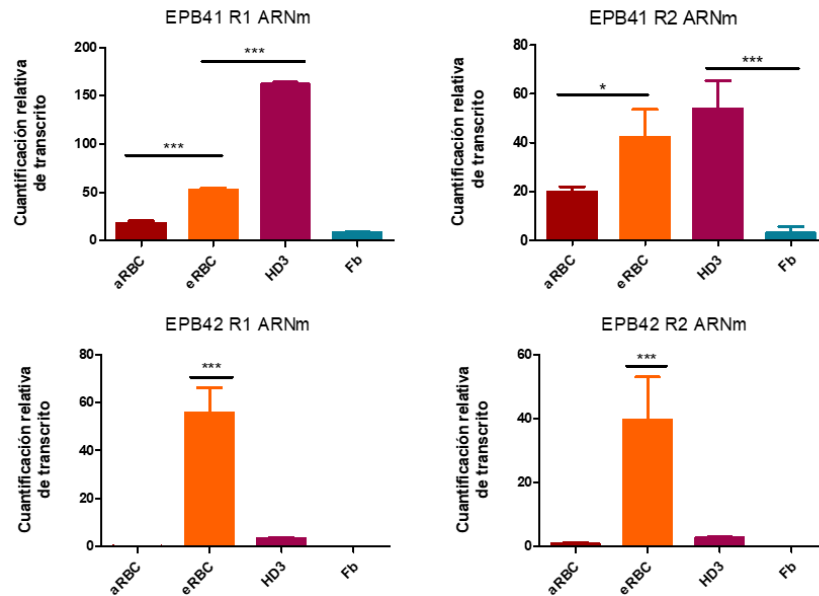


FIGURA27 GRÁFICOS DE EXPRESIÓN RELATIVA DE GENES ERITROIDES. EN EL PANEL DE ARRIBA SE MUESTRA EL TRASCRITO DE LA PROTEÍNA DE MEMBRANA ERITROIDE 4.1 CON LAS DOS RÉPLICAS BIOLÓGICAS MEDIDAS MEDIANTE QPCR. EN EL PANEL INFERIOR SE MUESTRA EL TRASCRITO DE PROTEÍNA DE MEMBRANA ERITROIDE 4.2. LAS PRUEBAS ESTADÍSTICAS UTILIZADAS FUERON ANOVA CON UNA PRUEBA DE TUKEY ($P < 0.05$).

En el grupo de los genes globina se eligieron un gen correspondiente al locus de las Globinas Alfa, las cuales se encuentran en el cromosoma 14 y un gen perteneciente a las Globinas Beta que se encuentran en el cromosoma 1. De esta forma se obtuvieron dos patrones de expresión presentes en cromosomas de tamaños muy distintos y alejados espacialmente.

Lo que se pudo observar en el patrón de expresión de las proteínas globinas fue que, en los dos casos, las únicas células que presentaron una transcripción significativa fueron las eRBC (Fig. 28), esto corroboró, con base en los datos de expresión reportados para estos genes, que los oligonucleótidos diseñados funcionaron de forma correcta ya que la expresión de las globinas en los fibroblastos y en las HD3 se encuentra reducida o es prácticamente nula. En el caso de las aRBC se pudo observar que prácticamente no existía ningún tipo de expresión de estos mensajeros lo que concuerda con lo observado tanto en los patrones anómalos de ARN como en la expresión de los genes eritroides. Esta fue una evidencia más

de que la cromatina de las células aRBC se encuentra en su mayor parte, silenciada transcripcionalmente.

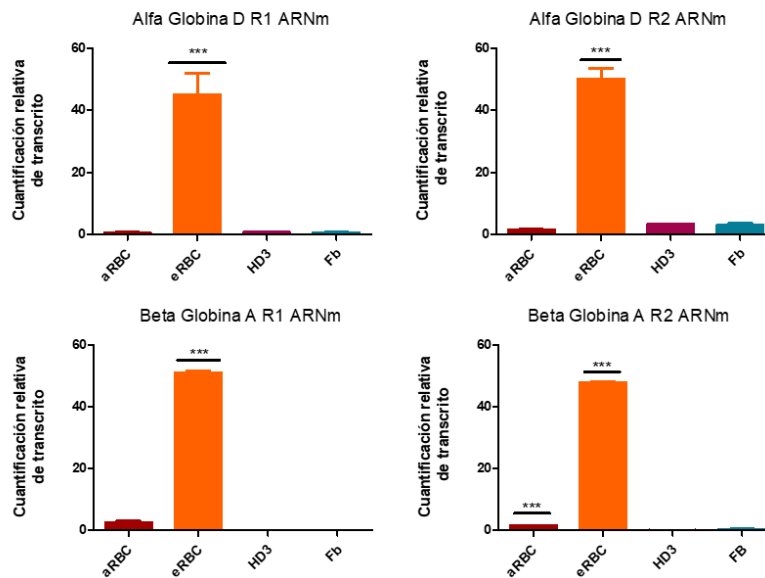


FIGURA 28 EXPRESIÓN RELATIVA DE LOS GENES GLOBINA. EN EL PANEL SUPERIOR SE MUESTRA LA EXPRESIÓN DE LA PROTEÍNA ALFA GLOBINA D MIDIENDO EL ARNm. EN EL PANEL INFERIOR SE MUESTRA LA EXPRESIÓN DEL ARNm MEDIDA DE BETA GLOBINA A. LAS PRUEBAS ESTADÍSTICAS UTILIZADAS FUERON ANOVA CON UNA PRUEBA DE TUKEY ($P < 0.05$).

En el caso de los genes constitutivos se observó una expresión diferente en los tipos celulares analizados principalmente reflejándose en una baja expresión notoria en las eRBC y aún más reducida en las aRBC (Fig. 29). Esto también es consistente con los patrones de ARN observados anteriormente en las extracciones con trizol donde el ARN de aRBC se mostraba prácticamente concentrado en una zona de bajo peso molecular y, en el caso de eRBC, aunque sí mostraban un patrón más similar a los otros dos tipos celulares, también se observaba una mayor cantidad de estos ARNs de tamaño pequeño que podrían significar algún tipo de degradación natural o también el enriquecimiento de alguna especie de ARN de tamaño pequeño.

Estos resultados indicaron que, en primer lugar, la expresión de los fibroblastos y de las HD3 es más parecida y corresponde a células que poseen un metabolismo activo y se encuentran ciclando. En segundo lugar, que los eritrocitos embrionarios parecen tener un patrón de expresión ya totalmente comprometido hacia su función específica relacionada con las globinas y los genes eritroides. Y, por último, en tercer lugar, que los eritrocitos adultos presentan una reducción generalizada de expresión tanto de genes constitutivos como específicos de células eritroides e incluso de las globinas.

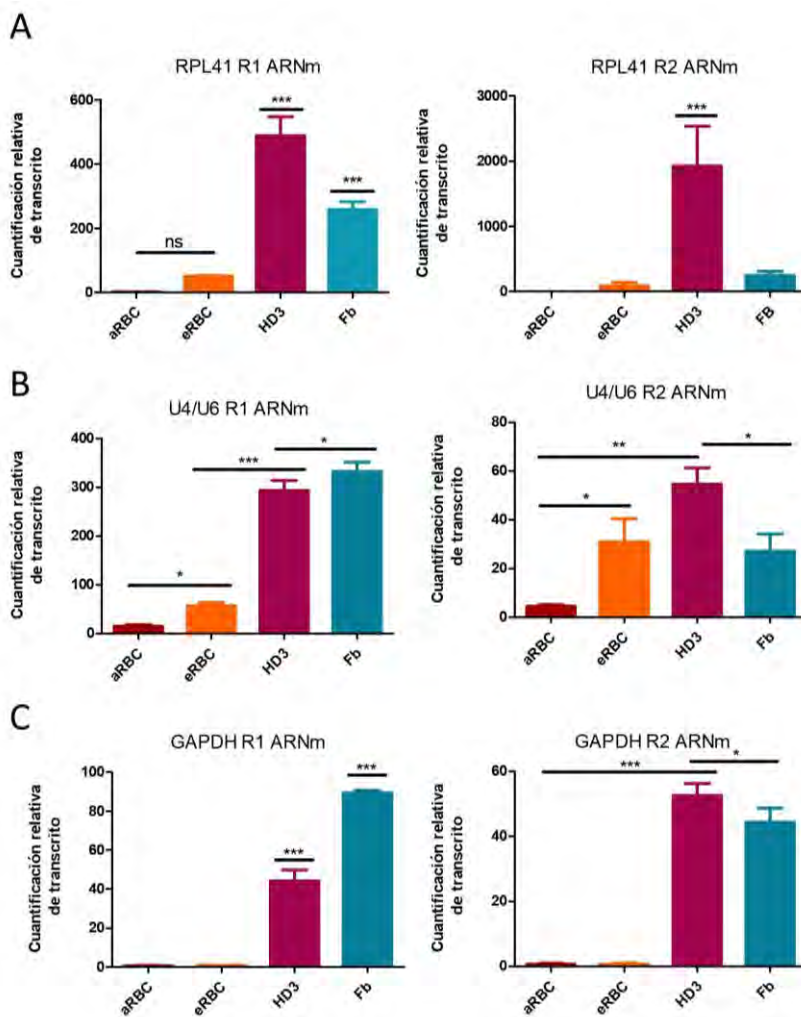


FIGURA 29 GRÁFICOS DE EXPRESIÓN RELATIVA DE ARNm CUANTIFICADA MEDIANTE EL MÉTODO DE CURVA ESTÁNDAR. A) GRÁFICO DE RPL4 EN LOS CUATRO TIPOS CELULARES, SE OBSERVA UNA DISMINUCIÓN DE LA EXPRESIÓN EN LOS ERITROCITOS. B) GRÁFICO DE EXPRESIÓN DE U4/U6 (PROTEÍNA RELACIONADA CON EL PROCESO DE SPLICING) SE OBSERVA UNA DISMINUCIÓN TAMBIÉN EN LOS ERITROCITOS. C) GRÁFICO DE EXPRESIÓN RELATIVA DE GAPDH, SE OBSERVA QUE PRÁCTICAMENTE SE ENCUENTRA AUSENTE DE LAS CÉLULAS CORRESPONDIENTES A ERITROCITOS. LAS PRUEBAS ESTADÍSTICAS UTILIZADAS FUERON ANOVA CON UNA PRUEBA DE TUKEY ($P < 0.05$).

De las pruebas anteriores se seleccionaron los pares de *primers* que amplificaran de forma más específica y respondieran a las muestras correspondientes a su rol biológico, es decir se esperaba observar amplificación de genes eritroide-específicos solamente en el ADNc de la línea celular HD3 y de los eritrocitos embrionarios de pollo. Así mismo se esperaba observar amplificación de las proteínas globinas en las muestras de eritrocitos embrionarios.

Los pares de *primers* que reconocen muchos sitios inespecíficos en el genoma o no amplificaron correctamente no fueron utilizados en las posteriores pruebas de PCR cuantitativa.

Determinación de la estructura tridimensional del genoma durante la maduración y diferenciación eritroides

Para estudiar la organización 3D del genoma en los tipos celulares de interés, se realizó el protocolo de Hi-C, descrito de manera general en los métodos. Este protocolo fue establecido por primera vez en el laboratorio como resultado del presente trabajo de investigación.

Para corroborar la eficiencia del protocolo a nivel de la digestión, ligación y generación de pares válidos de interacción, se eligieron dos pares de *primers*. Los primeros fueron diseñados en dos sitios aledaños de restricción de la enzima *MboI* en dirección sentido, de esa forma una amplificación dentro del genoma sin digerir no sería posible (Fig. 30).



FIGURA 30 DIAGRAMA DE DISEÑO DE CONTROLES PARA EL PROTOCOLO DE HI-C

El segundo par fue diseñado en dos zonas separadas por 20kpb aproximadamente. La primera zona corresponde a un enhancer llamado MRE que tiene capacidad regulatoria sobre la expresión de los genes globina alfa en el pollo. La segunda zona corresponde al gen globina alfa-D. Estas dos zonas se reportaron interaccionando en un estudio en realizado 2008 mediante la técnica de 3C en células eritroides de pollo (26) por lo cual fueron elegidos como un control de interacción positivo del protocolo.

Se probaron estos dos controles utilizando muestras 100ng de cada una de las reacciones de Hi-C y posteriormente realizando una amplificación por PCR (Fig. 31).

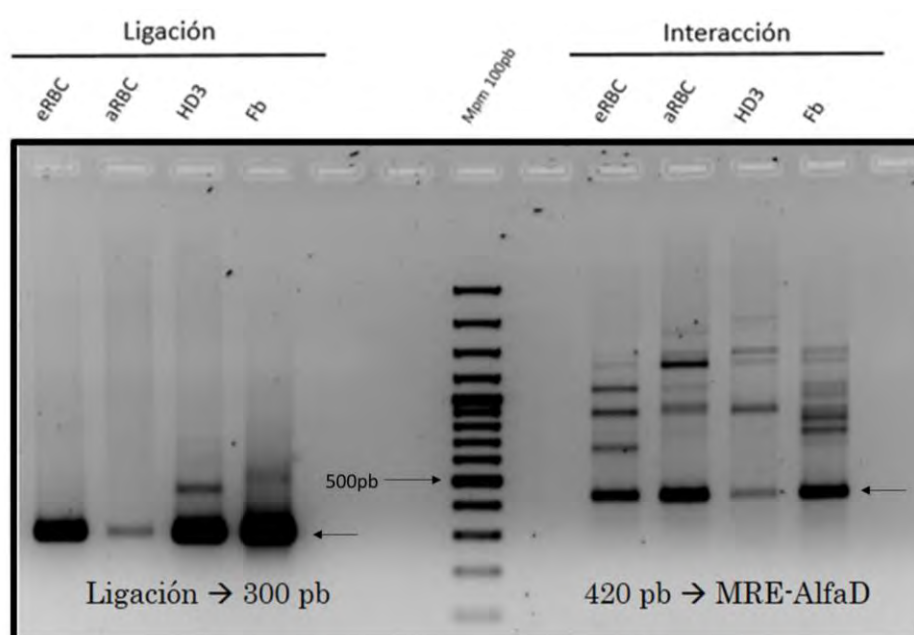


FIGURA 31 RESULTADOS DE CONTROLES DE PROTOCOLO DE HI-C. SE PUEDE OBSERVAR UNA AMPLIFICACIÓN DE REGIONES DEL TAMAÑO ESPERADO DE ACUERDO AL DISEÑO DE *PRIMERS* EN LOS DOS CASOS. ESO APUNTA A QUE LOS PASOS DE ENTRECruzAMIENTO, DIGESTIÓN Y LIGACIÓN SE REALIZARON EFICIENTEMENTE.

Observando los primeros controles realizados se pudo concluir ante la presencia de las dos amplificaciones del tamaño esperado que una proporción del ADN de las muestras había sido entrecruzado, digerido y ligado de la forma esperada. Posteriormente tomando como base estas amplificaciones (realizando una purificación de ADN de la banda del gel) se realizó una segunda PCR para ahora verificar la eficiencia en el proceso de rellenado, marcado con biotina y ligado de las muestras.

En este segundo control las muestras se dividieron en 4 tubos, de los cuales el primero no se digirió con ninguna enzima de restricción siendo la muestra sin digerir (S/d), el segundo se digirió con la enzima *Mbol*, el tercero con una combinación de las enzimas *Mbol* y *Clal* y el cuarto se digirió con la enzima *Clal* (Fig. 32A).

Este control se basa en una característica que brinda la enzima de restricción *Mbol*, la cual consiste en que durante el proceso de rellenado de los extremos cohesivos que deja la enzima, el sitio de *Mbol* se restablece y adicionalmente se forma un sitio de restricción para la enzima *Clal* (Fig. 32B).

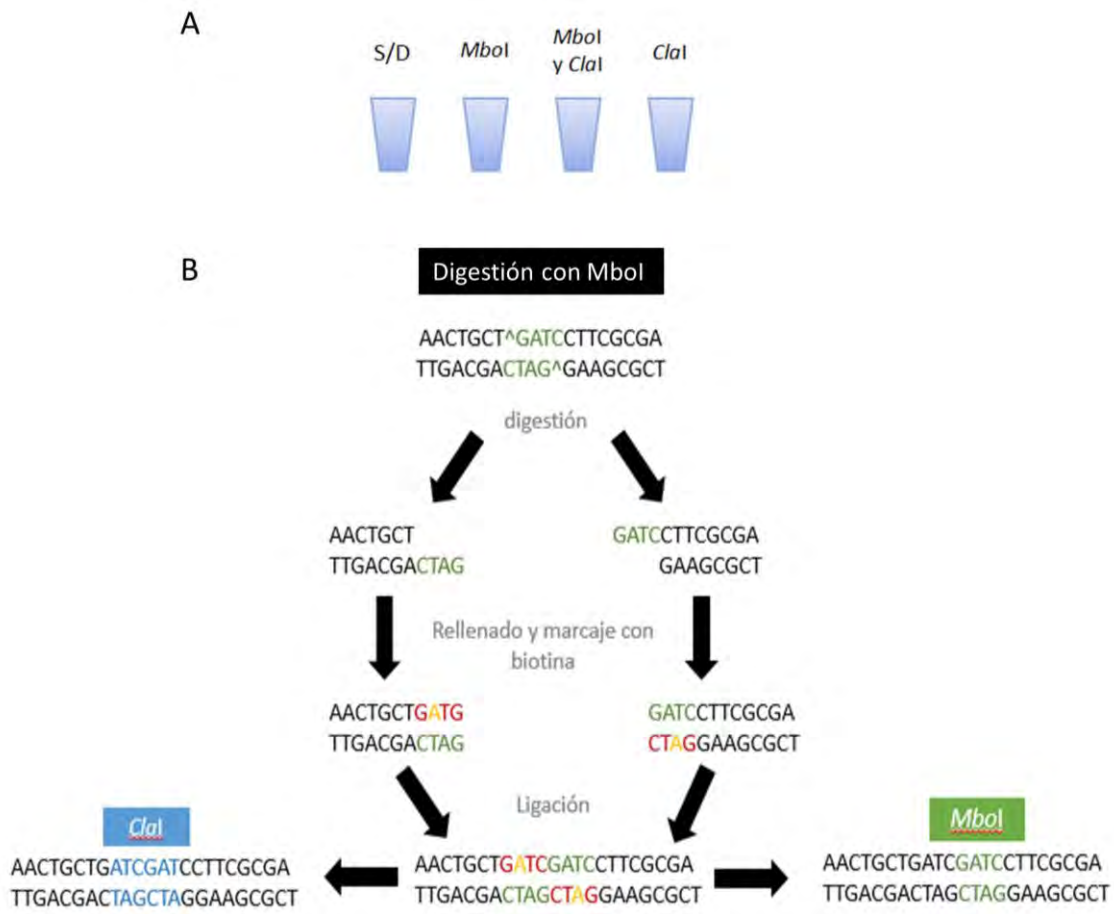


FIGURA 32 CONTROL DE DIGESTIÓN DE HI-C. A) ESQUEMA DE DIGESTIÓN DE MUESTRAS PARA CADA UNO DE LOS TIPOS CELULARES ANALIZADOS POR EL PROTOCOLO DE HI-C. B) PROCESO DE DIGESTIÓN Y RELLENADO DE NUCLEÓTIDOS QUE PERMITE LA REFORMACIÓN DEL SITIO *Mbol* Y LA FORMACIÓN DE UN SITIO DE RESTRICCIÓN PARA *Clal* QUE NO SE ENCONTRABA ANTERIORMENTE EN EL GENOMA SIN DIGERIR. LA LETRA EN AMARILLO REPRESENTA EL NUCLEÓTIDO MARCADO CON BIOTINA.

Utilizando esta característica que brinda la enzima se puede esperar que las muestras digeridas con las tres combinaciones de enzimas de restricción presente un patrón igual de bandeo en un gel de agarosa. Este control se realizó para las cuatro muestras obteniendo resultados que mostraban un patrón de digestión muy similar en los tres casos. Cabe resaltar que las bandas obtenidas de las digestiones presentaron tamaños esperados tanto de amplificación sin digestión como de las digestiones de acuerdo al diseño de los *primers* realizado en el *software* Primer3 y mapeados con ayuda de la plataforma Benchling (Fig. 33).

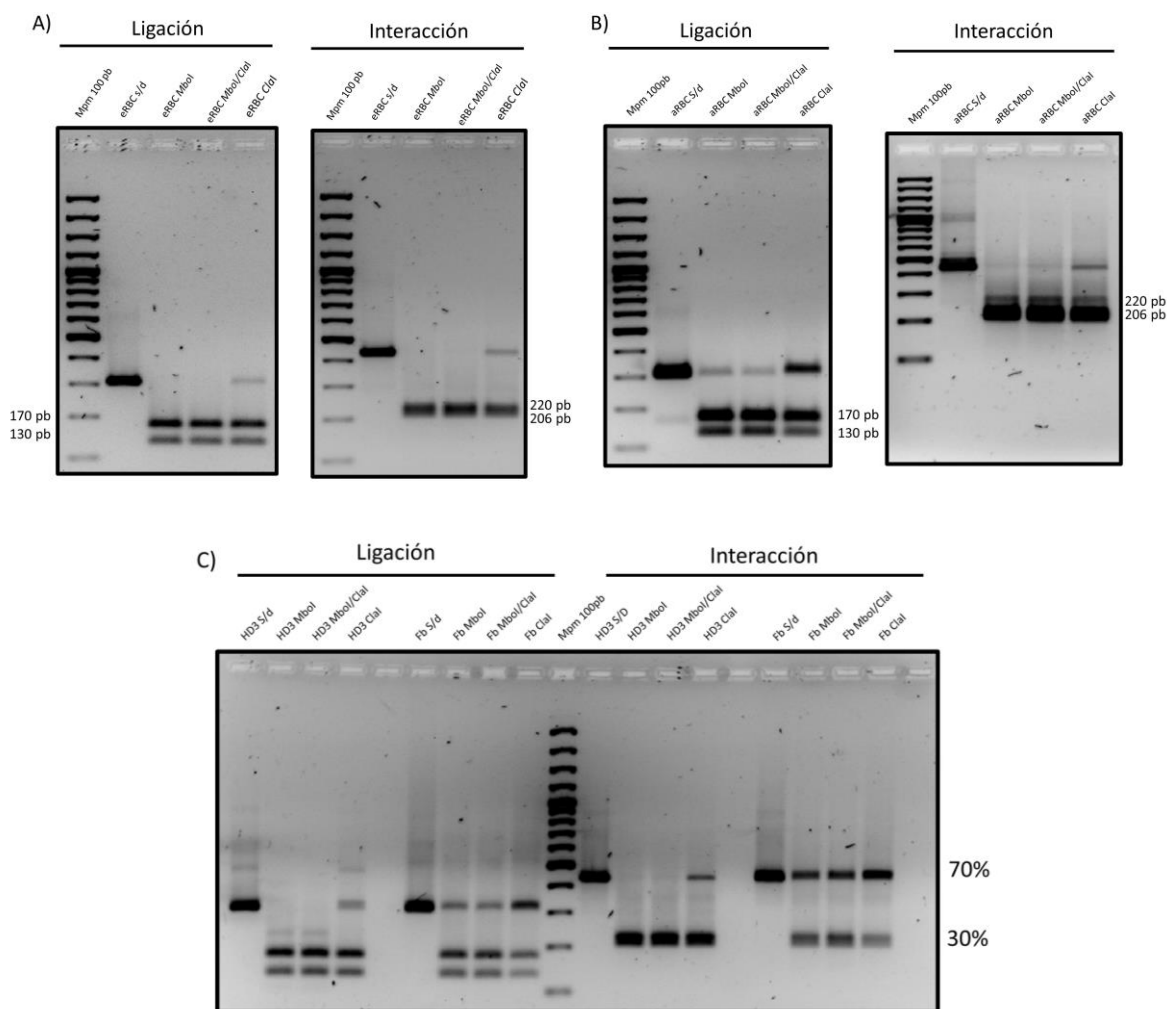


FIGURA 33 PATRÓN DE BANDEO DE DIGESTIONES REALIZADAS CON LAS ENZIMAS *MBOI* Y *CLAI* PARA OBSERVAR LA EFICIENCIA DEL PASO DE RELLENADO, MARCADO CON BIOTINA Y LIGACIÓN DEL PROTOCOLO DE HI-C. A) CORRESPONDE A ERITROCITOS EMBRIONARIOS DE POLLOS DE DIEZ DÍAS DE DESARROLLO. B) CORRESPONDE A ERITROCITOS DE POLLOS ADULTOS DE 38 SEMANAS. C) CORRESPONDE TANTO A LA LÍNEA HD3 COMO A CULTIVO PRIMARIO DE FIBROBLASTOS DE POLLO. EN EL CASO DE LOS FIBROBLASTOS SE OBSERVÓ UNA BAJA EFICIENCIA DE DIGESTIÓN CORRESPONDIENTE A UN 70% DE BANDAS DE ADN NO DIGERIDO CONTRA UN 30% DIGERIDO MEDIDO POR DENSITOMETRÍA.

Las bandas de tamaño similar a la muestra sin digerir observadas en varias de las muestras digeridas corresponden probablemente a regiones en las cuales existió una baja eficiencia en el proceso de rellenado de los extremos cohesivos o un error en la ligación. Esto se refiere a que, aunque es una región que se amplificó con los *primers* diseñados para el primer control, no presentan el patrón de digestión esperado. Sin embargo, debido a que se observaron bandas con del tamaño esperado posterior a la digestión con las enzimas de restricción, se decidió proseguir con el experimento de Hi-C.

Después del paso de selección de tamaño se realizó un gel para corroborar que se obtuvieran los tamaños requeridos para el proceso de sonicación de las muestras (Fig. 34) en donde se pudo observar que el proceso fue exitoso ya que se recuperaron tamaños de entre 300 y 500 pares de bases lo cuales eran adecuados para la sonicación.

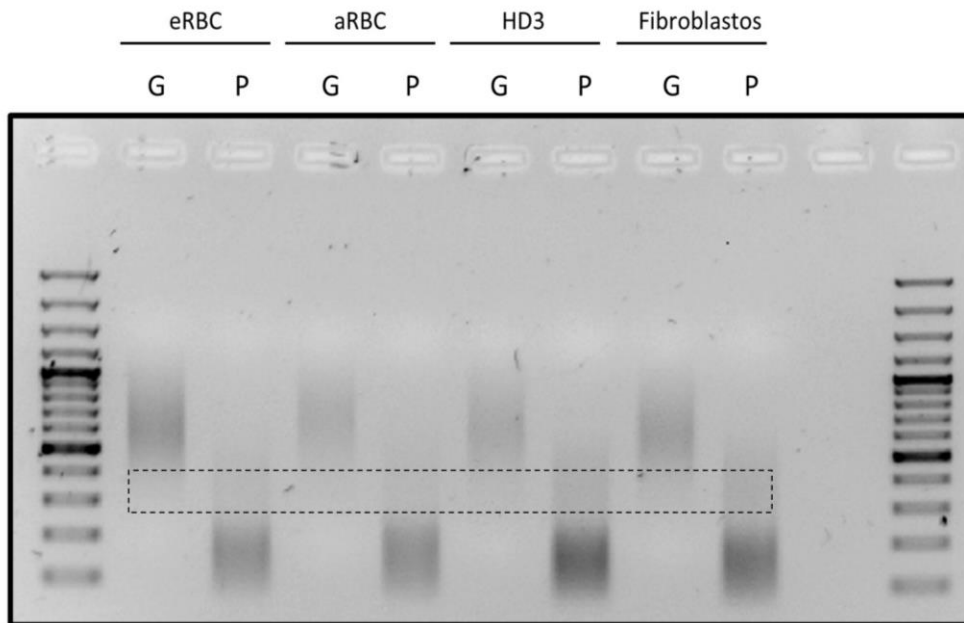


FIGURA 34 GEL DE 1% AGAROSA QUE MUESTRA LOS TAMAÑOS EXCLUIDOS DE LA LIBRERÍA DE HI-C EN EL PROCESO DE SELECCIÓN DE TAMAÑO CON SPRI BEADS. SE MUESTRAN TANTO LOS FRAGMENTOS GRANDES (G) COMO LOS PEQUEÑOS (P). EL ESPACIO QUE SE OBSERVA ENTRE LAS DOS PATRONES DE BANDEO, SEÑALADO POR LA LÍNEA PUNTEADA, CORRESPONDE A LA REGIÓN RECUPERADA PARA LA CONTINUACIÓN DEL PROTOCOLO.

