



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE MEDICINA

**Estudio de la participación de YY1 en la
estructuración del genoma durante la
diferenciación y maduración eritroide
en *Gallus gallus***

TESIS

Que para obtener el título de

Licenciado en Investigación Biomédica Básica

P R E S E N T A:

Juan Manuel Tzompantzi De Ita

DIRECTORA DE TESIS:

Dra. Mayra Furlan Magaril

Facultad de Medicina



Ciudad Universitaria, Cd. Mx., 2024



Universidad Nacional
Autónoma de México



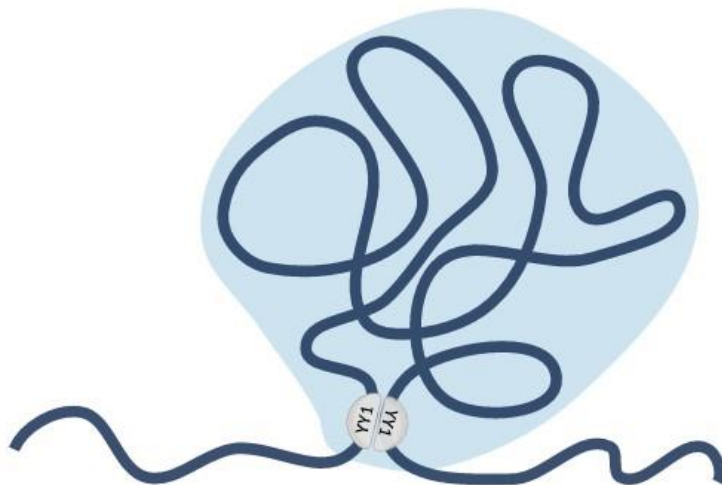
UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Estudio de la participación de YY1 en la estructuración del genoma durante la diferenciación y maduración eritroide en *Gallus gallus*



Juan Manuel Tzompantzi De Ita

*“From sprinkler splashes to fireplace ashes [...],
You've got no reason to be afraid.”*

Agradecimientos académicos

La presente tesis para obtener el grado de licenciado en Investigación Biomédica Básica se realizó bajo la dirección de la Dra. Mayra Furlan Magaril, en el laboratorio 202 Sur del Departamento de Genética Molecular de la División de Investigación Básica del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México (UNAM).

Se reconoce el asesoramiento y retroalimentación recibido en el presente trabajo por parte del Jurado de Examen de Grado constituido por:

| | |
|------------------------------------------------------|-----------------------------------------------|
| Presidenta: Dra. Erika Patricia Rendón Huerta | Facultad de Medicina, UNAM |
| Secretaria: Dra. Mayra Furlan Magaril | Instituto de Fisiología Celular, UNAM |
| Vocal: Dra. Rocío Ángeles García Becerra | Instituto de Investigaciones Biomédicas, UNAM |
| Suplente: Dr. Víctor Julián Valdés Rodríguez | Instituto de Fisiología Celular, UNAM |
| Suplente: Dra. Sylvia Patricia Garza Manero | Instituto de Fisiología Celular, UNAM |

El presente trabajo se realizó con financiamiento y apoyo del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, de la Dirección General de Asuntos del Personal Académico de la UNAM (PAPIIT; no. IN210323). Además del financiamiento otorgado por el Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT; no. 303068, no. 137721 y no. 15758).

Se reconoce a la Dra. Rosario Pérez Molina, técnica académica del laboratorio, por su ayuda y asesoramiento en la planeación y ejecución de experimentos, su organización logística para la secuenciación de bibliotecas genómicas y su apoyo administrativo.

Se reconoce a la M. en C. Giovana Sherlyn Claudio Galeana por su enseñanza, apoyo en la planeación y asesoramiento a lo largo de todos los experimentos realizados, así como en los análisis bioinformáticos y discusión de resultados.

Se reconoce el apoyo brindado por el Dr. Andrés Penagos Puig por su asesoramiento en los experimentos y discusión de resultados.

Se reconoce el apoyo de la Dra. Ruth Rincón Heredia y el personal de la Unidad de Imagenología del Instituto de Fisiología Celular por su enseñanza en el uso del microscopio confocal y asistencia en los experimentos de microscopia.

Se reconoce al personal académico de la Unidad de Biología Molecular del Instituto de Fisiología Celular de la UNAM. Agradezco a la Dra. Laura Ongay Larios, a la Biól. Guadalupe Códiz Huerta y a la M. en C. Minerva Mora Cabrera por su asesoramiento en uso de los equipo de qPCR de tiempo real, y por su apoyo en la síntesis de oligonucleótidos y asistencia en la electroforesis capilar de DNA de las bibliotecas genómicas.

Se reconoce al Dr. Augusto César Poot Hernández y al personal de la unidad de Bioinformática del Instituto de Fisiología Celular de la UNAM por su asesoramiento en el manejo, almacenamiento y análisis de datos genómicos generados.

Se reconoce el apoyo brindado por el personal de la Unidad del Bioterio del Instituto de Fisiología Celular de la UNAM, especialmente al M.V.Z Héctor Malagón Rivero, en el manejo de los animales empleados en este estudio.

Agradezco el apoyo brindado por parte del laboratorio 122 Norte, a cargo del Dr. Félix Recillas Targa, del Instituto de Fisiología Celular de la UNAM. Especialmente, se reconoce el apoyo por parte de la técnica académica, la Biól. Georgina Margarita Guerrero Avendaño, por su asesoramiento y enseñanza en el uso de equipos de qPCR tiempo real y sonicador.

Agradecimientos personales

A mi tutora, la Dra. Mayra Furlan Magaril, por sus valiosos consejos y guía a lo largo del final de mi carrera universitaria. Por brindarme la oportunidad de unirme a su laboratorio y darme la confianza para realizar los experimentos más desafiantes hasta ahora en mi trayectoria académica. Por siempre ayudarme a ver las enseñanzas en los resultados negativos y orientarme en la búsqueda de alternativas en momentos cuando yo dejaba de verlas.

A la Dra. Rosario (Charis), por siempre auxiliarme y escucharme en cualquier duda durante mi estancia en el laboratorio. Por su paciencia al enseñarme todos los procedimientos del laboratorio. También, por su calurosa compañía, y siempre hacerme sentir incluido y apreciado en el equipo de trabajo.

A mi “sensei”, Sherlyn, por su invaluable tiempo y apoyo en el laboratorio. Por siempre escuchar mis inquietudes, discutir ideas y ayudarme a simplificar problemas complejos. Agradezco por la gran amistad forjada a lo largo de este proyecto y por siempre ser un gran ejemplo de cómo debe ser un entorno de trabajo sano y colaborativo.

A todos mis amigos y compañeros del *3D Genome Lab* (laboratorio 202 Sur del IFC), Sherlyn Claudio, Ayerim Esquivel, Luis Tenorio, Aura Stephenson, Amaury Aguilar, Andrés Penagos, Eduardo Blanco, Abraham Román y Juan Carlos Gómora por amistad y cálido recibimiento. Gracias por las charlas extensas sobre la vida, consejos y valiosas enseñanzas dentro y fuera del laboratorio. Gracias por hacer de estos últimos dos años de mi trayectoria universitaria, los mejores años.

A mis extutores de la LIBB por su invaluable contribución en mi formación como científico. A la Dra. Leticia Rocha por su disposición, paciencia y compromiso en enseñarme como es el trabajo experimental del laboratorio después de la crisis sanitaria. Por mostrarme la calidez de su laboratorio, y lo más importante, por enseñarme el lado humano del científico. Al Dr. Luis B. Tovar y Romo por su interés y compromiso en formarme como científico en medio de una pandemia mundial, incluso estando a cientos de kilómetros de distancia. Le agradezco por siempre buscar nuevas estrategias de aprendizaje e inclusión en el trabajo del laboratorio. Al Dr. Luis A. Mendoza Sierra por ser mi primer tutor de la carrera y recibirme cordialmente en su laboratorio. Por guiarme en mi primer proyecto,

enseñarme a discutir artículos científicos y mostrarme la belleza detrás de la investigación interdisciplinaria.

Agradezco a mis padres, Patricia y Gabino, por mostrarme el apoyo más incondicional y puro que he conocido. Por ser mis mayores consejeros y siempre creer y animarme a lograr todo lo que me proponga. A mi madre por incondicionalmente confiar y apoyar los sueños de un niño que anhelaba en convertirse en científico, aun cuando las creencias de la sociedad que nos rodeaba nos invitaban a renunciar a eso y ser “realistas”. A mi papá por siempre aconsejarme y ayudarme a desenredar situaciones complejas en mi vida y mostrarme lo que, en ocasiones, me era difícil ver. La culminación de este capítulo en mi vida no se categoriza a un logro personal, sino a uno colectivo. Dedico este trabajo a ellos, pues es un logro resultado de todo su esfuerzo y sacrificios.

Agradezco a mis amigos de la LIBB por ser mi segunda familia y una red de apoyo ante todos los retos durante la carrera. Gracias por enseñarme y acompañarme a romper juntos los ciclos y patrones tóxicos presentes en las carreras de ciencia. Les agradezco por alentarme a tomar decisiones arriesgadas y nuevos retos que me hicieron crecer no solo como científico sino también como persona.

Al resto de mis amigos más viejos, por la bonita amistad y apoyo forjado a lo largo de una vida. Agradezco por siempre escuchar mis planes y confiar en ellos, incluso cuando ni yo mismo creía en ellos. Por estar presentes en mis crisis existenciales y en mis logros, por estar en tiempos duros y tiempos felices.

Finalmente, agradezco a cada una de las personas que a lo largo de esta trayectoria me brindaron su ayuda sincera. Parte importante de un individuo se forja de las interacciones con la gente que se vincula a lo largo de su vida. Agradezco la fortuna de haberme encontrado a personas increíbles a lo largo de este camino, que me han ayudado a convertirme en quien soy.

Índice

| | |
|-------------------------------------------------------------------------------------------------------------------|-----------|
| Lista de abreviaturas | 11 |
| Resumen | 13 |
| 1. Introducción | 15 |
| 1.1. Organización genómica | 15 |
| 1.2. Métodos y técnicas para el estudio de la topología genómica | 15 |
| 1.3. Niveles de organización del genoma | 18 |
| 1.4. Elementos que regulan la organización 3D del genoma | 21 |
| 1.4.1 CTCF | 21 |
| 1.4.2. Cohesina | 21 |
| 1.4.3. Mecanismo de extrusión | 22 |
| 1.4.4. Reguladores emergentes de la topología genómica | 22 |
| 1.5. Influencia de la topología genómica en la diferenciación | 24 |
| 2. Antecedentes | 26 |
| 2.1. Diferenciación y maduración eritroide en aves | 26 |
| 2.2. Arquitectura genómica durante la diferenciación y maduración eritroide en <i>Gallus gallus</i> | 28 |
| 2.2.1. Reorientación de la arquitectura nuclear en células eritroides terminalmente diferenciadas | 28 |
| 2.2.2. Los RBC terminalmente diferenciados retienen actividad transcripcional | 29 |
| 2.2.3. El pausado de la RNA polimerasa II mantiene accesibles promotores de genes pausados | 30 |
| 2.2.4. Pérdida generalizada de la estructura del genoma eritroide | 30 |
| 2.3. Pérdida de las principales proteínas arquitectónicas durante la diferenciación y maduración eritroide | 33 |
| 2.4. YY1 como proteína arquitectónica emergente | 34 |
| 3. Planteamiento | 36 |
| 4. Pregunta de investigación | 36 |
| 5. Hipótesis | 36 |
| 6. Objetivos | 37 |
| 6.1. Objetivo general | 37 |
| 6.2. Objetivos específicos | 37 |
| 7. Métodos | 38 |
| 7.1. Obtención de muestras y cultivo celular | 38 |
| 7.1.1. Extracción de sangre periférica de gallinas adultas (aRBC) | 38 |
| 7.1.2. Extracción de sangre periférica de embriones de pollo | 38 |
| 7.1.3. Cultivo primario de fibroblastos de pollo | 38 |
| 7.1.4. Cultivo celular de eritroblastos de pollo | 39 |
| 7.2. Western blot | 39 |
| 7.2.1. Extracción de proteínas | 39 |

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 7.2.2. Inmunoblot _____ | 40 |
| 7.3. Inmunofluorescencia _____ | 40 |
| 7.3.1. Análisis de imágenes _____ | 41 |
| 7.4. ChIP-seq _____ | 41 |
| 7.4.1 Diseño de oligonucleótidos control de regiones de unión de YY1 para ChIP-seq _____ | 41 |
| 7.4.2. Inmunoprecipitación de la cromatina _____ | 42 |
| 7.4.3. PCR punto final y qPCR _____ | 42 |
| 7.4.3. Preparación de bibliotecas con material de ChIP para secuenciación _____ | 43 |
| 7.5. Análisis bioinformático _____ | 43 |
| 7.5.1. Pipeline general de ChIP-seq y análisis subsecuentes _____ | 43 |
| 7.5.2. Integración de datos de topología, expresión y accesibilidad _____ | 44 |
| 8. Resultados _____ | 46 |
| 8.1. Los niveles de proteína YY1 disminuyen a lo largo de la diferenciación y maduración eritroide _____ | 46 |
| 8.2. Evaluación de los niveles de YY1 mediante inmunofluorescencia y localización nuclear _____ | 47 |
| 8.3. Análisis de distribución nuclear de YY1 y comparación con marcas de activación y represión _____ | 51 |
| 8.4. Inmunoprecipitación de la cromatina acoplada a secuenciación (ChIP-seq) _____ | 53 |
| 8.4.1. Establecimiento del protocolo de YY1 ChIP en pollo _____ | 53 |
| 8.4.2. YY1 se localiza en regiones promotoras e intergénicas en el genoma eritroide _____ | 53 |
| 8.4.3. Los picos de YY1 de eritrocitos maduros se preservan dentro de minidominios _____ | 58 |
| 8.4.4. YY1 no está enriquecido en fronteras de minidominios _____ | 58 |
| 8.4.5. Los picos de YY1 están enriquecidos en el compartimento A _____ | 60 |
| 8.4.6. Los picos de YY1 preservan accesibilidad y se enriquecen en promotores de genes no expresados en aRBC _____ | 62 |
| 8.4.7. YY1 se posiciona en promotores de genes que pierden expresión durante la diferenciación terminal eritroide _____ | 64 |
| 8.4.8. YY1 participa en procesos relacionados con control de la expresión génica, metabolismo y procesamiento de RNAs _____ | 65 |
| 8.4.9. Integración de escenarios genómicos _____ | 66 |
| 9. Discusión y conclusiones _____ | 77 |
| 9.1. La proteína YY1 se preserva en los eritrocitos terminalmente diferenciados _____ | 77 |
| 9.2. YY1 no presenta patrones de distribución periférica nuclear en eritrocitos terminalmente diferenciados _____ | 78 |
| 9.3. Las distribuciones nucleares de las marcas H3K27Ac y H3K27me3 son similares a la de YY1 en eritrocitos maduros _____ | 79 |
| 9.4. YY1 se enriquece en promotores y regiones intergénicas, pero no en fronteras de minidominios en eritrocitos maduros _____ | 81 |
| 9.5. YY1 se posiciona dentro y fuera de minidominios _____ | 83 |
| 9.6. La unión de YY1 a sus blancos está acompañada por cambios en la accesibilidad del genoma _____ | 84 |
| 9.7. YY1 podría ser importante para el mantenimiento del pausado de genes estructurados y la estructuración local de la cromatina _____ | 85 |

| | |
|--------------------------------------------------------------------------------------------------------------------------|------------|
| 10. Perspectivas | 89 |
| 11. Anexos | 90 |
| 11.1. Comparación de las proteínas homologas YY1 en humano y pollo | 90 |
| 11.2. Western blot e inmunofluorescencia contra YY1 | 91 |
| 11.3. Establecimiento de ChIP contra YY1 en eritrocitos embrionarios | 91 |
| 11.4. Análisis de calidad de secuenciación y mapeo | 93 |
| 11.5. Cálculo de factores de normalización Spike-In para ChIP-seq | 95 |
| 11.6. Resultados de llamado de picos de YY1 | 96 |
| 11.7. Análisis de picos llamados por BroadPeaks | 97 |
| 11.8. YY1 se preserva dentro de TADs, pero no se enriquece en fronteras (picos compartidos entre réplicas) | 97 |
| 11.9. Los picos compartidos de YY1 están enriquecidos en el compartimento A | 98 |
| 11.10. Los picos compartidos de YY1 preservan accesibilidad y se enriquecen en promotores de genes no expresados en aRBC | 99 |
| 11.11. Picos de YY1 posicionados en promotores de genes | 99 |
| 12. Referencias bibliográficas | 101 |

Lista de abreviaturas

3C: captura conformacional de cromosomas

4C: captura conformacional de cromosomas circular

4C-seq: captura conformacional de cromosomas circular acoplada a secuenciación

5C: captura conformacional de cromosomas en base de copia de carbono

aRBC: eritrocito adulto

ATAC-seq: ensayo de accesibilidad de la transposasa acoplada a secuenciación

bp: pares de bases

ChIA-PET: análisis de interacción de la cromatina mediante secuenciación de extremo pareada tagmentado.

ChIP-seq: inmunoprecipitación de la cromatina acoplada a secuenciación

CT: territorio cromosómico

CTCF: factor de unión a CCCTC

CUT&RUN: escisión bajo objetivos y liberación mediante nucleasa

CUT&Tag: escisión bajo objetivos y tagmentación

DAPI: 4',6-diamidino-2-fenilindo

DNA: ácido desoxirribonucleico

DNMT: DNA metiltransferasa

E1-20: día de incubación 1-20

eRBC: eritrocito embrionario

eRNA: RNA de *enhancer*

FISH: hibridación *in situ* de fluorescencia

FPKM: fragmentos de transcrito por kilobase por millón de secuencias mapeadas

GO: ontología de genes

H3K27Ac: acetilación de la lisina 27 de la histona 3

H3K27me3: trimetilación de la lisina 27 de la histona 3

H3K4me3: trimetilación de la lisina 4 de la histona 3

H3K9me3: trimetilación de la lisina 9 de la histona 3

HAT: acetiltransferasa de histonas

HDAC: deacetilasa de histonas
HH: estadio de desarrollo Hamburger–Hamilton
Hi-C: captura conformacional de cromosomas del genoma completo
HP1: proteína de heterocromatina 1
IDR: región intrínsecamente desordenada
IP: inmunoprecipitación
Kb: kilo base
LADs: dominios asociados a la lámina
LLPS: separación de fases líquido-líquido
lncRNA: RNA largo no codificante
Mb: mega base
PCR: reacción en cadena de la polimerasa
PRC2: complejo represivo de Polycomb 2
PTM: modificación postraduccional
qPCR: reacción en cadena de la polimerasa cuantitativa
RBC: glóbulos rojos
RNA: ácido ribonucleico
RNA-seq: técnica de análisis del transcriptoma
TAD: dominio topológicamente asociado
TF: factor transcripcional
TLR: receptor tipo toll
TSS: sitio de inicio de transcripción
YY1: proteína Yin Yang 1

Resumen

Durante los procesos de diferenciación terminal y maduración eritroide en aves, el núcleo del eritrocito se compacta progresivamente y el genoma es sujeto a mecanismos moleculares que resultan en su silenciamiento. La reciente caracterización del panorama transcripcional y de arquitectura genómica de eritrocitos de pollo reveló que, aunque existe una caída robusta en la transcripción, algunos genes mantienen su expresión. A medida que estos procesos progresan, los dominios cromatínicos conocidos como dominios topológicamente asociados (TADs), que dan estructura al genoma, se pierden. No obstante, se identificaron estructuras tipo TAD de menor tamaño denominadas como minidominios, las cuales organizan a promotores pausados en los eritrocitos. La pérdida de expresión y niveles proteicos de las principales proteínas arquitectónicas en vertebrados, como el factor de unión a CCCTC (CTCF) y subunidades del complejo cohesina a lo largo de la diferenciación y maduración eritroide, correlacionan con la pérdida global de estructura en el genoma y la hipercompactación su núcleo.

En el presente trabajo nos preguntamos si la proteína estructural Yin Yang 1 (YY1), cuyos niveles de expresión se mantienen a lo largo de la diferenciación terminal y maduración eritroide en *Gallus gallus*, podría ser responsable de estructurar a los minidominios de una forma análoga a la de CTCF. Los resultados muestran que los niveles proteicos de YY1 disminuyen significativamente a lo largo de la diferenciación terminal y maduración, pero aún son detectables en eritrocitos maduros. Además, describimos que YY1 está ampliamente distribuido por todo el núcleo eritroide y presente en regiones nucleares previamente identificadas como regiones accesibles y heterocromatinizadas en eritrocitos maduro, abriendo un panorama importante de una posible participación bivalente de activación y represión en el control global transcripcional durante la compactación del núcleo.

Adicionalmente, experimentos de ChIP-seq realizados en los en los diferentes estadios de diferenciación revelaron que, en eritrocitos maduros, YY1 no se enriquece en fronteras de minidominios, como lo hace CTCF en fronteras de TADs, pero sí en su interior. Dentro de los minidominios en eritrocitos, YY1 se enriquece en regiones promotoras de genes inactivos y, además, colocaliza mayoritariamente con regiones accesibles del genoma, reforzando la naturaleza bivalente de YY1. Específicamente, en el contexto de la

diferenciación terminal y maduración eritroide, YY1 podría desempeñar un papel importante en el pausado o represor transcripcional de loci específicos que permanecen estructurados. Complementariamente, YY1 podría ser responsable de mantener accesibles y promover estructura local de la cromatina de las regiones del genoma que conforman a los minidominios. Dicha accesibilidad y estructuración permitirían promover una configuración espacial necesaria para la reactivación de aquellos genes pausados y estructurados en respuesta a diversos estímulos.

Este trabajo representa un avance en el estudio de los elementos proteicos que regulan la estructuración del genoma durante el proceso de diferenciación terminal eritroide. También, permite la integración de nuevas hipótesis sobre los mecanismos ejecutados por otros factores nucleares que podrían permitir el mantenimiento de minidominios en eritrocitos terminalmente diferenciados. La unión de YY1 a regiones que se inactivan en el genoma a lo largo del proceso de diferenciación abre nuevas perspectivas para entender los mecanismos moleculares que gobiernan el silenciamiento global del genoma y aquellos mecanismos que gobiernan el pausado y manutención de accesibilidad de regiones genómicas importantes, ante la generación de un ambiente nuclear altamente represivo.

1. Introducción

1.1. Organización genómica

El núcleo es un organelo que resguarda la mayoría del material genético de una célula eucariota. El primer eslabón en la compleja organización nuclear comienza con la unidad básica estructural y funcional del genoma, conocida como nucleosoma. Un nucleosoma está comprendido por la interacción entre, aproximadamente, 150 pares de bases (bp) de DNA envueltas en un octámero proteico de dímeros de las histonas H2A, H2B, H3 y H4. La repetición del nucleosoma en todo el DNA genómico conforma a una estructura altamente dinámica conocida como cromatina¹. Gracias a las características físicas y bioquímicas resultantes de la asociación entre DNA y proteínas, el estado de compactación de la cromatina depende en parte de las modificaciones postraduccionales (PTM) en sus histonas. En cromosomas interfásicos, se denota como eucromatina a regiones que presentan un menor grado de compactación y son ricas en genes activos^{1,2}. Por otro lado, como heterocromatina a regiones altamente compactadas y con zonas transcripcionalmente inactivas^{1,2}.

Históricamente, la microscopía de campo claro en combinación con tinciones del núcleo sentó las bases del estudio de la topología genómica. La idea de que el material genético dentro del núcleo es un ambiente dinámico y bien estructurado ha tomado fuerza en las últimas décadas². En los 80s, el desarrollo de la técnica de microscopía conocida como hibridación *in situ* de fluorescencia (FISH), permitió confirmar la hipótesis de los territorios cromosómicos formulada a finales del siglo XIX e inicios del XX². La técnica de FISH, además, permitió realizar observaciones sobre el posicionamiento de los cromosomas en el núcleo. De forma general, se observó que regiones de cromosomas ricas en genes activos presentaban un posicionamiento preferencial hacia el centro nuclear, mientras que regiones ricas en genes inactivos estaban posicionadas a hacia la periferia nuclear^{2,3}. La integración de estas observaciones con su grado de compactación evidenció la existencia de una conformación espacial “especial” de la cromatina en el núcleo.

1.2. Métodos y técnicas para el estudio de la topología genómica

Actualmente, la organización 3D del genoma representa un mecanismo emergente cuya desregulación subyace en la aparición de enfermedades y la desregulación de procesos moleculares. Previamente, se creía que los cromosomas interfásicos no tenían

organización aparente definida. Sin embargo, el establecimiento y refinamiento de las tecnologías de captura conformacional de cromosomas (3C), de la mano con avances realizados mediante microscopía, ha permitido elucidar que la cromatina en interfase presenta organización particular en función a su escala genómica, la cual no es aleatoria.

La técnica 3C fue la primera en ser desarrollada, la cual utiliza el principio del ensayo de ligación de cromatina utilizando al formaldehído como agente entrecruzador para fijar todas las interacciones de la cromatina ocurriendo en el momento de realizar el ensayo⁴. De esta forma, es posible conservar las interacciones entre diferentes loci genómicos dando paso a un estudio sistemático de la topología del genoma. La técnica 3C, “uno contra uno”, permite estudiar y cuantificar mediante la técnica de reacción en cadena de la polimerasa (PCR) los contactos entre dos pares de secuencias previamente determinadas⁴. A pesar de la innovación de la técnica, esta presenta retos metodológicos cuando se requiere cuantificar productos de contactos raros e infrecuentes, haciendo que técnicas de cuantificación por PCR sean difíciles y mayormente viables para productos de contactos de largo alcance más representados en la población celular⁴. No obstante, modificaciones al protocolo de la técnica 3C dieron paso al desarrollo de métodos más robustos como la técnica de captura conformacional de cromosomas circular (4C).

La técnica 4C amplió el panorama de 3C al permitir estudiar interacciones “uno contra todos”. Es decir, dio paso a la identificación de regiones a nivel de genoma completo que contactan a una región de interés, mejor conocida como “*viewpoint*”⁴. La principal ventaja de esta técnica recae en que no se necesita de un conocimiento previo de candidatos que realicen contactos, a diferencia del 3C. El 4C hace uso de la PCR reversa, microarreglos y secuenciación de Sanger para analizar las secuencias de los contactos. El desarrollo y refinamiento de tecnologías de secuenciación de siguiente generación, permitió acoplar el 4C convirtiéndose en 4C-seq permitiendo generar perfiles robustos de interacciones realizadas por el “*viewpoint*” y explorar sus implicaciones regulatorias⁴.

La técnica de captura conformacional de cromosomas en base de copia de carbono (5C), “muchos contra muchos”, permite explorar los contactos que existen entre múltiples secuencias amplias⁴. Su desarrollo parte de una biblioteca convencional de 3C que se somete a amplificación mediada por un conjunto diverso de oligonucleótidos, generando una biblioteca 5C. La secuenciación de esta biblioteca permite obtener una matriz de

frecuencias de interacción de las regiones genómicas estudiadas. La técnica de 5C permitió estudiar la conformación de la cromatina sobre los locus de α -, β - globinas revelando su posicionamiento y configuración espacial dentro del núcleo⁴. Actualmente, una de las aproximaciones más ampliamente utilizadas para estudiar la topología genómica, derivadas del protocolo 3C modificado, es la captura conformacional de la cromatina de alto rendimiento, también conocida como captura conformacional de cromosomas del genoma completo (Hi-C). A diferencia de las técnicas 3-,4-,5-C, Hi-C es una aproximación de tipo “todo contra todo”, es decir, nos permite investigar los contactos e interacciones entre todos los loci comprendidos en todo el genoma⁴.

El campo actual de la topología genómica hace uso de las técnicas previamente descritas acopladas a secuenciación masiva que permiten la integración de la distribución ocupacional de proteínas que interactúan con el DNA. Una de las técnicas más utilizadas para elucidar la ocupación de una proteína en la hebra de cromatina a nivel de genoma es la inmunoprecipitación de cromatina acoplada a secuenciación (ChIP-seq)⁵. El ChIP-seq utiliza un anticuerpo específico contra la proteína de interés que presenta interacciones con el DNA. Una vez que la célula se fija y la cromatina es fragmentada mediante sonicación, se realiza una inmunoprecipitación que permite enriquecer los complejos DNA-proteína de interés. Luego, el DNA es purificado de los complejos y con el material recuperado se construyen bibliotecas de secuenciación. En los últimos años han sido desarrolladas tecnologías de escisión bajo objetivos y liberación mediante nucleasa (CUT&RUN) o tagmentación (CUT&Tag). Al igual que ChIP-seq, ambas tecnologías permiten elucidar la ocupación genómica de proteínas, ya sean marcas de histonas o factores de transcripción. Sin embargo, CUT&RUN y CUT&Tag han demostrado mayor robustez y ventajas importantes sobre la técnica de ChIP-seq^{6,7}.

Para poder obtener un panorama completo de como la topología genómica y dinámica nuclear influye en la funcionalidad y regulación de un contexto celular dado, se requiere evaluar la accesibilidad genómica presente y el transcriptoma celular. Para ello, se utilizan las metodologías de ensayo de accesibilidad de la transposasa acoplada a secuenciación (ATAC-seq) y la técnica de análisis del transcriptoma (RNA-seq). Los procesos como diferenciación y maduración celular son el reflejo de los cambios en los mecanismos epigenéticos y genéticos, que impactan de forma directa en cambios en la estructura de la cromatina⁸. De forma general, la cromatina tiende a mostrar cambios en su nivel de

compactación dependiente al fenómeno celular que lleve a cabo⁹. La técnica ATAC-seq nos permite estudiar el dinamismo de la cromatina respecto a su accesibilidad en diferentes fenómenos celulares⁹. Por su parte, la técnica de RNA-seq nos permite evaluar las secuencias de todos los RNAs presentes en una muestra y cuantificar su abundancia¹⁰.

1.3. Niveles de organización del genoma

El núcleo de las células eucariotas se encuentra altamente compactado en un espacio tan pequeño que requiere ser organizado y regulado a diferentes escalas para poder permitir la correcta respuesta transcripcional. Las tecnologías de captura de conformación de la cromatina acopladas a secuenciación en combinación con los últimos avances en microscopia de super resolución, han permitido resolver las múltiples capas de organización nuclear.

El nivel de organización topológica más grande, estudiado mediante microscopia y 3D FISH, comprende en tamaño a los cromosomas enteros. Los cromosomas interfásicos se acomodan en regiones discretas del núcleo conocidas como territorios cromosómicos (CTs, **Figura 1**)^{2,11}. El posicionamiento de los cromosomas dentro del núcleo está directamente relacionado con la densidad génica comprendida en un cromosoma. De esta forma, los cromosomas que contienen baja densidad génica se posicionan preferencialmente en la periferia nuclear y la región perinuclear, mientras que los que son ricos en genes se posicionan hacia el centro del núcleo¹¹. Además, se ha observado que el posicionamiento de los genes dentro de un territorio cromosómico correlaciona con su estado transcripcional y es altamente dinámico en función a necesidades de activación y represión¹¹.

A escala de cientos de megabases (Mb), experimentos de Hi-C y FISH acoplado a superresolución relevaron que la cromatina se segrega en dos compartimentos principales conocidos como A y B (**Figura 1**)^{2,12}. El compartimento A está asociado a regiones del genoma abiertas y las secuencias contenidas en general están enriquecidas de genes que se están transcribiendo y contienen marcas de histonas asociadas a activación transcripcional. Por el contrario, el compartimento B contiene genes inactivos y marcas de histonas asociadas a represión transcripcional¹². En las matrices de Hi-C

podemos observar a los compartimentos de un cromosoma entero como tableros de ajedrez.

A una escala menor que comprende regiones de entre 100 kilobases (Kb) a ~1 Mb podemos encontrar a los dominios topológicamente asociados (TADs) y asas de cromatina (**Figura 1**)². Los TADs son estructuras descritas a partir de experimentos de Hi-C, las cuales representan regiones del genoma, específicamente asas de cromatina de gran longitud, que se pliegan y asocian en el espacio nuclear y presentan una alta frecuencia de interacciones entre las secuencias que comprenden^{2,13}. Experimentalmente podemos observar a los TADs en matrices de interacción obtenidas mediante Hi-C como triángulos contiguos. Funcionalmente, los TADs facilitan la formación de asas de cromatina de menor tamaño, mejor conocidas como asas funcionales, que contienen elementos reguladores como *enhancers* y promotores, ayudando a propiciar sus interacciones (**Figura 1**)^{2,12,13}. Además, CTCF tiene la capacidad de delimitar a los TADs formando fronteras de CTCF, aislando a las regiones genómicas que comprenden de posibles contactos con elementos externos al TAD¹⁴.

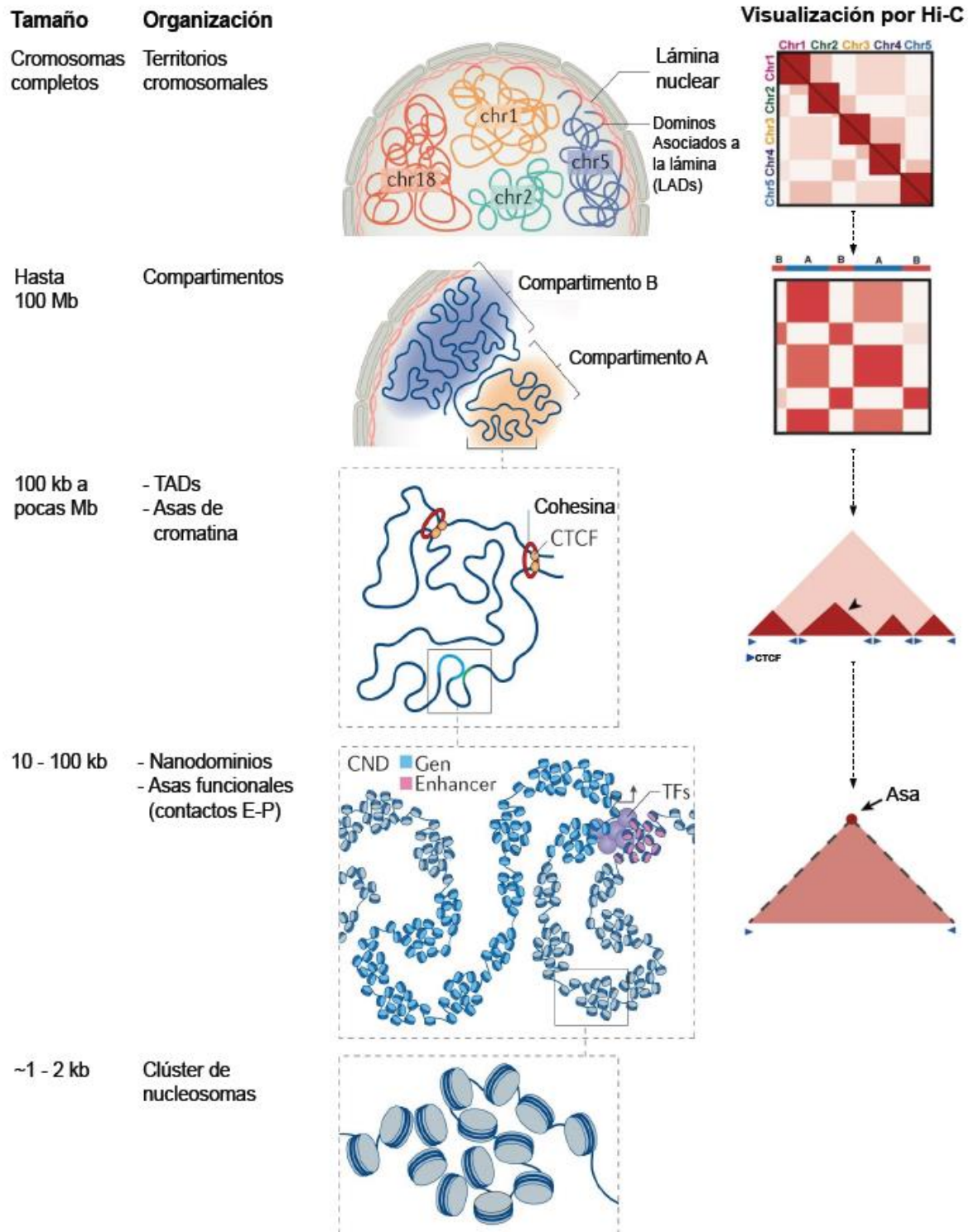


Figura 1. Niveles de organización de genoma. El genoma posee una estructura jerárquica. Se ilustran los niveles de esta jerarquía, ordenados de mayor a menor tamaño: territorios cromosomales, compartimentos, dominios topológicamente asociados (TADs), asas de cromatina y nucleosomas. Además, se muestran sus visualizaciones en mapas de interacción de Hi-C (derecha). Factor de unión a CCCTC (CTCF). Figura modificada de: Jerkovic', I. (2021)² y L Ayala-Guerrero (2023)³².

1.4. Elementos que regulan la organización 3D del genoma

La topología jerárquica 3D del genoma es estructurada mediante proteínas arquitectónicas y algunas especies de RNAs nucleares no codificantes. En vertebrados, las proteínas arquitectónicas más estudiadas son CTCF y el complejo de cohesina (Figura 1).

1.4.1 CTCF

CTCF es una proteína que posee 11 dedos de zinc, ubicuamente expresada en la mayoría de los tejidos, y altamente conservada en vertebrados. CTCF presenta alta afinidad por secuencias específicas de DNA y resulta esencial en diversos procesos celulares, como el desarrollo^{14,15}. Originalmente, CTCF fue descrita como un factor transcripcional con la capacidad de regular de forma negativa al gen *c-myc* y *MYC* en pollo y humano, respectivamente. Sin embargo, experimentos de ChIP-seq han revelado una extensa distribución de CTCF a lo largo del genoma, que además se encuentran altamente conservados entre vertebrados y tipos celulares^{15,16}. En combinación con Hi-C, se observó que CTCF muestra un enriquecimiento preferencial en sitios de fronteras de TADs y en las anclas de las asas de cromatina. Sistemas de degradación de proteínas (degron) inducidos por auxina, que permiten evaluar el efecto funcional de una proteína *in vivo*, han demostrado que la depleción de CTCF resulta en la pérdida de asas de cromatina en los dominios y la pérdida del aislamiento de los TADs en células troncales embrionarias de ratón¹⁷.

Aunado a esta evidencia, se ha detectado que CTCF se encuentra enriquecido en los bordes de dominios topológicos que interactúan frecuentemente con elementos de la lámina nuclear, los cuales se han denominado dominios asociados a la lámina (LADs)¹⁸. Los LADs correlacionan con regiones del genoma con baja o nula actividad transcripcional y tiempos tardíos de replicación¹⁸. Por tanto, CTCF es un factor que cumple una función organizadora del genoma formando asas de cromatina y fronteras de TADs, así como permitir el confinamiento de los LADs a la periferia nuclear.

1.4.2. Cohesina

Cohesina es un complejo multiproteico anular con cuatro subunidades conservadas en vertebrados: SMC1A, SMC3, SA1/2 y RAD21^{15,19}. Su función más conocida es durante la fase S del ciclo celular, en donde se une a las cromátidas hermanas y permite la

correcta segregación de cromosomas durante la división celular¹⁹. No obstante, el complejo cohesina ha demostrado tener un papel relevante en el establecimiento de la arquitectura genómica de la mano de CTCF, pues muchos sitios de cohesina en la cromatina coinciden con este factor de transcripcional¹⁵. Estudios funcionales han demostrado que la degradación aguda de la subunidad RAD21 provoca la pérdida generalizada de asas de cromatina y dominios cromatínicos, mientras que el rescate de la subunidad resulta en el restablecimiento de estos^{20,21}. Adicionalmente, experimentos en ovocitos fertilizados de ratón sometidos a *knockout* condicional de RAD21 mostraron una ausencia generalizada de TADs y asas de cromatina²². En conjunto, múltiples evidencias demuestran que el complejo cohesina desempeña un papel fundamental en el establecimiento y formación de asas de cromatina y TADs.

1.4.3. Mecanismo de extrusión

La colocalización de cohesina y CTCF, además de la influencia de CTCF sobre la distribución de cohesina en la cromatina, evidencian la mutua influencia y cooperación entre ambos elementos para regular la topología genómica. El mecanismo de extrusión de cromatina explica el mecanismo por el cual CTCF y cohesina cooperan para el establecimiento de asas de cromatina y TADs. En este modelo, una pequeña asa de DNA pasa a través del anillo del complejo cohesina, el cual funciona como factor de extrusión, que se desliza en la fibra de cromatina permitiendo la formación total del asa. La extrusión de la cromatina continua hasta que la cohesina se encuentra con un factor CTCF unido a la cromatina y correctamente orientado. El asa de cromatina es estabilizada con la formación de un ancla cuando dos moléculas de CTCF orientadas de forma convergente forman un dímero¹⁵.

1.4.4. Reguladores emergentes de la topología genómica

Si bien CTCF y cohesina son las principales proteínas arquitectónicas en vertebrados, en los últimos años ha surgido evidencia relacionada a proteínas arquitectónicas emergentes y otros mecanismos relevantes en la organización 3D del genoma. En particular, las proteínas que conforman la matriz nuclear han demostrado tener un papel relevante como proteínas de andamiaje para la cromatina y su organización en diversos procesos celulares¹⁶. Podemos clasificar a los elementos de la matriz nuclear como internos y periféricos. Ejemplos de componentes periféricos abarcan a las proteínas que forman parte del complejo del poro y la lámina nuclear, las cuáles permiten secuestrar y

mantener a regiones heterocromatinizadas del genoma adosadas a la periferia nuclear¹⁶. Por otro lado, estudios funcionales de componentes internos de la matriz, como Matrin3, SAFB y HNRNPK, han revelado su importancia en la estructuración del genoma pues su depleción conlleva a fenómenos como la reorganización de asas de cromatina, incremento en las interacciones de DNA con la lámina nuclear e incluso cambios en la fuerza de aislamiento de las fronteras de los compartimentos^{16,23}.