Posterior a la sonicación se continuó el protocolo hasta la amplificación por PCR de punto final después de la recuperación de los fragmentos marcados con biotina. Para este paso se realizaron pruebas con distinto número de ciclos de PCR debido a que se debe tener suficiente material amplificado para poder mandar a secuenciar la librería, pero también se debe evitar una sobre amplificación que brindaría una información redundante (Fig. 35). Se utilizaron 20 ciclos en la primera amplificación en la cual se pudo corroborar que la mayor parte de la librería se encontraba dentro de los tamaños adecuados, sin embargo, se observó que estaba sobre amplificada. Posteriormente se utilizaron 10 y 15 ciclos de amplificación y se tomó la decisión de utilizar 9 ciclos para evitar sobre amplificar las muestras.

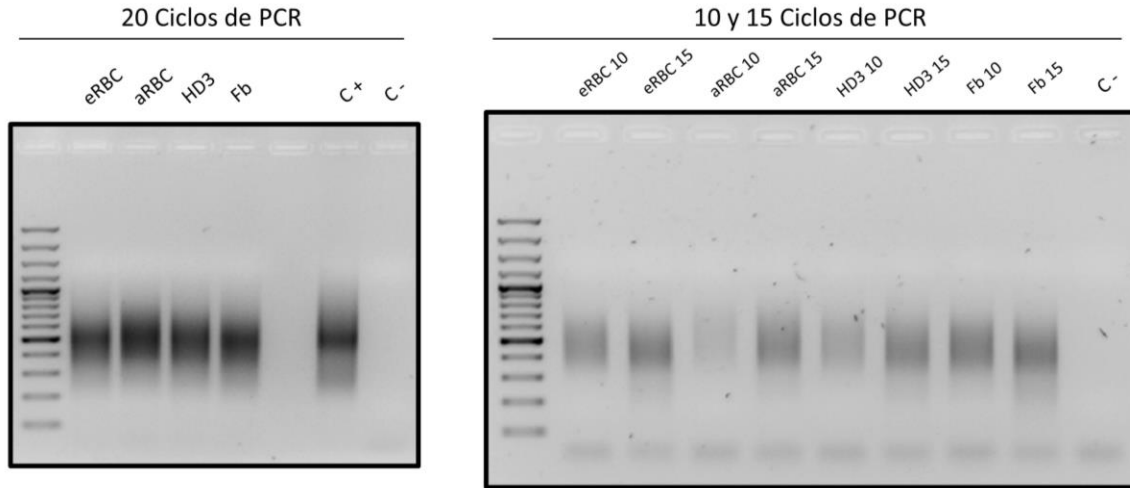


FIGURA 35 CICLOS DE PCR PUNTO FINAL PROBADOS PARA LA AMPLIFICACIÓN DE LAS LIBRERÍAS DE HI-C ANTES DEL PROCESO DE SECUENCIACIÓN.

El último control experimental que se realizó sobre las muestras consistió en amplificar por PCR punto final dos alícuota de las librerías 10 ciclos. Posteriormente se tomó una de las dos alícuotas y se digirió con la enzima *Clal* esperando que la banda de amplificación en el gel bajara debido a que la librería se encuentra enriquecida con fragmentos de interacción de Hi-C (Fig. 36). Se pudo observar que en el caso de las tres células del linaje eritroide el barrido de amplificación indicó que había una gran cantidad de fragmentos que eran cortados por la enzima de restricción. Sin embargo, esto no fue así en el caso de los fibroblastos por lo que, tomando en cuenta también los resultados de los controles anteriores, se decidió que para estas células era pertinente realizar algunos cambios en el protocolo para hacer más eficientes los procesos de corte, rellenado y ligación.

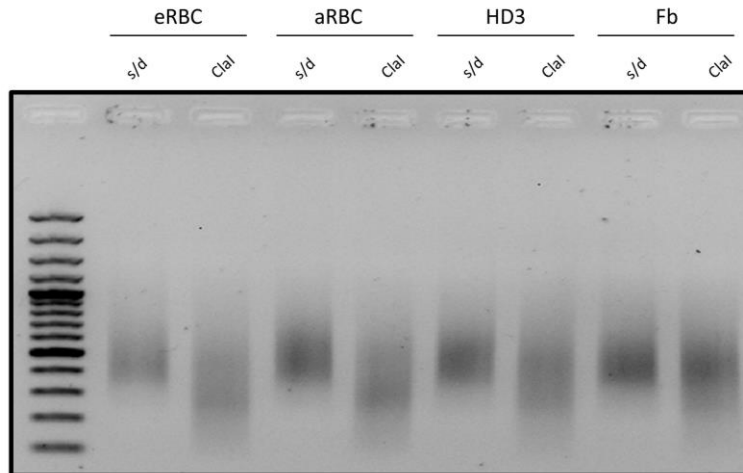


FIGURA 36 GEL CONTROL DE HI-C CON AMPLIFICACIÓN DE LAS LIBRERÍAS FINALES POR PCR DE PUNTO FINAL UTILIZANDO 10 CICLOS. SE OBSERVA LA COMPARACIÓN ENTRE LAS LIBRERÍAS SIN DIGERIR (S/D) Y DIGERIDAS CON LA ENZIMA CLAL.

Después de este control, los Hi-C de los tres tipos celulares eritroides se mandaron a secuenciar utilizando un equipo Illumina Hiseq 2500 por el método de secuenciación pareada y lecturas de 50pb. La secuenciación se llevó a cabo en el Babraham Institute utilizando adaptadores marcados con una secuencia distinta que funciona como un código de barras y permite diferenciar las muestras. Se utilizó una línea de secuenciación que genera alrededor de 200 millones de lecturas obteniendo un aproximado de 66 millones de lecturas por cada experimento de Hi-C.

La estructura tridimensional del genoma cambia drásticamente entre los eritroblastos y los eritrocitos

Para filtrar las lecturas útiles en la detección de interacciones y corroborar la calidad de los experimentos realizados, después de obtener los datos secuenciados se utilizó el programa HiCUP (50) desarrollado en el Instituto Babraham. HiCUP calcula la cantidad de lecturas válidas que están formadas por productos de ligación entre secuencias del genoma y elimina lecturas inválidas como son los fragmentos circulares, las re-ligaciones, fragmentos que tengan tamaños incorrectos, entre otras (Fig. 37).

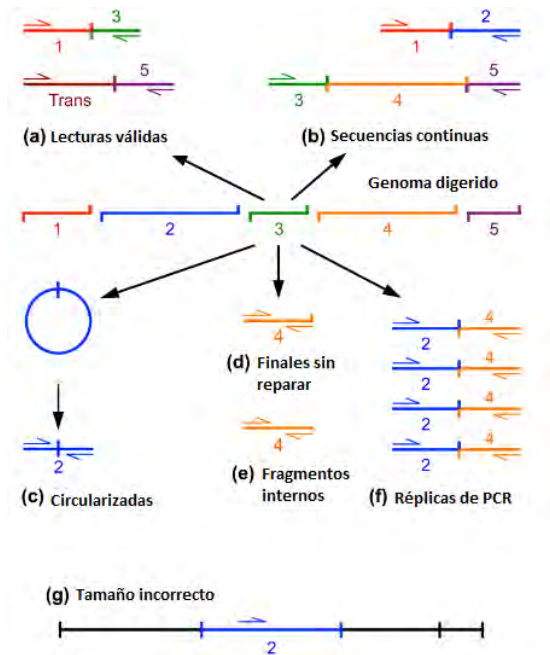


FIGURA 37 DEPURACIÓN DE DATOS DE SECUENCIACIÓN QUE REALIZA EL PROGRAMA HICUP. ELIMINA LOS FRAGMENTOS QUE NO BRINDAN INFORMACIÓN DE INTERACCIONES MEDIANTE RECONOCER LAS SECUENCIAS EN CADA LECTURA Y MAPEARLAS EN EL GENOMA, DE ESTA FORMA REÚNE SOLAMENTE LAS LECTURAS QUE ESTÁN FORMADAS POR UNA INTERACCIÓN DE REGIONES QUE NO SE ENCUENTRAN LINEALMENTE CONTINUAS. (MODIFICADA DE WINGETT, *ET AL.* 2016)

Los tres experimentos fueron secuenciados en la misma línea de secuenciación utilizando etiquetas genómicas para poder mapearlos posteriormente. Con ayuda del programa HiCUP, de las lecturas totales obtenidas, se obtuvieron las lecturas válidas, las cuales corresponden a fragmentos que muestran una interacción representada por la sección “a” de la figura 37 y se descartaron todas las lecturas que no brindaran información para la construcción de los mapas de interacción. Aproximadamente se obtuvieron 65 millones de lecturas para cada experimento de las cuales se pudieron obtener 34.3 millones de pares válidos en las HD3, 38.1 millones en las eRBC y 40.8 millones en las aRBC representando el 68.2%, 74.9% y 84.3% respectivamente de las lecturas totales obtenidas de cada experimento (Fig. 38, 39 y 40). Con esta cantidad de lecturas válidas se prosiguió a la construcción de mapas de interacción para cada una de las muestras.

Estadística de experimento de Hi-C utilizando Hicup para eritrocitos de gallinas adultas (aRBC)

Lecturas alineadas al genoma de referencia (A)		
aRBC	Lectura 1	Lectura 2
Alineamientos únicos	57,429,579	56,914,407
Alineamientos múltiples	5,212,869	5,175,315
No alineados	2,103,000	2,662,500
Pareados	48,438,147	48,438,147

Pares válidos e inválidos reconocidos por Hicup (B)	
aRBC	Número de lecturas ("Di-tags")
Pares válidos	40,834,962
Pares inválidos	7,603,185
Mismo fragmento circular	15,037
Mismo fragmento no ligado	119,672
Mismo fragmento interno	1,358,308
Re-ligación	1,344,443
Secuencias continuas	841,031
Tamaño incorrecto	3,924,694
Pares totales	48,438,147

Interacciones cercanas y lejanas identificadas por Hicup (C)		
aRBC	Todos Di-Tags	Di-Tags únicos
Lecturas pares	40,834,962	39,715,205
Cis-cercano (< 10Kbp)	3,488,003	3,391,692
Cis-lejano (> 10Kbp)	26,057,281	25,340,392
Trans	11,289,678	10,983,121

Composición de lecturas completas y de tamaño adecuado para alineación (A)		
aRBC	Lectura 1	Lectura 2
Lecturas totales	67,237,050	67,237,050
No truncadas	61,039,271	61,012,863
Truncadas	6,197,779	6,224,187
Demasiado cortas para mapear	1,080,204	1,073,430
Largo aproximado de secuencias truncadas	30.25	30.23

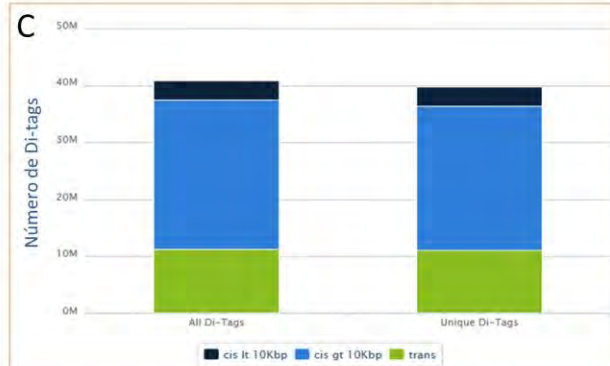
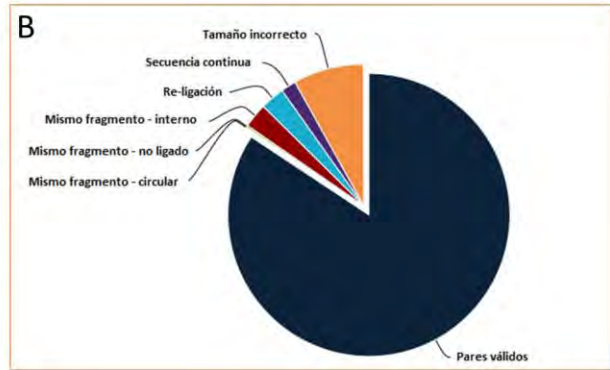
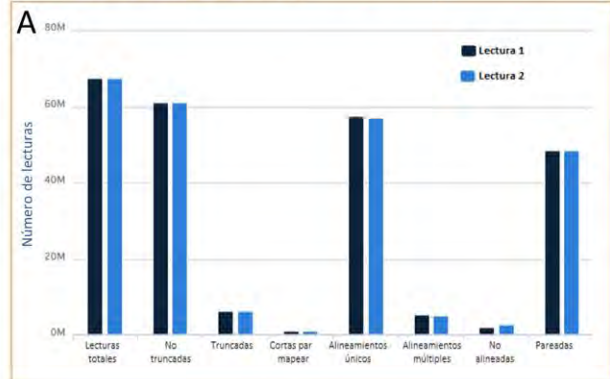


FIGURA 38 GRÁFICOS CONSTRUIDOS POR EL PROGRAMA HICUP DE BABRAHAM BIOINFORMARTICS, EXPERIMENTO HIC RÉPLICA 1 DE ERITROCITOS DE GALLINAS ADULTAS. A) GRÁFICO QUE MUESTRA LA CANTIDAD DE LECTURAS DE FRAGMENTOS DE INTERACCIÓN IDENTIFICADAS POR EL PROGRAMA. PARA QUE LAS LECTURAS SEAN ÚTILES PARA LA CONSTRUCCIÓN DEL MAPA DE INTERACCIONES DEBEN SER NO TRUNCADAS, LINEAMIENTOS ÚNICOS Y PAREADAS. B) CLASIFICACIÓN DE FRAGMENTOS PAREADOS DE ACUERDO A ÉL MAPEO SOBRE EL GENOMA DE REFERENCIA. C) CLASIFICACIÓN DE LEJANÍA DE INTERACCIONES DE LOS PARES VÁLIDOS IDENTIFICADOS POR EL PROGRAMA (CIS CERCANO = < 10 KB, CIS LEJANO = >10 KB Y TRANS = CONTACTO ENTRE DOS CROMOSOMAS DISTINTOS).

Estadística de experimento de Hi-C utilizando Hicup para Eritrocitos de embriones de 10 de desarrollo

Lecturas alineadas al genoma de referencia (A)		
eRBC	Lectura 1	Lectura 2
Alineamientos únicos	59,213,819	58,838,790
Alineamientos múltiples	5,011,179	4,982,422
No alineados	1,937,752	2,330,667
Pareados	50,859,246	50,859,246

Pares válidos e inválidos reconocidos por Hicup (B)	
eRBC	Número de lecturas ("Di-tags")
Pares válidos	38,110,239
Pares inválidos	12,749,007
Mismo fragmento circular	20,090
Mismo fragmento no ligado	343,180
Mismo fragmento interno	3,511,393
Re-ligación	3,165,486
Secuencias continuas	1,657,658
Tamaño incorrecto	4,051,200
Pares totales	50,859,246

Interacciones cercanas y lejanas identificadas por Hicup (C)		
eRBC	Todos Di-Tags	Di-Tags únicos
Lecturas pares	38,110,239	37,426,912
Cis-cercano (< 10Kbp)	3,679,398	3,613,243
Cis-lejano (> 10Kbp)	25,073,647	24,622,640
Trans	9,357,194	9,191,029

Composición de lecturas completas y de tamaño adecuado para alineación (A)		
eRBC	Lectura 1	Lectura 2
Lecturas totales	68,854,802	68,854,802
No truncadas	63,179,756	63,147,114
Truncadas	5,675,046	5,707,688
Demasiado cortas para mapear	1,074,300	1,085,171
Largo aproximado de secuencias truncadas	29.09	29.05

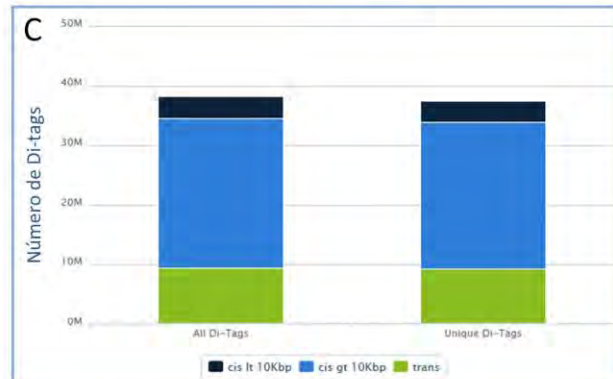
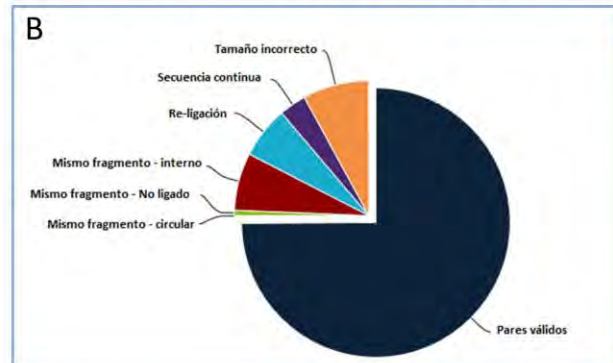
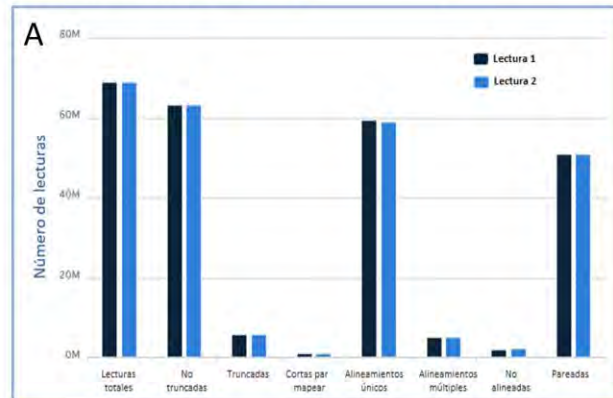


FIGURA 39 GRÁFICOS CONSTRUIDOS POR EL PROGRAMA HICUP DE BABRAHAM BIOINFORMARTICS, EXPERIMENTO HIC RÉPLICA 1 DE ERITROCITOS DE EMBRIONES DE POLLO. A) GRÁFICO QUE MUESTRA LA CANTIDAD DE LECTURAS DE FRAGMENTOS DE INTERACCIÓN IDENTIFICADAS POR EL PROGRAMA. PARA QUE LAS LECTURAS SEAN ÚTILES PARA LA CONSTRUCCIÓN DEL MAPA DE INTERACCIONES DEBEN SER NO TRUNCADAS, LINEAMIENTOS ÚNICOS Y PAREADAS. B) CLASIFICACIÓN DE FRAGMENTOS PAREADOS DE ACUERDO A ÉL MAPEO SOBRE EL GENOMA DE REFERENCIA. C) CLASIFICACIÓN DE LEJANÍA DE INTERACCIONES DE LOS PARES VÁLIDOS IDENTIFICADOS POR EL PROGRAMA (CIS CERCANO = < 10 KB, CISLEJANO = >10 KB Y TRANS = CONTACTO ENTRE DOS CROMOSOMAS DISTINTOS).

Estadística de experimento de Hi-C utilizando Hicup para la línea celular HD3 (eritroblastos de pollo)

Lecturas alineadas al genoma de referencia (A)		
HD3	Lectura 1	Lectura 2
Alineamientos únicos	59,612,509	59,060,355
Alineamientos múltiples	5,406,764	5,392,987
No alineados	2,631,395	3,185,376
Pareados	50,330,212	50,330,212

Pares válidos e inválidos reconocidos por Hicup (B)	
HD3	Número de lecturas ("Di-tags")
Pares válidos	34,330,618
Pares inválidos	15,999,594
Mismo fragmento circular	34,723
Mismo fragmento no ligado	499,525
Mismo fragmento interno	4,324,384
Re-ligación	3,964,510
Secuencias continuas	2,195,075
Tamaño incorrecto	4,981,377
Pares totales	50,330,212

Interacciones cercanas y lejanas identificadas por Hicup (C)		
HD3	Todos Di-Tags	Di-Tags únicos
Lecturas pares	34,330,618	33,597,826
Cis-cercano (< 10Kbp)	4,526,959	4,430,069
Cis-lejano (> 10Kbp)	23,442,042	22,940,161
Trans	6,361,617	6,227,596

Composición de lecturas completas y de tamaño adecuado para alineación (A)		
HD3	Lectura 1	Lectura 2
Lecturas totales	65,825,652	65,825,652
No truncadas	59,602,017	59,650,021
Truncadas	6,223,635	6,175,631
Demasiado cortas para mapear	1,204,134	1,216,084
Largo aproximado de secuencias truncadas	30.19	30.20

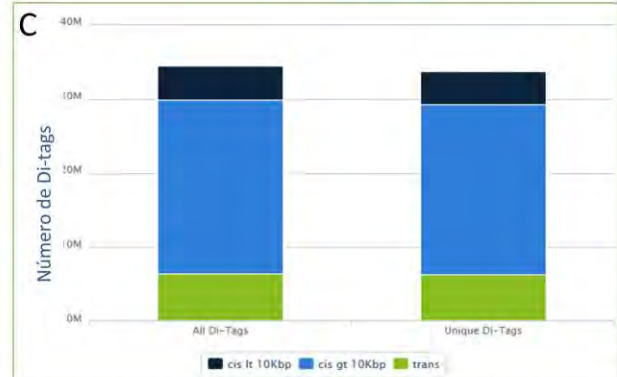
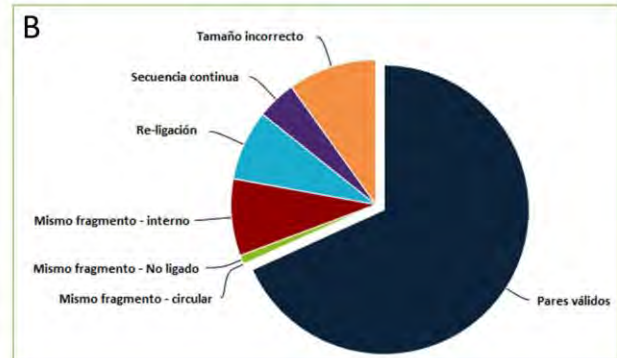
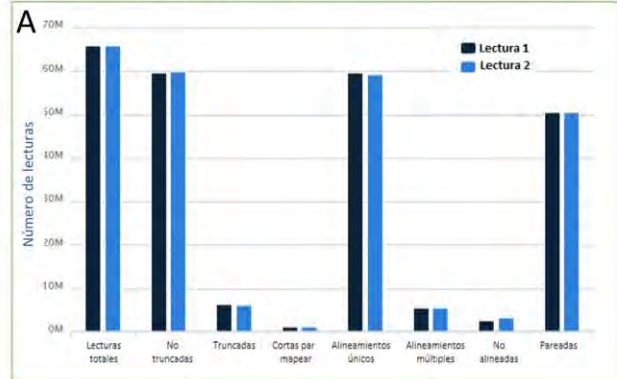


FIGURA 40 GRÁFICOS CONSTRUIDOS POR EL PROGRAMA HICUP DE BABRAHAM BIOINFORMARTICS, EXPERIMENTO HIC RÉPLICA 1 DE HD3. A) GRÁFICO QUE MUESTRA LA CANTIDAD DE LECTURAS DE FRAGMENTOS DE INTERACCIÓN IDENTIFICADAS POR EL PROGRAMA. PARA QUE LAS LECTURAS SEAN ÚTILES PARA LA CONSTRUCCIÓN DEL MAPA DE INTERACCIONES DEBEN SER NO TRUNCADAS, LINEAMIENTOS ÚNICOS Y PAREADAS. B) CLASIFICACIÓN DE FRAGMENTOS PAREADOS DE ACUERDO A ÉL MAPEO SOBRE EL GENOMA DE REFERENCIA. C) CLASIFICACIÓN DE LEJANÍA DE INTERACCIONES DE LOS PARES VÁLIDOS IDENTIFICADOS POR EL PROGRAMA (CIS CERCANO = < 10 KB, CIS LEJANO = >10 KB Y TRANS = CONTACTO ENTRE DOS CROMOSOMAS DISTINTOS).

Con el análisis de HiCUP también se puede obtener el número de lecturas únicas o “alineamientos únicos” los cuáles se refieren a la cantidad de las secuencias reconocidas por el programa que representan datos no provenientes de réplicas de PCR. Cabe resaltar que uno de los últimos pasos en el experimento de Hi-C consiste en la amplificación de la librería para su posterior secuenciación, esto puede resultar en la obtención de información redundante obtenida por fragmentos iguales amplificados en la PCR, así HiCUP permite identificar estas secuencias y eliminarlas de los datos ya que no representan información biológica. En el caso de las tres librerías se obtuvo alrededor de un 90% de lecturas únicas, lo que indica que no existió una sobre amplificación del material y la mayor parte de la información obtenida corresponde con datos biológicos.

El análisis de HiCUP también brinda información acerca de la cercanía que tienen los contactos encontrados en el genoma. Dado que los cromosomas se encuentran distribuidos en territorios discretos dentro del núcleo (20) se espera que la mayoría de los contactos recuperados del experimento sean dentro del mismo cromosoma (en *cis*) y no entre cromosomas (en *trans*). Si un experimento posee un mayor número de contactos en *trans* podría implicar un error en el paso del entrecruzamiento.

En los tres casos los experimentos brindaron una mayor cantidad de contactos en *cis* teniendo cerca del 20% de los contactos encontrados en *trans* (Fig. 38C, 39C y 40C). Por esta razón se decidió que los experimentos eran adecuados para proseguir con la construcción de los mapas de interacción.

Una característica que se pudo observar entre la proporción de contactos *trans* fue que ésta era directamente proporcional al nivel de diferenciación y maduración de las células. Se pensó que quizá este fenómeno puede deberse a que mientras mayor nivel de compactación presenta la cromatina es posible que por cercanía espacial se enriquezcan los contactos presentes entre cromosomas.

Posteriormente se utilizó el programa HiC-Pro el cual permite realizar un análisis de los datos obtenidos en la secuenciación mapeándolos al genoma de referencia e identificando las interacciones entre secuencias cercanas o lejanas del genoma(58).

El programa permite la construcción de mapas desde genoma completo hasta regiones pequeñas específicas de cada cromosoma utilizando como método de normalización la Corrección Iterativa y Descomposición de Vectores Propios o ICE por sus siglas en inglés. Este método toma como punto de partida que todos los loci deberían tener una visibilidad igual y utiliza los alineamientos únicos y la estructura de genoma completo de los datos de Hi-C para descomponer los mapas de contacto y calcular una probabilidad relativa de interacción entre cualquier par de loci en el genoma, alcanzando una visibilidad equivalente a lo largo de todas las regiones (59).Utilizando este programa, en primer lugar, se construyeron las matrices de interacción de genoma completo para observar el paisaje general de cada uno de los tipos celulares (Fig. 41).

En los mapas se pudieron observar varias características del genoma de *Gallus gallus*. En la línea celular HD3 se mapearon varias zonas que destacan en colores más oscuros y se encuentran lejanas a la diagonal. Estas zonas muy probablemente corresponden alteraciones cromosómicas que se han desarrollado en las células dado su cultivo prolongado.

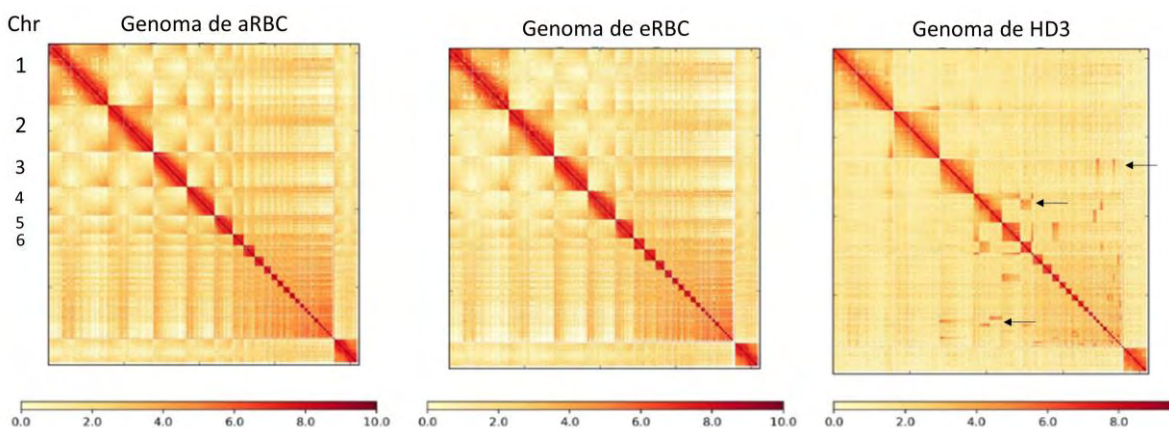
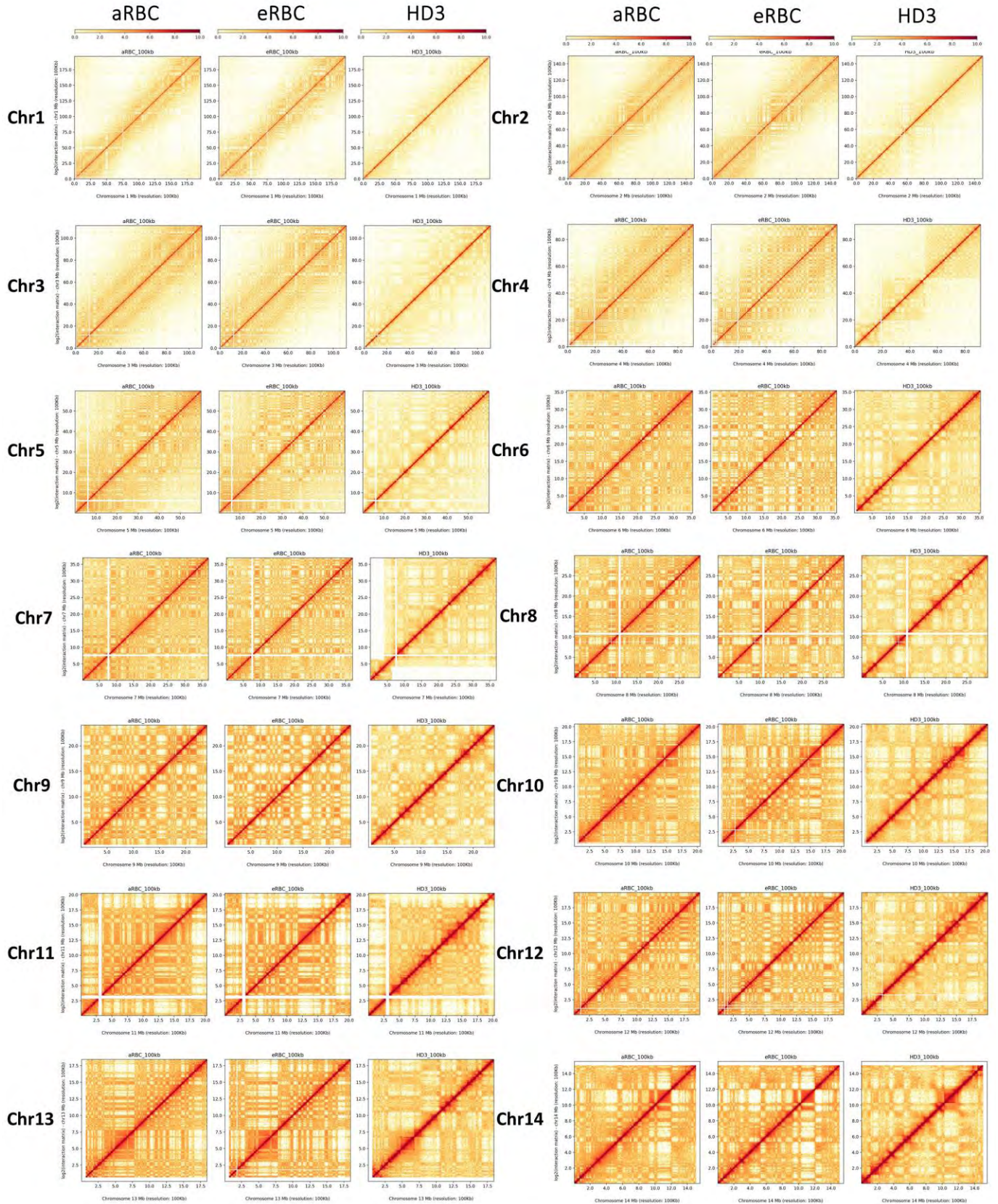


FIGURA 41 MATRICES DE INTERACCIONES GENÓMICAS EN aRBC, eRBC Y HD3 CONSTRUIDOS CON EL PROGRAMA HiC-PRO. LOS CROMOSOMAS SE ENCUENTRAN ORDENADOS POR NÚMERO COMENZANDO POR LA ESQUINA SUPERIOR IZQUIERDA. LOS ÚLTIMOS DOS CROMOSOMAS CORRESPONDEN AL W Y Z. LAS INTERACCIONES SE ENCUENTRAN REPRESENTADAS EN UNA ESCALA DE COLOR QUE VA DEL ROJO AL BLANCO SIENDO EL COLOR ROJO EL NÚMERO MÁS ALTO DE INTERACCIONES ENCONTRADAS (REPRESENTADO EN LA ESCALA NORMALIZADA COMO EL NÚMERO 10) Y EL BLANCO REPRESENTANDO UNA NULA INTERACCIÓN ENTRE LAS REGIONES (REPRESENTADO POR EL NÚMERO 0 EN LA ESCALA).

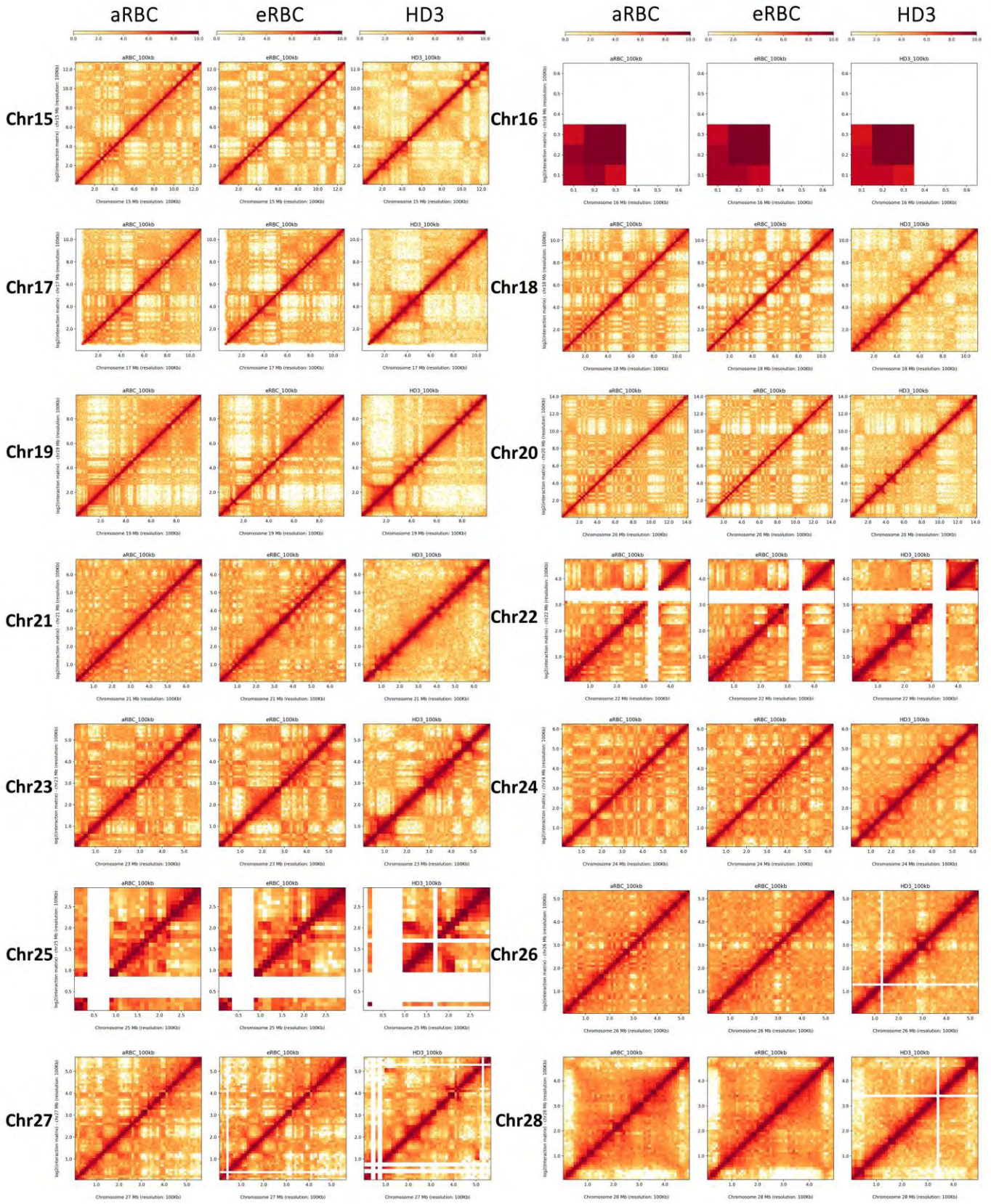
Por otro lado, en los mapas de eritrocitos se puede observar el fenómeno detectado por el programa de HiCUP. Las interacciones en *trans* se encuentran enriquecidas de acuerdo a la escala de color en primer lugar en el mapa de aRBC y en segundo lugar en el de eRBC los dos en comparación con la línea celular de eritroblastos HD3. Así mismo se puede observar un patrón de bandas o líneas que muestran interacciones entre los extremos de los cromosomas lo cual podría dar indicios del arreglo cromosómico que poseen estos tipos celulares que atraviesan por un proceso de silenciamiento transcripcional.

Después de la construcción de los mapas de genoma completo se obtuvieron las matrices de interacción de todos los cromosomas de manera individual (Fig. 42). Debido a que es la primera vez que se utiliza el genoma de *Gallus gallus* en un experimento de Hi-C se tuvo que probar el programa en distintas resoluciones para poder observar cuál resolución podía brindar mayor información. Se realizaron mapas de 100kb (Fig. 42), 50kb, 20kb (Fig. S1) y 10kb de cada uno de los cromosomas. Se observó que en las resoluciones de 100 y 50kb no fue posible la construcción de los mapas de todos los cromosomas debido a que el tamaño de los microcromosomas es muy pequeño y requeriremos de un mayor número de lecturas para poder construirlos en alta resolución. En el caso de 20kb fue posible observar las matrices de interacción, pero no fue posible sacar conclusiones de la información brindada debido a que los mapas cuentan tan sólo con un par de píxeles.

100kb de resolución



100kb de resolución



100kb de resolución

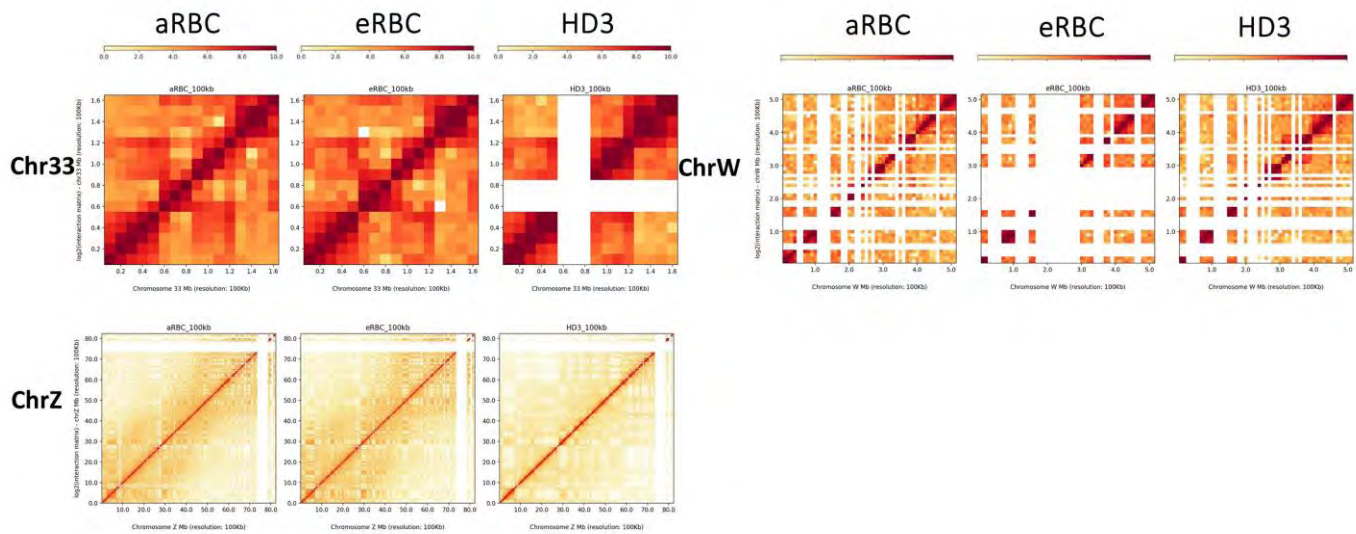


FIGURA 42 MATRICES DE INTERACCIÓN CONSTRUIDAS CON LOS DATOS DEL EXPERIMENTO DE HiC. SE MUESTRAN TODOS LOS CROMOSOMAS DE LOS CUALES SE OBTUVO SUFICIENTE INFORMACIÓN PARA LA CONSTRUCCIÓN DE LAS MATRICES. LOS MAPAS TIENEN UNA RESOLUCIÓN DE 100kb POR PIXEL. EN ALGUNOS MAPAS SE PRESENTAN LÍNEAS BLANCAS LAS CUALES CORRESPONDEN CON REGIONES DEL GENOMA QUE NO SE ENCUENTRAN ANOTADAS POR LO QUE NO ES POSIBLE MAPEAR LA INFORMACIÓN OBTENIDA. LAS INTERACCIONES SE ENCUENTRAN REPRESENTADAS EN UNA ESCALA DE COLOR QUE VA DEL ROJO AL BLANCO SIENDO EL COLOR ROJO EL NÚMERO MÁS ALTO DE INTERACCIONES ENCONTRADAS (REPRESENTADO EN LA ESCALA NORMALIZADA COMO EL NÚMERO 10) Y EL BLANCO REPRESENTANDO UNA NULA INTERACCIÓN ENTRE LAS REGIONES (REPRESENTADO POR EL NÚMERO 0 EN LA ESCALA).

Una característica que se pudo observar en los mapas del genoma de las células eritroides en proceso de maduración (eRBC y aRBC) es que el genoma se encuentra dividido en regiones muy compartimentadas y que hay una aparente ganancia de interacciones a largo alcance en comparación con las células HD3 que corresponden a eritroblastos.

Esto podría estar ocurriendo debido a que, como se ha mencionado anteriormente, las células eritroides atraviesan por un proceso de silenciamiento y compactación del genoma y es posible que este proceso sea gradual por lo tanto se podría esperar que las eRBC aun siendo inmaduras presenten ya un mayor nivel de silenciamiento en su genoma y probablemente también una mayor compactación en comparación con los eritroblastos los cuales aún no se encuentran diferenciados.

Además del programa de HiC-Pro se implementó la utilización de un segundo programa de construcción de matrices y análisis de Hi-C llamado Juicer (60) el cual fue desarrollado por el laboratorio de Lieberman Aiden y permite la construcción de matrices de Hi-C y algunos análisis adicionales como el Análisis de Componentes Principales (PCA) y el reconocimiento de TADs entre otros. Con este programa se construyeron los mapas tanto de genoma completo como por cromosoma (Fig.43) y se observaron patrones similares a los observados con el programa HiC-Pro donde las interacciones de larga distancia se encuentran enriquecidas tanto en los eritrocitos adultos como en los embrionarios en comparación con las HD3.

Con el programa Juicer es posible observar un mejor patrón de interacción de los cromosomas más cortos ya que la resolución llega hasta 5kb, sin embargo, aún en los cromosomas más pequeños la información no es suficiente para poder observar estructuras definidas.

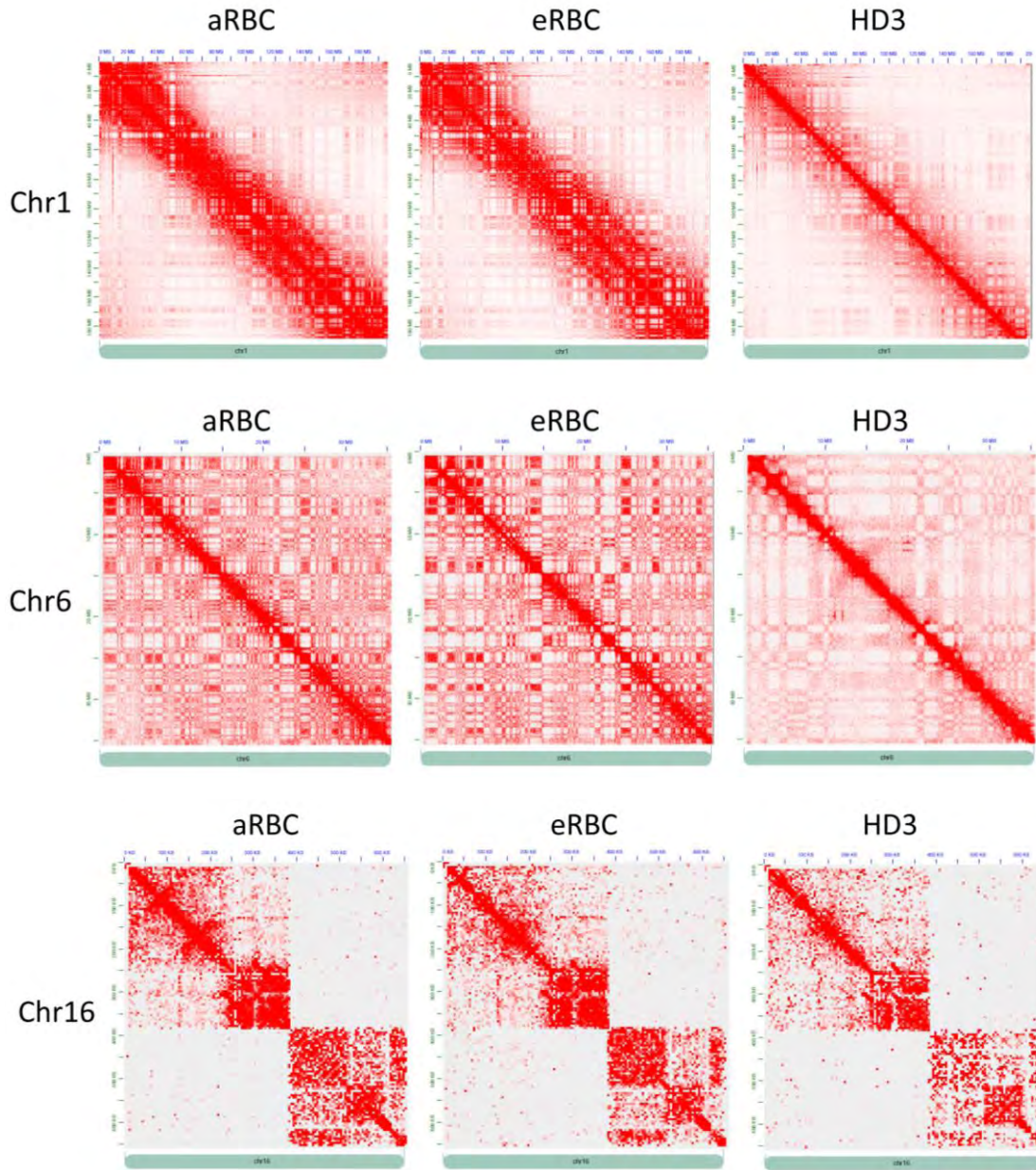


FIGURA 43 MATRICES DE INTERACCIÓN CONSTRUIDAS CON DATOS DE HI-C DE ARBC, eRBC Y HD3 UTILIZANDO EL PROGRAMA JUICER Y VISUALIZADAS CON JUICE-BOX. SE MUESTRAN TRES EJEMPLOS DE MACRO (CHR1), MEDIO (CHR6) Y MICRO-CROMOSOMAS(CHR16) A UNA RESOLUCIÓN DE 100KB EN EL CASO DEL CHR1 Y CHR6 Y 5KB PARA EL MAPA DE CHR16. ESCALA DE COLOR 1-5. LOS MAPAS MUESTRAN LOS DATOS CRUDOS OBTENIDOS DE LOS EXPERIMENTOS DE HI-C. JUICER CONSTRUYE LOS MAPAS DE HI-C UTILIZANDO UN GRADIENTE DE COLOR QUE VA DEL BLANCO AL ROJO DONDE EL BLANCO REPRESENTA UNA NULA INTERACCIÓN ENCONTRADA EN LOS DATOS Y EL ROJO REPRESENTA LA MAYOR CANTIDAD DE INTERACCIONES ENCONTRADAS EN LOS DATOS Y MAPEADAS.

Dado que la observación de aparente enriquecimiento en interacciones a largo alcance dentro de los cromosomas se pudo observar en la mayoría de los mapas construidos, se

decidió utilizar una de las funciones presentes en el programa HiCPlotter para construir matrices comparadas. Para esto se tuvieron que normalizar el número de pares válidos de cada una de las muestras, se tomó el experimento con un menor número de pares válidos, el cual era el de las HD3 con 34, 330,618 lecturas válidas y con ayuda del HiCPro se tomó una muestra aleatoria de pares válidos de los otros dos experimentos. De esta forma la frecuencia de interacciones observada en las matrices podría ser comparada.

Se tomaron tres ejemplos de cromosomas los cuales se mapearon con HiCPlotter y posteriormente el programa realizó una resta entre las interacciones de los tres mapas involucrados (Fig. 44, 45 y 46). Posteriormente se construyeron las matrices normalizadas de las diferencias entre las interacciones, si las interacciones muestran color rojo esto significa un enriquecimiento en la frecuencia de interacción del primer mapa sobre el segundo, por otro lado, si las interacciones son de color azul significa una disminución de la frecuencia de interacción del primer mapa sobre el segundo.

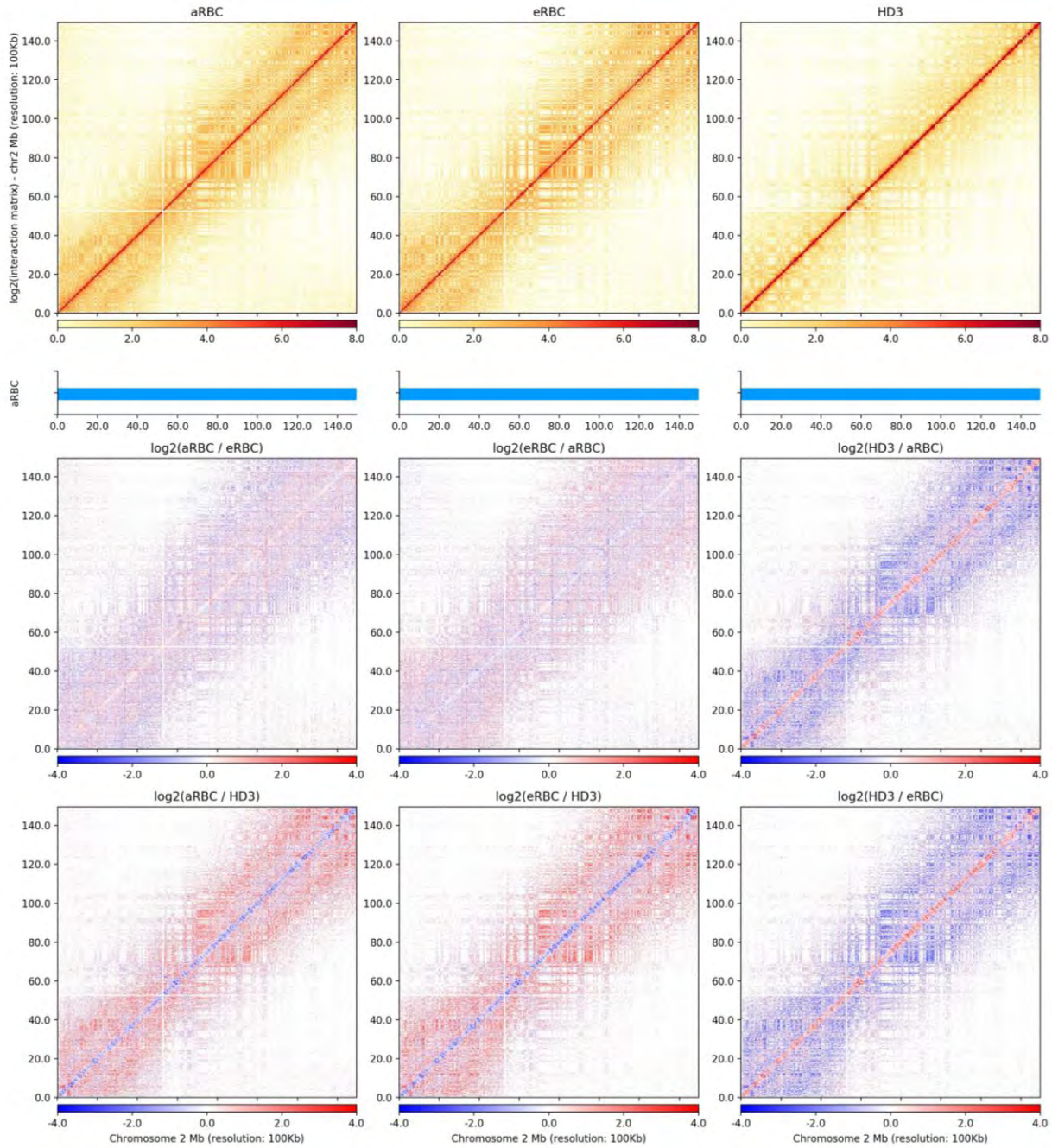


FIGURA 44 MATRICES DE FRECUENCIA DE INTERACCIÓN DE HI-C DEL CROMOSOMA 2 DE ARBC, eRBC Y HD3. EN LA PRIMERA FILA SE MUESTRAN LAS MATRICES NORMALIZADAS POR EL MÉTODO DE ICE Y TAMBIÉN IGUALADAS EN NÚMERO DE PARES VÁLIDOS PARA LOS TRES TIPOS CELULARES. EN LA SEGUNDA Y TERCERA FILA SE MUESTRAN LAS MATRICES RESTADAS. LAS INTERACCIONES SE MARCAN EN COLOR ROJO HAY UN ENRIQUECIMIENTO, EN COLOR BLANCO SI SON SIMILARES ENTRE LOS MAPAS O EN COLOR AZUL SI HAY UNA DISMINUCIÓN DE LAS MISMAS.