Es bien sabido que los factores transcripcionales (TFs) permiten regular diferentes programas transcripcionales dado un estímulo o como parte constitutiva en diversos procesos biológicos. Al interactuar con el DNA, muchos TFs tienen la capacidad de alterar la organización de regiones de la cromatina en el espacio nuclear al establecer y promover interacciones de corto o largo alcance¹⁶. Los receptores olfatorios, por ejemplo, necesitan ser exclusivamente expresados uno a la vez en cada neurona olfatoria adulta; sin embargo, existen más de mil genes en la familia de receptores olfatorios²⁴. La manera en la que se selecciona que gen debe expresarse en una neurona está dictaminada en función a los contactos inter-cromosomales que se presenten²⁴. En este contexto, se ha observado que el TF conocido como LHX2 tiene la capacidad de establecer un ambiente en el que un gen particular para un receptor olfatorio es acercado espacialmente a un clúster de *enhancers*, permitiendo su expresión. Por otro lado, tenemos a TFs que median interacciones dentro del mismo cromosoma al establecer contactos más pequeños mediante asas de cromatina entre elementos regulares como promotores y *enhancers*²⁵. Dentro de esta función contamos con TFs ubicuamente expresados en diferentes tejidos y tipos celulares, como Yin Yang (YY1)²⁶; y también a TFs específicos para un tipo celular que participan en la identidad celular, como FOXP3²⁷, GATA1 o MyoD^{16,23}.

Adicionalmente, los RNAs no codificantes han ganado relevancia significativa en la topología genómica. Se ha observado que RNAs no codificantes pueden regular la expresión de genes al mediar la remodelación de la cromatina entre promotores y *enhancers*. Por ejemplo, cuando se activa el receptor alfa de estrógenos, inmediatamente se comienza a transcribir el RNA de su *enhancer* (eRNA). Este eRNA media la formación de un asa de DNA a través de la asociación con el complejo mediador provocando la interacción entre *enhancer*-promotores de sus genes blanco²⁸. Por otro parte, un ejemplo clásico de regulación topológica esta dado por *Xist*, un RNA largo no codificante (lncRNA) que tiene la capacidad de silenciar a uno de los dos cromosomas X en hembras. *Xist*

recubre al cromosoma X para silenciarlo al reclutar diversas proteínas remodeladoras, como el complejo represor de Polycomb y DNA metiltransferasas (DNMTs), con el propósito de establecer el estado heterocromático del cromosoma X-inactivo. Finalmente, el cromosoma X inactivo es relocalizado a la periferia nuclear y *Xist* permite anclarlo a la lámina nuclear²⁸.

Otro escenario que ha cobrado gran relevancia es el fenómeno de separación de fases dentro del núcleo como un potencial organizador del genoma. Muchas proteínas nucleares, factores de transcripción y especies de RNAs tienen propiedades fisicoquímicas que propician la formación de agregados, concentración-dependientes, a través de interacciones multivalentes²⁹. Los agregados nucleares forman estructuras no membranosas en forma de gotas, por lo que a este fenómeno se le ha denominado separación de fases líquido-líquido (LLPS)²⁹. La LLPS participa en la formación de organelos sin membrana con función específica dentro del núcleo como el nucleolo y gránulos P³⁰. Además, se han relacionado directamente con enfermedades como el cáncer³¹. Uno de los fenómenos más representativos en el campo de la organización 3D del genoma y la LLPS es la formación de la heterocromatina, mediante la proteína heterocromatina 1 (HP1), y la compartimentalización de la cromatina²⁹. Esto se debe a que, en la separación de fases se favorece el agregamiento de componentes similares y se excluyen moléculas no relacionadas, propiciando un ambiente favorable para ciertas reacciones fisicoquímicas. Dicho ambiente creado tiene la capacidad de acercar espacialmente elementos reguladores y promover su interacción mediante asas de cromatina propiciando una respuesta transcripcional²⁹.

1.5. Influencia de la topología genómica en la diferenciación

Durante la diferenciación celular y el desarrollo, se ha observado que la topología del genoma exhibe cambios drásticos consecuencia del dinamismo epigenético respecto a la expresión diferencial de variantes de histonas y sus PTMs³². Esto concede la capacidad de una regulación espacio-temporal concisa de regiones genómicas que resguardan programas genéticos pluripotenciales o linaje-específicos^{8,32}.

Durante el desarrollo embrionario, la reprogramación epigenética y los cambios en los niveles de organización topológica del genoma se encuentran intrínsecamente ligados y su dinamismo es evidente desde la fertilización y a través de todos los estadios del

desarrollo³². Por su parte, a medida que las células troncales se diferencian a células especializadas de un linaje específico, se requiere que la cromatina sea remodelada de forma global^{8,32}. Dicha remodelación orquestada por una reprogramación epigenética que permite la ganancia de marcas de histonas de represión como la trimetilación de la lisina 27 de la histona 3 (H3K27me3), y la pérdida de marcas de activación como la trimetilación de la lisina 4 de la histona 3 (H3K4me3)⁸. Esto da paso a que las células troncales pierdan la configuración de cromatina abierta, a una compacta y represiva, con el objetivo de silenciar genes de pluripotencia y activar aquellos linaje-específicos⁸.

De forma reciente, un modelo de diferenciación terminal de músculo esquelético en células C2C12 demostró que la proteína Matrin3, componente principal de la matriz nuclear interna, participa en la regulación de la estructura 3D del genoma²³. A través de sistemas de degradación aguda de proteínas dirigido a Matrin3, fue posible evidenciar que esta proteína tiene la capacidad de mediar el proceso de diferenciación estabilizando sitios de accesibilidad e interacciones de asas de cromatina en dominios²³. Además, se observó que tiene la capacidad de mediar la formación de asas de cromatina de tipo promotor-*enhancer* formadas por YY1²³. Un fenómeno interesante pues previamente se había reportado la capacidad de Matrin3 de asociarse a la cromatina a través de interacciones con las proteínas arquitectónicas CTCF y Rad21³³.

Otro ejemplo que ilustra el impacto de la reconfiguración de la topología genómica en la diferenciación ocurre en el establecimiento del linaje de las células T y su maduración. Experimentos de Hi-C de estadios de diferenciación del linaje T, que abarcan desde células troncales hematopoyéticas hasta timocitos dobles positivos CD4+CD8+, evidenciaron cambios sustanciales a nivel de compartimentos y TADs³⁴. Por su parte, en células troncales hematopoyéticas y progenitores linfoides comunes, los compartimentos A que contienen genes de pluripotencia cambian al compartimento represivo B cuando estas células se diferencian a timocitos dobles negativo estadio 3 (DN3 CD4-CD8-)³⁴. En contra parte, aquellos compartimentos B que contienen genes linaje-T específicos, cambian al compartimento A al ocurrir la diferenciación a células DN3³⁴. El estadio DN3 marca el establecimiento definitivo al linaje T. En este estadio, la actividad intra-TAD incrementa y está asociado a una amplia reorganización de la arquitectura de la cromatina que asegura y facilita la activación del programa transcripcional asociado con el linaje T, y restringiendo otros el acceso a programas correspondientes a otros linajes³⁴.

2. Antecedentes

2.1. Diferenciación y maduración eritroide en aves

La eritropoyesis terminal es un proceso que involucra el establecimiento de células troncales hematopoyéticas hacia unidades formadoras de colonias eritroides, y subsecuentemente a glóbulos rojos (RBCs) maduros³⁵. Durante la eritropoyesis terminal, diversas cascadas de señalización inducidas por eritropoyetina producen cambios en la transcripción de genes necesarios para la diferenciación y regulación de procesos como la inhibición de la apoptosis³⁵. A medida que la diferenciación avanza, la cromatina sufre una condensación gradual pero dramática dando lugar a un silenciamiento progresivo de genes³⁶. Sí bien muchos mecanismos que provocan la condensación aún son desconocidos, se conoce la influencia de ciertas deacetilasas de histonas (HDACs), variantes de histonas y PTMs asociadas^{36,37}. En mamíferos, la eritropoyesis terminal culmina con la expulsión del núcleo (enucleación) altamente condensado, otorgando a los RBC la capacidad de almacenar más hemoglobina y flexibilidad física^{35,36}. A diferencia de los mamíferos, la diferenciación y maduración de RBC en aves no involucra la expulsión del núcleo, pero sí un alto grado de compactación de su núcleo³⁷⁻⁴⁰ (**Figura 2a**).

La eritropoyesis aviar puede dividirse en dos grandes etapas: primitiva y definitiva. La eritropoyesis primitiva comienza en el saco vitelino durante el desarrollo embrionario, en donde las células eritroides progenitoras y células endoteliales forman agregados conocidos como “*blood islands*” en la etapa de desarrollo 6 de Hamburger–Hamilton (HH6). Inmediatamente en el día de incubación 1 (E1; en estadio HH7) ocurre la señal más temprana de diferenciación eritroide dada por el inicio de la expresión de los genes de hemoglobina φ , ρ y ϵ ³⁹. La subsecuente formación de vasculatura permite que, en HH10-11, los RBC primitivos comiencen a segregarse unos de otros para iniciar circulación en E2 (HH12-13). Los RBC primitivos circulantes se caracterizan por su gran tamaño y morfología circular, en comparación con los RBC maduros³⁹. En la etapa E5 de desarrollo, aparecen en circulación los primeros RBC de la etapa definitiva, y para E7 representan a la mayoría de RBC circulantes. A diferencia de los primitivos, los RBC definitivos expresan principalmente a las hemoglobinas α -A, α -D y β -A; además tienen forma oval y son más pequeños³⁹.

Ambos linajes de células rojas primitivas y definitivas, producidas por el saco vitelino, maduran a lo largo de los estadios de megaloblastos, proeritroblastos, eritroblastos, reticulocitos, hasta llegar a RBC. Desde su establecimiento y hasta E10, el saco vitelino es la fuente principal de producción de RBC. En E12, la médula ósea comienza a funcionar como fuente de producción eritropoyética embrionaria, y se posiciona como el nicho principal de producción de RBC en E13-15³⁹. Como consecuencia, la producción del saco vitelino recae prominentemente para E15 y cesa en E19³⁹ (**Figura 2b**).

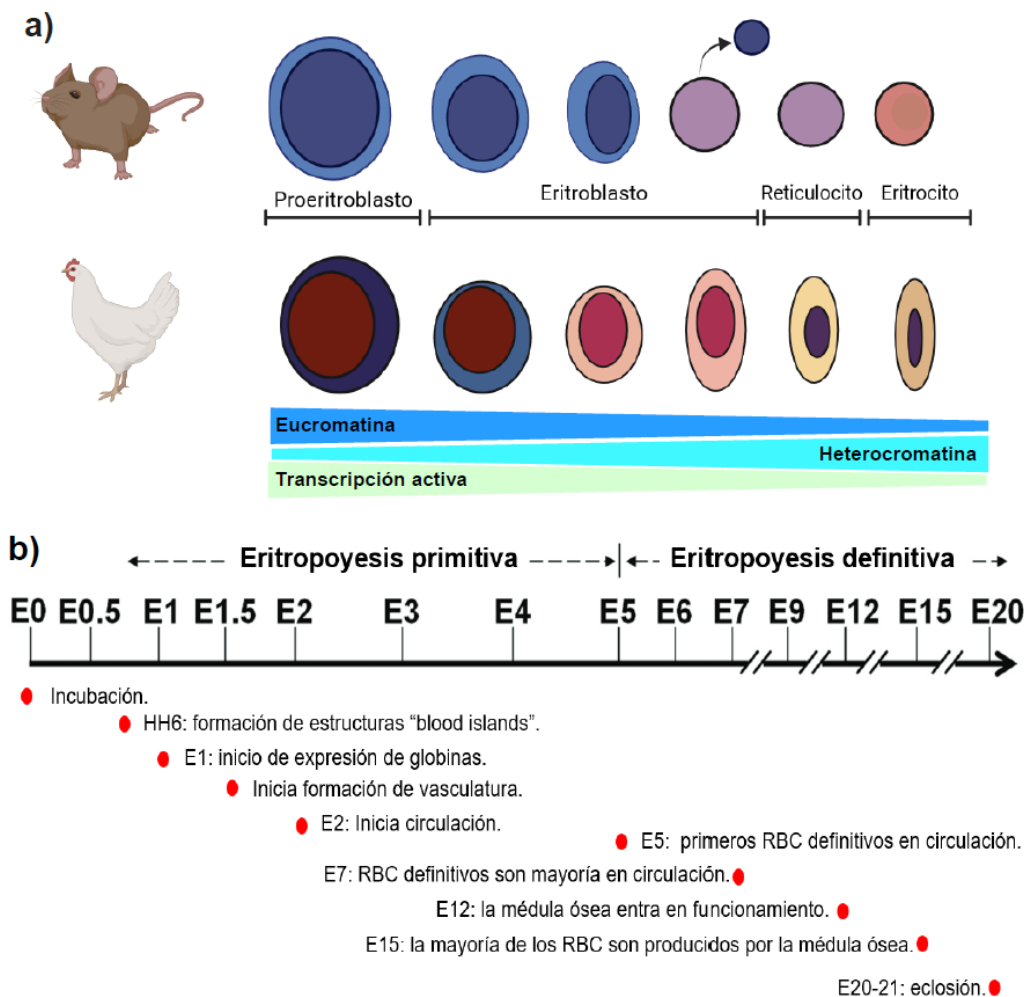


Figura 2. Diferenciación y maduración eritroide en aves. **a)** El proceso de eritropoyesis involucra una pérdida generalizada de eucromatina y aumento de heterocromatina, acompañado de un silenciamiento generalizado de genes. Durante este proceso, los eritrocitos sufren cambios morfológicos importantes, de los que destaca la extensa compactación su núcleo. En aves, a diferencia de los mamíferos, los eritrocitos terminalmente diferenciados y maduros circulantes conservan su núcleo. **b)** Eventos fundamentales de la eritropoyesis en aves durante el desarrollo embrionario. E0-20 representan el día de incubación, mientras que HH el estadio Hamburger–Hamilton de desarrollo. Figuras modificadas de: a) Campbell, Terry (2013)⁴⁰; y b) Guojun Sheng (2010)³⁹.

Aunque en aves la compactación del núcleo es un proceso poco estudiado, el fenómeno en células sanguíneas ha sido asociado a la histona H5 *linker* o la proteína MENT en RBC y células blancas, respectivamente³⁷. La compactación global provoca un cese total de la producción de DNA y, presuntamente, los RBC se transforman en células transcripcionalmente inertes³⁷.

En RBC terminalmente diferenciados maduros, hasta el 60% de las histonas H1 *linker* son reemplazadas por H5, cuya deposición provee una mayor estabilidad a la cromatina produciendo regiones genómicas altamente reprimidas. El dominio globular de H5 representa un factor indispensable para la compactación debido a su tendencia a la dimerización, facilitando la agregación de nucleosomas y la subsecuente formación de estructuras de tallo permitiendo el plegamiento de la cromatina³⁷.

2.2. Arquitectura genómica durante la diferenciación y maduración eritroide en *Gallus gallus*

La eritropoyesis terminal es un modelo ampliamente utilizado para estudiar mecanismos que acompañan a la diferenciación celular. En los últimos años, este proceso se ha posicionado como una oportunidad de investigar cambios en la topología genómica. Especialmente, los RBC circulantes y funcionales de las aves representan un modelo interesante que, al retener su núcleo, permite estudiar los cambios de la topología del genoma durante la generación de un ambiente altamente represivo. En los últimos años, nuestro grupo de investigación ha estudiado y caracterizado la organización 3D del genoma en RBC de *Gallus gallus* a través de la diferenciación y maduración eritroide.

2.2.1. Reorientación de la arquitectura nuclear en células eritroides terminalmente diferenciadas

Para investigar los cambios que ocurren en la condensación del núcleo durante los procesos de diferenciación terminal y maduración eritroides, fueron empleados eritrocitos embrionarios (eRBC) provenientes de embriones de pollo de 10 días, y eritrocitos adultos (aRBC) obtenidos de gallinas de 30 semanas. Ambas poblaciones pertenecen a células en donde la cromatina ha sido extensamente compactada y, además, en donde la transcripción global ha disminuido. Además, la línea celular HD3 fue empleada como un punto temporal correspondiente a un eritroblasto temprano en la eritropoyesis aviar⁴¹.

Mediante ATAC-see se elucidó un reposicionamiento de la cromatina accesible hacia la periferia nuclear en aRBC, siendo menos abundante en centro del núcleo. Este hallazgo contrasta directamente con la organización convencional de la cromatina dentro del núcleo, la cual se encuentra presente en eRBC y eritroblastos. La reorientación de la arquitectura nuclear fue validada mediante la inmunofluorescencia y análisis de distancia núcleo-periferia contra la trimetilación de la lisina 9 de la histona 3 (H3K9me3), una marca correspondiente a heterocromatina constitutiva que se enriquece preferencialmente en la periferia nuclear. En eRBC y aRBC, H3K9me3 se posiciona preferencialmente en focos localizados hacia el centro nuclear en contraste con los eritroblastos, en donde H3K9me3 decrece gradualmente desde la periferia nuclear hacia el centro. Además, mediciones restringidas al centro y la periferia nuclear revelaron que H3K9me3 exhibe un enriquecimiento preferencial únicamente en el centro. Esta observación era concordante con el análisis de la distribución de señal de DAPI, que mostraba mayor enriquecimiento al centro nuclear en RBCs, en comparación del enriquecimiento periférico encontrado en eritroblastos⁴¹.

Además, fueron detectados cambios en la composición de la membrana nuclear. Específicamente, la lámina A/C se pierde en la envoltura nuclear de los aRBC. En este trabajo discutió que su pérdida podría estar relacionada con la reorientación de la cromatina accesible hacia la periferia nuclear pues, en células fotorreceptoras de bastón en mamíferos de vida nocturna se observa una reorganización similar de su arquitectura nuclear consecuencia de la pérdida de la laminas A/C y B⁴¹. Por otro lado, los cambios en la composición de la membrana nuclear han sido relacionados en otros estudios con alteraciones de la forma del núcleo. Un ejemplo interesante es lo que ocurre con Matrin3. En células U2OS fue descrito que la desregulación de la lámina A/C afecta directamente a la expresión de Matrin3, provocando una alteración en la morfología nuclear⁴².

2.2.2. Los RBC terminalmente diferenciados retienen actividad transcripcional

Los RBCs de pollo han sido referidos principalmente como células transcripcionalmente inertes debido a la dramática disminución en la producción de RNA⁴³. No obstante, a raíz de la observación de cromatina accesible mediante ATAC-see, se caracterizó el estado transcripcional de RBC de pollo.

Experimentos de cuantificación de RNA total y perfiles electroforéticos confirmaron un decremento drástico en la transcripción en eRBC y aRBC, concordante con estudios previos. Al realizar experimentos de Click-iT RNA para marcar el RNA naciente, mediante la incorporación de 5-etinil urudina a RNA naciente, se pudo confirmar la presencia de transcripción activa en estas células⁴¹.

A través de RNA-seq se encontró una desregulación a la baja de la expresión genética en eRBC y aRBC, comparado con los eritroblastos. El pequeño subconjunto de genes que mantenían su expresión en eRBC (1354) y aRBC (804) corresponde a genes asociados a funciones eritroides, como el transporte de oxígeno, diferenciación eritroide y respuesta inmune⁴¹.

2.2.3. El pausado de la RNA polimerasa II mantiene accesibles promotores de genes pausados

Derivado de estos experimentos, se concluyó que los focos transcripcionales observados en RBCs que no presentan un enriquecimiento preferencial en la periferia nuclear, no pueden explicar la presencia de cromatina accesible en la periferia nuclear de RBCs de pollo. Sin embargo, se hipotetizó que el pausado proximal al promotor de la RNA polimerasa II podría mantener accesibles regiones del genoma y estructuración local⁴¹. El análisis de RNA-seq mostraba que genes como HEXIM1 y 7SK, involucrados en el pausado proximal de la RNA pol II, retienen su expresión en eRBC y aRBC haciendo plausible dicha hipótesis. Experimentos adicionales CHIP-seq de la RNA pol II pausada (Ser5-P), de la mano con ATAC-seq, confirmaron la presencia de RNA pol II en su estado de pausado proximal permitiendo mantener la accesibilidad de promotores de genes sin expresión en eRBC y aRBC⁴¹.

2.2.4. Pérdida generalizada de la estructura del genoma eritroide

Estudios de Hi-C en distintos estadios de la diferenciación eritroide en pollo permitieron evidenciar la reconfiguración global de genoma e integrarlos con los cambios observados en la arquitectura nuclear, accesibilidad de la cromatina y expresión genética.

Conforme la diferenciación terminal y maduración transcurren, ocurre un incremento en la segregación en los compartimentos A y B visible en los mapas de contactos de Hi-C. A diferencia de los eritroblastos, los eRBC y aRBC presentan un patrón de ajedrez más prominente, indicativo de la progresiva compactación de su genoma. Además, las

interacciones de largo alcance (>10 Mb) incrementan en los RBCs. Aunado a esto, la identificación de los compartimentos A y B mostró la formación de nuevos compartimentos Bs acompañada de la disrupción de As en RBC. Por tanto, se evidenció una esperada hiper-compartimentalización del genoma en eRBC y aRBC, concordante con la disminución de su estado transcripcional (**Figura 3a**)⁴¹.

Adicionalmente, la identificación de TADs reveló una reducción generalizada en el número de dominios cromatínicos en eRBC (916) y aRBC (903), en contraste con los eritroblastos (2,562) (**Figura 3b**)⁴¹. La pérdida generalizada de TADs en eritrocitos maduros coincide con reportes previos⁴⁴. Los dominios identificados presentan un tamaño promedio de 241 Kbs, haciéndolos más pequeños al compararlos con los TADs encontrados en eritroblastos. Por tal motivo, se han referido a estos dominios cromatínicos tipo TAD de menor tamaño como “minidominios”. En general, se encontró que, aunque la mayoría de estructura se pierde en RBC, alrededor de un 24.18% y 20.4% del genoma de eRBC y aRBC, respectivamente, se mantiene estructurado en minidominios (**Figura 3c-d**)⁴¹.