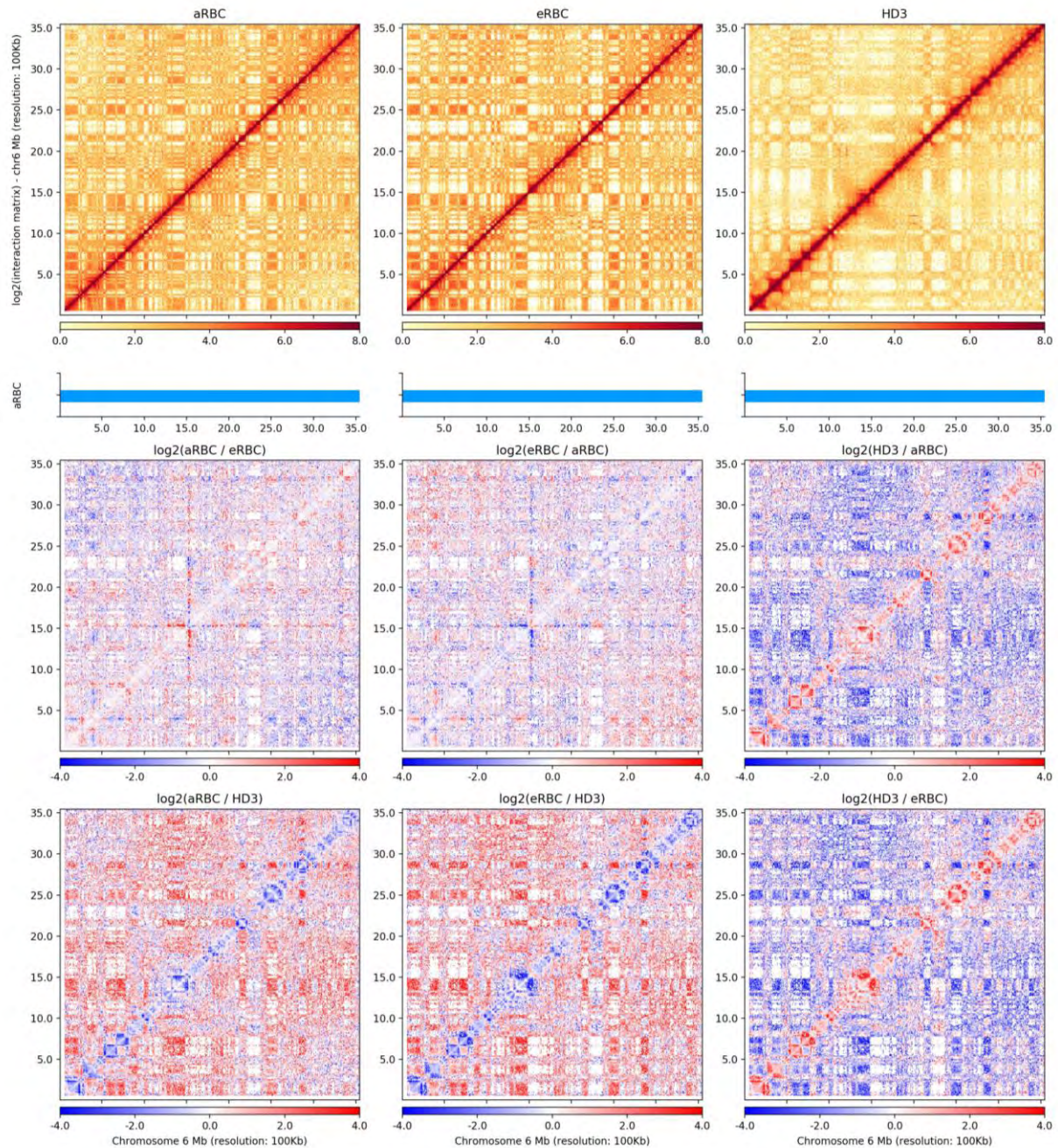


FIGURA 45 MATRICES DE FRECUENCIA DE INTERACCIÓN DE HI-C DEL CROMOSOMA 6 DE ARBC, eRBC Y HD3. EN LA PRIMERA FILA SE MUESTRAN LAS MATRICES NORMALIZADAS POR EL MÉTODO DE ICE Y TAMBIÉN IGUALADAS EN NÚMERO DE PARES VÁLIDOS PARA LOS TRES TIPOS CELULARES. EN LA SEGUNDA Y TERCERA FILA SE MUESTRAN LAS MATRICES RESTADAS. LAS INTERACCIONES SE MARCAN EN COLOR ROJO HAY UN ENRIQUECIMIENTO, EN COLOR BLANCO SI SON SIMILARES ENTRE LOS MAPAS O EN COLOR AZUL SI HAY UNA DISMINUCIÓN DE LAS MISMAS.

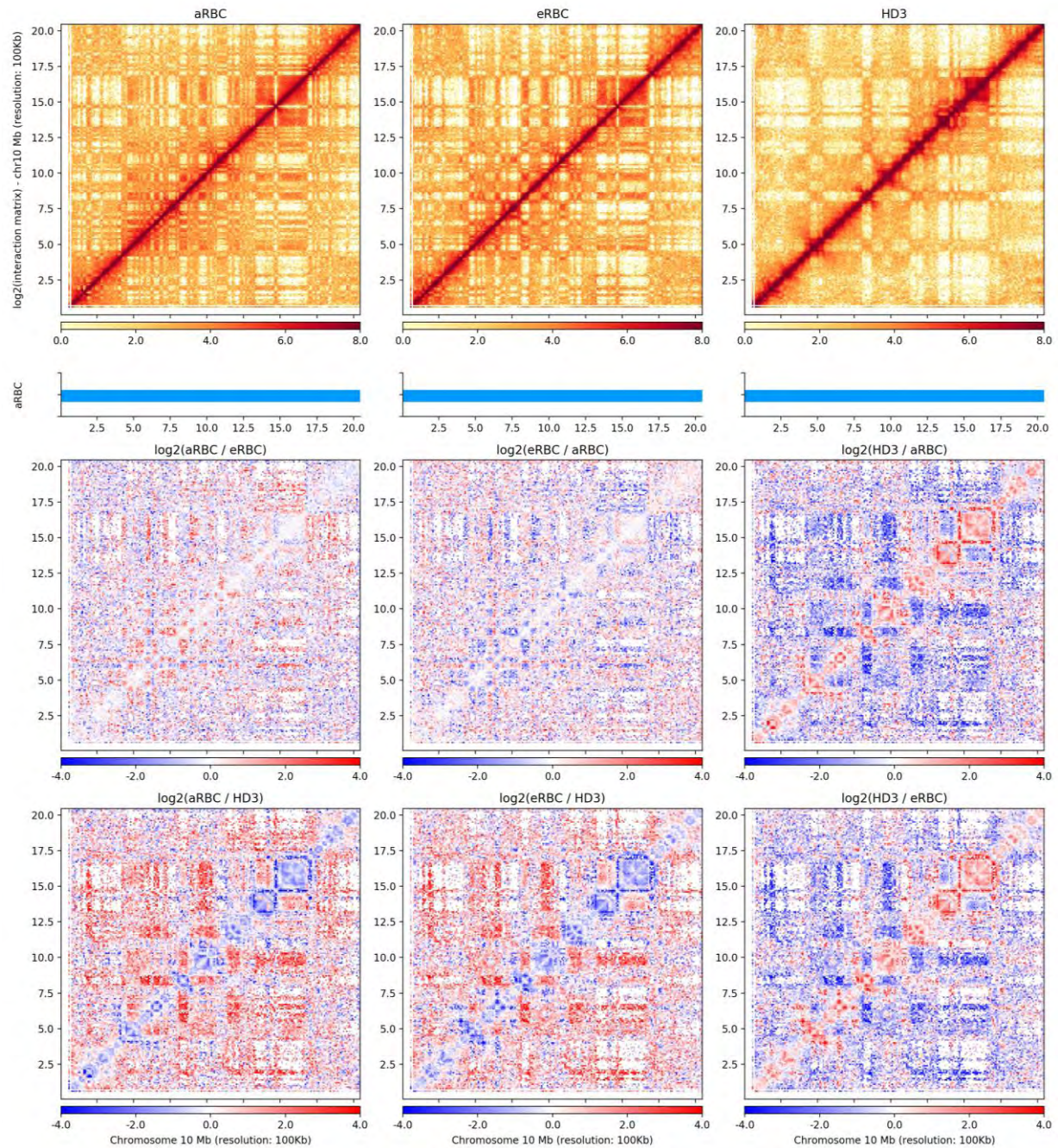


FIGURA 46 MATRICES DE FRECUENCIA DE INTERACCIÓN DE HI-C DEL CROMOSOMA 10 DE ARBC, ERBC Y HD3. EN LA PRIMERA FILA SE MUESTRAN LAS MATRICES NORMALIZADAS POR EL MÉTODO DE ICE Y TAMBIÉN IGUALADAS EN NÚMERO DE PARES VÁLIDOS PARA LOS TRES TIPOS CELULARES. EN LA SEGUNDA Y TERCERA FILA SE MUESTRAN LAS MATRICES RESTADAS. LAS INTERACCIONES SE MARCAN EN COLOR ROJO HAY UN ENRIQUECIMIENTO, EN COLOR BLANCO SI SON SIMILARES ENTRE LOS MAPAS O EN COLOR AZUL SI HAY UNA DISMINUCIÓN DE LAS MISMAS.

Las observaciones apuntadas en las primeras matrices acerca del enriquecimiento en interacciones de largo alcance observado en eritrocitos en comparación con los eritroblastos se hicieron evidentes con la construcción de las matrices comparadas. Se pudo observar que desde los macro cromosomas (Fig. 44) las interacciones en *Cis* pero de largo alcance se encontraban enriquecidas en los dos eritrocitos (aRBC y eRBC). Esto puede notarse en las matrices de interacción aRBC/HD3 y eRBC/HD3 donde, en los dos casos hay claramente una dominancia del color rojo en todo el mapa exceptuando las regiones cerca de la diagonal. Esta característica también fue observada en cromosomas de otros tamaños como el 6 (Fig. 45) y el 10 (Fig. 46). Lo que se puede observar en las regiones cerca de la diagonal es una disminución generalizada de las interacciones en *Cis* pero cercanas desde la perspectiva lineal.

Hasta este punto se puede afirmar que existe una disminución generalizada en los contactos *Cis* a corto alcance y un aumento en los contactos *Cis* a largo alcance observado en los eritrocitos en comparación con las HD3. Las diferencias entre los aRBC y eRBC no eran tan generalizadas ni notorias por lo menos en un patrón observable a simple vista en las matrices. El siguiente paso consistió en caracterizar estas observaciones y relacionarlas con los compartimentos y los Dominios Topológicamente Asociados.

Dominios Topológicos Asociados (TADs)

Como se mencionó con anterioridad, existen dos características que definen en gran parte la estructura tridimensional de un genoma y las ayudan a observar las diferencias o similitudes entre la misma comparando distintos estadios, células o tejidos. La primera de estas características corresponde a la presencia de Dominios Topológicamente Asociados o TADs por sus siglas en inglés los cuales se presentan generalmente como triángulos observables sobre la diagonal de las matrices de interacción construidas por medio de experimentos de Hi-C.

Para el estudio de los TADs en los resultados de Hi-C de los tres tipos celulares se utilizó un programa bioinformático llamado TADtool (53) el cual permite la identificación de los dominios dentro datos de Hi-C en todo el genoma completo mediante la métrica de grado de aislamiento o IS por su siglas en inglés.

Observando los patrones de las matrices de Hi-C se pudo notar una aparente disminución de contactos de corto alcance y un aumento de contactos de largo alcance en los dos eritrocitos en comparación con las HD3 por lo que el primer interrogante fue si existía alguna diferencia también entre el número de TADs que conformaban la estructura del genoma.

Lo que se encontró obteniendo el número de TADs presentes en el genoma completo de los tres Hi-C fue un aumento en la cantidad de dominios encontrados en el caso de los eritrocitos adultos (Fig.47). Esto fue una sorpresa ya que, observando los mapas de cromosomas completos se podía observar una disminución o aparente carencia de estructuras reconocidas como TADs, sin embargo, aparentemente el genoma de los adultos presentaba un mayor número de estas estructuras.

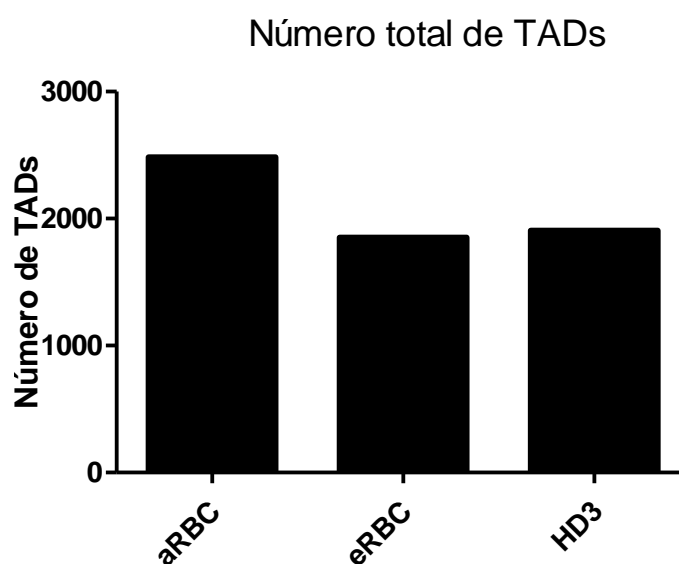


FIGURA 47 GRÁFICO DE BARRAS MOSTRANDO LA DIFERENCIA ENTRE EL NÚMERO TOTAL DE TADS RECONOCIDAS EN LAS TRES MUESTRAS ANALIZADAS CON EL PROGRAMA TADTOOL.

Después de realizar el conteo de forma general en el genoma se decidió observar si existía quizá un patrón que definiera el número de TADs. Se sabe que los cromosomas de *Gallus gallus* poseen una gran variación de tamaño por lo cual se quiso observar el patrón de organización en TADs de cada uno de los cromosomas (Tabla 4). Se pudo observar que no existía ninguna distribución preferencial de TADs definida por el tamaño de los cromosomas. Los patrones variaron mucho de cromosoma a cromosoma pudiéndose observar en algunos de ellos un mayor número presente en las HD3 o en los eritrocitos embrionarios, como es el caso del cromosoma 7 y el 19 respectivamente, sin embargo, en la mayoría de los cromosomas, tal como había indicado el gráfico anterior, se observó un número mayor de TADs en los eritrocitos adultos.

TABLA 4 SE MUESTRA EL NÚMERO DE TADS ENCONTRADO CON EL PROGRAMA TADTOOL EN CADA UNO DE LOS HI-C DE LOS TRES TIPOS CELULARES ANALIZADOS.

Número de TADs por cromosoma			
Cromosoma	aRBC	eRBC	HD3
chr1	587	401	244
chr2	229	156	334
chr3	251	185	208
chr4	246	137	161
chr5	180	164	99
chr6	93	83	70
chr7	50	46	92
chr8	82	81	48
chr9	74	66	41
chr10	40	31	61
chr11	60	41	39
chr12	41	33	52
chr13	53	53	34
chr14	48	38	30
chr15	25	23	33
chr17	31	24	26
chr18	22	19	28
chr19	25	26	21
chr20	27	23	36
chr21	9	9	24
chr22	6	8	9
chr23	19	20	10
chr24	19	13	17
chr26	14	13	16
chr27	10	11	14
chr28	13	12	5
chrZ	230	136	153
Total	2484	1852	1905

Lo siguiente que surgió como pregunta fue referente al tamaño de las TADs observadas dentro de los datos de los tres Hi-C (Fig. 48). Dado que el programa TADtool brinda los resultados marcando las coordenadas de inicio y término de cada uno de los Dominios, fue posible utilizar esos datos para realizar una comparación de los tamaños observados

promedio tanto en la totalidad de los datos como en cada uno de los cromosomas analizados.

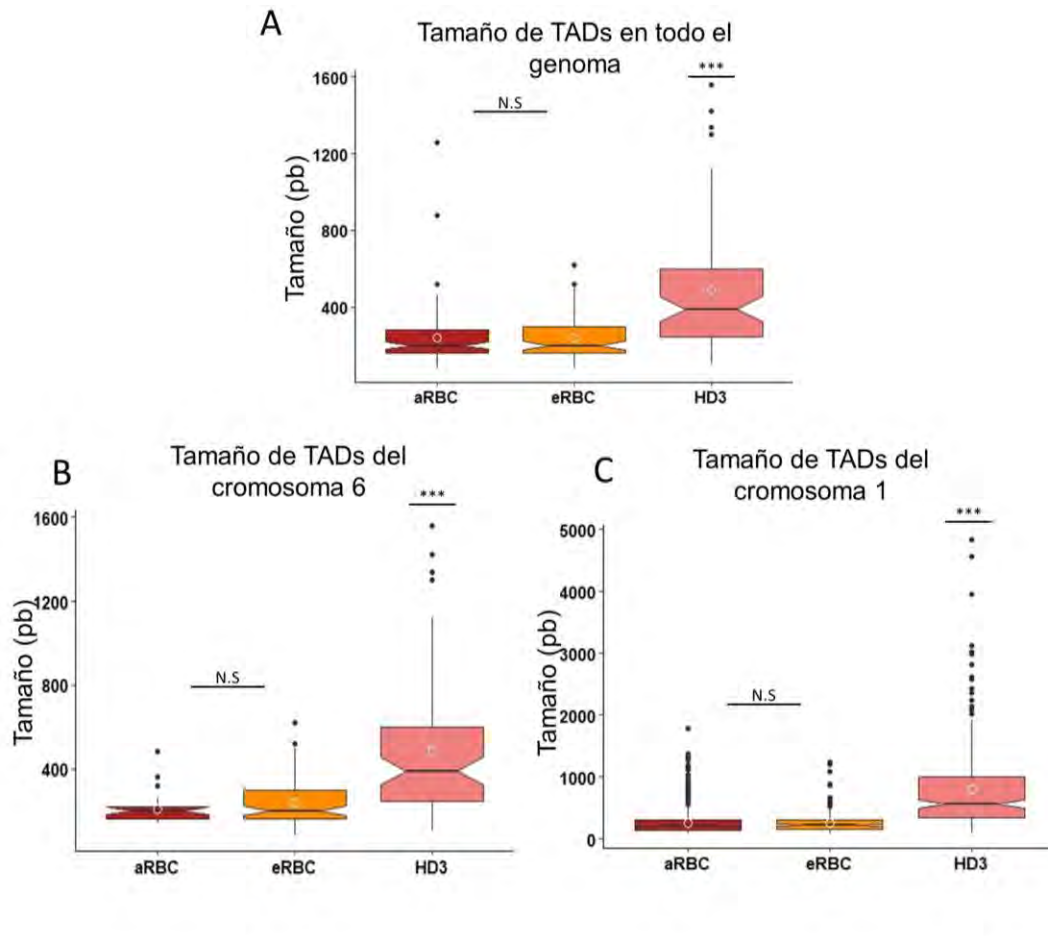
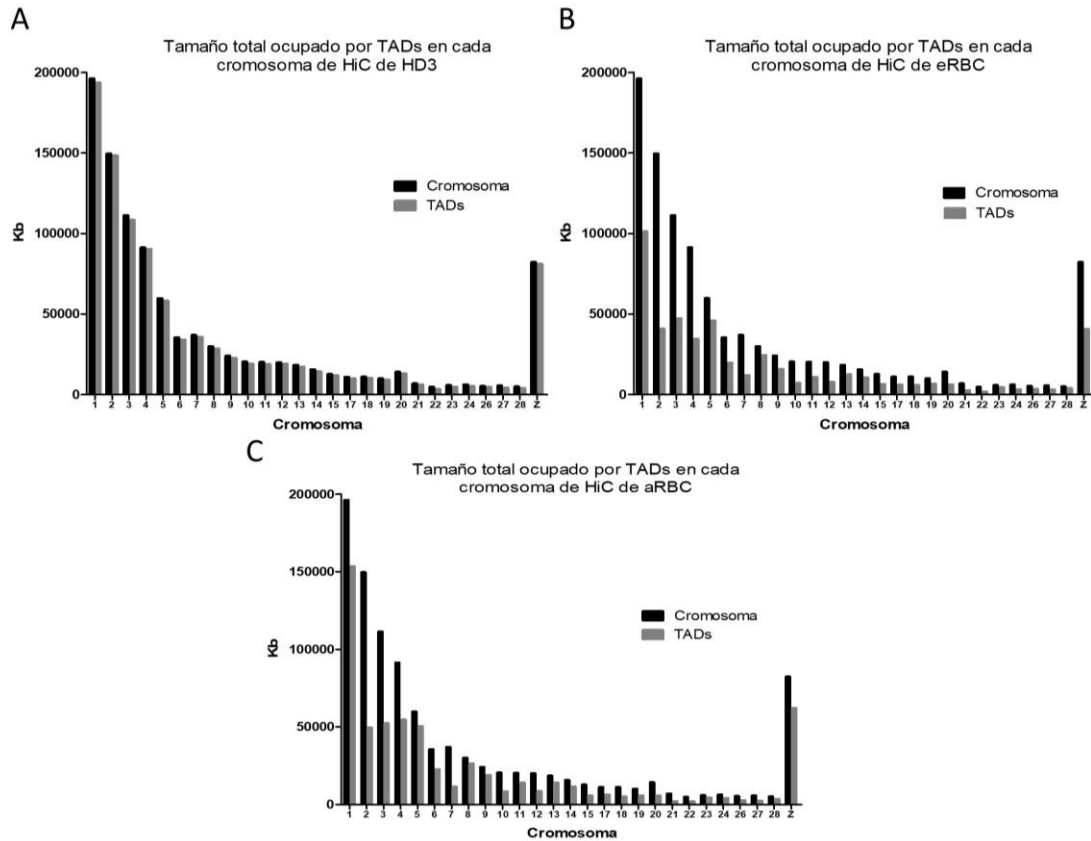


FIGURA 48 GRÁFICOS MOSTRANDO EL TAMAÑO DE LOS TADs ENCONTRADOS EN LOS TRES HI-C ANALIZADOS CON EL PROGRAMA TADTOOL. A) TAMAÑO TOTAL DE LOS TADs. B) TAMAÑO DE LOS TADs ENCONTRADOS EN EL CROMOSOMA 6 (TAMAÑO MEDIO). C) TAMAÑO DE LOS TADs ENCONTRADOS EN EL CROMOSOMA 1 (MACRO-CROMOSOMA). LAS PRUEBAS ESTADÍSTICAS UTILIZADAS FUERON ANOVA CON UNA PRUEBA DE TUKEY ($P < 0.05$).

A diferencia de lo que se observó en el número total de TADs, en este caso el mayor tamaño de los TADs fue encontrado en las células HD3. Esto fue un patrón que se mantuvo en la mayoría de los cromosomas como puede observarse en los gráficos. Los TADs de las células eritroides poseen un tamaño que varía entre 150 y 300 pb, sin embargo, el tamaño observado dentro de las células HD3 es mucho más variable encontrándose la mayoría entre las 300 y 700 pb y habiendo Dominios de hasta más de una megabase. Estas

características concuerdan más con lo reportado en la literatura donde los TADs en los genomas de mamíferos son descritos con un tamaño promedio de 880kb (29) en algunos casos y de poco menos de una megabase en otros (36). Recordando que el genoma de *Gallus gallus* mide aproximadamente 3 veces menos que el genoma del ratón o del humano podría esperarse quizá que el tamaño promedio de los TADs fuera también menor, aunque este es un dato que deberá corroborarse mediante el uso de otros tejidos y tipos celulares del pollo.

Sin embargo, la diferencia entre el promedio de tamaño de las TADs en las células HD3 y en los eritrocitos fue una característica notoria por lo que la siguiente pregunta consistió en cuantificar cuánto del genoma se encuentra estructurado en TADs en los tres tipos celulares. Para responder esta pregunta se decidió tomar cada una de las coordenadas de los TADs reportadas por TADtool y sumarlas, posteriormente comparar esa suma, que representa el total de pares de bases ocupadas por TADs, con la longitud total de cada uno de los cromosomas del genoma del pollo (Fig. 49). Esto esclareció la razón de los resultados anteriores mostrando que, aunque en las HD3 había un menor número de TADs, éstos ocupaban el 97% de la longitud del genoma que se analizó mediante HiCPro y TADtool. En comparación, los eritrocitos mostraron un 48% de ocupación de TADs en el caso de eRBC y cerca de un 60% en el caso de los aRBC.



los TADs ya que estos parecen estar acentuados respecto a las los eritroblastos. A continuación, se formalizará el análisis de los compartimentos para redondear dicha observación.

Posteriormente se observó si existían diferencias o similitudes entre estructuración de los dominios en los tres tipos celulares. Para esto se utilizó un programa diseñado para la comparación específica entre las coordenadas de la lista de TADs brindadas por TADtool. Se consideró que, dado que las matrices utilizadas para llamar los TADs tenían una resolución de 20kb, se debía dar una ventana de la misma longitud como límite para la comparación de dominios compartidos en los tres tipos celulares (Fig. 50).

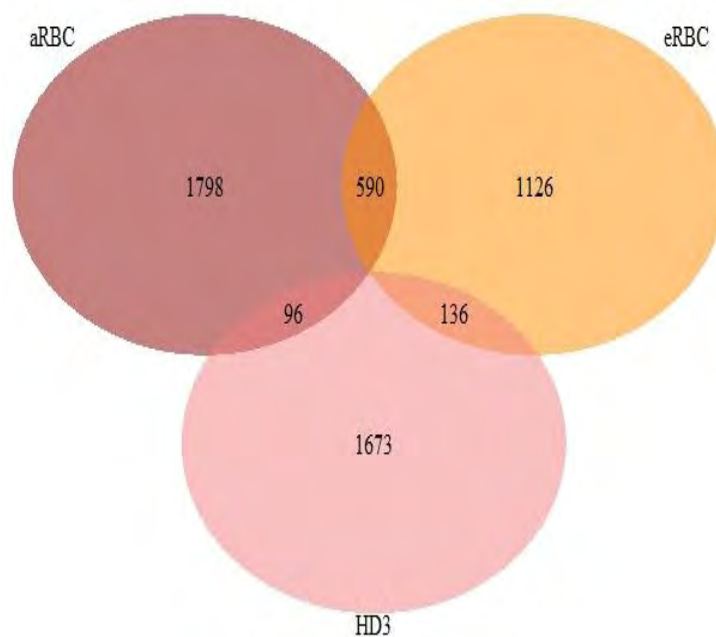


FIGURA 50 DIAGRAMA DE VENN DONDE SE MUESTRAN LOS TADs RECONOCIDOS POR EL PROGRAMA BIOINFORMÁTICO TADTOOL QUE SE COMPARTEN ENTRE LOS TRES TIPOS CELULARES. SE TOMÓ UNA VENTANA DE 20KB ENTRE EL INICIO Y TÉRMINO DE LOS DOMINIOS DEBIDO A QUE LA RESOLUCIÓN DE LAS MATRICES UTILIZADAS FUE DE 20KB POR CADA PIXEL.

En esta comparación se observó que el mayor número de TADs compartidas se encontraron entre las dos células eritroides, sin embargo el porcentaje de dominios compartidos entre ellas no rebasa el 30% del total de los TADs encontrados en cada una de ellas por lo que se puede considerar que la estructuración, por lo menos en una escala a

nivel de Dominios Topológicamente Asociados, es aparentemente muy distinta entre estos dos estadios celulares y definitivamente diferente a la observada en las HD3 las cuales presentan un patrón mucho más similar a lo reportado con anterioridad en genomas de otros organismos.

Dadas las observaciones referentes a los dominios se decidió tomar algunas regiones de interés y observar de cerca el patrón de interacción presente en cada una de ellas. Para la elección de regiones se tomaron como base las zonas donde se encuentran los genes cuya expresión fue medida debido a que se tenían los datos referentes a las diferencias en la transcripción de los mismos en los tres tipos celulares. El procedimiento consistió en buscar la localización de estos genes dentro del genoma y posteriormente, con ayuda de la herramienta que se utilizó anteriormente para llamar TADs en todo el genoma, analizar exclusivamente la zona. Cada zona constó aproximadamente de 3Mb debido a que, de acuerdo a la documentación de TADtool es la mínima zona que puede ser analizada.

Para la calibración del valor de corte o "*cutoff*" en el programa TADtool se utilizó la zona observada en HD3. Posteriormente se construyó la matriz de interacción de cada una de las zonas a estudiar con ayuda del programa HiCPlotter y se colocaron las coordenadas de los genes y de los TADs en cada una de las regiones.

Por último, debido a que la zona a estudiar era amplia y para poder apreciar las interacciones de una forma más específica de cada uno de los genes, se realizó un estudio de 4C virtual con ayuda del programa SeqMonk, el cual permitió construir gráficos de interacción utilizando los datos producidos por los experimentos de Hi-C. Esto se realizó tomando como punto vista solamente un punto en el genoma y graficando todas las interacciones que esa región elegida realiza con el resto del genoma.

La primera zona analizada por estos programas corresponde al locus Alfa Globina D, el cual se encuentra en el cromosoma 14 muy cercano a los otros dos genes de Alfa Globinas (Alfa

Globina Pi y Alfa Globina A) (Fig. 51). Lo que se pudo observar comenzando desde el panel superior de las tres figuras, es una clara diferencia en la estructuración. Si bien las líneas blancas que corresponden a la estructuración en compartimentos se comparten entre las tres matrices esto no ocurre con los Dominios Topológicamente Asociados.