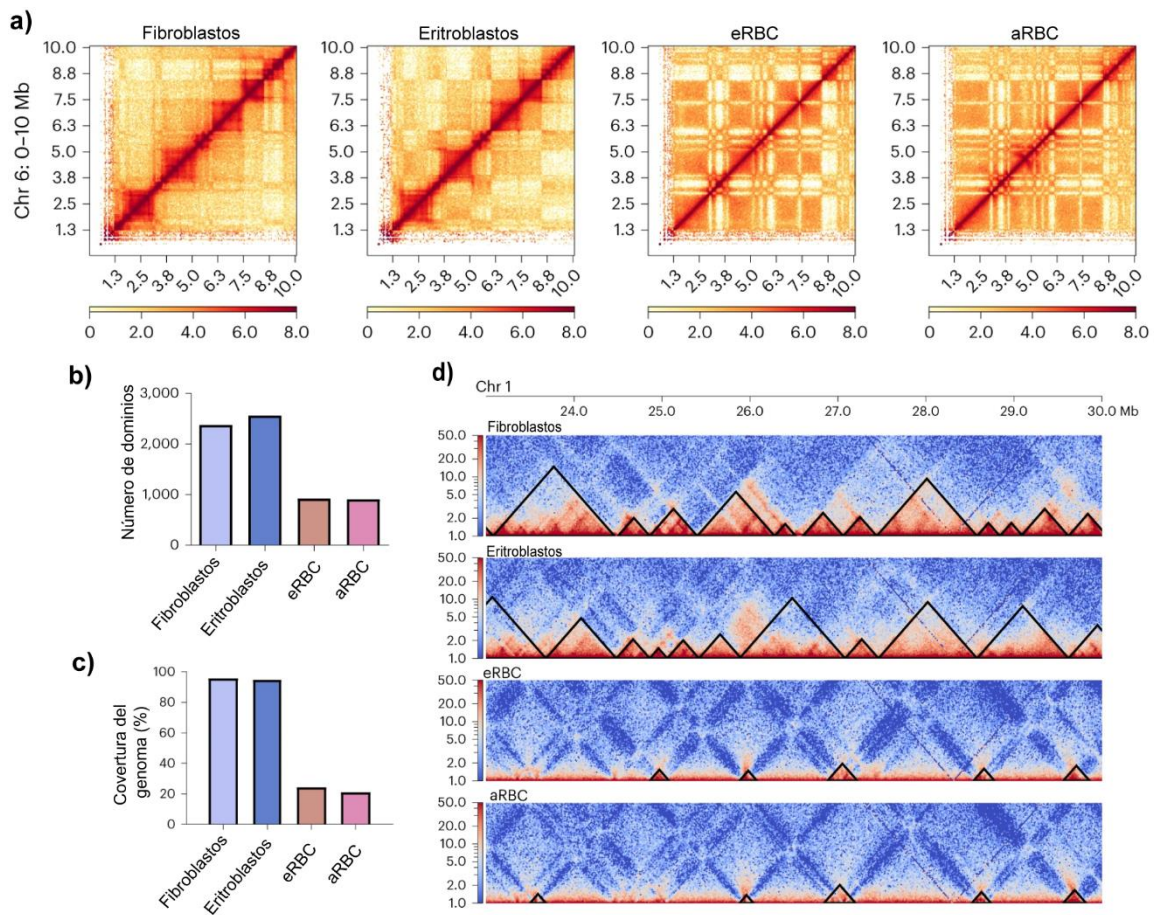


Figura 3. Reorganización global del genoma eritroide durante la diferenciación terminal y maduración. a) Mapas de contactos de Hi-C del cromosoma seis donde es posible observar los TADs a simple vista en fibroblastos y eritroblastos. En eRBC y aRBC los TADs son imperceptibles. Además, se observa el patrón de ajedrez correspondiente a compartimentos, lo que refleja la hiper-compartmentalización que los genomas eritroides sufre en estadios terminalmente diferenciados. **b)** Número de dominios. **c)** Porcentaje del genoma que permanece estructurado. **d)** Ejemplos de TADs y los mini dominios. Figura modificada de: Penagos-Puig, A. (2023)⁴¹.

De forma interesante, los promotores de genes pausados con accesibilidad involucrados con procesos como el procesamiento de RNA, regulación de la expresión genética y transporte de RNA, se encuentran estructurados dentro de los minidominios en eRBC y aRBC⁴¹. La estructuración en mini-dominios de loci pausados con la presencia de la RNA polimerasa II pausada podrían preparar a los eritrocitos de pollo para montar una respuesta transcripcional rápida ante estímulo externos (**Figura 4**)⁴¹.

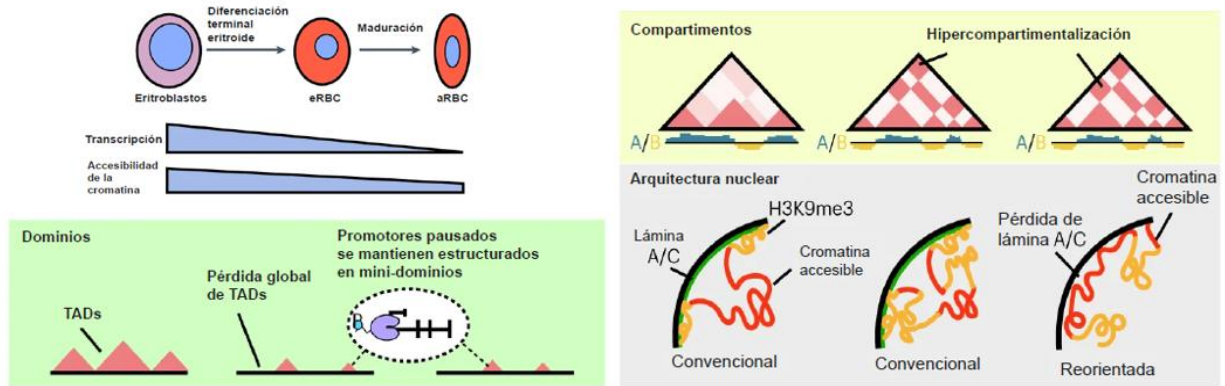


Figura 4. Modelo de los cambios en la organización del genoma durante la diferenciación y maduración eritroide en *Gallus gallus*. La compactación del núcleo a través de la diferenciación terminal resulta en una dramática, pero no total, pérdida en la accesibilidad y la transcripción. A medida que el eritrocito madura, estos fenómenos se exageran. A nivel topológico, los genomas de eritrocitos embrionarios (eRBC) y eritrocitos adultos (aRBC) pierden TADs de forma global y se hipercompartimentalizan. El genoma que permanece estructurado en minidominios se posiciona alrededor de promotores pausados y accesibles y con presencia de la RNA polimerasa II pausada. Además, la pérdida de la lámina A/C provoca una reorientación de la arquitectura nuclear en aRBC. Figura modificada de: Penagos-Puig, A. (2023)⁴¹.

2.3. Pérdida de las principales proteínas arquitectónicas durante la diferenciación y maduración eritroide

Como se mencionó previamente, experimentos funcionales han demostrado la relevancia de CTCF y proteínas del complejo cohesina en el establecimiento de TADs^{15,17,20,21}. Por tanto, la pérdida generalizada de TADs en eRBC y aRBC es concordante con la disminución en los niveles de expresión de los transcritos de proteínas arquitectónicas como CTCF, la lámina nuclear y subunidades del complejo cohesina^{41,45}.

A su vez, se propuso que la estructuración remanente del genoma en eRBC y aRBC podría relacionarse con la ocupación residual de otras proteínas arquitectónicas. De todos los niveles de expresión estudiados de proteínas estructurales en vertebrados, YY1 es la única proteína que conserva su expresión tanto en eRBC y aRBC (**Figura 5a**)⁴⁵. Al evaluar los niveles proteicos mediante western blots se reveló que las proteínas CTCF, SMC1 y lamina B1 disminuyen en estadios avanzados de la diferenciación y maduración eritroide a niveles casi indetectables o indetectables (**Figura 5b**)^{41,45}.

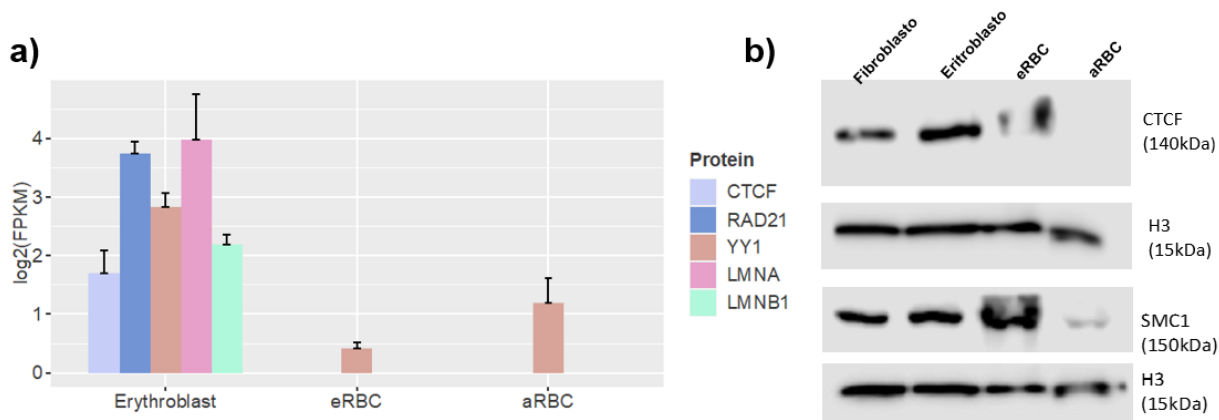


Figura 5. Los eritrocitos terminalmente diferenciados y maduros pierden a las principales proteínas estructurales. **a)** Niveles de expresión obtenidos mediante RNA-seq de las proteínas arquitectónicas más importantes en vertebrados. Sólo YY1 presenta expresión en eRBC y aRBC. **b)** Westerns blots de CTCF y SMC1 (complejo cohesina) demuestran la disminución o pérdida de proteínas arquitectónicas en eritrocitos. Figuras tomadas de: Claudio Galeana, G. S. (2020)⁴⁵.

2.4. YY1 como proteína arquitectónica emergente

YY1 es un factor transcripcional ubicuamente expresado en células de mamíferos con capacidad contexto-dependiente de funcionar como activador o represor transcripcional²⁶. Este TF está implicado en la regulación de múltiples procesos relacionados con el desarrollo, mecanismos de reparación de DNA, proliferación celular, replicación y transcripción²⁶. Además, la alteración en su funcionalidad ha sido ligada a procesos carcinogénicos^{26,46}.

YY1 presenta distintos dominios que le confieren sus funciones duales activador-represor. En humanos, YY1 está compuesta de 414 aminoácidos, mientras que en pollo presenta 420⁴⁷. Aunque la proteína YY1 del pollo no ha sido ampliamente estudiada, sus dominios preservan alta similitud en la secuencia de ambos organismos (**Figura suplementaria 1**)⁴⁸. En el extremo C-terminal podemos encontrar a 4 dedos de zinc de tipo C2H2 que le permiten la unión a DNA²⁶. En la región central se encuentran los dominios REPO y GK, ambos asociados a su función de represión transcripcional al interactuar con proteínas de la familia de Polycomb y HDACs, respectivamente²⁶. Por otro lado, el extremo N-terminal de la proteína que media la activación transcripcional, se encuentra constituido principalmente por aminoácidos ácidos y un clúster de histidinas²⁶.

De forma reciente, se caracterizó la capacidad de YY1 de regular la transcripción a través de la modulación de la topología genómica⁴⁹. Dado que CTCF no es el principal

responsable de formar asas de cromatina que permiten la interacción entre elementos reguladores, se comenzó una búsqueda de candidatos que mediaran la formación de asas funcionales. Se propuso que uno de los candidatos podría ser YY1 ya que, al igual que CTCF, es indispensable para la supervivencia celular y se encuentra ubicuamente expresado en muchos tejidos⁴⁹. Además, a nivel estructura, YY1 posee dedos de zinc que le permiten unirse a secuencias de DNA hipometilado, análogamente a CTCF^{26,49,50}.

A través de experimentos de dimerización *in vitro* y ChIA-PET, se concluyó que YY1 tienen la capacidad de promover interacciones entre promotores y *enhancers* al homodimerizarse⁴⁹. El mecanismo encontrado es análogo al de CTCF. Además, la perturbación de sitios de unión a YY1 resulta en la pérdida de contactos promotor-*enhancer* y alteraciones en la expresión génica. También, fue posible demostrar que YY1 se posiciona preferencialmente en regiones de *enhancers* y promotores, a diferencia de CTCF, al comparar los perfiles de ocupación genómica de YY1 con marcas correspondientes a elementos reguladores⁴⁹.

Como se mencionó anteriormente, la separación de LLPS es un fenómeno con implicaciones emergentes en la estructuración del genoma⁵¹. En una investigación reciente, fue descrita la capacidad de YY1 de formar LLPS *in vitro* e *in vivo*, gracias a la presencia del clúster de histidinas ubicado en la región N-terminal⁴⁶. Además, las predicciones de AlphaFold2 sobre la estructura terciaria de YY1 sugieren la presencia de distintas regiones intrínsecamente desordenadas (IDRs) (**Figura suplementaria 1**), indispensables para la LLPS^{29,51,52}. En general, se demostró que los condensados formados por la LLPS de YY1 pueden reclutar a coactivadores transcripcionales como EP300, BRD4, MED1 y la RNA polimerasa II⁴⁶. En consecuencia, los condensados provocan un acercamiento espacial de elementos reguladores que culminan en la activación de genes como FOXM1⁴⁶.

3. Planteamiento

A pesar de la pérdida generalizada de estructura del genoma en eRBC y aRBC, parte de este permanece estructurado en minidominios. Además, la estructuración de promotores accesibles de genes sin expresión con la RNA polimerasa II en estado pausado sugieren un papel funcional relevante, posiblemente permitiendo montar respuestas transcripcionales ante estímulos externos.

Considerando que los niveles de expresión génica y proteica de las principales proteínas arquitectónicas en vertebrados están disminuidos en eRBC y aRBC, resulta indispensable evaluar si YY1 pudiera contribuir en la estructuración remanente del genoma eritroide.

4. Pregunta de investigación

¿De qué forma contribuye YY1 en la estructuración del genoma de eritrocitos terminalmente diferenciados y maduros en *Gallus gallus*?

5. Hipótesis

La proteína arquitectónica YY1 se preserva en la cromatina y contribuye en la estructuración del genoma durante la diferenciación terminal y maduración eritroide en *Gallus gallus*.

6. Objetivos

6.1. Objetivo general

Estudiar la participación de YY1 en la estructuración del genoma en la diferenciación y maduración eritroide en *Gallus gallus*.

6.2. Objetivos específicos

1. Analizar la abundancia de la proteína estructural YY1 en eritrocitos de *Gallus gallus*.
2. Determinar la localización nuclear de la proteína YY1.
3. Elucidar la ocupación genómica de YY1 en la fibra de cromatina.
4. Análisis de datos de ChIP-seq de YY1 e integración de los resultados con datos genómicos de ATAC-seq, RNA-seq y Hi-C mediante análisis bioinformáticos.

7. Métodos

7.1. Obtención de muestras y cultivo celular

7.1.1. Extracción de sangre periférica de gallinas adultas (aRBC)

Para obtener muestras de aRBC se sacrificaron dos gallinas (replicados biológicos) de 38 semanas. Cada gallina fue inmovilizada y colocada en una guillotina. Una vez decapitada, se recuperan aproximadamente 50 mL de sangre en 200 mL de buffer de colecta (PBS 1X/EDTA 0.5M) manteniendo la muestra siempre en hielo. Posteriormente, 50 mL de sangre fueron centrifugados a 1000 rpm por 10 min a 4°C y se descartó el sobrenadante. El pellet de eritrocitos fue resuspendido en 30 mL de buffer de colecta y se contaron las células utilizando una cámara de Neubauer. Al finalizar el conteo, se realizaron pellets de 200 millones de células para extracción de proteínas y se almacenaron a -80°C. También se fijaron con formaldehído pellets de células y células sembradas sobre cubreobjetos en cajas de 6 pozos de acuerdo con los protocolos de ChIP-seq e inmunofluorescencia, respectivamente.

7.1.2. Extracción de sangre periférica de embriones de pollo

Se obtuvieron muestras de eRBC de pollo de 10 días de desarrollo mediante punción apical de venas superficiales. Para esto, el cascarón de un huevo se limpió con alcohol y posteriormente, con ayuda de pinzas estériles, se retiró el cascarón del lado de la cámara de aire y la membrana que separa la clara de la cámara de aire. Esto permitió dejar al descubierto los vasos sanguíneos del saco vitelino y realizar una punción con una aguja. La sangre fue inmediatamente recolectada con una pipeta de 200 uL y depositada en 10 mL de buffer de colecta (PBS 1X/EDTA 0.5 M) mantenido en hielo. La sangre fue centrifugada y lavada dos veces en buffer de colecta y se realizó el conteo en la cámara de Neubauer. Se realizaron pellets de 200 millones de células para la extracción de proteínas y se fijaron pellets de 150 millones de células con formaldehído para ChIP-seq. También se sembraron y fijaron células en cubreobjetos para inmunofluorescencia.

7.1.3. Cultivo primario de fibroblastos de pollo

Se extrajo de su huevo a un embrión de pollo de 10 días de desarrollo y con ayuda de fórceps y un estereoscopio, se obtuvo el tejido del torso y las extremidades. El tejido muscular fue disgregado en pequeños fragmentos en una caja de petri con PBS 1X, excluyendo todo el tejido óseo y piel. Se colectó el tejido disgregado en un tubo falcon de

50 mL, y se adicionaron 15-20 mL de tripsina/EDTA 1X para ser incubados a 37°C por 15 minutos o hasta que la mayoría de los agregados de tejido desaparecieron. Seguidamente, el sobrenadante se transfirió a un nuevo tubo falcon y se inactivó la tripsina/EDTA adicionando 1:1 de suero fetal bovino. El sobrenadante fue centrifugado a 1000 g por 5 minutos a temperatura ambiente (RT). El pellet de células fue resuspendido en medio de cultivo HD3 (DMEM/10% FBS/2% suero de pollo/1% penicilina/estreptavidina) y transferido a una caja de cultivo. Las células se mantuvieron en la incubadora en 5% de CO₂ a 37°C. Al día siguiente se hizo cambio del medio de cultivo para eliminar restos celulares muertos. Una vez que el cultivo alcanzó la confluencia, se obtuvieron pellets de células para extracción de proteínas o cultivaron sobre cubreobjetos para realizar inmunofluorescencia.

7.1.4. Cultivo celular de eritroblastos de pollo

Se empleó la línea celular HD3 correspondiente a eritroblastos de pollo adulto. Las células se cultivaron en suspensión utilizando medio HD3 y se mantuvieron en incubación en 5% CO₂ a 37°C. Una vez alcanzada la confluencia, las células fueron procedas para obtener pellets para extracción de proteínas, fijadas para ChIP-seq o sembradas en cubreobjetos para inmunofluorescencias.

7.2. Western blot

7.2.1. Extracción de proteínas

Para este fin, se utilizaron pellets de células HD3, eRBC y aRBC. La fracción nuclear de proteínas se obtuvo mediante un fraccionamiento núcleo-citoplasma. Primero, los núcleos fueron aislados utilizando buffer A (HEPES 1M, KCL 1M, MgCl₂ 1M, sacarosa 1M, glicerol y H₂O destilada), suplementado con un coctel de inhibidores de proteasas, que permite lisar la membrana celular. Una vez aislados mediante centrifugación, el pellet de núcleos fue resuspendido en buffer RIPA (Tris-HCl 1M, EDTA 0.5M, EGTA 0.5M, Tritón X-100 10%, deoxicolato de sodio 10%, SDS 10%, NaCl 5M, y H₂O destilada) suplementado con inhibidores de proteasas, y se incubó en hielo por 15 min pasándolo por vortex cada 3 min. Luego, las muestras fueron sonicadas por 1 min 30 seg ON y 30 seg OFF a 30% de amplitud. Finalmente, las muestras se centrifugaron a 14,000 g por 15 min a 4°C y se recolectó el sobrenadante; las proteínas fueron cuantificadas a través del método de Bradford. Los extractos de proteínas fueron almacenados a -80°C.

7.2.2. Inmunoblot

Con la finalidad de evaluar la abundancia de YY1 a nivel proteínica se realizó WB. Se cargaron 50 µg de los extractos de proteínas nucleares de los tipos celulares estudiados en un gel de poliacrilamida para ser separados por tamaños mediante electroforesis. Posteriormente, las proteínas en el gel fueron transferidas a una membrana de PVDF y esta se incubó con buffer de bloqueo (TBST 1X/Leche 5% w/v) por 2 h en agitación constante. Terminado el tiempo de bloqueo, la membrana se lavó tres veces con TBST. Consecutivamente, la incubación se realizó con el anticuerpo primario anti-YY1 (D5D9Z, Cell Signaling) a 1:1000 en buffer de dilución (TBS 1X/Tween 20 0.1%/BSA 5% w/v) toda la noche (ON) en agitación suave a 4°C. Al día siguiente, la membrana se lavó tres veces con TBST y se incubó el anticuerpo secundario (HRP) en agitación durante 1 h a RT. Finalmente, la membrana se lavó tres veces con TBST y se reveló utilizando luminol en el equipo C-DiGit Blot Scanner.