En el caso de las HD3 pueden observarse claramente 6 o 7 estructuras triangulares correspondientes a los enriquecimientos en las frecuencias de interacción relacionados con los TADs. Sin embargo, en el caso de las eRBC y las aRBC esto no es tan fácilmente apreciable. Las estructuras que se logran observar son mucho más pequeñas y difíciles de definir.

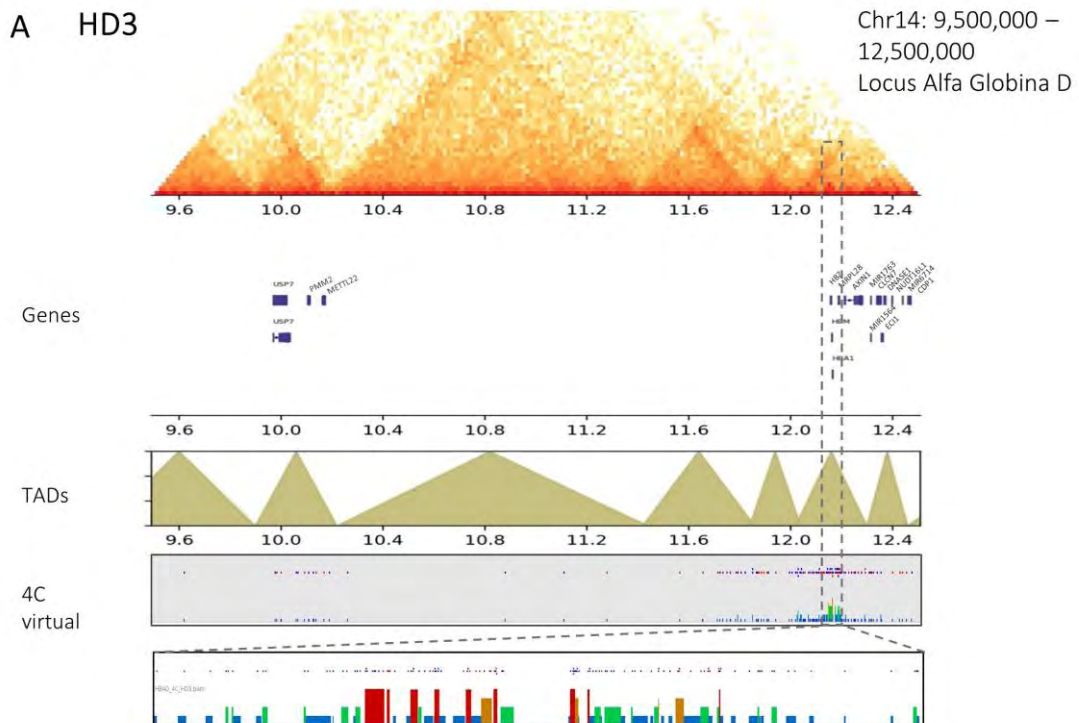
Una observación que parece mantenerse en los tres casos es el enriquecimiento de interacciones justo localizado en la zona donde se encuentran los tres genes que codifican para las proteínas Alfa Globinas (Fig. 51 marcado por la línea punteada). El enriquecimiento no parece ser un TAD definido, pero es evidente que la interacción entre esta zona, probablemente entre los promotores de los genes es frecuente en los tres tipos celulares.

Esto concuerda con lo descrito en la literatura donde se observó mediante el uso de la técnica de Captura de Conformación Cromosómica (3C) que, en células HD3, la zona de Alfa Globina D contiene varios elementos regulatorios como son un enhancer río arriba llamado MRE y una zona sensible a DNAsa y estos elementos presentaron interacciones frecuentes especialmente con los genes Alfa D y Alfa A (26). Esto se corroboró con los análisis de 4C virtual, los cuales consisten en utilizar los datos obtenidos de los experimentos Hi-C y, utilizando el programa SeqMonk, tomar un solo punto de vista o “ancla” dentro de los datos y obtener todas las interacciones que se encontraron en el experimento con la zona elegida.

Esto permite obtener información más precisa acerca de una zona específica del genoma y todas las regiones que, dentro del experimento de genoma completo, se encontraron

cercanas a un punto represando posibles interacciones. En los 4C virtuales se observa que las interacciones de la zona se encuentran principalmente concentradas en un espacio lineal pequeño, correspondiente a la zona cercana a los genes Globina y especialmente orientadas hacia 5' donde se encuentran los elementos regulatorios de la zona (26). Esta puede ser la razón del enriquecimiento de señal en esa zona, sin embargo, en cuanto a las TADs los tres modelos celulares no parecen presentar más similitudes.

Al igual que con el conteo, tamaño y ocupación de TADs, las HD3 parecen ser las que correlacionan de mejor manera con lo anteriormente reportado, lo cual concuerda con su estatus transcripcionalmente activo. Los TADs reconocidos por TADtool presentaron una estructura muy similar a la observada en estas mismas células en un estudio de 5C(61) realizado en el año 2016 lo cual puede indicar que el experimento es consistente con lo observado en este tipo celular y resalta las diferencias observadas con los eritrocitos maduros.



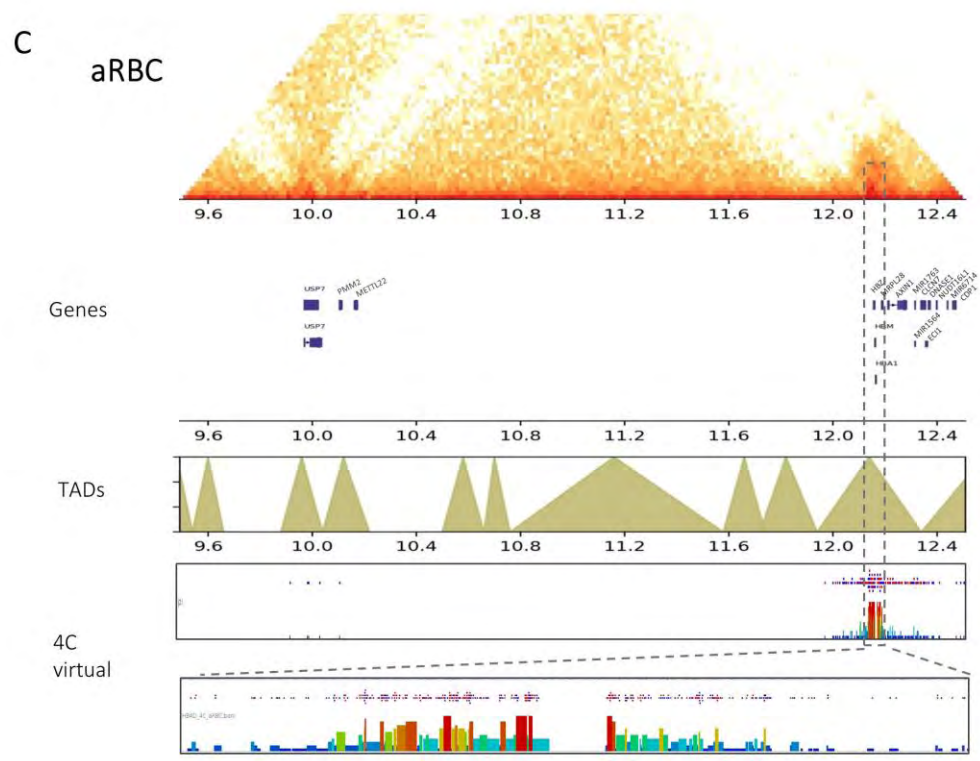
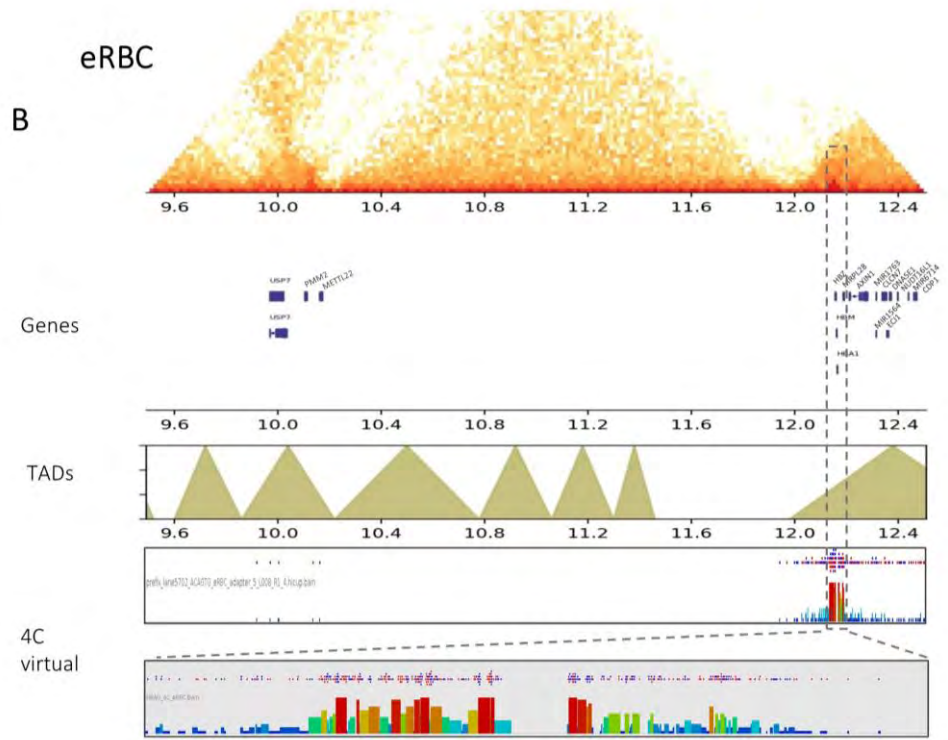


FIGURA 51 GRÁFICOS DONDE SE MUESTRA UNA ZONA DE 3MB DEL CROMOSOMA 14 DONDE SE ENCUENTRAN LOS TRES GENES QUE CODIFICAN PARA LAS PROTEÍNAS ALFA GLOBINAS (GEN CANDIDATO DEL GRUPO DE LAS GLOBINAS). SE MUESTRA LA ZONA ANALIZADA POR HICPRO Y GRAFICADA COMO UNA MATRIZ DE INTERACCIÓN CON EL PROGRAMA HICPLOTTER A 20KB. SE MUESTRAN TAMBIÉN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA ANOTADAS EN LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA. SE MUESTRAN LOS TADS RECONOCIDOS POR EL PROGRAMA TADTOOL Y LOCALIZADOS DENTRO DE LA ZONA REPRESENTADOS COMO TRIÁNGULOS. POSTERIORMENTE SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL REALIZADO CON EL PROGRAMA SEQMONK, DONDE SE ELIGIÓ COMO ANCLA LA ZONA QUE CONTIENE A LOS TRES GENES ALFA GLOBINA, TAMBIÉN SE MUESTRA UN ACERCAMIENTO A LA ZONA MARCADA POR LAS DOS LÍNEAS PUNTEADAS, LAS BARRAS REPRESENTAN LA INTERACCIÓN DE ESA ZONA CON EL ANCLA, LA CUAL ESTÁ MARCADA POR EL RECTÁNGULO DE LÍNEAS PUNTEADAS. A) ZONA DE ALFA GLOBINA D EN HD3. B) ZONA DE ALFA GLOBINA D EN ERBC. C) ZONA DE ALFA GLOBINA D EN ARBC.

En el caso de los genes eritroides los patrones observados fueron similares. El gen EPB42 que codifica para la proteína de la membrana eritroide 4.2 y que presentó una expresión casi exclusiva en eRBC encontrándose de forma muy baja en HD3, presentó una estructuración similar a la encontrada en la zona de Alfa Globina D (Fig.52).

En el caso de las matrices de interacción visiblemente pueden apreciarse la presencia de pequeños y medianos triángulos formando los TADs en las células HD3, sin embargo, esto no es claro en los otros dos tipos celulares. De forma interesante, observando los TADs encontrados en el análisis el gen de interés se encuentra dentro de uno de estos dominios que es compartido tanto por las HD3 como por las eRBC. Esto no ocurre en las aRBC donde el dominio es ligeramente más pequeño. Cabe resaltar también que esta zona presenta un número mucho mayor de TADs en las células HD3 en comparación con los eritrocitos.

Los compartimentos nuevamente parecen comportarse de manera similar pero la estructuración a un nivel más fino está alterada. El 4C virtual corrobora el dominio topológico en el que se encuentra el gen especialmente en eRBC y aRBC donde puede apreciarse que las interacciones se encuentran principalmente limitadas dentro del mismo. Así mismo por este análisis puede apreciarse que la estructuración de toda la zona con referencia al gen que se utilizó como ancla es más compleja o está mayormente extendida en las HD3 donde, aunque la mayor parte de las interacciones se encuentran cerca del gen,

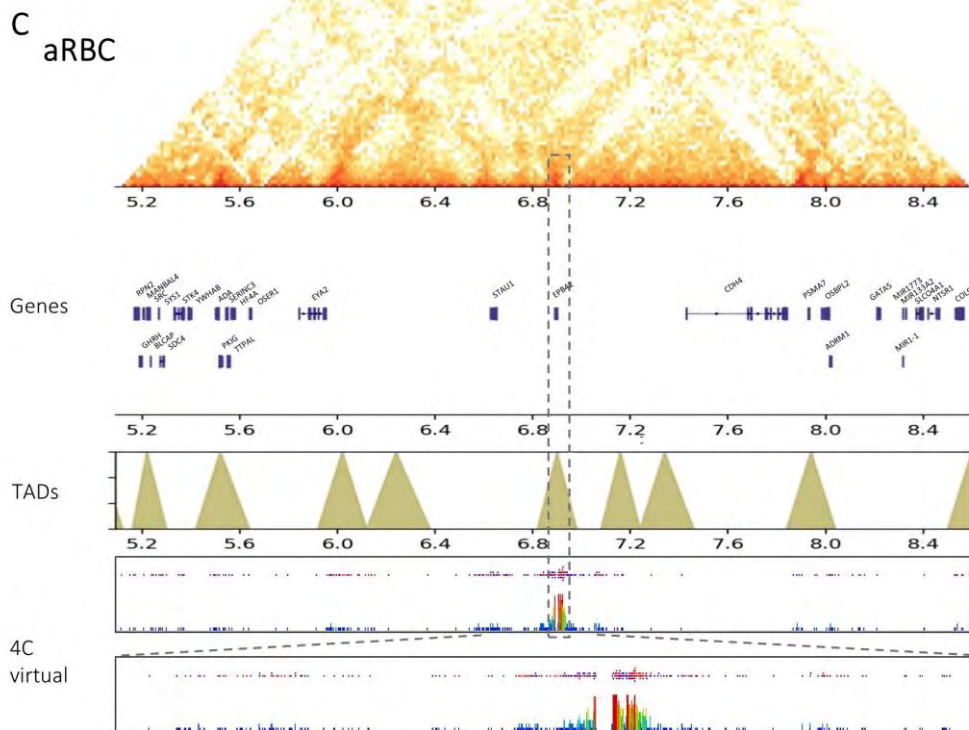
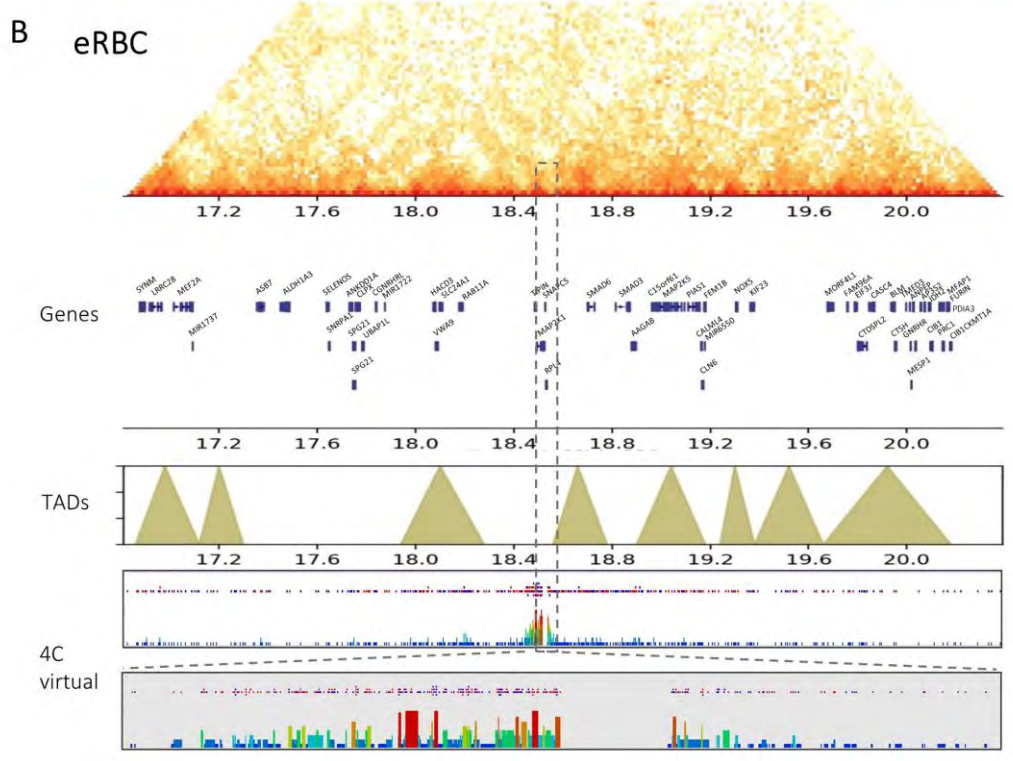
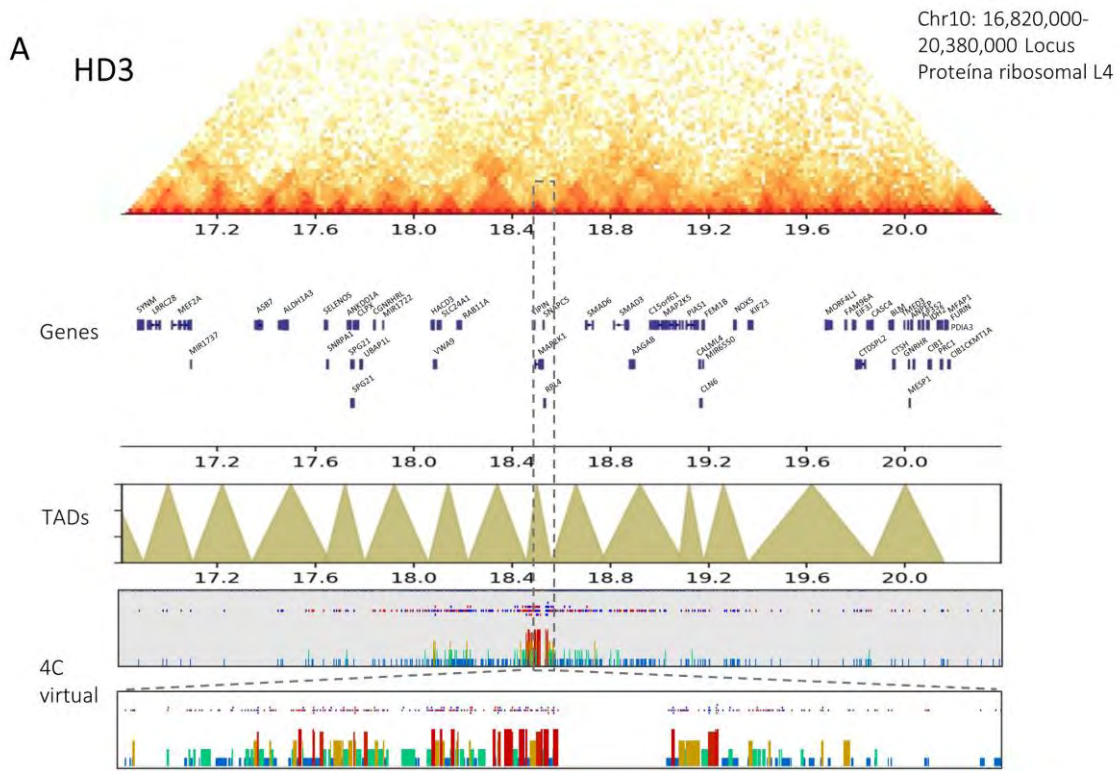


FIGURA 52 GRÁFICOS DONDE SE MUESTRA UNA ZONA DE 3.5MB DEL CROMOSOMA 20 DONDE SE ENCUENTRA EL LOCI DEL GEN EPB42 (GEN CANDIDATO DEL ERITROIDE). SE MUESTRA LA ZONA ANALIZADA POR HICPRO Y GRAFICADA COMO UNA MATRIZ DE INTERACCIÓN CON EL PROGRAMA HICPLOTTER A 20KB. SE MUESTRAN TAMBIÉN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA ANOTADAS EN LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA. SE MUESTRAN LOS TADS RECONOCIDOS POR EL PROGRAMA TADTOOL Y LOCALIZADOS DENTRO DE LA ZONA REPRESENTADOS COMO TRIÁNGULOS. POSTERIORMENTE SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL DE LA ZONA REALIZADO CON EL PROGRAMA SEQMONK DONDE SE TOMÓ COMO ANCLA LA ZONA DONDE SE ENCUENTRA EL GEN EPB42, Y UN ACERCAMIENTO A LA ZONA MARCADA POR LAS DOS LÍNEAS PUNTEADAS, LAS BARRAS REPRESENTAN LA INTERACCIÓN DE ESA ZONA CON EL ANCLA, LA CUAL ESTÁ MARCADA POR EL RECTÁNGULO DE LÍNEAS PUNTEADAS. A) ZONA DE EPB42 D EN HD3. B) ZONA DE EPB42 EN ERBC. C) ZONA DE EPB42 EN ARBC.

El patrón nuevamente es observado en el caso de un gen candidato que corresponde al grupo de los genes constitutivos. RPL4 es un gen ribosomal, que presentar expresión en HD3 y también, aunque de forma mucho menor, se expresa en eRBC. Los TADs son observados en las interacciones de la zona como pequeños triángulos y aunque algunos TADs se conservan entre los tres tipos celulares las interacciones reportadas por el 4C virtual se observan de forma aislada en el caso de los eritrocitos y mayormente distribuidas por la zona en el caso de las HD3 (Fig. 53).



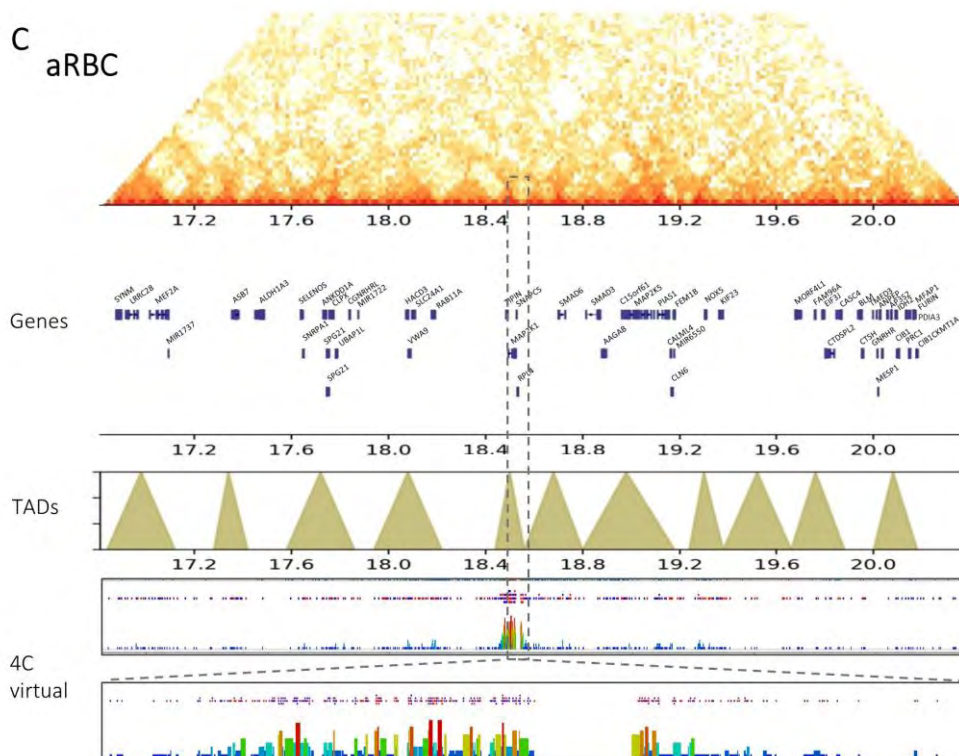


FIGURA 53 GRÁFICOS DONDE SE MUESTRA UNA ZONA DE 3.56MB DEL CROMOSOMA 10 DONDE SE ENCUENTRA EL LOCI DEL GEN RPL4 (GEN CANDIDATO DEL GRUPO CONSTITUTIVO). SE MUESTRA LA ZONA ANALIZADA POR HICPRO Y GRAFICADA COMO UNA MATRIZ DE INTERACCIÓN CON EL PROGRAMA HICPLOTTER A 20KB. SE MUESTRAN TAMBIÉN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA ANOTADAS EN LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA. SE MUESTRAN LOS TADS RECONOCIDOS POR EL PROGRAMA TADTOOL Y LOCALIZADOS DENTRO DE LA ZONA REPRESENTADOS COMO TRIÁNGULOS. POSTERIORMENTE SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL DE LA ZONA REALIZADO CON EL PROGRAMA SEQMONK DONDE SE TOMÓ COMO ANCLA LA ZONA DONDE SE ENCUENTRA EL GEN RPL4, Y UN ACERCAMIENTO A LA ZONA MARCADA POR LAS DOS LÍNEAS PUNTEADAS, LAS BARRAS REPRESENTAN LA INTERACCIÓN DE ESA ZONA CON EL ANCLA, LA CUAL ESTÁ MARCADA POR EL RECTÁNGULO DE LÍNEAS PUNTEADAS. A) ZONA DE RPL42 D EN HD3. B) ZONA DE RLP42 EN ERBC. C) ZONA DE RPL42 EN ARBC.

Con lo observado, los datos de los TADs indican que las diferencias entre las células HD3 y los eritrocitos tanto embrionarios como adultos, son significativas. Sin embargo, la estructuración de los eRBC y los aRBC parece no tener características tan claras como para poder definir si los pequeños pero numerosos TADs encontrados por TADtool corresponden a algún tipo de estructuración particular o si quizá el programa se encuentra reconociendo interacciones ligeramente más frecuentes que no forzosamente representan lo que se ha definido como Dominios Topológicamente Asociados.

Para poder tener una aproximación más clara a este nivel de estructuración se deben utilizar diferentes programas con distintos algoritmos que puedan reconocer Dominios cómo puede el Arrow Head, el cual es la herramienta de Juicer para el llamado de TADs. Sin embargo, este tipo de algoritmos requieren de una gran cantidad de datos por lo que, se espera que se puedan utilizar una vez que las réplicas experimentales de Hi-C de estos tres tipos celulares se encuentren completas.

La observación de la estructura que encuentre Arrow Head y la comparación con lo analizado por TADtool puede brindar mayor información acerca del patrón que se observa en estos tipos celulares.

El genoma de los eritrocitos muestra un patrón altamente compartimentado

Dentro de las observaciones realizadas gracias a los experimentos de Hi-C se encontró que además de que los dominios o TADs se encontraban aislados en interacciones entre ellos tenían una preferencia por interactuar, en una escala más grande, con otros dominios. Las interacciones encontradas se relacionaron con el estado transcripcional o cromatínico de los dominios, es decir, los TADs activos tenían una preferencia por interactuar con otros TADs activos y viceversa (24). Estas interacciones dieron lugar a los llamados compartimentos A y B los cuales se observan en las matrices de interacción como franjas claras y oscuras de mayor tamaño que los TADs que atraviesan por completo todo el mapa.

En los resultados obtenidos de los Hi-C se pudo observar a simple vista que los eritrocitos tanto embrionarios como adultos, presentaban una compartimentación bastante pronunciada ya que las franjas de zonas extensas en el genoma están mucho más delimitadas que en las matrices observadas de las HD3. Sin embargo, para corroborar estos datos y observar las diferencias entre los compartimentos de los experimentos se utilizó el programa Homer el cual está diseñado para analizar datos de Hi-C. Una de las funciones que tiene es realizar el análisis de componentes principales, el cual brinda la separación en

compartimentos A y B. Posteriormente se utilizó el programa IGV para poder observar los datos y compararlos entre ellos.