7.3. Inmunofluorescencia

Las células aRBC, eRBC y HD3 fueron fijadas con formaldehído al 4% sobre cubreobjetos tratados con poli-L-lisina. Por otro lado, los fibroblastos de pollo fueron sembrados y cultivados sobre cubreobjetos sin tratar por 3 días y después fueron fijados de la misma manera. Las células se incubaron con buffer de bloqueo (BSA 5%/PBS 1X) por 1 h a RT. Terminado el periodo de bloqueo, el sobrenadante se descartó y se agregó el anticuerpo primario anti-YY1 (D5D9Z, Cell Signaling) diluido 1:800 para incubarse a 4°C ON. Al día siguiente, el anticuerpo se retiró y se lavó tres veces con buffer de lavado. Inmediatamente, se agregó la dilución del anticuerpo secundario acoplado al fluoróforo FITC y se incubó por 1 h a RT. Al terminar, la dilución del anticuerpo se retiró y las células se lavaron 3 veces más para eliminar cualquier remanente del anticuerpo secundario no unido. Se procedió a incubar a las células con DAPI por 20 min a RT. Transcurrido el tiempo, el montaje de los cubreobjetos se realizó sobre portaobjetos utilizando Fluoromount (Sigma) y permitiendo su secado por 30 min a RT. Las preparaciones fueron analizadas mediante el uso del microscopio confocal Zeiss LSM800 a 60X. Las imágenes obtenidas fueron analizadas utilizando el software Fiji.

7.3.1. Análisis de imágenes

Todas las imágenes de inmunofluorescencia fueron procesadas utilizando el software de Fiji. Para el análisis de intensidad de fluorescencia se realizó una proyección en Z utilizando el método Sum Slices. Subsecuentemente, los bordes del núcleo se delimitaron mediante una máscara binaria generada utilizando el canal DAPI y la aplicación de filtro gaussiano. Una vez delimitadas todas las regiones de interés (ROIs; i.e. núcleos), la intensidad de la señal de cada uno de los núcleos se cuantificó. Por tipo celular fueron considerados 90 núcleos.

Por otro lado, los análisis de colocación entre las regiones de mayor intensidad de píxeles entre los canales de YY1 y DAPI se realizaron mediante el plugin JaCoP de Fiji. Adicionalmente, las máscaras binarias de los mismos canales que fueron sobrepuestas fueron generadas a través de un filtro para seleccionar aquellos píxeles que fueron más brillantes que el promedio.

Finalmente, se generaron mapas de distancia radial núcleo-periferia para evaluar la distribución preferencial de YY1. Estos mapas de distancia se basan en la segmentación del núcleo en coronas desde la periferia hacia el centro del núcleo. Para cada corona se calculó la intensidad de fluorescencia promedio. De esta forma, fue posible graficar la intensidad de fluorescencia en función a la distancia hacia la periferia nuclear. Para poder integrar y comparar los resultados entre tipos celulares, se realizaron normalizaciones en función al tamaño celular y la intensidad máxima de cada fluoróforo estudiado. Por cada tipo celular consideramos 90 núcleos. El análisis de estos datos y su procesamiento se llevaron a cabo en RStudio. Las gráficas muestran los trazos individuales de la distribución de señal de cada tipo celular y su respectivo trazo promedio.

7.4. ChIP-seq

7.4.1 Diseño de oligonucleótidos control de regiones de unión de YY1 para ChIP-seq

Se diseñaron y sintetizaron oligonucleótidos de una región ocupadas por YY1 y otra de un sitio donde no se uniera, previamente reportadas en la literatura (**Tabla 1**)⁵³.

Tabla 1. Oligonucleótidos empleados como regiones control para CHIP-seq

| Oligonucleótido | Secuencia | Descripción | Tamaño |
|------------------------------|----------------------------------------------------|---------------------------------------------------------------------------------------------|--------|
| YY1 control positivo (pollo) | A120P_Forward: GACGTGGGCAGCAGATAGCCTCG | Sitio de unión de YY1 en el <i>enhancer</i> de la región 3' del locus de α -globina. | 169 pb |
| | A120P_Reverse: GACCCGTGGTAACCCAGGCC | | |
| YY1 control negativo (pollo) | $\beta^{A/\epsilon}$ _Forward: CTGGGTGGGGGCAGGT | Región de una porción del <i>enhancer</i> β -globina. | 303 pb |
| | $\beta^{A/\epsilon}$ _Reverse: CTGCTTTTGCTGCCCTGTG | | |
| IFN- α | IFNA_Forward: CAACGACACCATCCTGGACA | Primer de expresión contra IFN- α . Usado como control interno de la qPCR. | 147 pb |
| | IFNA_Reverse: ATCCGGTTGAGGAGGCTTTG | | |

7.4.2. Inmunoprecipitación de la cromatina

Dos replicados de pellets HD3, eRBC, aRBC e hígado de ratón (control externo) fueron fijados con 1% de formaldehído durante 10 min para los experimentos de inmunoprecipitación de la cromatina (ChIP). Adicionalmente, se fijaron pellets de las células S2R+ para controles del ChIP. Una prueba inicial del ChIP se realizó únicamente en eRBC utilizando los anticuerpos grado ChIP anti-YY1 (D5D9Z, Cell Signaling) y anti-YY1 (C15410345, Diagenode) con el objetivo de determinar cuál era mejor para el experimento a secuenciar. Las muestras fueron lisadas y sonicadas de 3-7 ciclos a 5 s ON/9.9 s OFF a 30% de amplitud. Antes de iniciar cada ChIP se adicionó 1% de cromatina de *D. megalogaster*, proveniente de la línea celular S2R+, como *Spike-In*. Cada ChIP fue ejecutado utilizando 10 μ l de anticuerpo anti-YY1 (D5D9Z, Cell Signaling) por cada inmunoprecipitación (IP), y perlas magnéticas Dynabeads Protein G (Invitrogen, 10003D). Terminada el periodo de IP, se revirtió el entrecruzamiento incubando las muestras con 5 μ l de RNAsa A y 22 μ l de PK, a 37°C por 2 h y 65°C ON, respectivamente. Finalmente, el DNA se purificó utilizando el kit de purificación ChIP DNA Clean and Concentrator (Zymo Research, D5205) y el material fue cuantificado mediante Qubit.

7.4.3. PCR punto final y qPCR

Los oligonucleótidos sintetizados (**Tabla 1**) fueron estandarizados mediante PCR punto final y PCR cuantitativa (qPCR). Se utilizaron 20 ng de DNA genómico de pollo para cada prueba. Subsecuentemente, estos oligonucleótidos se emplearon como control positivo y negativo del experimento ChIP mediante qPCR. Para ambos pares de oligonucleótidos, las temperaturas de *melting* (T_m) determinadas fueron 60°C. Para estas pruebas fueron

utilizados 20 ng de input de DNA genómico y 4 µl de DNA purificado del experimento de CHIP. Las qPCRs se realizaron utilizando el kit KAPA SYBR Fast qPCR. Cada experimento fue sometido a 40 ciclos de PCR y los valores de Cts (del inglés *threshold cycle*) fueron obtenidos. Finalmente, el enriquecimiento se determinó mediante el método de porcentaje de input.

7.4.3. Preparación de bibliotecas con material de CHIP para secuenciación

Finalmente, las bibliotecas de CHIP-seq se prepararon utilizando el kit de NEBNext Ultra™ DNA Library Prep Kit (Illumina) utilizando los barcodes de la **Tabla 2**. La calidad de cada biblioteca y el tamaño de los fragmentos a secuenciar se validaron mediante electroforesis capilar utilizando el TapeStation System. Subsecuentemente, las bibliotecas fueron enviadas a secuenciación.

Tabla 2. Barcodes utilizados en las bibliotecas para secuenciación.

| Muestra | Barcodes i5-i7 | |
|------------------------|----------------|-------|
| eRBC réplica 1 | i-501 | i-701 |
| eRBC réplica 2 | i-502 | i-701 |
| aRBC réplica 1 | i-501 | i-705 |
| eRBC réplica 2 | i-502 | i-705 |
| HD3 réplica 1 | i-501 | i-704 |
| HD3 réplica 2 | i-502 | i-704 |
| Input pollo | i-501 | i-703 |
| Hígado ratón réplica 1 | i-501 | i-702 |
| Hígado ratón réplica 2 | i-502 | i-702 |
| Input ratón | i-502 | i-703 |

7.5. Análisis bioinformático

7.5.1. Pipeline general de CHIP-seq y análisis subsecuentes

Una vez obtenidas las secuencias, los archivos FASTQ recibidos fueron sometidos a análisis de calidad de secuenciación usando el programa FASTQC. Los resultados fueron compactados en un solo reporte mediante MultiQC. Adicionalmente, los replicados biológicos fueron sometidos a un análisis de correlación usando la herramienta

plotCorrelation de deepTools. También, las secuencias fueron sometidas a un proceso de remoción de duplicas mediante el algoritmo de TrimGalore.

Subsecuente a los análisis de calidad, las lecturas fueron mapeadas mediante Bowtie2 a su genoma de referencia correspondiente de pollo GRCg6a o ratón GRCm39, obtenidos de Ensembl. Las duplicas presentes en los archivos de alineamientos fueron removidas.

Los archivos sin duplicas fueron normalizadas a través del método Spike-In. Para esto, las lecturas de los experimentos de pollo y ratón fueron mapeadas contra el genoma de *Drosophila melanogaster* BDGP6.46, para obtener el número de lecturas únicas que alinean a este genoma. Esto permite calcular un factor “*downsampling*” para cada una de las muestras al dividir el número de lecturas únicas encontradas contra el número menor encontrado en cada grupo. Finalmente, el archivo de alineamiento normalizado se obtuvo para cada experimento al multiplicar el archivo sin duplicas por el factor “*downsampling*”⁵⁴.

El llamado de picos se realizó mediante los métodos *BroadPeaks* y *NarrowPeaks* utilizando MACS2 con $q = 0.05$. Los picos identificados fueron filtrados para remover falsos positivos mediante valores individuales de lecturas mayores a 5, mediante SeqMonk. Las listas de picos obtenidos fueron sometidas a anotación y búsqueda de motivos mediante HOMER. Finalmente, los picos compartidos se buscaron mediante InterVene.

7.5.2. Integración de datos de topología, expresión y accesibilidad

Se integraron los resultados del ChIP seq de YY1 con los datos de topología, transcripción y accesibilidad. Se generaron los archivos necesarios, principalmente archivos bed, dependiendo del tipo de datos a integrar y se utilizaron las herramientas InterVene o Intersect de bedtools. El objetivo primordial era conocer el posicionamiento de YY1 respecto a la estructuración del genoma, estado cromatínico y el estado transcripcional de las regiones que ocupa.

Los comandos generales fueron los siguientes.

Intersect:

```
bedtools intersect -wa -a A1.bed -b A2.bed > A1xA2.bed
```

Intervene:

```
intervene venn -i A1.bed A2.bed --type genomic --names=A1,A2 -bedtools -options wa --output  
Intersect --save-overlaps
```

A1 y A2 son los archivos bed de interés. Realizamos las siguientes intersecciones en nuestros sets de datos:

- TADsxChIPseqYY1
- ChIPseqYY1xTADs
- FronterasxYY1ChIPseq
- ChIPseqYY1xFronteras
- ATACseqxChIPseqYY1
- ChIPseqYY1xATACseq
- CompartimentosxChIPseqYY1
- ChIPseqYY1xCompartimentos
- ChIPseqYY1xRNAseq
- RNAseqxChIPseqYY1
- SetAleatorioxTADs/Compartimentos/Fronteras

Los datos de expresión fueron cruzados dividiendo a los genes en “features” como promotores, intrones, exones y regiones intergénicas.

Para la visualización de los datos integrados utilizamos el software pyGenomeTracks⁵⁵. Después de generar los archivos .ini que contenían los datos de Hi-C, RNA-seq, ChIP-seq, ATAC-seq y las listas de genes de interés utilizamos el siguiente comando para generar el gráfico:

```
pyGenomeTracks --tracks archivo.ini --region chr#:00000000-00000000 --dpi 300  
-o region.pdf
```

8. Resultados

8.1. Los niveles de proteína YY1 disminuyen a lo largo de la diferenciación y maduración eritroide

Basados en observaciones previas de experimentos de RNA-seq y Western blot donde se encontró que la expresión o abundancia de las principales proteínas arquitectónicas disminuían a lo largo de la diferenciación o eran prácticamente indetectables. Por ello, se decidió investigar si otra proteína remanente en estadios terminalmente diferenciados y maduros podría estar relacionada con la estructuración del genoma en mini dominios. La presencia de YY1 en la cromatina podría, además, estar relacionada con los sitios de accesibilidad encontrados en el genoma. De acuerdo con los antecedentes conocidos de YY1 como proteína estructural y que sus niveles de expresión no desaparecían como las del resto de proteínas arquitectónicas, se decidió investigar si, a nivel de proteína, YY1 estaba presente en los estadios de diferenciación estudiados. Por tanto, realizamos experimentos de Western blot contra YY1 utilizando extractos nucleares para determinar su abundancia.

Se encontró que, conforme los procesos de diferenciación terminal y la maduración eritroide avanzan, los niveles de la proteína YY1 disminuyen significativamente, pero aún son detectables en todos los estadios. El estadio de eritroblasto representado por la línea celular HD3 exhibió los niveles más altos de proteína. En los estadios correspondientes a eritrocitos terminalmente diferenciados, eRBC, y terminalmente diferenciados maduros, aRBC, se observa una disminución dramática en los niveles de YY1 con respecto a los eritroblastos (**Figura 6a y Figura Suplementaria 2a**). Observamos que los niveles de proteína de YY1 en eRBC se reducen un 83.66%, y un 87.8% en aRBC, con respecto a los eritroblastos (**Figura 6b**). Entre ambos tipos de eritrocitos, eRBC exhibe niveles ligeramente mayores de proteína comparado con aRBC. En conjunto, los niveles de proteína de YY1 aún son detectables en eRBC y aRBC.

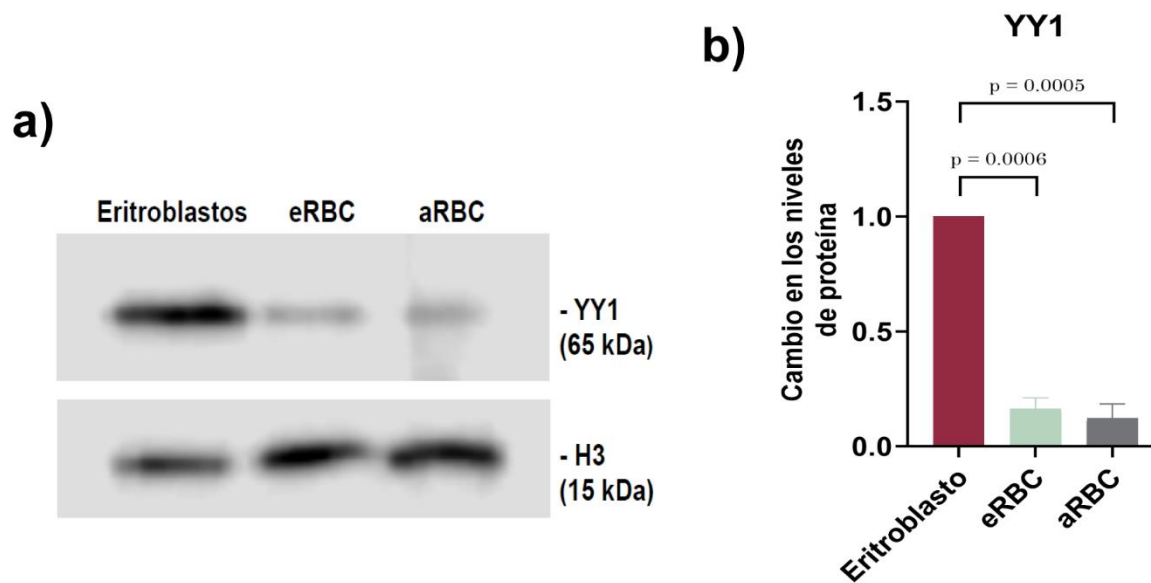


Figura 6. Los niveles de la proteína YY1 disminuyeron en aRBC y eRBC. a) Western blot de extractos nucleares contra YY1 en eritroblastos (células HD3), eRBC y aRBC. Una disminución notable de la proteína YY1 se observa en los estadios de diferenciación terminal y maduración. Como control de carga se utilizó a la histona 3 (H3) b) Cuantificación de los niveles de proteína. Se encontró una disminución estadísticamente significativa entre los eritroblastos y las células eritroides. Se realizó una prueba ANOVA de un factor como estadístico seguido de una prueba post-hoc de Dunnett's partiendo de dos replicados experimentales.

8.2. Evaluación de los niveles de YY1 mediante inmunofluorescencia y localización nuclear

Dado el posible papel estructural de YY1 en el genoma eritroide, decidimos explorar mediante inmunofluorescencia como se distribuye su señal dentro del núcleo y corroborar la abundancia de la proteína. La cuantificación de los niveles de intensidad de fluorescencia realizada en los diferentes tipos celulares muestra una tendencia de disminución en los niveles de YY1 en eRBC y aRBC, respecto a los eritroblastos. Este resultado respalda los hallazgos encontrados por western blot. Para este experimento, se emplearon fibroblastos de pollo como control no eritroide y se estudiaron 90 núcleos provenientes de 3 réplicas para cada tipo celular (**Figura 7a y Figura Suplementaria 2b**).

Previas evaluaciones mediante inmunofluorescencia de la marca H3K9me3, asociada a heterocromatina constitutiva, mostraron un posicionamiento preferencial al centro del núcleo en focos en aRBC y eRBC, contrastante con el patrón convencional periférico encontrado en eritroblastos. Además, la reorientación de la cromatina accesible a la periferia nuclear en aRBC evaluada por ATAC-see y el enriquecimiento de la señal de

DAPI al centro del núcleo, son observaciones que hacen relevante estudiar si YY1 mostraba algún patrón de distribución similar. Las imágenes obtenidas mediante microscopia confocal mostraron que, en todos los tipos celulares estudiados, la señal de YY1 está distribuida en todo el núcleo. Sin embargo, en eRBC y aRBC es posible observar un patrón donde YY1 se encuentra concentrado en estructuras similares a foci (focos) que se localizan principalmente en zonas donde la intensidad de la señal de DAPI es menor (ver *merge*). No obstante, este patrón no se observa en los eritroblastos, los cuales exhiben una distribución más similar a los fibroblastos (**Figura 7b**).

Si bien la señal de YY1 está por todo el núcleo, dado a que la acumulación en focos de este factor transcripcional parece ser excluyente de sitios donde la cromatina es más compacta, se decidió generar máscaras binarias de los canales de YY1 y DAPI. El análisis de máscara binaria nos permite crear una caricatura del núcleo conservando únicamente los píxeles más brillantes de cada fluoróforo en cada uno de los canales. De esta forma, podemos sobreponer las máscaras de ambos canales, y explorar una potencial exclusión de las zonas con mayor intensidad de cada tinción. Así, fue posible confirmar que tanto en eRBC y aRBC, los píxeles más brillantes del canal YY1 (verde) presentan un posicionamiento principalmente excluyente de aquellos lugares donde se enriquecen los de DAPI (azul) (**Figura 7c**).

Adicionalmente, realizamos análisis de colocalización basado en el método de máscaras binarias empleando el plug-in de Fiji, JaCoP⁵⁶, permitiendo obtener perfiles de colocalización mediante la intersección de las máscaras binarias de los canales DAPI y YY1. Confirmamos que eRBC y aRBC presentan un menor grado de solapamiento entre los canales. Además, al obtener la distribución de píxeles dependientes (que se solapan) entre ambos canales mediante fluorogramas, es posible observar una mayor proporción de píxeles dependientes en fibroblastos y eritroblastos, a diferencia de las células eritroides (**Figura 7d**). El coeficiente de colocalización M1, basado en el coeficiente de correlación de Pearson, representa una medida del grado de colocalización entre los canales de YY1 (verde) y DAPI (azul). Este coeficiente reveló que las aRBC tienen el menor grado de colocalización entre ambos canales ($M = 0.471$) (Inferior **Figura 7d**). Es importante aclarar que también fueron detectadas pequeñas regiones en el núcleo donde YY1 se encuentra en zonas donde la cromatina está más compacta (enriquecimiento mayor de la señal de DAPI).

En conjunto, estos resultados indican que los niveles de YY1 en células eritroides son menores respecto al eritroblasto. Además, aunque YY1 está distribuido en todo el núcleo, las zonas con mayor enriquecimiento dentro del núcleo parecen ser regiones de eucromatina.

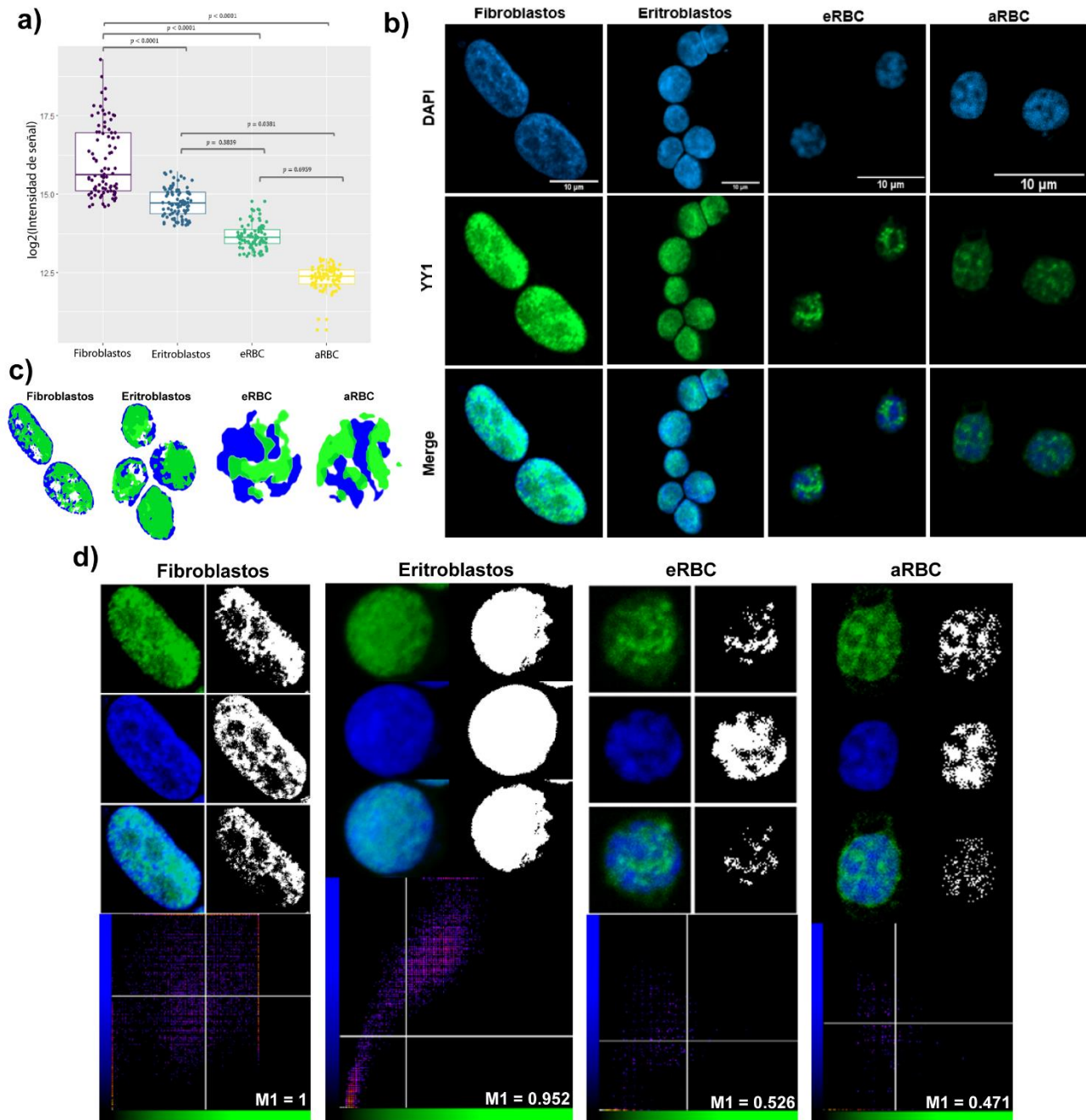


Figura 7. Análisis de la distribución de YY1 mediante inmunofluorescencia. a) Cuantificación de los niveles de intensidad de fluorescencia de YY1. En los eritrocitos se observa una tendencia a la disminución de los niveles de YY1, respecto a los eritroblastos o fibroblastos de pollo. $n=90$ núcleos por tipo celular. Se empleó una prueba ANOVA de un factor seguido de la prueba post-hoc Tukey. b) Inmunofluorescencia contra YY1 de los tres tipos celulares estudiados y fibroblastos (control no eritroide). La señal de YY1 se encuentra en todo el núcleo; sin embargo, en eRBC y aRBC, las regiones más intensas se posicionan en focos. c) Máscara binaria de los píxeles más brillantes de YY1 (verde) o DAPI (azul). En eRBC y aRBC, la señal de píxeles más brillantes de cada canal se segrega de forma independiente. d) Análisis de colocalización de señales. Los canales de DAPI y YY1 exhiben su respectiva máscara binaria. En el canal *merge*, la máscara binaria está constituida por los píxeles que se sobrelapan en ambos canales únicamente. Al inferior, se muestra los fluorogramas de los píxeles dependientes en cada tipo celular. M1 es un coeficiente de sobrelape basado en el coeficiente de Pearson, donde valores cercanos a 0 representan nula colocalización y 1 colocalización total.

8.3. Análisis de distribución nuclear de YY1 y comparación con marcas de activación y represión

Para investigar posibles patrones no evidentes de distribución de señal de YY1, se realizaron análisis de distancia centro-periferia del núcleo (**Figura 8a**). Noventa núcleos fueron considerados para cada tipo celular. La intención de estos análisis era detectar si existía alguna distribución nuclear preferencial en el núcleo, tal como se detectaron en la marca H3K9me3 o la señal de ATAC-seq, especialmente en eRBC y aRBC.

En estos mapas, 0 denota a la periferia nuclear y 100 el centro. Se grafican los trazos individuales de cada célula (líneas suaves) y el trazo promedio (línea fuerte). En concordancia con resultados previos, la señal de DAPI en fibroblastos y eritroblastos muestra un posicionamiento evidente a la periferia nuclear. Mientras que en las células eritroides terminalmente diferenciadas y maduras, la señal de DAPI es menos intensa en la periferia y mayor en el centro. Por otro lado, YY1 en fibroblastos y eritroblastos muestra un posicionamiento mayor hacia zonas céntricas y menor en zonas periféricas. Esta distribución es concordante a la arquitectura nuclear convencional. Si bien, en las células eRBC y aRBC no se observó un posicionamiento preferencial de YY1 a regiones periféricas, podemos apreciar ligeras diferencias entre eRBC y aRBC. Las eRBC presentan un patrón más similar a los eritroblastos, donde el trazo promedio de la señal de intensidad se encuentra por debajo del trazo promedio de DAPI en la periferia nuclear, e incrementa hacia el centro. Sin embargo, para aRBC la intensidad de la señal de YY1 se incrementaba ligeramente por arriba de la de DAPI en la periferia nuclear (ver **Figura 8a-aRBC**). Los patrones en la distribución nuclear de YY1 podrían estar relacionado con marcas asociadas a silenciamiento o activación transcripcional. Especialmente, podría jugar un papel importante en la reorientación de la arquitectura nuclear encontrada en aRBC.

Dado que YY1 es un factor transcripcional con funciones de activador o represor dependiendo de las proteínas con la que interactúe, se decidió comparar su perfil de distribución en eritroblastos y aRBC, contra las marcas de acetilación de la lisina 27 de la histona 3 (H3K27Ac) y H3K27me3 (**Figura 8b-c**). Para esto, se hizo uso de imágenes previamente obtenidas en otros experimentos en el laboratorio, las cuales fueron

procesadas para obtener sus respectivas máscaras binarias. De forma interesante, encontramos que las marcas H3K27me3 y H3K27Ac presentan un patrón de distribución similar al de YY1 en aRBC. Es decir, ambas marcas muestran un posicionamiento principal hacia regiones de cromatina menos compactas y con pequeñas regiones que sobrelapan con sitios más compactos. Por otro lado, H3K27me3 en eritroblastos muestra un enriquecimiento en zonas cercanas a la periferia, lo que es consistente con el posicionamiento periférico de heterocromatina en este tipo celular. Además, la marca H3K27Ac presenta un posicionamiento versátil entre regiones compactas y laxas.

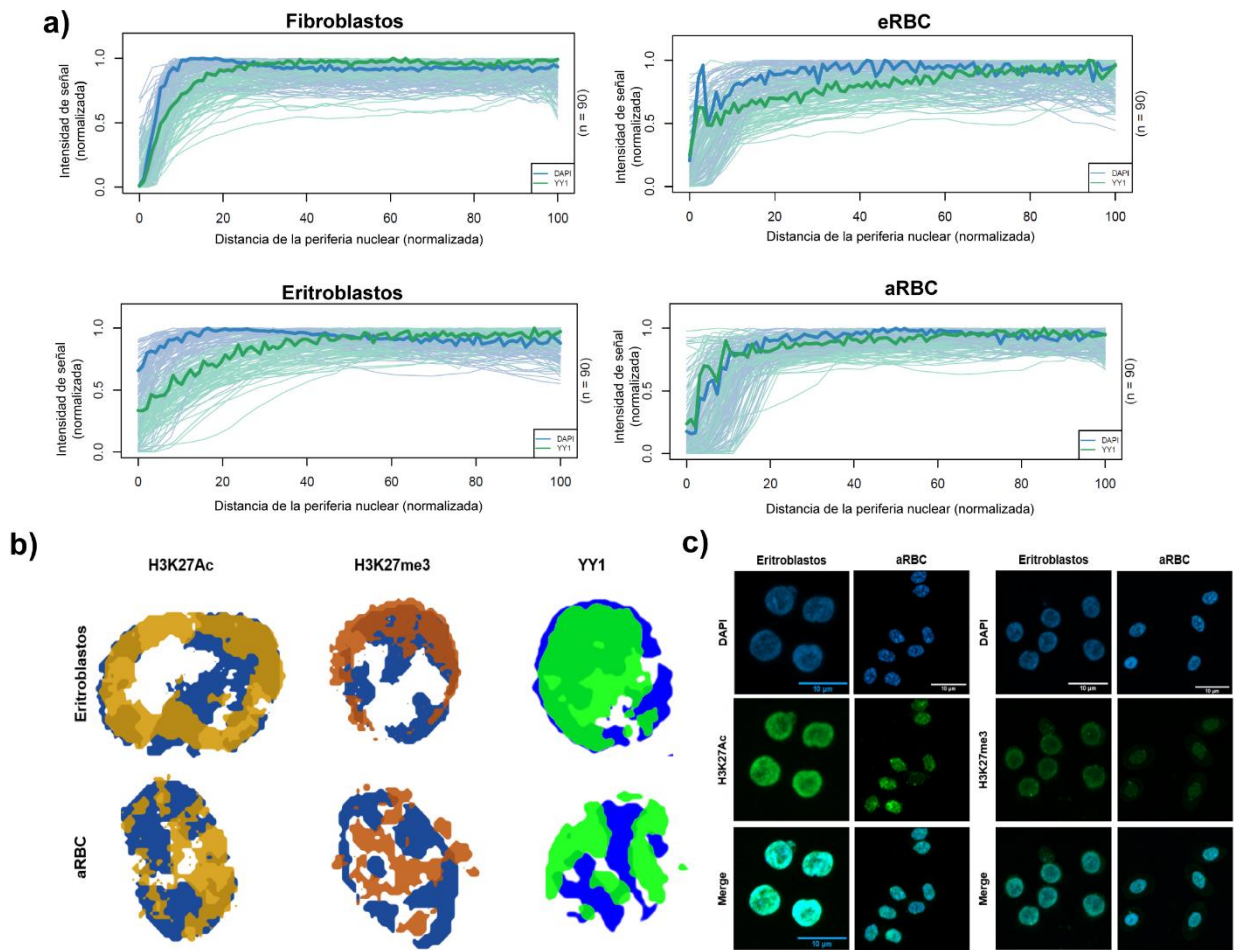


Figura 8. Análisis de distancias y comparación con marcas de histonas. a) Mapas de distancia periferia-centro de la intensidad de la señal de YY1 y DAPI. El trazo promedio se muestra como una línea intensa. b) Máscaras binarias de los píxeles más intensos que el promedio. Se muestra la distribución de las marcas H3K27Ac (amarillo), H3K27me3 (naranja) y YY1 (verde) en comparación a la señal de DAPI (azul). c) Imágenes representativas de las marcas H3K27Ac y H3K27me3 evaluadas por inmunofluorescencia.

8.4. Inmunoprecipitación de la cromatina acoplada a secuenciación (ChIP-seq)

Dada la necesidad de caracterizar la participación de YY1 en la organización del genoma durante los procesos de diferenciación y maduración eritroide, requerimos elucidar en alta resolución la ocupación genómica de YY1 en la fibra de cromatina. Para esto, se estableció un protocolo de ChIP para YY1, con el objetivo de generar bibliotecas de ChIP para secuenciación.

8.4.1. Establecimiento del protocolo de YY1 ChIP en pollo

Debido a que el modelo de estudio es muy poco explorado en el campo de la topología genómica, para realizar un experimento como ChIP-seq, primero debíamos establecer los controles y condiciones apropiados. Decidimos replicar el único trabajo en el cual se ha realizado un experimento de ChIP contra YY1 en eritrocitos embrionarios de pollo⁵³, con el propósito de estandarizar las condiciones apropiadas para los anticuerpos con los que contábamos (**Figura Suplementaria 3**). El anticuerpo de mayor eficiencia fue Cell Signaling (D5D9Z), el cual estaba respaldado por buenos antecedentes en western blot e inmunofluorescencias. Una vez determinados las condiciones apropiadas para YY1, se realizaron los experimentos de ChIP para los tipos celulares eRBC, aRBC y HD3. Además, como control adicional se añadió el experimento ChIP de hígado de ratón, pues el anticuerpo originalmente se encontraba validado para humano, ratón, rata y mono. De esta manera podríamos tener un control más robusto de la calidad del experimento dada por el anticuerpo al obtener los resultados. Todos los experimentos fueron realizados simultáneamente y con sus respectivos replicados. Una vez obtenido el material resultante del ChIP, se prepararon las bibliotecas de DNA con el kit NEBNext de Illumina y se analizaron mediante electroforesis capilar para su subsecuente secuenciación (**Figuras Suplementarias 4 y 5**).

8.4.2. YY1 se localiza en regiones promotoras e intergénicas en el genoma eritroide

Una vez secuenciadas las muestras, las lecturas obtenidas fueron sometidas a un análisis bioinformático general para ChIP-seq (**Figura 9a** y métodos) iniciando por los controles de calidad de secuenciación, y su alineamiento a su respectivo genoma de referencia (**Figuras Suplementarias 6 y 7**).

Realizamos un análisis de correlación de Spearman en nuestras muestras entre réplicas para definir la similitud entre ellas (**Figura 9b**); se espera que los replicados muestren un alto grado de correlación. El análisis evidenció que el experimento menos eficiente pertenecía a los eRBC. Se procedió a realizar un llamado de picos mediante los algoritmos *BroadPeaks* y *NarrowPeaks*. El método *NarrowPeaks* se recomienda para la búsqueda de picos generados por factores de transcripción. Al realizar una inspección visual, decidimos realizar un filtrado de los picos (ver métodos) para eliminar falsos positivos los cuales pueden aparecer por factores como el ensamblado del genoma. De forma esperada, los picos de YY1 eran menos en eRBC y aRBC comparados con los eritroblastos (**Figura 9c y Tabla Suplementaria 3**). Procedimos a visualizar la distribución de las lecturas obtenidas mediante el experimento de ChIP-seq YY1 para cada una de las réplicas y el input genómico procesados (**Figura 9d**). Las lecturas de los experimentos, especialmente en la réplica 2, mostraron una distribución esperada para un experimento de ChIP-seq, exhibiendo enriquecimiento de lecturas *formando* picos en regiones discretas del genoma. Las lecturas obtenidas de las réplicas de eRBC exhibían niveles de ruido parecidos a los encontrados en el input genómico. En conjunto, se determinó que, para todos los tipos celulares, las réplicas 1 fueron la menos eficiente al exhibir el menor número de picos llamados y mayor ruido en la distribución de lecturas.

Posteriormente, los picos encontrados fueron sometidos a un análisis de anotación. Éste reveló que la mayoría de los picos caen en las regiones de promotores e intergénicas (**Figura 9e y Figura Suplementaria 8a**). Dichos resultados están en concordancia con reportes previos que demostraron que YY1 se posiciona de forma preferencial en promotores y *enhancers*⁴⁹.

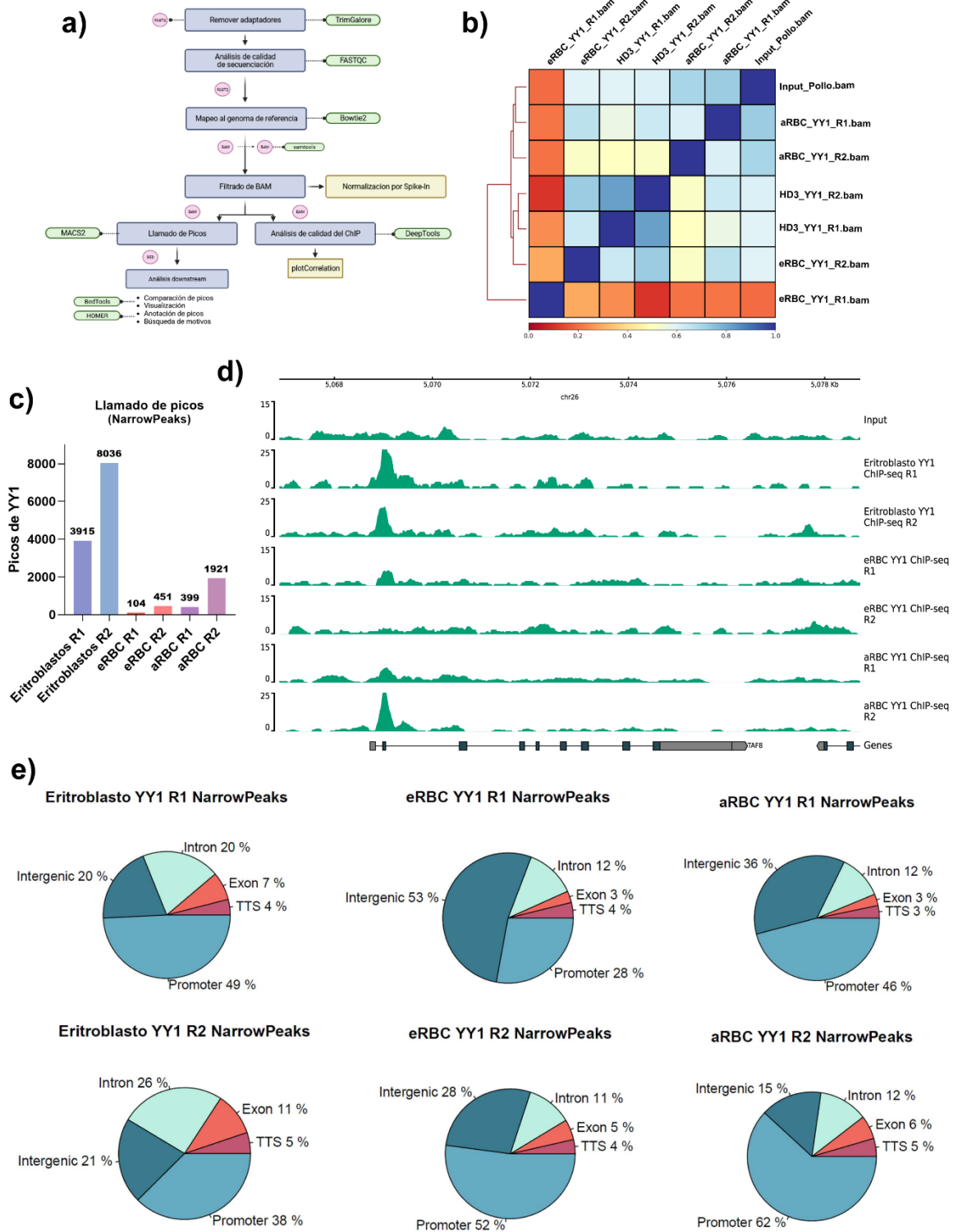


Figura 9. Análisis de ChIP-seq. **a)** Pipeline general de análisis de datos secuenciados para ChIP-seq. **b)** Mapa de calor del análisis de correlación de Spearman entre réplicas de las muestras de ChIP-seq. eRBC es la muestra que presenta menor grado de correlación entre réplicas. **c)** Número de picos llamados en las réplicas de cada muestra. Los picos disminuyen conforme la diferenciación eritroide avanza. **d)** Visualización en el locus TAF8 de la distribución de lecturas obtenidas de cada experimento ChIP-seq y su input genómico. **e)** Anotado genómico de los picos de YY1 llamados por NarrowPeaks de ambas réplicas. Las categorías más representadas son regiones promotoras e intergénicas.

Mediante análisis de búsqueda de motivos utilizando HOMER se encontró a YY1 como el motivo más enriquecido en los picos identificados (**Figura 10a y Figura Suplementaria 8b**). Al intersecar las listas de picos entre sus correspondientes réplicas para buscar picos comunes, observamos que las listas de picos resultantes se reducían drásticamente; sin embargo, aún encontramos a YY1 como motivo principal (**Figura 10b-c y Figura Suplementaria 8c**). Detectamos que más del 50% de los picos de R1 son un subconjunto de R2 (**Figura 10d**). Dado que los experimentos de la réplica 2 fueron los que mostraron mayor eficiencia, decimos investigar si los picos que no se compartían con la réplica 1 tenían el motivo de YY1. El análisis reveló que estos picos no compartidos en R2 también contenían el motivo (**Figura Suplementaria 8d**). No obstante, los picos no compartidos provenientes de R1 llamados por ambos métodos en eRBC y aRBC, no contenían a YY1 como motivo principal. Considerando estos resultados y una exhaustiva visualización de los picos en SeqMonk, determinamos que el método *NarrowPeaks* proporcionaba los picos más confiables y tomamos la decisión de considerar y reportar los resultados de los picos de la réplica 2 (R2) y los compartidos entre réplicas para análisis subsecuentes.