Para la asignación del signo en el análisis de PCA se utilizó una lista de los inicios de la transcripción (TSS por sus siglas en inglés) obtenida de los datos del genoma del pollo versión 5 (galGal5). En otros modelos como el ratón o el humano se pueden utilizar anotaciones de marcas de histonas que correspondan a zonas de eucromatina o heterocromatina para la asignación del signo (Positivo normalmente para compartimento A y negativo para el B), también los datos de ensayos de hipersensibilidad a la DNAsa pueden ayudar a relacionar el componente principal con la eucromatina. Desgraciadamente, debido a que *Gallusgallus* es un modelo poco estudiado en referencia al estado de la cromatina de sus células y específicamente de los eritrocitos, se tomó la decisión de asignar el signo con base en la posición de los TSS. Sin embargo, en el modelo estudiado, los TSS no representan quizá el mejor parámetro para dividir la cromatina por su estado transcripcional, esto se debe a que, como se ha mostrado con anterioridad, el perfil transcripcional de las células es altamente variable y en el caso de los eritrocitos tanto adultos como embrionarios, muy reducido. Dado que los TSS se encuentran en el mismo lugar en los tres tipos celulares no pueden representar con precisión la distinción de eucromatina y heterocromatina de cada tipo, pero ayudan a establecer el compartimento A con un parámetro conocido.

Después de realizar el análisis de componentes principales y obtener los datos positivos y negativos para cada cromosoma, se decidió observar si el cambio entre zonas positivas y negativas correspondía con el patrón de rallas observado en los mapas de Hi-C (Fig. 54). La comparación del componente 1 arrojado por el análisis y las interacciones “compartimentadas” en las matrices correlacionó de forma adecuada lo que significó que el arreglo cromático que se estaba observando a simple vista en las matrices era recuperado por el PCA.

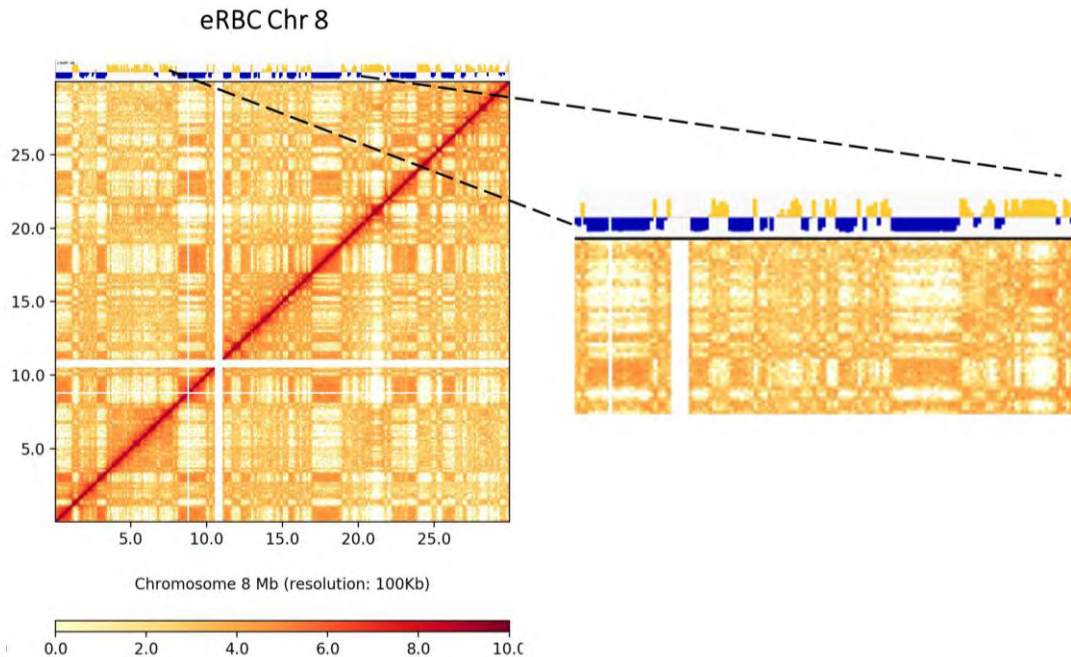


FIGURA 54 MATRIZ DE INTERACCIONES DEL CROMOSOMA 8 EN ERITROCITOS EMBRIONARIOS CONSTRUIDA CON DATOS DE HI-C. SE MUESTRA TAMBIÉN EL VALOR DEL PC1 ARROJADO POR EL ANÁLISIS DE COMPONENTES PRINCIPALES QUE REPRESENTA LA COMPARTIMENTACIÓN OBSERVADA EN LAS INTERACCIONES DE REGIONES GRANDES DEL GENOMA QUE SE MUESTRAN COMO FRANJAS A LO LARGO DEL MAPA.

Posteriormente se realizó la comparación de los compartimentos entre las muestras analizadas por Hi-C (Anexos). En la mayoría de los casos las diferencias más notorias entre compartimentos fueron observadas entre los dos eritrocitos (aRBC y eRBC) y las HD3. Este patrón se mantuvo a lo largo de la mayoría de los cromosomas sin importar su tamaño o posición (Fig. 55, 56 y 57). Este resultado, por un lado, reflejó y apoyó lo observado con anterioridad donde, tanto en los mapas como en las TADs, los dos eritrocitos mostraron parecerse más entre ellos que comparados con los eritroblastos. Sin embargo, cuando se analizó con base en los niveles de expresión observados no parecía correlacionar con lo reportado en la literatura donde se menciona que estas franjas observadas, así como los compartimentos arrojados por el PCA 1 representan la cromatina activa e inactiva de la cromatina (29,33,62).

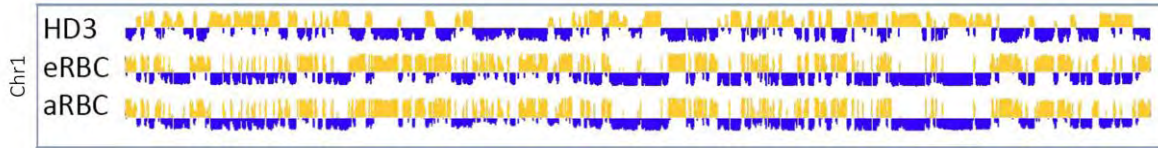


FIGURA 55 GRÁFICO MOSTRANDO EL ANÁLISIS DE COMPONENTES PRINCIPALES REALIZADO SOBRE LOS DATOS DE HI-C. SE OBSERVA LA ORGANIZACIÓN DE LOS DATOS PRESENTE A LO LARGO DE TODO EL CROMOSOMA 1 VISTO DE FORMA LINEAL. LOS DATOS EN AMARILLO REPRESENTAN REGIONES QUE FUERON CATEGORIZADAS DENTRO DEL COMPARTIMENTO A O POSITIVO Y LOS DATOS EN AZUL REGIONES DENTRO DEL COMPARTIMENTO B O CON SIGNO NEGATIVO.

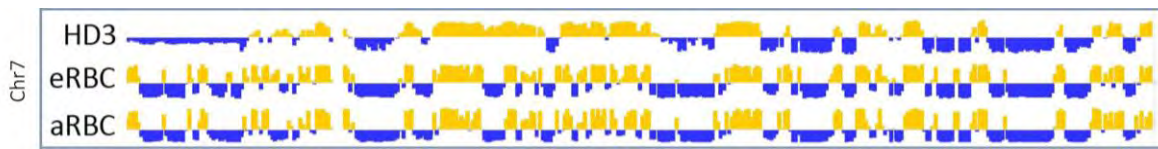


FIGURA 56 GRÁFICO MOSTRANDO EL ANÁLISIS DE COMPONENTES PRINCIPALES REALIZADO SOBRE LOS DATOS DE 7HI-C. SE OBSERVA LA ORGANIZACIÓN DE LOS DATOS PRESENTE A LO LARGO DE TODO EL CROMOSOMA 7 VISTO DE FORMA LINEAL. LOS DATOS EN AMARILLO REPRESENTAN REGIONES QUE FUERON CATEGORIZADAS DENTRO DEL COMPARTIMENTO A O POSITIVO Y LOS DATOS EN AZUL REGIONES DENTRO DEL COMPARTIMENTO B O CON SIGNO NEGATIVO.

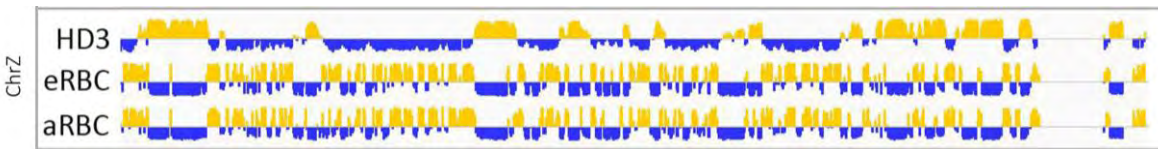


FIGURA 57 GRÁFICO MOSTRANDO EL ANÁLISIS DE COMPONENTES PRINCIPALES REALIZADO SOBRE LOS DATOS DE HI-C. SE OBSERVA LA ORGANIZACIÓN DE LOS DATOS PRESENTE A LO LARGO DE TODO EL CROMOSOMA Z VISTO DE FORMA LINEAL. LOS DATOS EN AMARILLO REPRESENTAN REGIONES QUE FUERON CATEGORIZADAS DENTRO DEL COMPARTIMENTO A O POSITIVO Y LOS DATOS EN AZUL REGIONES DENTRO DEL COMPARTIMENTO B O CON SIGNO NEGATIVO.

Dado que las mediciones de expresión por qPCR mostraron una aparente disminución generalizada en la tasa de transcripción de aRBC la hipótesis acerca del estado de la cromatina en estas células consistía en encontrar un genoma mayoritariamente en el compartimento B. Sin embargo, al medir la cantidad de regiones encontradas pertenecientes a cada uno de los compartimentos se observó que la diferencia entre los compartimentos encontrados por el programa no se encontraba preferencialmente hacia

el compartimento B en las células eritroides, más bien era similar entre los tres tipos celulares (Fig. 58)

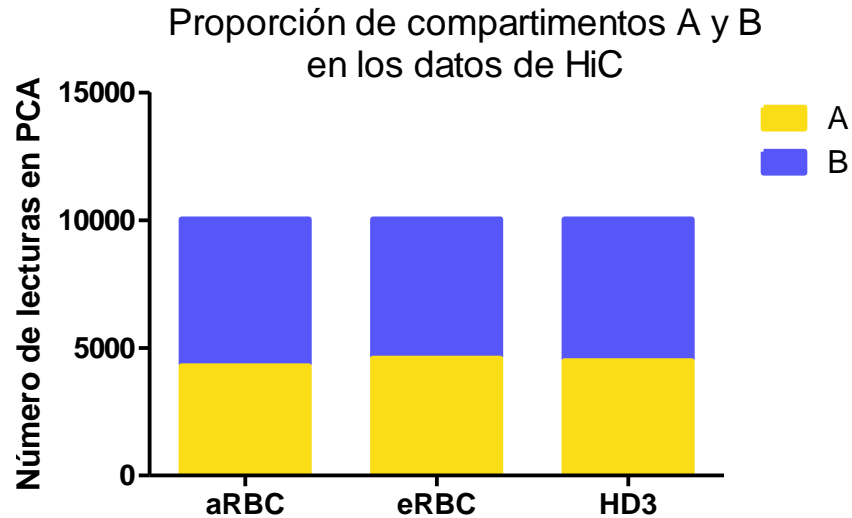


FIGURA 58 GRÁFICO DE BARRAS MOSTRANDO LA PROPORCIÓN DE ZONAS RECONOCIDAS POR EL ANÁLISIS DE COMPONENTES PRINCIPALES PERTENECIENTES A LECTURAS POSITIVAS (A EN AMARILLO) O CON VALOR NEGATIVO (B EN AZUL) ENCONTRADAS EN EL TOTAL DE DATOS DE HI-C OBTENIDOS PARA LOS TRES TIPOS CELULARES ANALIZADOS.

Esta observación indica que posiblemente la estructuración en compartimentos que se aprecia no refleje necesariamente el estado transcripcional de la cromatina. Lo que la información indica es que el genoma dentro del núcleo se encuentra organizado de tal forma que existen interacciones bastante bien delimitadas entre regiones extensas del genoma, sin embargo, con la información que se tiene actualmente acerca del estado transcripcional y de la cromatina de este modelo, no es posible aseverar qué representa esta organización o si se encuentra relacionada con el proceso de silenciamiento nuclear o si es una característica presente en el genoma de este linaje.

Comparación de datos obtenidos con la línea celular DT40

Los resultados de los experimentos de Hi-C realizados en las tres células de linaje eritroide mostraron características que parecen ser propias, ya sea del genoma de esta especie o del

linaje celular. Sin embargo, para encontrar la respuesta a la aseveración anterior se requiere la comparación de los datos obtenidos con un tejido o un linaje celular distinto. Esto se planteó en un inicio utilizando como control no eritroide al cultivo primario de fibroblastos, sin embargo, el experimento de Hi-C no resultó eficiente dentro de este linaje celular.

Dado que no se había publicado ningún experimento de Hi-C realizado en el pollo los resultados se limitaron a los tres tipos celulares antes analizados, sin embargo, en el presente año se publicó un estudio de estructura tridimensional del genoma realizado en una línea celular de pollo proveniente del linaje linfoide (DT40) de *Gallus gallus*, donde realizaron experimentos de Hi-C, uno de ellos en células en interfase(63).

Para poder realizar una comparación de la estructura con células no eritroides de pollo se tomaron los datos de Hi-C sin analizar publicados por Gibcus J. *et al.* y con ayuda del programa HiCPro se analizaron los datos para la construcción de las matrices de interacción. Dado que las células HD3 fueron las que presentaron características más similares a lo antes descrito en la literatura se tomó la decisión de realizar a comparación con esta línea celular. Se construyeron mapas de distintos cromosomas (Fig. 59), específicamente del cromosoma 1, 6 y 17 para tener un representante de los tres grupos de tamaños presentes en el genoma del pollo. Como se puede observar en los mapas aparentemente existe una menor compartimentación en las matrices de las células DT40 donde las interacciones observadas se concentran mayoritariamente cerca de la diagonal y las marcas de los compartimentos no se encuentran tan claras. Cabe resaltar que el experimento de Gibcus J. *et al.* fue realizado con la enzima de restricción *HindIII* la cual reconoce un sitio de restricción de 6 pares de bases realizando cortes menos frecuentes en el genoma y, por lo tanto, obteniendo una menor resolución en el experimento. Por esto se consideró que si bien, la comparación con las DT40 puede dar una pista de las características estructurales del linaje eritroide, se debe realizar una comparación con un Hi-C con las mismas condiciones experimentales.

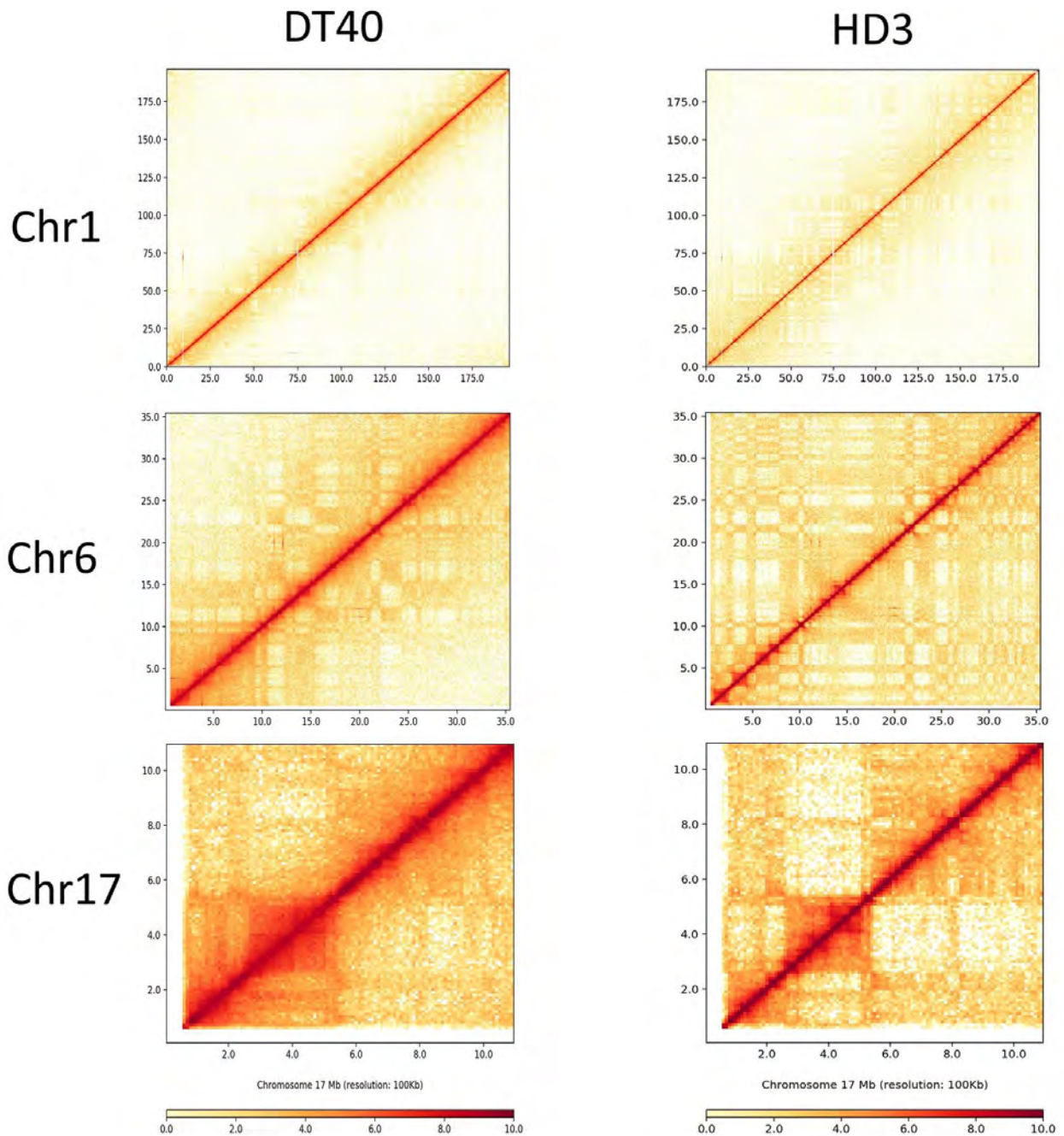


FIGURA 59 MATRICES DE INTERACCIÓN REALIZADAS CON DATOS DE GIBBUS *ET AL.* 2018 A 100KB DE RESOLUCIÓN. COMPARACIÓN CON MATRICES DE HD3 REALIZADAS CON DATOS DE HI-C. SE MUESTRA EN LA FILA SUPERIOR EL CROMOSOMA 1, EN LA FILA MEDIA EL CROMOSOMA 6 Y EN LA FILA INFERIOR EL CROMOSOMA 17. LA ESCALA DE COLOR UTILIZADA VA DEL ROJO, REPRESENTANDO EL MAYOR NÚMERO DE INTERACCIONES ENCONTRADAS (EN LA ESCALA NORMALIZADA = 10) AL BLANCO, REPRESENTADO NULA INTERACCIÓN ENTRE LAS REGIONES MAPEADAS DEL GENOMA (EN LA ESCALA NORMALIZADA = 0)

Así mismo, con ayuda del programa TADtool y utilizando un *cutoff* calibrado para los mapas de 100kb construidos tanto de HD3 como de DT40 se analizaron los TADs compartidos entre las células (Fig. 60).

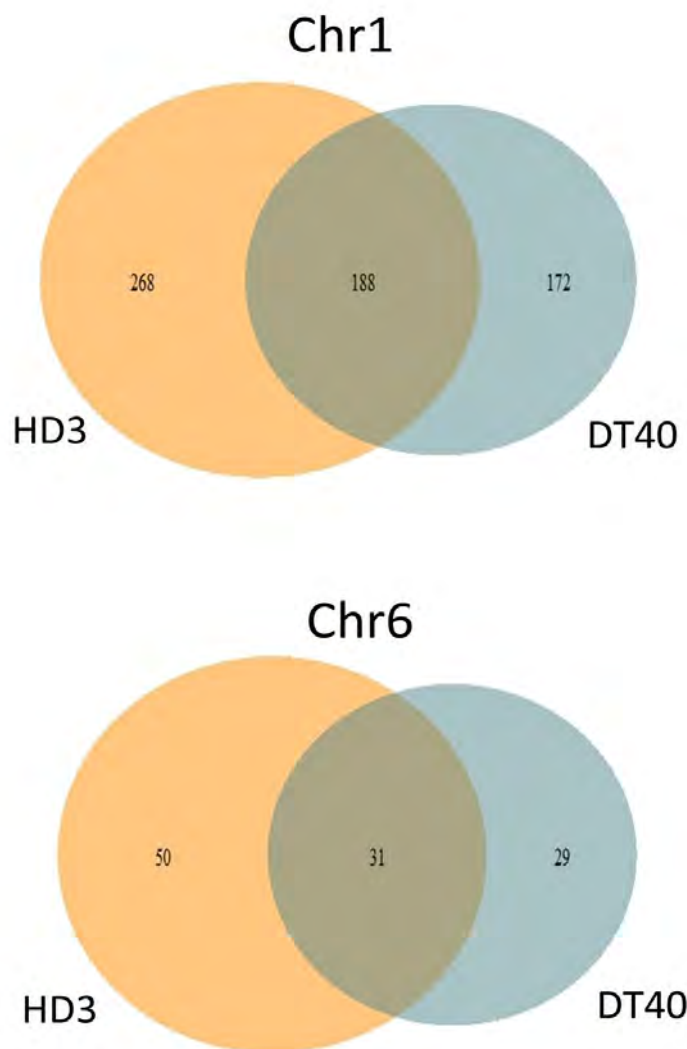


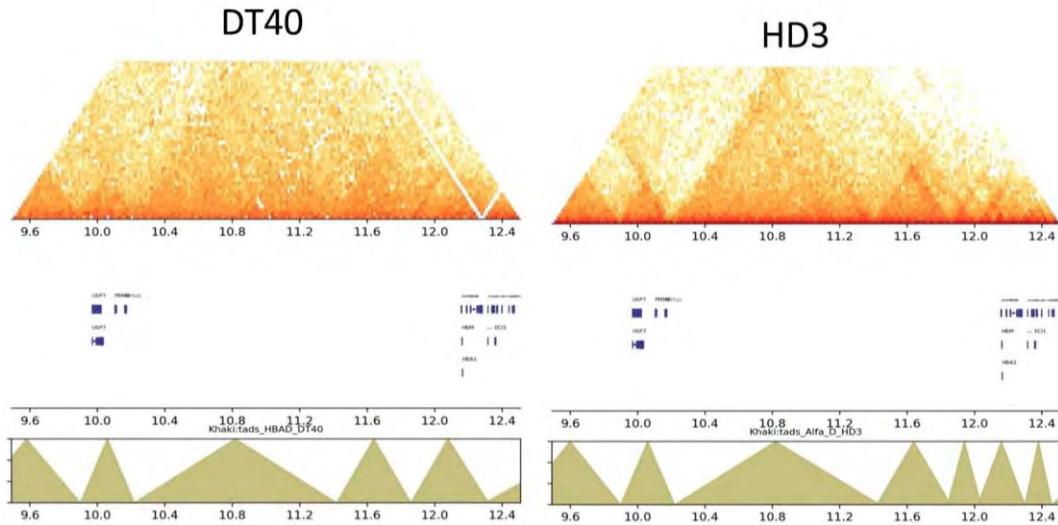
FIGURA 60 DIAGRAMA DE VENN MOSTRANDO LOS DOMINIOS TOPOLÓGICAMENTE ASOCIADOS QUE SE COMPARTEN ENTRE LAS CÉLULAS HD3 Y LAS DT40 A 50KB DE RESOLUCIÓN UTILIZANDO COMO VENTANA PARA COMPARAR LAS COORDENADAS DE LOS TADS -50KB +50KB.

Se observó que la proporción de TADs que se compartía entre los experimentos iba del 40 al 50% dependiendo del número original encontrado en cada tipo celular. Estos datos indican que, preliminarmente, las células HD3 y DT40 son más parecidas en su estructuración en comparación con la estructura de los eritrocitos observada

anteriormente. Sin embargo, cabe recalcar que, dado que el experimento no fue realizado con la misma resolución, es posible que esta proporción de dominios compartidos se altere si, en un futuro, se realiza un experimento bajo las mismas condiciones.

Para ejemplificar más claramente la cuestión de la resolución y observar puntualmente la formación de TADs en las células DT40, se construyeron mapas tomando dos zonas ejemplo anteriormente utilizadas en la comparación con los eritrocitos (Fig. 61).

Locus Alfa Globina D



Locus de Histonas

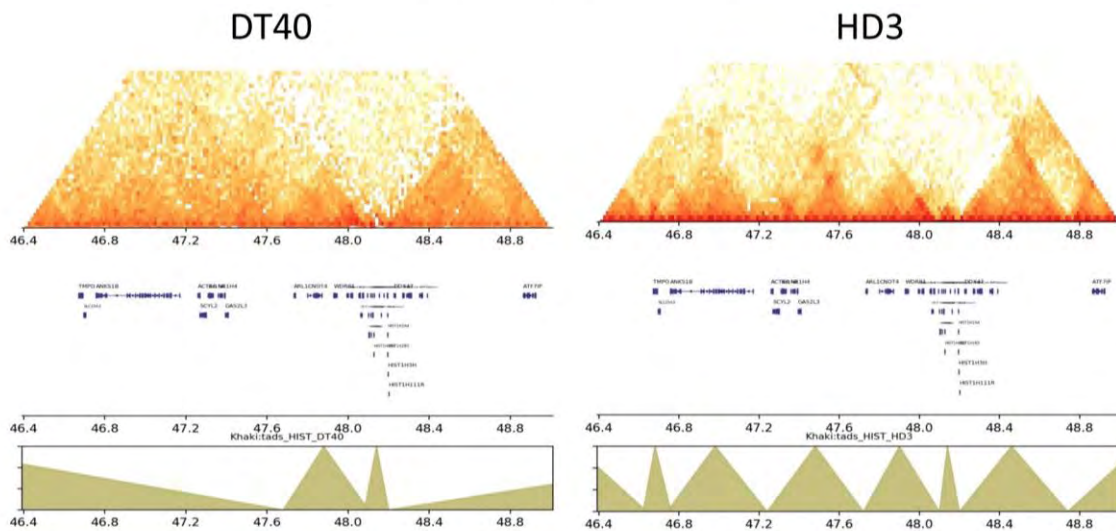


FIGURA 61 MATRICES DE INTERACCIÓN CONSTRUIDAS CON EL PROGRAMA HICPLOTTER DE DATOS DE HI-C. LOS DATOS DE LAS CÉLULAS DT40 FUERON TOMADOS DEL EXPERIMENTO REALIZADO POR GIBCUS *ET AL.* 2018 EN COMPARACIÓN CON LOS DATOS DE HI-C DE LAS CÉLULAS HD3. LOS MAPAS TIENEN 40KB DE RESOLUCIÓN. LOS TADS FUERON RECONOCIDOS CON EL PROGRAMA TADTOOL UTILIZANDO EL MISMO *CUTOFF* EN LAS DOS MUESTRAS. LA FILA SUPERIOR MUESTRA LA ZONA DONDE SE ENCUENTRAN LOS GENES ALFA GLOBINA UTILIZANDO 3MB DE EXTENSIÓN. LA FILA INFERIOR MUESTRA LA ZONA DONDE SE ENCUENTRAN LOS GENES DE LAS PROTEÍNAS HISTONAS. LOS GENES SE MAPEARON CON INFORMACIÓN OBTENIDA DEL BUSCADOR DE LA UNIVERSIDAD DE CALIFORNIA EN LA VERSIÓN DEL GENOMA GALGAL5.

Dentro de las zonas estudiadas se puede observar que existe una formación de TADs similar en las células DT40 en comparación con las HD3, sin embargo, también es claro que la claridad de los dominios es considerablemente menor en el Hi-C de DT40. Esto puede deberse a la menor resolución que brinda la enzima *HindIII* en comparación con *MboI* utilizada en el experimento de HD3. Sin embargo, principalmente en la zona de Alfa globina, se puede observar que los TADs son similares.