Aunado a esto, ya que el análisis de correlación entre réplicas del ChIP-seq indicó que eRBC fue la menos parecida entre réplicas, la mayoría de las observaciones serán consideradas únicamente entre los eritroblastos y los aRBC. Con esto en mente, buscamos a los picos comunes entre eritroblastos y aRBC. De forma esperada, la mayoría de los picos de YY1 pertenecientes a aRBC también están en eritroblastos (**Figura 10e**).

En conjunto, se concluye que los sitios de unión de YY1 disminuyen a medida que la diferenciación y maduración de los eritrocitos avanza, pasando de 8036 en eritroblastos a solo 1921 en aRBC, pertenecientes a R2. Además, se demostró que YY1 se posiciona mayoritariamente en regiones promotoras e intergénicas. Los experimentos de ChIP-seq son confiables pues, en todos los estadios de la diferenciación estudiados, logramos recuperar al motivo de YY1 como el top número uno en los análisis de búsqueda de motivos.

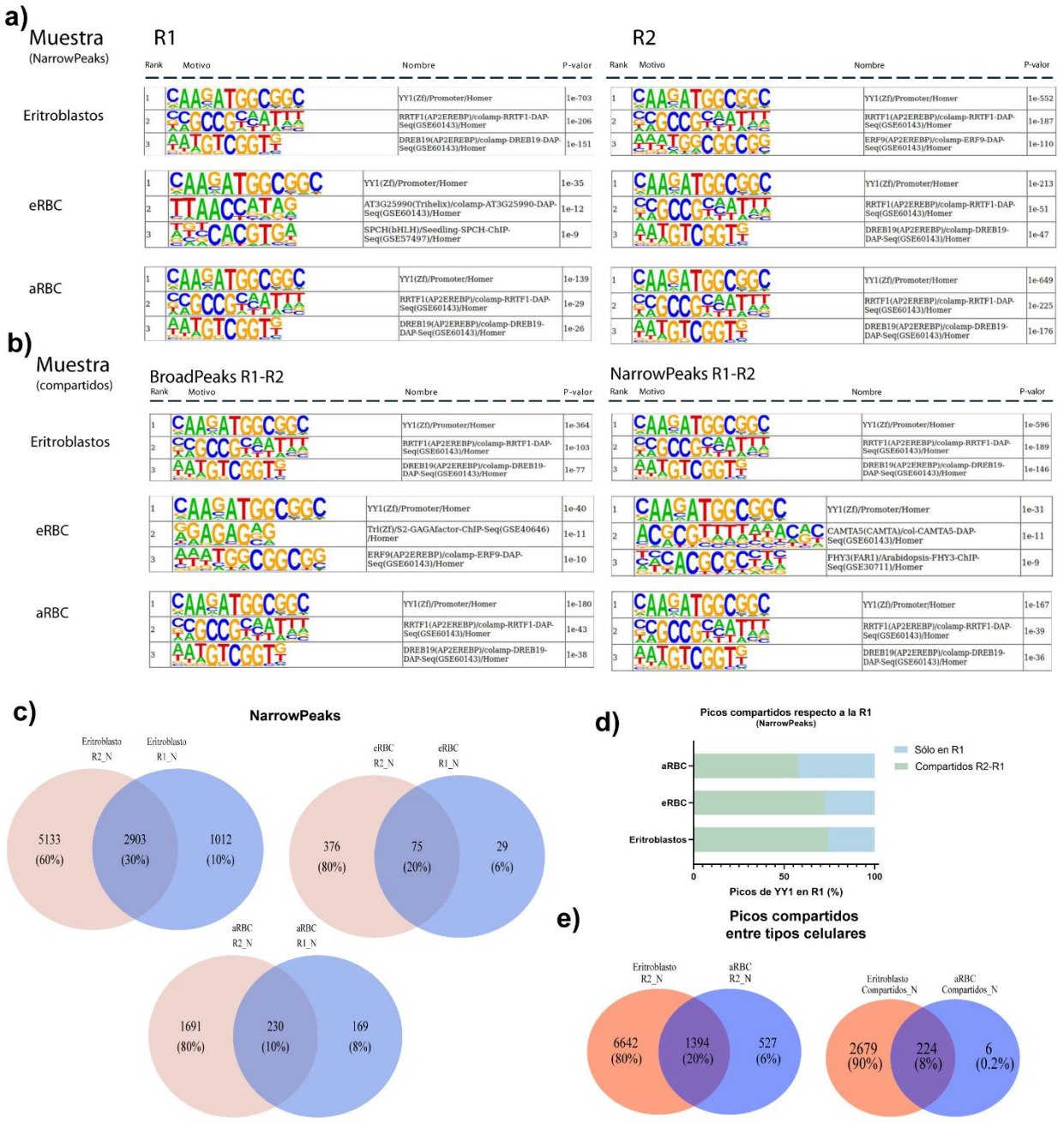


Figura 10. Análisis de picos de YY1. a-b) Resultados del análisis de búsqueda de motivos en los picos llamados por NarrowPeaks y los picos compartidos entre réplicas para cada tipo celular. **c)** Diagramas de Venn de las intersecciones de los picos compartidos entre réplicas. **d)** Comparación de picos compartidos respecto al total de picos en R1 (NarrowPeaks). **e)** Picos compartidos entre eritroblastos y aRBC.

Es importante mencionar que los experimentos de ChIP en los eritrocitos de pollo han exhibido dificultades experimentales en el laboratorio posiblemente asociadas a su alto grado de compactación de su genoma. Sin embargo, una de las principales perspectivas

futuras del proyecto contempla la ejecución de dos réplicas adicionales para cada tipo celular con el objetivo de robustecer los resultados.

8.4.3. Los picos de YY1 de eritrocitos maduros se preservan dentro de minidominios

Con la finalidad de contextualizar el paisaje genómico de los picos de YY1 encontrados en el proceso de diferenciación eritroide, se utilizaron datos generados en el laboratorio y previamente publicados de Hi-C, ATAC-seq y RNA-seq que comprenden los mismos estadios estudiados de la diferenciación y maduración eritroide⁴¹.

En primera instancia se anotaron los picos de YY1 a los TADs descritos por medio de Hi-C. De los TADs y minidominios previamente reportados (2,562 en eritroblastos, 916 en eRBC, y 903 aRBC)⁴¹, encontramos que el 67% y 55% de los dominios cromatínicos contienen picos de YY1 en eritroblastos y aRBC, respectivamente (**Figura 11a**). Al hacer el análisis recíproco, observamos que alrededor de 80% de los picos de YY1 en eritroblastos se ubican dentro de TADs. Por otro lado, el 38 y 47% de los picos de YY1 se ubican en minidominios en eRBC y aRBC, respectivamente (**Figura 11b**). Utilizando un set de datos de regiones aleatorias en el genoma tomando un número y longitudes equivalentes a los picos de cada tipo celular, descartamos que el posicionamiento de YY1 dentro o fuera de dominios fuese un fenómeno aleatorio. Definimos a la distribución de las regiones del set aleatorio como “esperadas” y a la distribución dada por YY1 como la “observada” (**Figura 11c**). Ya que el genoma en eRBC y aRBC tiene menor número de dominios⁴¹, si el posicionamiento de YY1 fuese aleatorio, entonces encontraríamos una mayor proporción de picos fuera de minidominios (Figura 11c; eRBC y aRBC: esperada). No obstante, con YY1 observamos lo opuesto. Al realizar un análisis estadístico entre las distribuciones de cada tipo celular, confirmamos que el posicionamiento de YY1 en minidominios es un evento dado por el factor transcripcional *per se* y no por aleatoriedad.

8.4.4. YY1 no está enriquecido en fronteras de minidominios

Considerando la pérdida generalizada de estructura en los genomas de eritrocitos terminalmente diferenciados y maduros, acompañada de la pérdida de proteína de las principales proteínas arquitectónicas, se consideró la posibilidad de que YY1 fuera la responsable de promover la estructuración en minidominios. Este escenario es plausible considerando que YY1 ha sido reportado como estructurador del genoma,

específicamente de asas de cromatina de tipo *enhancer*-promotor⁴⁹. Para ello, los picos de YY1 deberían encontrarse posicionados en fronteras de minidominios de forma análoga a CTCF. Con el fin de abordar este escenario, intersecamos las fronteras de los dominios y los picos.

Encontramos que menos del 10% de las fronteras en aRBC contienen picos de YY1 comparado con 30% en los eritroblastos. Del total de las fronteras en aRBC, solo alrededor del 10% de ellas tienen picos de YY1 (**Figura 11d**). Desde la perspectiva de los picos de YY1, encontramos que alrededor del 20% de ellos se posicionan en fronteras en eritroblastos, mientras que para aRBC, este porcentaje disminuye a solo 10% (**Figura 11e**). Realizamos un análisis análogo a la Figura 11c, pero considerando a las regiones de fronteras de dominios. Con este análisis determinamos que el posicionamiento no preferencial de YY1 en fronteras es independiente a fenómenos aleatorios (**Figura 11f**).

Dado que los archivos de fronteras consideran a una frontera como la ventana genómica de +/- 20 Kb respecto a la coordenada de inicio o fin de un dominio cromatínico determinada por el llamador de TADs, los picos de YY1 que caen en fronteras podrían estar dentro o fuera de su respectivo dominio. Para refinar aún más los hallazgos previos a esta pregunta, consideramos únicamente los picos que residen dentro de los dominios, y los intersecamos con las fronteras. De esta forma, encontramos que alrededor del 13% de los picos contenidos en minidominios en aRBC se encuentran en fronteras (**Figura 11g**) o un 6% del total de picos en R2 (117 de 1921 picos totales de YY1). Considerando estos resultados (**Figura 11 y Figura Suplementaria 9**), mostramos que YY1 no ejecuta un papel de delimitar dominios en el genoma eritroide a lo largo de la diferenciación.

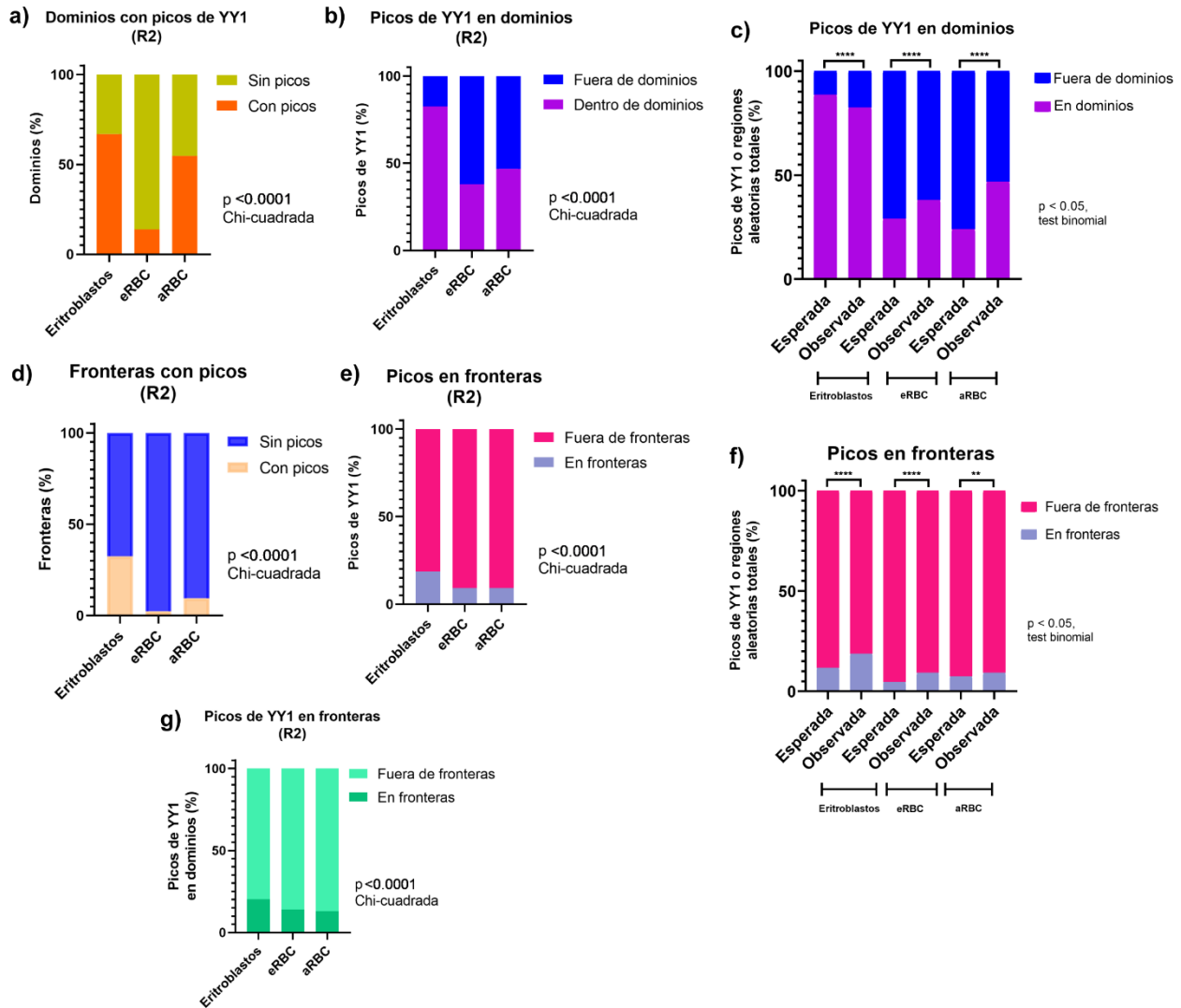


Figura 11. YY1 se preserva en minidominios, pero no está enriquecido en sus fronteras. **a)** Dominios cromatinicos que contienen picos de YY1. **b)** Picos de YY1 que se encuentran dentro de dominios. Los picos dentro de fronteras disminuyen en aRBC respecto a los eritroblastos. **c)** Proporción de picos esperados (set aleatorio) y observados (picos de YY1) mapeados a dominios o fuera de dominios. ($p < 0.05$, prueba de binomial; **** indica $p < 0.0001$). **d)** Porcentaje de las fronteras que contiene picos de YY1. **e)** Porcentaje de picos contenidos en las fronteras. **f)** Proporción de picos esperados (set aleatorio) y observados (picos de YY1) mapeados a fronteras de dominios o fuera de fronteras de dominios. ($p < 0.05$, prueba de binomial; **** indica $p < 0.0001$ y ** indica $p = 0.018$) **g)** Picos que están contenidos dentro de dominios y además están en sus fronteras. Se encontraron diferencias estadísticamente significativas entre los estadios de eritroblasto, eRBC y aRBC en a), b), d), e) y g) ($p < 0.0001$, prueba de Chi-cuadrada).

8.4.5. Los picos de YY1 están enriquecidos en el compartimento A

Ya que alrededor del 50% de los picos de YY1 no están contenidos en los minidominios aRBC y eRBC, decidimos investigar su posicionamiento a nivel de compartimentos. Para iniciar este análisis, se consideró al universo de picos de YY1 que estaban dentro de TADs y minidominios. En los tres estadios de diferenciación eritroide explorados encontramos que los picos de YY1 están enriquecidos (>70%) en el compartimento A y,

una menor proporción se localizaban en B o zonas identificadas como A/B. De forma más detallada, en los eritroblastos el 74% de picos caen en compartimento A, el 17% en B y el 9% en A/B. En eRBC, 72% están en el compartimento A, 27% en B y el 1% no pudo ser asignado a un compartimento (compartimento C). En el caso de aRBC, el 78% de los picos residen en el compartimento A, y el 22% en B. De forma interesante observamos que, en los eritroblastos, la proporción de picos en el compartimento B es menor respecto a los aRBC y hay picos que mapean el compartimento A/B. Considerando ambas observaciones, es posible sugerir un posible posicionamiento de los picos de YY1 en regiones de transición de compartimento A a B. Así, en el estadio de aRBC, la fracción del compartimento B aumenta (**Figura 12a**).

Al investigar el mismo panorama para aquellos picos de YY1 que caían fuera de dominios cromatínicos, encontramos el mismo comportamiento. Es decir, la mayoría de los picos de YY1 fuera de dominios se enriquecen en compartimentos identificados como A (**Figura 12b**). Sin embargo, notamos que la fracción de este subconjunto de picos que residían dentro del compartimento B y C incrementa, especialmente en eRBC y aRBC.

Al considerar el total de los picos encontrados y mapearlos a compartimentos sin importar si estaban estructurados o no, se evidenció que YY1 se posiciona preferencialmente en el compartimento A en todo el genoma, en todos los estadios de diferenciación estudiados (**Figura 12c**). Para descartar estadísticamente que este enriquecimiento no fuese meramente aleatorio, construimos un set de datos de regiones aleatorias en el genoma tomando un número y longitudes equivalentes a los picos de cada tipo celular. Los cruces de estos sets aleatorios se realizaron con los compartimentos y la distribución obtenida fue considerada como el evento esperado, mientras que la distribución de YY1 fue considerada como la distribución observada (**Figura 12d**). Dado que el genoma de aRBC está altamente compacto, sabemos que existe una mayor proporción de compartimento B⁴¹. Si el posicionamiento de YY1 fuese aleatorio, es más probable que YY1 este en el compartimento B que en el A (ver **Figura 12d** - aRBC: Esperada), pero YY1 está mayoritariamente en A. Al ejecutar la prueba estadística de Chi-cuadrada para cada tipo celular con su respectivo set aleatorio, encontramos diferencias estadísticamente significativas entre la distribución esperada y encontrada de cada tipo celular, demostrando que YY1 se enriquece preferencialmente en el compartimento A del genoma. Estos resultados contrastan con la versatilidad de YY1 observada en las

inmunofluorescencias respecto a su localización a regiones del núcleo previamente identificadas como activas y, en menor medida con regiones compactas en las células eritroides (**Figura 12 y Figura Suplementaria 10**).

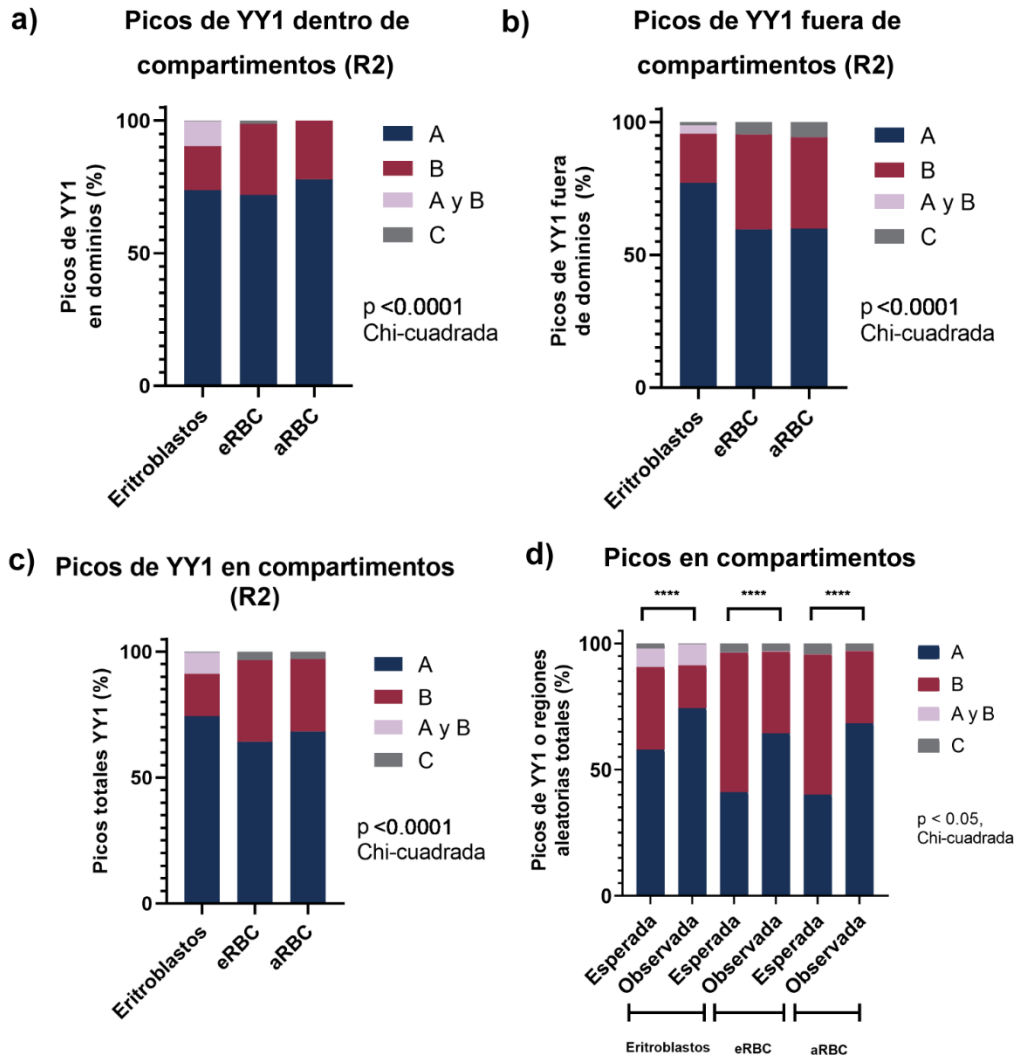


Figura 12. YY1 está enriquecido en el compartimento A. Los picos de YY1 dentro (a) y fuera (b) de dominios cromatínicos exhiben un posicionamiento preferencial en el compartimento A en los tres estadios de diferenciación. c) Picos totales de R2 mapeados a compartimentos. Las diferencias entre los estadios de eritroblasto, eRBC y aRBC en a), b) y c) son estadísticamente significativas ($p < 0.0001$, prueba de Chi-cuadrada). d) Proporción de picos esperados (set aleatorio) y observados (picos de YY1) mapeados a compartimentos A, B, A y B, y C. ($p < 0.05$, prueba de Chi-cuadrada; **** indica $p < 0.0001$).