Estas observaciones apuntan a que posiblemente las células HD3 son similares a otras células somáticas que no poseen las características propias observadas en los eritrocitos y será de gran importancia obtener datos de tejidos distintos de este organismo para afinar las observaciones.

Discusión

El patrón preliminar observado en las matrices de interacción del primer experimento de Hi-C realizado en células eritroides de pollo muestra que no existe un cambio radical en la estructura entre las células eritroides adultas y las embrionarias, sin embargo, sí se observa una diferencia de estas dos en comparación a los eritroblastos de la línea celular HD3. Este patrón puede tener relación con la disminución observada en las cantidades ARN tanto en eritrocitos adultos como en embrionarios. En un estudio realizado en el 2002 midieron la cantidad de ARN presente en los eritrocitos embrionarios a lo largo del desarrollo del embrión y observaron que había un decremento en la cantidad encontrada desde el día 3 de desarrollo hasta el día 17 (13) y postularon que la razón para este fenómeno residía en que los NTPs funcionan como reguladores alostéricos de la hemoglobina. Aunque no hay datos aún de una medición realizada en eritrocitos adultos en la literatura, el patrón indica que la disminución es gradual y puede ser comparada tanto con la línea celular HD3 como con los fibroblastos, de los cuales se obtuvo una cantidad similar de ARN que ronda los 100 μg comparados con los eritrocitos de 10 días de desarrollo donde hay un quinto de la cantidad presente en las células antes mencionadas y posteriormente se puede observar

una disminución mucho más pronunciada en la cantidad encontrada en los eritrocitos adultos. Si se compara este patrón con lo observado anteriormente por Baumann *et al.* (8) durante el desarrollo y se consideran también las características del proceso de maduración eritroide donde se ha reportado una condensación generalizada de la cromatina acompañada de un silenciamiento transcripcional (14) se puede encontrar una explicación al patrón atípico de ARN observado para los eritrocitos adultos. En la observación de la similitud en la estructura entre las aRBC y las eRBC, una hipótesis podría consistir en que, si existe un re-arreglo topológico, éste se dé anteriormente en el desarrollo o la diferenciación eritroide ya que la tasa de transcripción en los eRBC parece estar ya disminuida.

Por otro lado, la observación del enriquecimiento en las interacciones de largo alcance observado en las figuras 40, 41 y 42 indican que en efecto existe un proceso que promueve que zonas del genoma que se encuentran más lejanas de se encuentren a distancias más cortas dentro del genoma de los eritrocitos. Esta característica podría relacionarse de forma directa con el proceso de condensación de la cromatina donde al encontrarse más compacta se promueven las interacciones entre regiones lejanas. Sin embargo, también se debe tomar en cuenta el panorama proteico para establecer las hipótesis sobre el incremento de interacciones lejanas en *Cis*. Dado que el análisis de ocupación de proteínas estructurales representa parte de las perspectivas que acompañan a este proyecto no se tiene información de cuáles de éstas podrían estar involucradas en el proceso, sin embargo, hay una observación en la literatura cuyo patrón es similar a lo que se está observando. En un estudio realizado en el 2017 por Haarhuis J. *et al.* observaron que la deficiencia del complejo SCC4, el cual se encarga de cargar la proteína cohesina sobre el ADN, genera un incremento en las interacciones lejanas en *Cis* así como una disminución del tamaño de las asas de cromatina a corto alcance(40). Ellos explican este fenómeno con la teoría de que la cohesina posee funciones variadas e importantes en el establecimiento de la estructura tridimensional de la cromatina, en primer lugar promueve la extensión de las asas de cromatina y por lo tanto la correcta formación de los TADs (40) esto se podría relacionar con la disminución observada en el tamaño de los TADs en los eritrocitos de

pollos. En segundo lugar mencionan que el incremento en la interacciones de largo alcance en *Cis* se debe a que las “cohesinas restringen la flexibilidad de los cromosomas” (40). Esta característica no se encuentra completamente descrita, sin embargo, corresponde bien con lo observado también dentro de los mapas de interacción de los eritrocitos.

Con estas observaciones se hace evidente la necesidad de caracterizar tanto la presencia como la ocupación de proteínas de interés dentro del núcleo celular de los modelos celulares analizados. No existen muchas evidencias de patrones similares a los que se observan en los mapas de cromosomas completos de las células eritroides de pollo, pero con el antecedente anterior se puede construir una lista de proteínas tanto estructurales como relacionadas con la transcripción que brinde mayor información acerca del proceso por el que atraviesa la cromatina de estas células.

En el caso más puntual de las observaciones de zonas específicas y retomando la hipótesis planteada con anterioridad en este proyecto, se puede considerar que se cumple la desorganización del genoma en el nivel estructural de dominios. Esto se afirma ya que en los eritrocitos tanto embrionarios como adultos, se observa un patrón atípico formando estructuras mucho menos apreciables a simple vista y aparentemente significativamente más pequeñas y numerosas que las observadas en las células HD3. Las células HD3 presentan una organización en Dominios Topológicamente Asociados similar a lo observado en otros tipos celulares eucariontes aunque el promedio de tamaño es más pequeño en comparación con lo reportado en Hi-C de ratón y de humano (entre 880 y 1Mb de tamaño(64)) esta característica podría deberse a la diferencia de tamaño que existe entre los genomas de estos organismos.

De estas observaciones se debe considerar que el programa utilizado TADtool ocupa un algoritmo basado en el aislamiento de interacciones que encuentra en una ventana de un tamaño brindando al programa (53), esta es una de las formas que existen para el reconocimiento de TADs dentro de un experimento de Hi-C, sin embargo no es la única. Será necesario corroborar la identificación de los dominios pequeños observados en los eritrocitos mediante el uso de otros programas y utilizando también la réplica biológica del

experimento para observar si el lugar donde se encuentran se mantiene similar y también relacionar estos dominios con la ocupación proteica de factores estructurales como CTCF o YY1.

Con referencia a los compartimentos observados las observaciones de la variación entre las células son similares a las estructurales. A grandes rasgos no hay diferencias notorias entre la estructuración en compartimentos de eRBC y aRBC, las diferencias más contundentes pueden encontrarse cuando se comparan con las células HD3. Esto podría ser esperado debido a la diferencia en el patrón transcripcional de estas células si los compartimentos reflejaran el estado transcripcional de la cromatina, sin embargo, en estos datos la estructuración no parece representar esa característica ya que en células con baja tasa transcripcional se esperaría una mayor proporción de zonas ocupando el compartimento B, como se planteó en la hipótesis (Fig. 58). No obstante, esto no ocurre así por lo que no se sabe precisamente qué representa la estructuración en zonas de gran tamaño que parecen interaccionar de forma muy delimitada entre ellas mientras se encuentran aisladas de otras regiones del genoma.

Una hipótesis podría relacionarse con la estructuración tridimensional que reflejan las dos células eritroides donde, de forma similar a la estructura característica del cromosoma X silenciado que forma dos grandes dominios fácilmente apreciables en los mapas de Hi-C (39), las líneas delimitadas de los mapas de Hi-C se relacionen con la organización necesaria para la para el correcto funcionamiento de estas células o relacionarse con otras características del genoma que no necesariamente hablen del estado transcripcional.

Estas observaciones dejan un panorama amplio de estudio para estos modelos celulares. Es necesario obtener mayor información acerca de las proteínas y del estado de la cromatina para poder relacionar las características estructurales con el proceso biológico por el que están atravesando estas células y así generar hipótesis acerca de la relación de la estructura tridimensional y el estado transcripcional de los modelos biológicos, así como describir el proceso de silenciamiento generalizado del genoma por el que atraviesan estas células. Así mismo será de vital importancia establecer un modelo celular distinto del linaje eritroide y

realizar los experimentos de Hi-C para tener dentro de las comparaciones un modelo que permita descartar si las observaciones descritas en este proyecto son características del linaje eritroide del pollo o relacionadas directamente con el genoma de la especie con la que se está trabajando.

Conclusiones

- Los niveles de transcripción de los genes candidato (eritroides, globinas y constitutivos) se encuentran disminuidos de forma significativa en los eritrocitos embrionarios y casi completamente silenciados en los eritrocitos adultos en comparación con la línea celular de eritroblastos HD3.
- La estructura tridimensional de los eritrocitos (eRBC y aRBC) muestra diferencias notorias comparada con la línea celular de eritroblastos HD3.
- Los mapas de interacciones con datos de Hi-C muestran un aumento generalizado en interacciones *CIS* de largo alcance en los eritrocitos (eRBC y aRBC) y aparentemente en interacciones en *TRANS* comparados con los mapas de la línea celular de eritroblastos HD3.
- El patrón de ARN presente en los eritrocitos adultos muestra especies pequeñas de ARN que podrían corresponder a degradación y una reducción notoria de ARN ribosomales.
- La estructuración en Dominios Topológicamente Asociados (TADs) parece estar disminuida en los eritrocitos ocupando una extensión reducida del genoma en comparación con la línea celular de eritroblastos HD3.
- La estructuración en compartimentos parece mantenerse en los tres tipos celulares en proporciones similares, sin embargo, se requieren más análisis para esclarecer a qué corresponde esta organización.
- La estructura tridimensional particular del genoma y la reducción de la transcripción podrían estar relacionados con el papel biológico específico que desempeñan estas células en la fisiología de *Gallus gallus* y abren la posibilidad a variadas preguntas

acerca de la regulación transcripcional y topología en un proceso de silenciamiento de la expresión genética.

Perspectivas

Dadas las observaciones acerca de la estructuración en pequeños dominios y compartimentos bien delimitados observados en los modelos celulares eRBC y aRBC, además de la aparente disminución transcripcional generalizada, las perspectivas de este proyecto incluyen la realización de experimentos de ensayo de accesibilidad a la transposasa (ATAC-seq por sus siglas en inglés) que permitirán mapear las zonas del genoma que se encuentran accesibles así como tener un panorama completo de la ocupación proteica presente en la cromatina en cada uno de los modelos. Estos experimentos están siendo realizados en un proyecto paralelo por el Biol. Andrés Penagos Puig bajo la tutoría de la Dra. Mayra Furlán por lo que podrán complementar las observaciones del proyecto.

Así mismo, como se ha discutido con anterioridad, el análisis por microscopía de proteínas como CTCF, Polimerasa de ARN y YY1, junto con algunas otras candidatas podrá brindar información para entender de forma más global el proceso biológico que ocurre dentro del núcleo de estas células.

También es de vital importancia corroborar la información observada con réplicas biológicas de estas células, así como con un modelo celular no perteneciente al linaje eritroide del pollo para realizar las comparaciones pertinentes.

Por último resultará muy informativo poseer el patrón transcripcional global de estas células ya que los genes elegidos en este proyecto representan sólo a algunos grupos característicos de ciertas funciones celulares, sin embargo, podrían existir ARNs que se encuentren siendo expresados y se relacionen con los procesos específicos de estas células y también de la especie particular del pollo que brinden información de nuevos procesos

biológicos relacionados con el silenciamiento transcripcional y la estructuración de la cromatina.

Referencias

1. Dzierzak E, Philipsen S. Erythropoiesis: Development and Differentiation. *Cold Spring Harb Perspect Med* [Internet]. abril de 2013 [citado el 23 de noviembre de 2016];3(4). Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3684002/>
2. Chen K, Liu J, Heck S, Chasis JA, An X, Mohandas N. Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. *Proc Natl Acad Sci*. el 13 de octubre de 2009;106(41):17413–8.
3. Glomski CA, Pica A. *The Avian Erythrocyte: Its Phylogenetic Odyssey*. CRC Press; 2016. 644 p.
4. An X, Schulz VP, Li J, Wu K, Liu J, Xue F, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. el 29 de mayo de 2014;123(22):3466–77.
5. Sheng Guojun. Primitive and definitive erythropoiesis in the yolk sac: a bird's eye view. *Int J Dev Biol*. 2010;54:1033–43.
6. Mechanism of developmental regulation of alpha pi, the chicken embryonic alpha-globin gene. [Internet]. [citado el 2 de diciembre de 2018]. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC360089/>
7. Geerts Willie J.C., Lambers Wouter H., Moorman Antoon F. M. Differences in erythropoiesis in normal chicken and quail embryos. *Histochemical Journal*. 1993;25:280–90.
8. Bergman Mathias, Ringertz Nils. Gene expression pattern of chicken erythrocyte nuclei in heterokaryons. *Journal of Cell Science*. 1990;97:167–75.
9. Pica A. *The Avian Erythrocyte: Its Phylogenetic Odyssey* [Internet]. Science Publishers; 2011 [citado el 23 de noviembre de 2016]. Disponible en: <http://www.crcnetbase.com/doi/book/10.1201/b11052>
10. Ji P, Murata-Hori M, Lodish HF. Formation of mammalian erythrocytes: Chromatin condensation and enucleation. *Trends Cell Biol*. julio de 2011;21(7):409–15.
11. Migliaccio AR. Erythroblast enucleation. *Haematologica*. diciembre de 2010;95(12):1985–8.
12. Wiersma PA, Cox GS. Synthesis of messenger-like RNA in avian erythrocyte nuclei. *Arch Biochem Biophys*. octubre de 1985;242(1):90–103.
13. Baumann R, Gotz R, Dragon S. NTP pattern of avian embryonic red cells: role of RNA degradation and AMP deaminase/5'-nucleotidase activity. *Am J Physiol Regul Integr Comp Physiol*. marzo de 2003;284(3):R771-779.

14. Ney PA. Normal and disordered reticulocyte maturation. *Curr Opin Hematol*. mayo de 2011;18(3):152–7.
15. Popova EY, Krauss SW, Short SA, Lee G, Villalobos J, Etzell J, et al. Chromatin condensation in terminally differentiating mouse erythroblasts does not involve special architectural proteins but depends on histone deacetylation. *Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol*. 2009;17(1):47–64.
16. Kowalski A, Pałyga J. Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non-histone protein MENT. *Chromosome Res Int J Mol Supramol Evol Asp Chromosome Biol*. julio de 2011;19(5):579–90.
17. Axelsson Erik, Webster Matthew T., Smith Nick G. C., Burt David W., Ellegren Hans. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Research*. 2005;15.
18. Burt D. W. Origin and evolution of avian microchromosomes. *Cytogenet Genome Res*. 2002;96:97–112.
19. Dekker J., Marti-Renom M. A, Mirny L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature reviews*. 2013;14:390–403.
20. Cremer Thomas, Cremer Marion. Chromosome Territories. *Cold Spring Harb Perspect Biol*. 2010;2:1–22.
21. T. Cremer, C. Cremer. Chromosome Territories, Nuclear Architecture and Gene Regulation in Mammalian Cells. *Nature reviews*. 2001;2:293–301.
22. Narasimharao V. Marella, Brandon Seifert, Priyadharsini Nagarajan, Satrajit Sinha, Ronald Berezney. Chromosomal Rearrangements During Human Epidermal Keratinocyte Differentiation. *J Cell Physiol*. 2009;221:139–46.
23. Dekker Job, Rippe Karsten, Dekker Martijn, Kleckner Nancy. Capturing Chromosome Conformation. *Science*. 2002;295:1306–11.
24. Dekker Job, Misteli Tom. Long-Range Chromatin Interactions. *Cold Spring Harb Perspect Biol*. 2015;7.
25. de Laat Wouter, Dekker Job. 3C-based Technologies To Study The Shape Of The Genome. *Methods*. 2012;58(3):189–91.
26. Gavrillov AA, Razin SV. Spatial configuration of the chicken alpha-globin gene domain: immature and active chromatin hubs. *Nucleic Acids Res*. agosto de 2008;36(14):4629–40.
27. Denker Annette, de Laat Wouter. The second decade of 3C technologies detailed insights into nuclear organization. *Genes and Development*. 2016;30:1357–82.

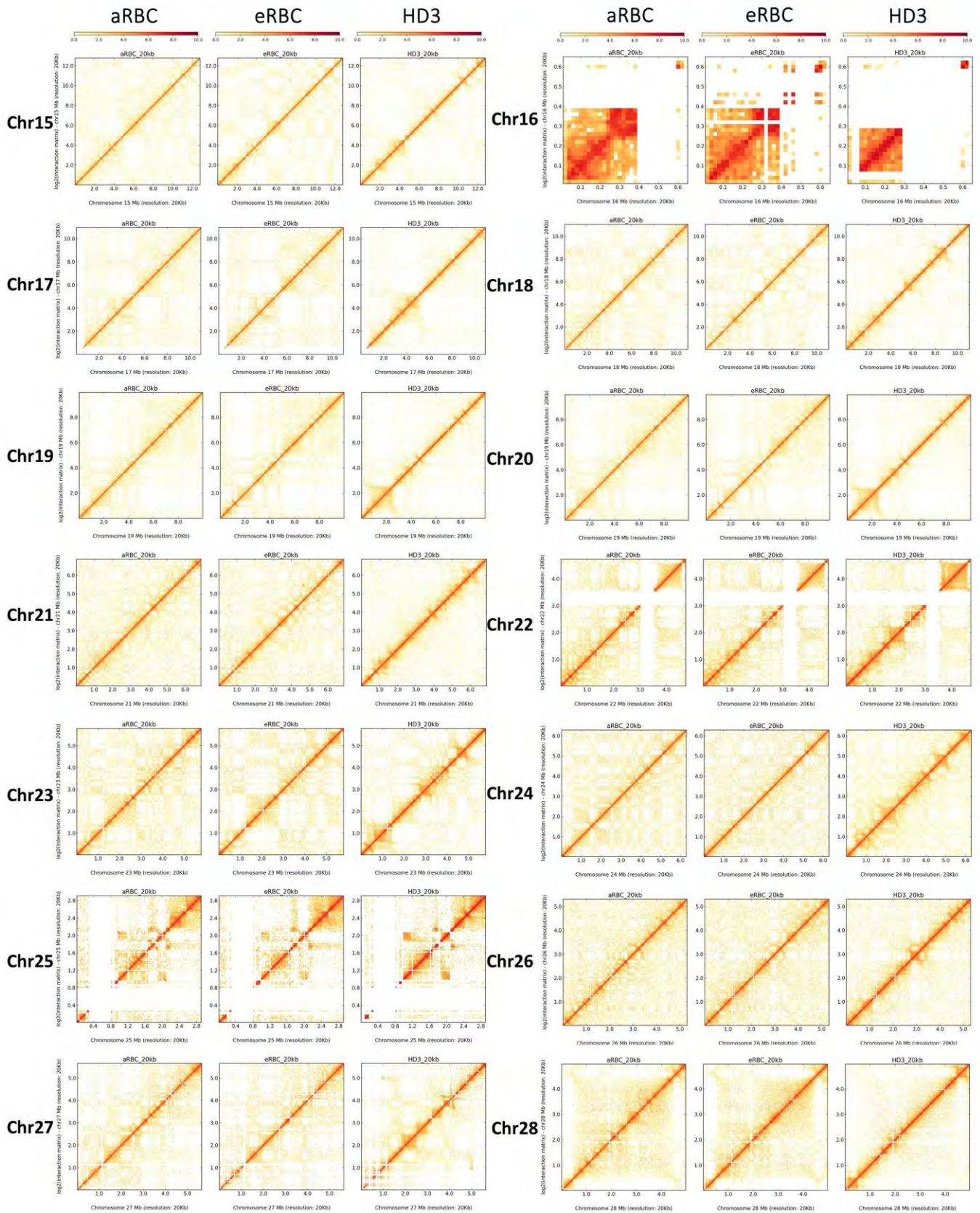
28. Lieberman-Aiden Erez, van Berkum Nynke L., Williams Louise, Ragozy Tobias, Telling Agnes, Amit Ido, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*. 2009;326:289–93.
29. Dixon Jesse R., Gorkin David U., Ren Bing. Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell*. 2016;62:668–680.
30. Fraser Peter, Bickmore Wendy. Nuclear organization of the genome and the potential for gene regulation. *Nature*. 2007;447:413–7.
31. Gonzalez-Sandoval Adriana, Gasser Susan M. On TADs and LADs: Spatial Control Over Gene Expression. *Trends in Genetics*. 2016;32(8).
32. Chung Ill-Min, Ketharnathan Sarada, Kim Seung-Hyun, Thiruvengadam Muthu, Rani Mari Kavitha, Rajakumar Govindasamy. Making Sense of the Tangle: Insights into Chromatin Folding and Gene Regulation. *Genes*. 2016;7(71):1–11.
33. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. el 19 de febrero de 2015;518(7539):331–6.
34. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from higher-order genomic compartmentalization. *bioRxiv*. el 9 de enero de 2017;095802.
35. Dekker Job, Heard Edith. Structural and functional diversity of Topologically Associating Domains. *FEBS Letters*. 2015;589:2877–84.
36. Furlan-Magaril Mayra, Várnai Csilla, Nagano Takashi, Fraser Peter. 3D genome architecture from populations to single cells. *Current Opinion in Genetics and Development*. 2015;31:36–41.
37. Wang Siyuan, Su Jun-Han, Beliveau Brian J., Bintu Bogdan, Moffitt Jeffrey R., Wu Chao-ting, et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science*. 2016;353:598–602.
38. Dixon Jesse R., Selvaraj Siddarth, Yue Feng, Kim Audrey, Li Yan, Shen Yin, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485:376–80.
39. Giorgetti Luca, Lajoie Bryan R., Carter Ava C., Attia Mikael, Zhan Ye, Xu Jin, et al. Structural organization of the inactive X chromosome in the mouse. *Nature*. 2016;535:575–9.
40. Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell*. el 4 de mayo de 2017;169(4):693-707.e14.
41. Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell*. el 26 de junio de 2009;137(7):1194–211.

42. Haering CH, Farcas A-M, Arumugam P, Metson J, Nasmyth K. The cohesin ring concatenates sister DNA molecules. *Nature*. el 17 de julio de 2008;454(7202):297–301.
43. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell*. el 14 de diciembre de 2017;171(7):1573–1588.e28.
44. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. el 21 de mayo de 2015;161(5):1012–25.
45. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*. el 27 de febrero de 2015;347(6225):1017–21.
46. Darrow Emily M., Huntley Miriam H., Dudchenko Olga, Stamenova Elena K., Durand Neva C., Sun Zhuo, et al. Deletion of DXZ4 on the human inactive X chromosome.pdf. 2016;31. Disponible en: www.pnas.org/cgi/doi/10.1073/pnas.1609643113
47. Kantidze Omar L., Iarovaia Olga V., Philonenko Elena S., Yakutenko Irina I., Razin Sergey V. Unusual compartmentalization of CTCF and other transcription factors in the course of terminal erythroid differentiation. *BBA*. 2007;1773:924–33.
48. Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, et al. HiCUP: pipeline for mapping and processing Hi-C data. *F1000Research*. 2015;4:1310.
49. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. el 1 de diciembre de 2015;16:259.
50. Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*. julio de 2016;3(1):95–8.
51. Akdemir KC, Chin L. HiCPlotter integrates genomic data with interaction matrices. *Genome Biol*. el 21 de septiembre de 2015;16:198.
52. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, et al. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst*. julio de 2016;3(1):99–101.
53. Kruse K, Hug CB, Hernández-Rodríguez B, Vaquerizas JM. TADtool: visual parameter identification for TAD-calling algorithms. *Bioinforma Oxf Engl*. 15 de 2016;32(20):3190–2.
54. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods*. octubre de 2012;9(10):999–1003.

55. Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*. el 9 de julio de 2015;523(7559):240–4.
56. Philipsen S, Hardison RC. Evolution of hemoglobin loci and their regulatory elements. *Blood Cells Mol Dis*. mayo de 2018;70:2–12.
57. Seki R, Li C, Fang Q, Hayashi S, Egawa S, Hu J, et al. Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. *Nat Commun*. el 6 de febrero de 2017;8:14229.
58. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. el 1 de diciembre de 2015;16:259.
59. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods*. octubre de 2012;9(10):999–1003.
60. Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*. julio de 2016;3(1):95–8.
61. Ulianov SV, Galitsyna AA, Flyamer IM, Golov AK, Khrameeva EE, Imakaev MV, et al. Activation of the alpha-globin gene expression correlates with dramatic upregulation of nearby non-globin genes and changes in local and large-scale chromatin spatial structure. *Epigenetics Chromatin*. 11 de 2017;10(1):35.
62. Bonev B, Cavalli G. Organization and function of the 3D genome. *Nat Rev Genet*. 2016;17(12):772.
63. Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, et al. A pathway for mitotic chromosome formation. *Science*. 09 de 2018;359(6376).
64. Gonzalez-Sandoval A, Gasser SM. On TADs and LADs: Spatial Control Over Gene Expression. *Trends Genet TIG*. 2016;32(8):485–95.

Anexos

20kb de resolución



20kb de resolución

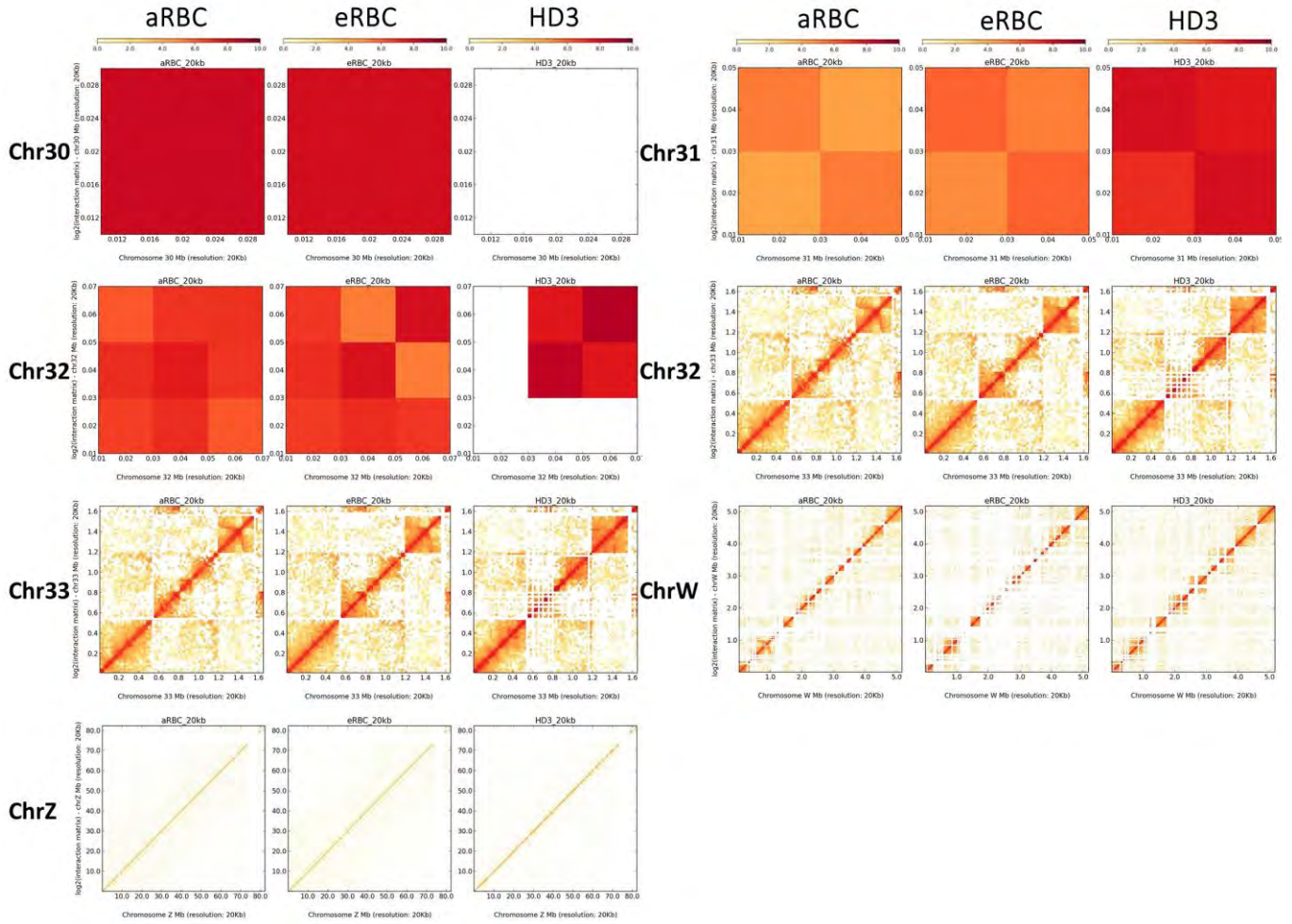
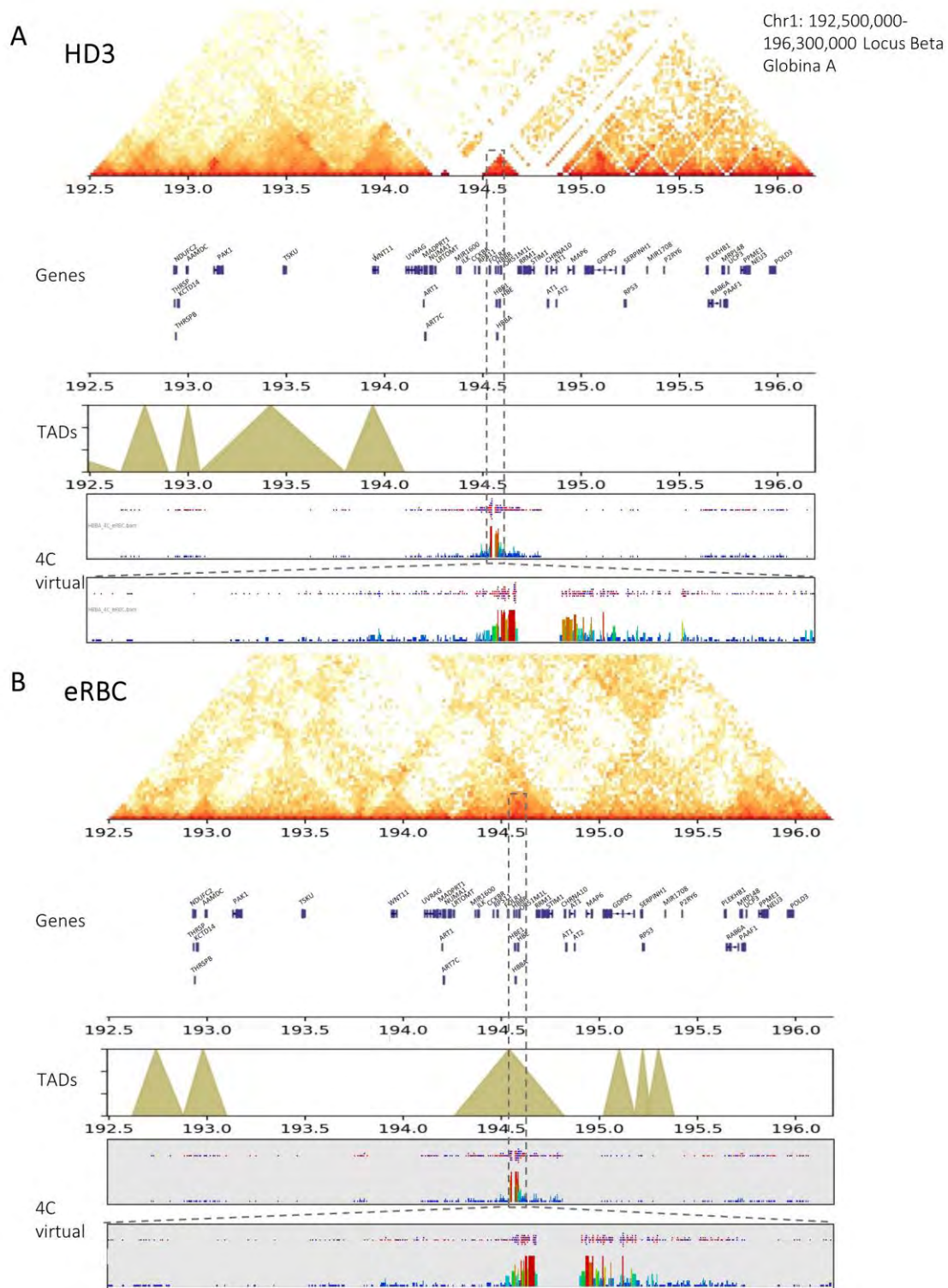


FIGURA S 1 MATRICES DE INTERACCIONES POR CROMOSOMA DE EXPERIMENTO DE HIC EN CÉLULAS ERITROIDES DE POLLO (ARBC, ERBC Y HD3) CONSTRUIDAS CON EL PROGRAMA HIC-PRO. 20KB DE RESOLUCIÓN. EL CROMOSOMA 29 NO SE ENCUENTRA ANOTADO EN EL GENOMA DE GALGAL5. LOS CROMOSOMAS 30, 31 Y 32 MUESTRAN POCA INFORMACIÓN DEBIDO A QUE SON MUY CORTOS PARA PODER CONSTRUIR UN MAPA CON ESTA RESOLUCIÓN.



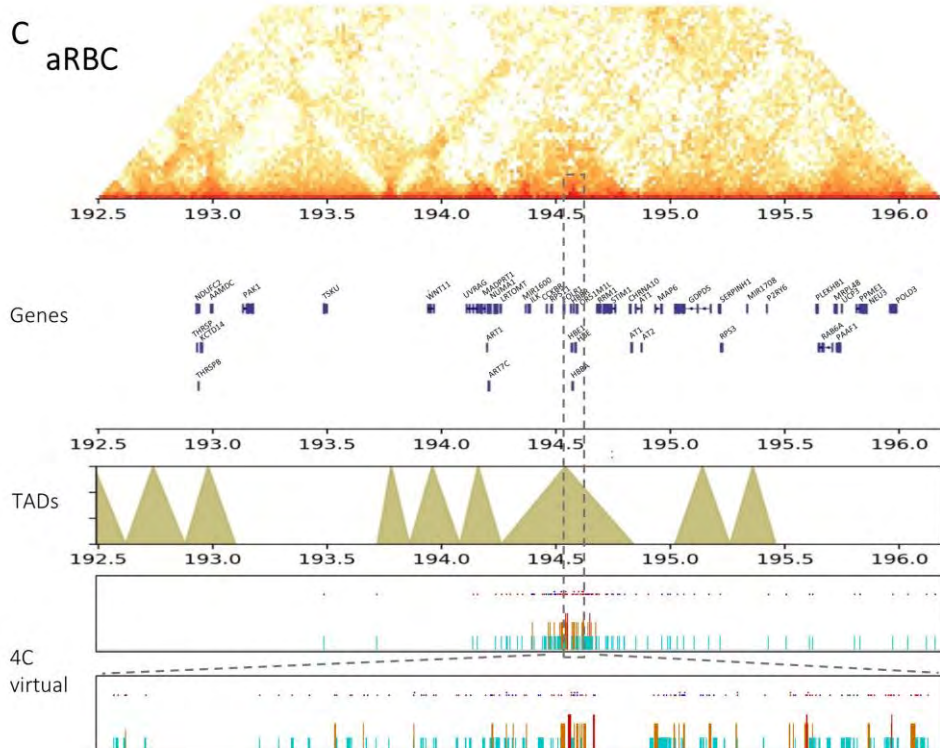
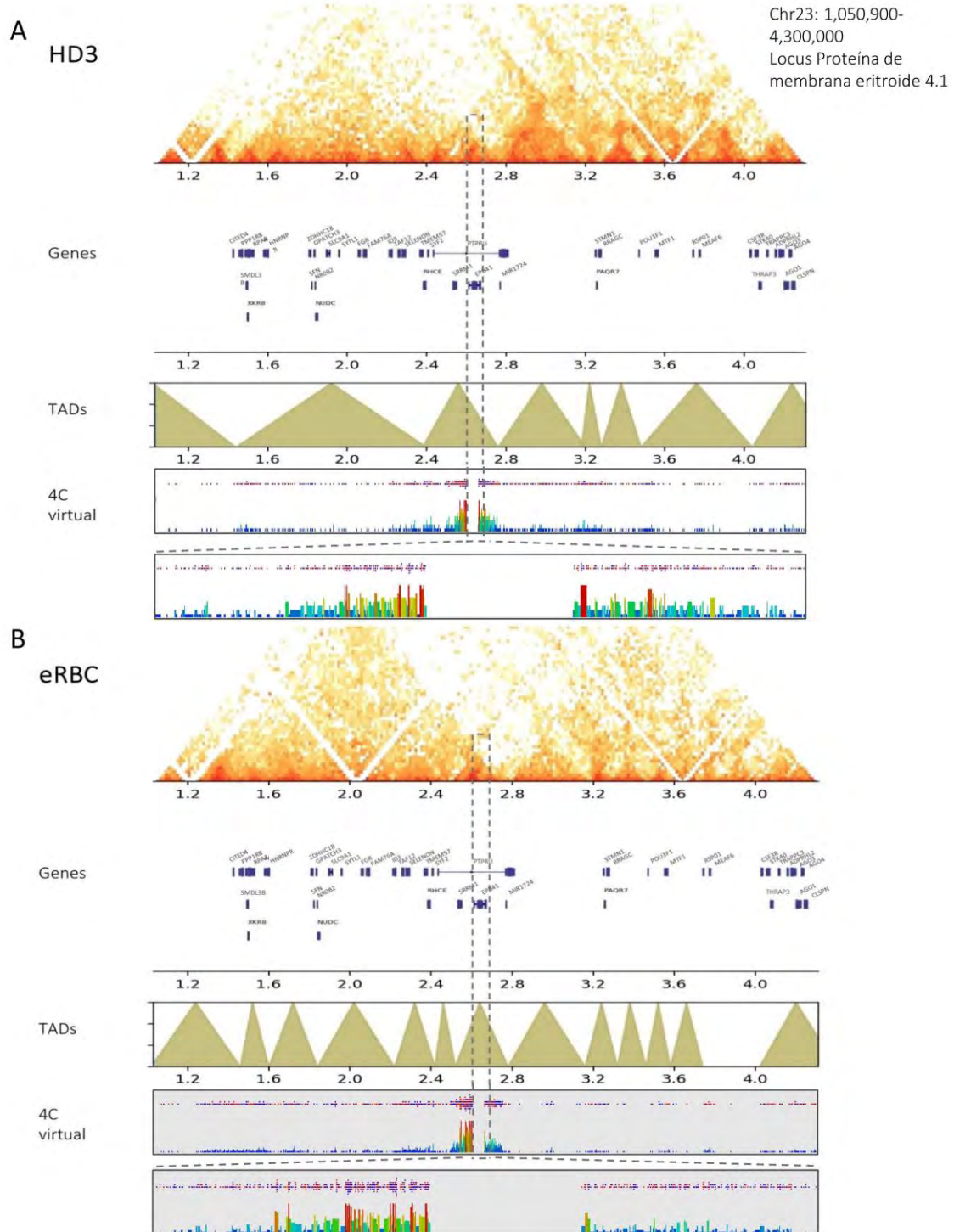


FIGURA S 2 DIAGRAMAS DEL LOCUS BETA GLOBINA A ANALIZADO DE LOS DATOS OBTENIDOS DE HI-C EN LOS TRES TIPOS CELULARES. EN EL PRIMER PANEL SE MUESTRA LA MATRIZ DE INTERACCIÓN CONSTRUIDA CON AYUDA DE HICPLOTTER. EN EL SEGUNDO PANEL SE MUESTRAN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA OBTENIDAS DE LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA (UCSC). EN EL TERCER PANEL SE MUESTRAN LAS COORDENADAS DE LOS TADs OBTENIDOS MEDIANTE EL ANÁLISIS CON TADTOOL. EN EL CUARTO PANEL SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL REALIZADO CON EL PROGRAMA SEQMONK EN DONDE LAS LÍNEAS PUNTEADAS MUESTRAN LA ZONA UTILIZADA COMO ANCLA Y LAS BARRAS MUESTRAN LAS INTERACCIONES ENCONTRADAS CON ESA ZONA. EN EL CUARTO PANEL SE MUESTRA UN ACERCAMIENTO DE LA ZONA A ESTUDIAR REPRESENTADO POR LAS LÍNEAS PUNTEADAS HORIZONTALES. A) DIAGRAMAS DE HD3. B) DIAGRAMAS DE ERBC. C) DIAGRAMAS DE ARBC.



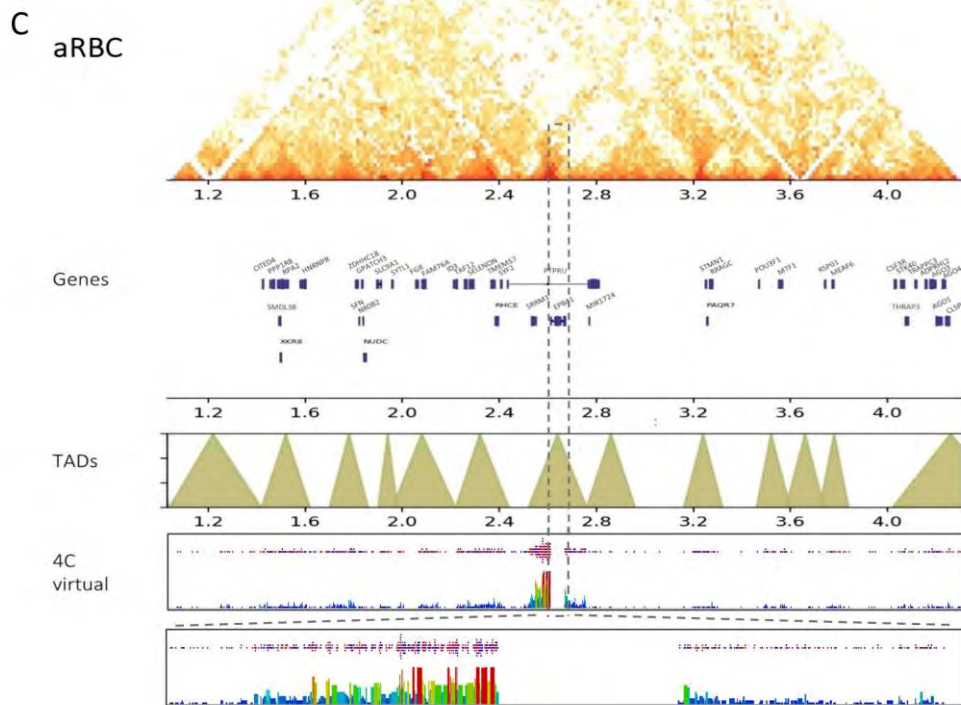


FIGURA S 3 DIAGRAMAS DEL LOCUS DE EPB41 (PROTEÍNA DE MEMBRANA ERITROIDE 4.1) ANALIZADO DE LOS DATOS OBTENIDOS DE HI-C EN LOS TRES TIPOS CELULARES. EN EL PRIMER PANEL SE MUESTRA LA MATRIZ DE INTERACCIÓN CONSTRUIDA CON AYUDA DE HICPLOTTER. EN EL SEGUNDO PANEL SE MUESTRAN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA OBTENIDAS DE LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA (UCSC). EN EL TERCER PANEL SE MUESTRAN LAS COORDENADAS DE LOS TADs OBTENIDOS MEDIANTE EL ANÁLISIS CON TADTOOL. EN EL CUARTO PANEL SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL REALIZADO CON EL PROGRAMA SEQMONK EN DONDE LAS LÍNEAS PUNTEADAS MUESTRAN LA ZONA UTILIZADA COMO ANCLA Y LAS BARRAS MUESTRAN LAS INTERACCIONES ENCONTRADAS CON ESA ZONA. EN EL CUARTO PANEL SE MUESTRA UN ACERCAMIENTO DE LA ZONA A ESTUDIAR REPRESENTADO POR LAS LÍNEAS PUNTEADAS HORIZONTALES. A) DIAGRAMAS DE HD3. B) DIAGRAMAS DE ERBC. C) DIAGRAMAS DE ARBC.

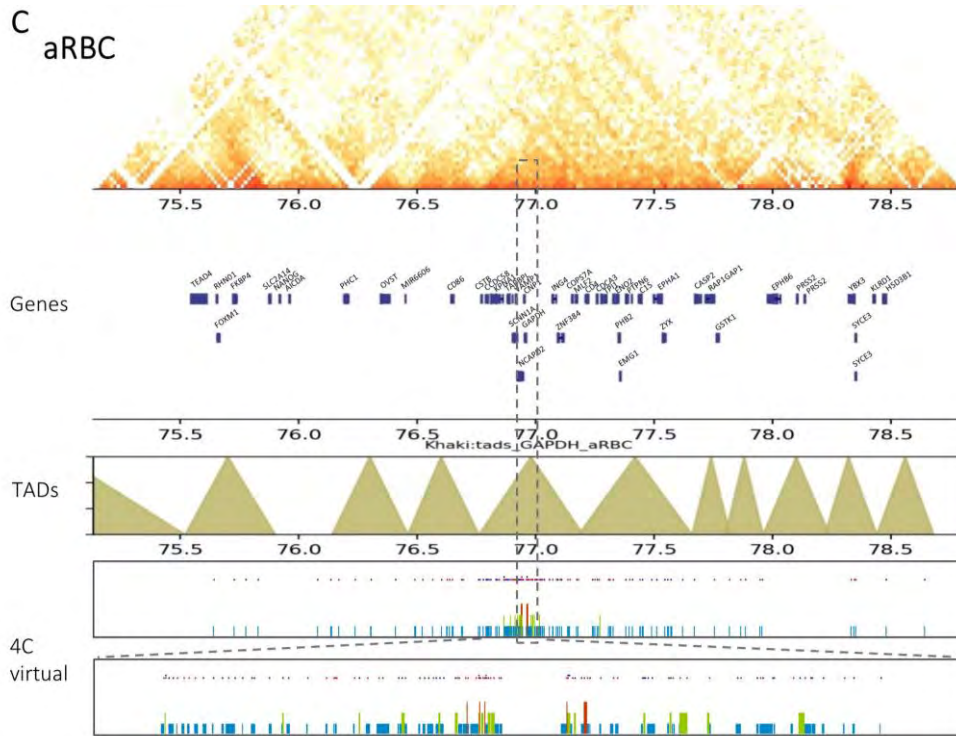
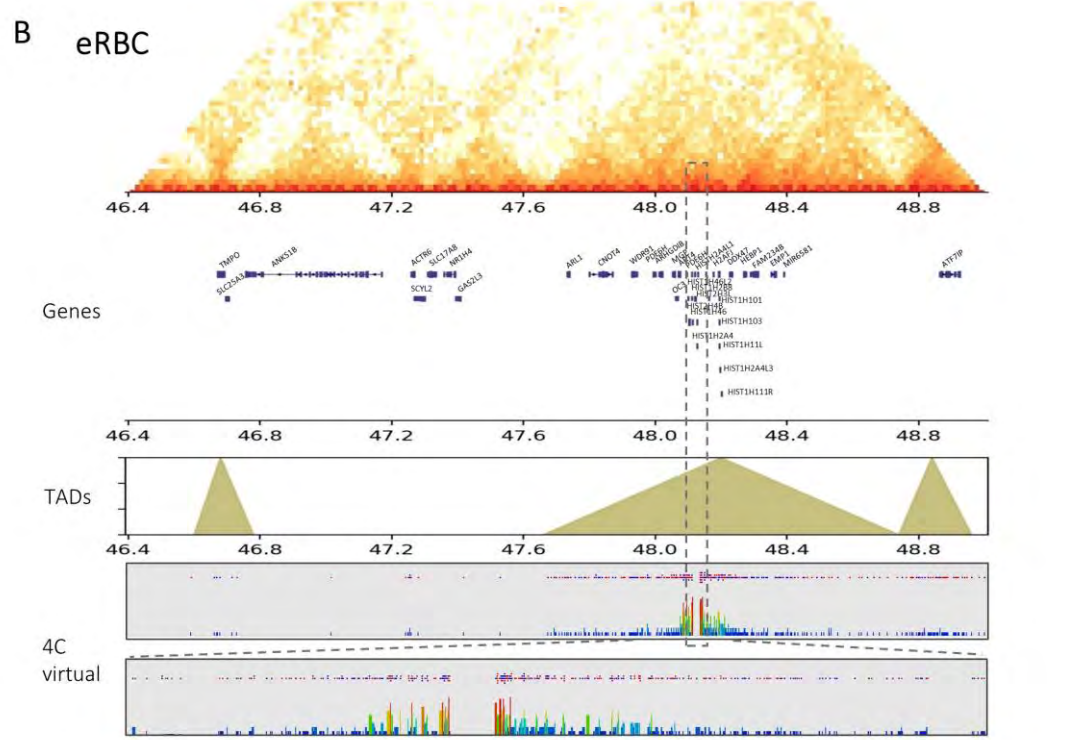
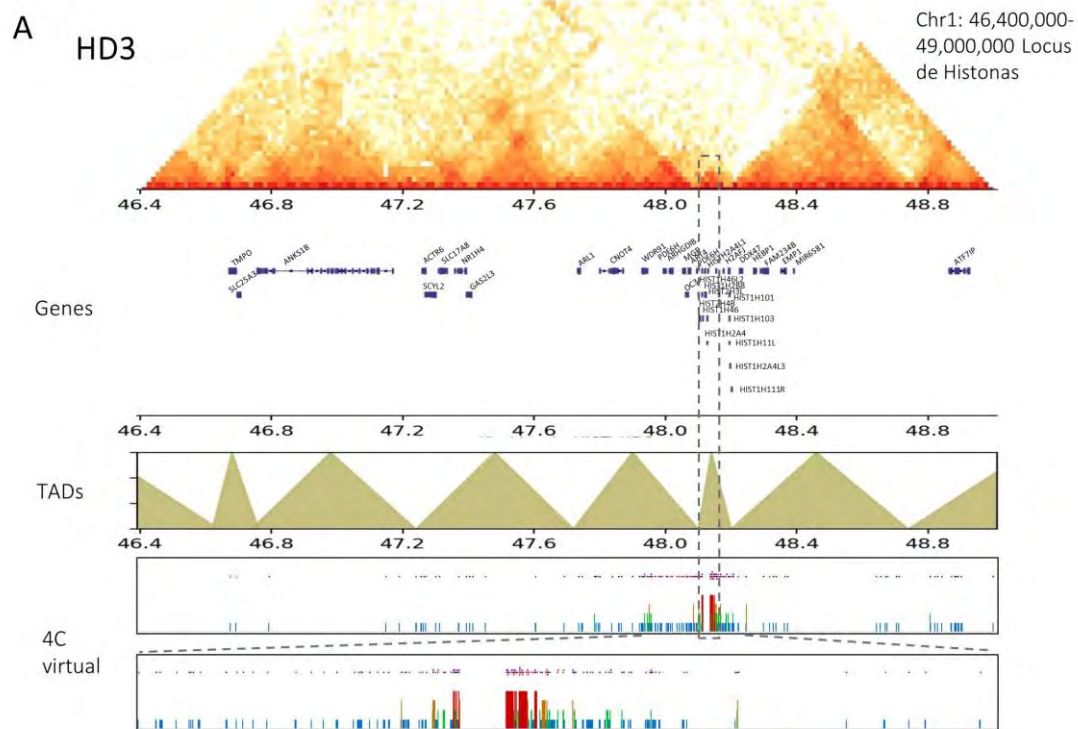


FIGURA S 4 DIAGRAMAS DEL LOCUS DE GAPDH ANALIZADO DE LOS DATOS OBTENIDOS DE HI-C EN LOS TRES TIPOS CELULARES. EN EL PRIMER PANEL SE MUESTRA LA MATRIZ DE INTERACCIÓN CONSTRUIDA CON AYUDA DE HICPLOTTER. EN EL SEGUNDO PANEL SE MUESTRAN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA OBTENIDAS DE LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA (UCSC). EN EL TERCER PANEL SE MUESTRAN LAS COORDENADAS DE LOS TADS OBTENIDOS MEDIANTE EL ANÁLISIS CON TADTOOL. EN EL CUARTO PANEL SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL REALIZADO CON EL PROGRAMA SEQMONK EN DONDE LAS LÍNEAS PUNTEADAS MUESTRAN LA ZONA UTILIZADA COMO ANCLA Y LAS BARRAS MUESTRAN LAS INTERACCIONES ENCONTRADAS CON ESA ZONA. EN EL CUARTO PANEL SE MUESTRA UN ACERCAMIENTO DE LA ZONA A ESTUDIAR REPRESENTADO POR LAS LÍNEAS PUNTEADAS HORIZONTALES. A) DIAGRAMAS DE HD3. B) DIAGRAMAS DE eRBC. C) DIAGRAMAS DE aRBC.



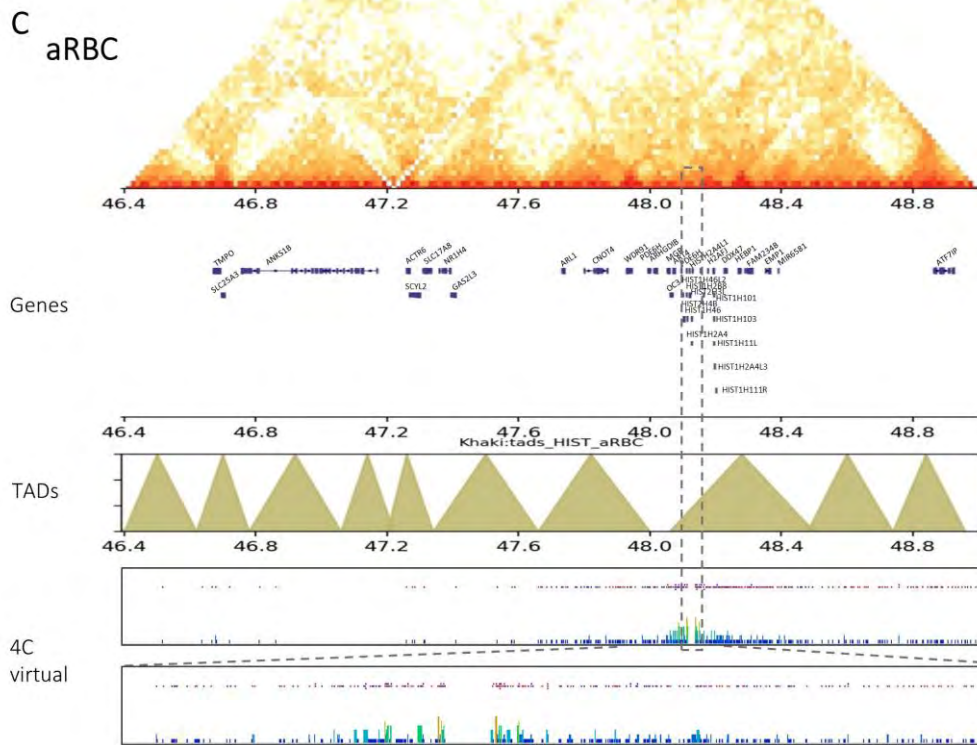


FIGURA S 5 DIAGRAMAS DEL LOCUS DE PROTEÍNAS HISTONAS ANALIZADO DE LOS DATOS OBTENIDOS DE HI-C EN LOS TRES TIPOS CELULARES. EN EL PRIMER PANEL SE MUESTRA LA MATRIZ DE INTERACCIÓN CONSTRUIDA CON AYUDA DE HICPLOTTER. EN EL SEGUNDO PANEL SE MUESTRAN LAS COORDENADAS DE LOS GENES PRESENTES EN LA ZONA OBTENIDAS DE LA BASE DE DATOS DE LA UNIVERSIDAD DE CALIFORNIA (UCSC). EN EL TERCER PANEL SE MUESTRAN LAS COORDENADAS DE LOS TADs OBTENIDOS MEDIANTE EL ANÁLISIS CON TADTOOL. EN EL CUARTO PANEL SE MUESTRA UN ANÁLISIS DE 4C VIRTUAL REALIZADO CON EL PROGRAMA SEQMONK EN DONDE LAS LÍNEAS PUNTEADAS MUESTRAN LA ZONA UTILIZADA COMO ANCLA Y LAS BARRAS MUESTRAN LAS INTERACCIONES ENCONTRADAS CON ESA ZONA. EN EL CUARTO PANEL SE MUESTRA UN ACERCAMIENTO DE LA ZONA A ESTUDIAR REPRESENTADO POR LAS LÍNEAS PUNTEADAS HORIZONTALES. A) DIAGRAMAS DE HD3. B) DIAGRAMAS DE ERBC. C) DIAGRAMAS DE ARBC.