8.4.6. Los picos de YY1 preservan accesibilidad y se enriquecen en promotores de genes no expresados en aRBC

Con el fin de inferir posibles papeles funcionales en la estructuración del genoma eritroide y su relación con el estado cromatínico integramos picos de accesibilidad, obtenidos mediante ATAC-seq, con los picos de YY1. Nuestros análisis indican que alrededor del

75% de los picos de YY1 encontrados en eRBC y aRBC comparten picos de accesibilidad en la cromatina (**Figura 13a y Figura Suplementaria 11a**). En aRBC, más de la mitad de los picos de estos picos de YY1 accesibles están posicionados dentro de minidominios. De forma contrastante, el porcentaje de picos accesibles de YY1 encontrados en eritroblastos es de sólo el 31%, donde la mayoría de ellos se encuentran contenidos dentro de dominios topológicos.

Para caracterizar el estado transcripcional que acompañaba a los picos de YY1, los picos se dividieron de acuerdo con su anotación en regiones genómicas (i.e. promotores, exones, intrones y regiones intergénicas) y se cruzaron con los genes con y sin expresión identificados por medio de RNA-seq en su respectivo estadios de diferenciación eritroide.

En los eritroblastos encontramos que los picos de YY1 se posicionan preferencialmente en promotores de genes con expresión. La segunda categoría más enriquecida fue la de intrones activos, lo que está en congruencia con reportes que indican la importancia de la unión de YY1 a secuencias promotoras e intrónicas para controlar la expresión génica y procesos de splicing²⁶. De forma interesante, esta relación se invierte en los eritrocitos terminalmente diferenciados y maduros. Encontramos que los picos de YY1 están enriquecidos en promotores de genes inactivos en aRBC. En resumen, los picos de YY1 se enriquecen en promotores de genes que se están transcribiendo en eritroblastos, y en los inactivos en aRBC (**Figura 13b y Figura Suplementaria 11b**). Para refinar estos resultados, realizamos el mismo análisis considerando únicamente los picos YY1 accesibles. Las mismas relaciones fueron encontradas, un resultado esperado considerando que los picos de YY1 con accesibilidad son un subconjunto de los generales.

Adicionalmente, se realizaron análisis de correlación entre los picos de YY1 y los valores de FPKM de los genes que tuvieran a YY1 en sus promotores (**Figura 13c-e**). De forma congruente a los resultados previos, observamos que la mayoría de los picos de YY1 en promotores se encuentran en genes con valores de FPKM cercanos a cero en eRBC y aRBC. Por el contrario, en eritroblastos los picos de YY1 se posicionan preferencialmente en promotores de genes con valores altos de FPKM. Para ambos estadios de eritrocitos terminalmente diferenciados, además, pudimos observar que existen genes que presentan alto enriquecimiento de YY1 en su promotor, pero tienen valores bajos de

FPKM. Al mismo tiempo, encontramos a genes donde el enriquecimiento de YY1 es menor pero los valores de FPKM más altos. Ambos escenarios refuerzan la naturaleza bivalente, previamente reportada, de YY1 como activador y represor transcripcional. En el estadio de eritroblasto, observamos en menor medida ambos escenarios. Los análisis de correlación son negativos pero muy cercanos al cero, evidenciando una relación no directa entre el estado transcripcional y la presencia de YY1.

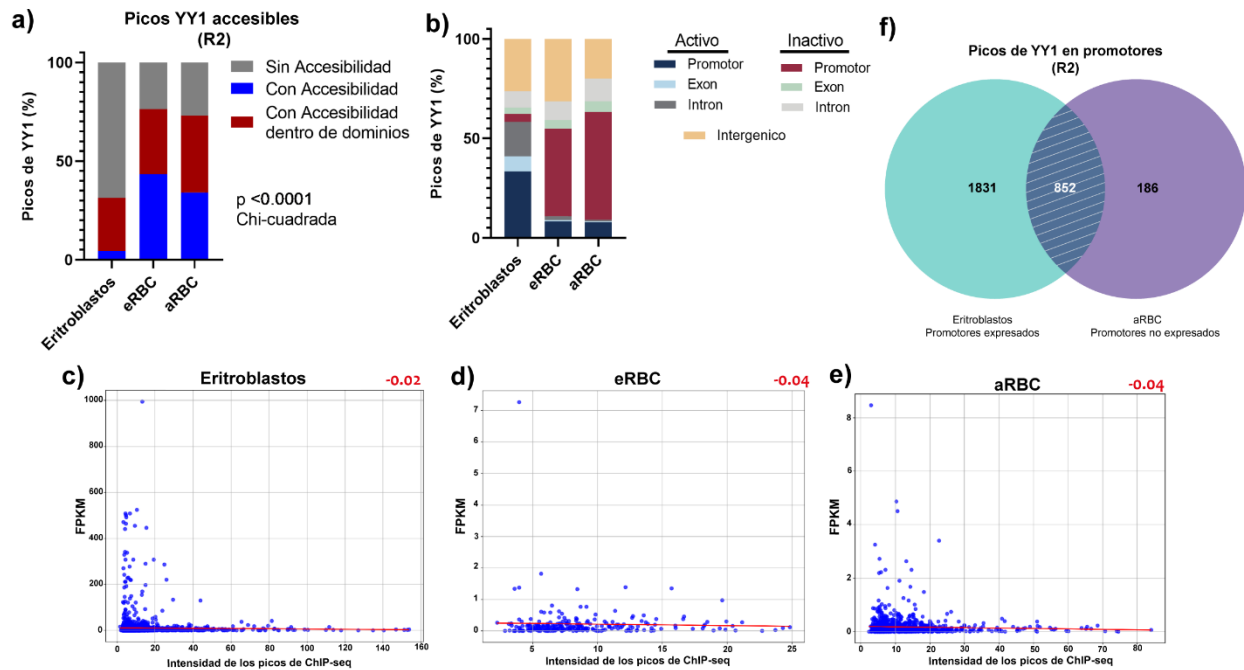


Figura 13. YY1 preserva accesibilidad y se enriquece en promotores de genes inactivos. a) Porcentaje de picos de YY1 que preservan accesibilidad. La mayoría de los picos en aRBC tienen accesibilidad ($p < 0.0001$, prueba de Chi-cuadrada). b) Anotado de picos YY1 ChIP-seq de promotores, exones, intrones y regiones intergénicas respecto a su estado transcripcional. YY1 se enriquece en promotores de genes inactivos en aRBC. c-e) Análisis de correlación entre los picos de YY1 y los valores de FPKM de genes con YY1 en sus promotores para cada tipo celular. f) Picos de YY1 en promotores que se pierden expresión en eritrocitos maduros.

8.4.7. YY1 se posiciona en promotores de genes que pierden expresión durante la diferenciación terminal eritroide

Es posible que los promotores de genes inactivos en los que se posiciona YY1 en aRBC sean un subconjunto de aquellos promotores con picos de YY1 de genes que se estaban transcribiendo de forma activa en eritroblastos. Estos genes estarían perdiendo expresión por mecanismos de silenciamiento o pausado génico que ocurren a lo largo de la diferenciación terminal eritroide. Para investigar esta posibilidad, decidimos intersecar a las listas de promotores con picos de YY1 que están activos en eritroblastos e inactivos en aRBC. Podemos observar que más del 80% de los picos que caen en promotores de

genes no expresados en aRBC eran promotores de genes expresados en eritroblastos (**Figura 13f y Figura Suplementaria 11c**). Por tanto, se hace evidente que YY1 se posiciona y preserva en promotores de genes que pierden su expresión en el estadio maduro eritroide.

8.4.8. YY1 participa en procesos relacionados con control de la expresión génica, metabolismo y procesamiento de RNAs

Los resultados previos revelaron la existencia de un conjunto considerable de picos de YY1 accesibles y estructurados los cuales se enriquecen preferencialmente en regiones promotoras. Específicamente, observamos que aquellos promotores de genes activos en eritroblastos se inactivaban conservando a YY1 en su promotor.

En estudios previos, la categorización de minidominios dependiente del estado transcripcional y cromatínico de sus sitios de inicio de la transcripción (TSS) elucidó que, los minidominios de eRBC y aRBC contienen y organizan principalmente a promotores de genes pausados y activos. Es decir, los minidominios contiene a TSS que presentan a cromatina abierta y pueden tener o no tener transcripción activa. Aquellos TSS sin transcripción y con cromatina abierta fueron referidos como promotores “pausados”, y están presentes en la mayoría de minidominios⁴¹. Tomando en cuenta que aquellos promotores de genes inactivos con YY1 en su promotor encontrados en aRBC eran activos en el estadio de eritroblasto, decidimos comparar las listas de promotores de genes pausados de ese estudio contra la lista de genes inactivos con YY1 en su promotor y que conservan estructura. Observamos que muchos de los genes en ambas listas se comparten. Por tal motivo, decidimos investigar las posibles funciones biológicas asociadas de este subconjunto de picos a través de análisis de ontologías de genes (GO). Mediante gProfiler⁵⁷ obtuvimos las GOs de genes con picos de YY1 en su promotor que están dentro de dominios para eritroblastos y aRBC. Para los eritroblastos utilizamos la lista de genes expresados dado que fue la categoría más enriquecida, y para aRBC la lista de genes no expresados (ver Figura 13b).

Los resultados de las GOs mostraron que el subconjunto de picos de YY1 en promotores de genes activos-estructurados en eritroblastos enriquecen procesos relacionados con la traducción de proteínas, la expresión génica, metabolismo de RNAs y organización del genoma (**Figura 14a**). Por otro lado, los promotores con picos de YY1 de genes inactivos-

estructurados, además de enriquecerse en procesos biológicos similares, enriquecen procesos involucrados con la modificación postraduccional de proteínas (**Figura 14b**). El enriquecimiento de procesos biológicos similares de este subconjunto de genes inactivos en aRBC, comparados con el eritroblasto, refuerza la observación previa respecto a su presunta inactivación acompañada por YY1 a lo largo de la diferenciación terminal eritroide.

Ya que una cantidad considerable de genes con picos de YY1 en sus promotores no está dentro de minidominios, decidimos investigar cuáles procesos biológicos enriquecían dichos genes en forma global respecto a su estado transcripcional. Tanto para los genes con promotores activos de los eritroblastos como para los genes con promotores inactivos en aRBC, fueron detectados los mismos procesos previamente encontrados y la adición de algunas otras funciones (**Figura Suplementaria 12**). Estos resultados son congruentes pues el conjunto de genes globales evaluados contiene al conjunto de genes estructurados. Respecto a las nuevas funciones biológicas encontradas tenemos a la biogénesis de ribosomas para los eritroblastos, mientras que en aRBC encontramos la biogénesis de complejos ribonucleoproteicos, regulación de daño al DNA y la transcripción orquestada por la RNA polimerasa III. No fue posible realizar las GOs en los eRBC dado al bajo número de picos previamente detectado.

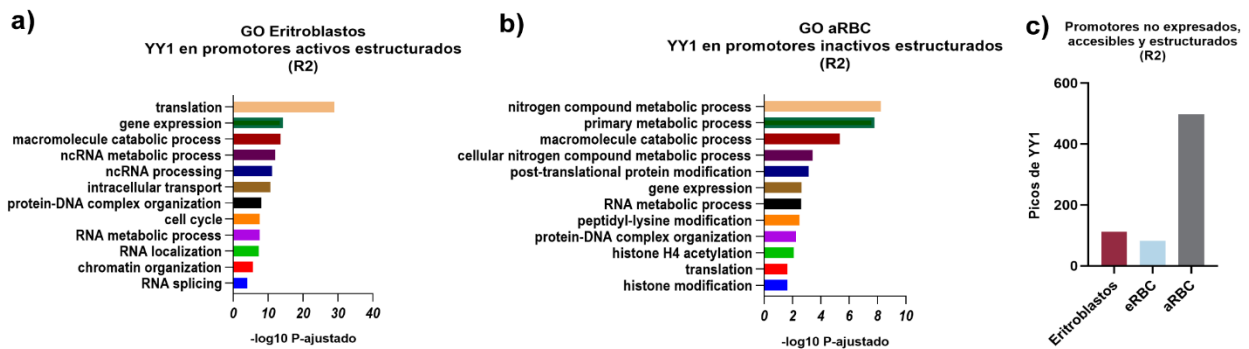


Figura 14. Ontologías de genes (GO) estructurados con YY1 en su promotor. a) GO en eritroblastos de genes activos (R2). Los procesos biológicos más enriquecidos están relacionados con la traducción, expresión génica y procesos metabólicos. **b)** GO en aRBC (R2) de genes inactivos. Los procesos biológicos más enriquecidos se relación con procesos metabólicos, modificación postraduccional de proteínas y expresión génica. **c)** Número de picos en promotores génicos no expresados, accesibles y estructurados.

8.4.9. Integración de escenarios genómicos

Finalmente, ya que la integración de los picos de YY1 con los datos de topología, expresión y accesibilidad de la cromatina permitió reconocer diferentes panoramas genómicos, decidimos realizar la visualización de estos utilizando el software pyGenomeTracks⁵⁵. Iniciamos refinando el set de picos de YY1 localizados en promotores de genes estructurados obtenidos en los análisis de la sección anterior. Dependiendo de su perfil transcripcional, expresados para eritroblastos o no expresados para aRBC, fueron filtrados y conservados solo los genes con picos en sus promotores que compartieran picos de accesibilidad. De esta manera, obtuvimos listas de picos en promotores de genes expresados (eritroblastos) o no expresados (aRBC), accesibles y estructurados. Para la selección de ejemplos a visualizar decidimos escoger genes que se compartían entre las listas recién obtenidas y los procesos biológicos con los valores-*p* más significativos de las GOs realizadas (**Figura 14c**).

El locus de alfa globinas del pollo fue previamente identificado como una región del genoma que se mantenía estructurada⁴⁴ a lo largo del proceso de diferenciación y maduración eritroide⁴⁵. Este dominio comprende a tres genes que codifican para las globinas π , α^D y α^A , y ha sido ampliamente utilizado como modelo para estudiar mecanismos de control transcripcional en eucariontes⁵⁸. Además, esta región representa una zona de alta expresión en eritrocitos⁴⁵. Como se mencionó previamente, en la región 3' del locus de alfa globina en el pollo, aproximadamente 450 pb “rio abajo” del gen alfa globina A, están localizados elementos reguladores en donde fue descrita una región de 120 pb donde YY1 se une para regular la actividad de su *enhancer* a través de la interacción con GATA-1⁵³. Por estos motivos, decidimos seleccionar esta zona pues, además de ser la única reportada en el pollo donde se ha validado la unión de YY1, permitía presentar una región con alta expresión, accesible y estructurada a lo largo de toda la diferenciación terminal eritroide (**Figura 15**).

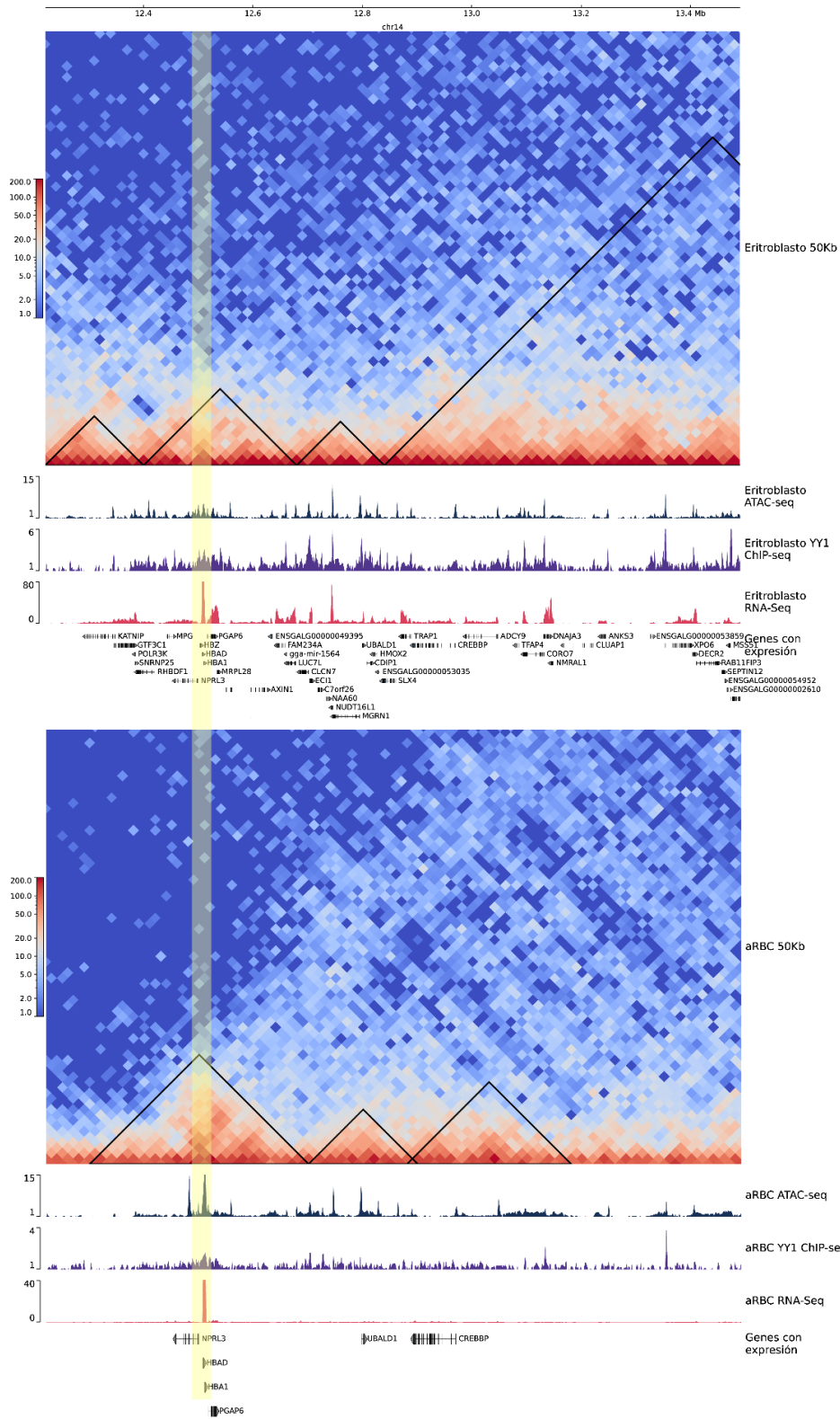


Figura 15. Visualización del dominio alfa globina del pollo. Mapas de Hi-C con los paneles de ATAC-seq, YY1 ChIP-seq y RNA-seq en eritroblastos y aRBC. Esta región genómica representa una zona de alta expresión que mantienen su estructura a través del todo el proceso de diferenciación terminal y maduración eritroide en el pollo. YY1 se une río abajo del HBA1.

El siguiente panorama que visualizamos fue aquel donde existiera una reestructuración a nivel de dominios a lo largo de la diferenciación terminal eritroide. Específicamente, encontramos zonas donde los TADs en eritroblastos reducían su tamaño, reestructurándose en minidominios en aRBC. En este contexto, encontramos genes que están expresados en los eritroblastos y se inactivan en aRBC. Además, estos genes poseen picos de accesibilidad y YY1 notables y que se reducen en el estadio de aRBC. Algunos ejemplos de esto son los genes RPS19BP1, TAB1 y TOMM22 (**Figura 16**). El gen RPS19BP1 codifica para una subunidad de precursor primario de la subunidad pequeña del ribosoma eucarionte y necesario para su biogénesis, al que se le conoce como complejo procesoma⁵⁹. TOMM22 codifica para una proteína integral de la membrana mitocondrial externa con función de translocasa⁶⁰. La inactivación de estos dos genes concuerda con el hecho que los eritrocitos adultos carecen casi su totalidad de mitocondrias y ribosomas^{39,45}. Por su parte, TAB1 genera a una proteína adaptadora que tiene el potencial de activar la vía MAP3K7/TAK1 y contribuir en la activación de NF-kappa-B y promover la producción de citocinas proinflamatorias en respuesta a estímulos mediante los receptores tipo toll (TLRs)⁶¹. Lo último hace relevante la participación de YY1 en un posible pausado de este gen y su rápida reactivación cuando estas células están sometidas a estímulos inmunológicos que activen a los TLRs.

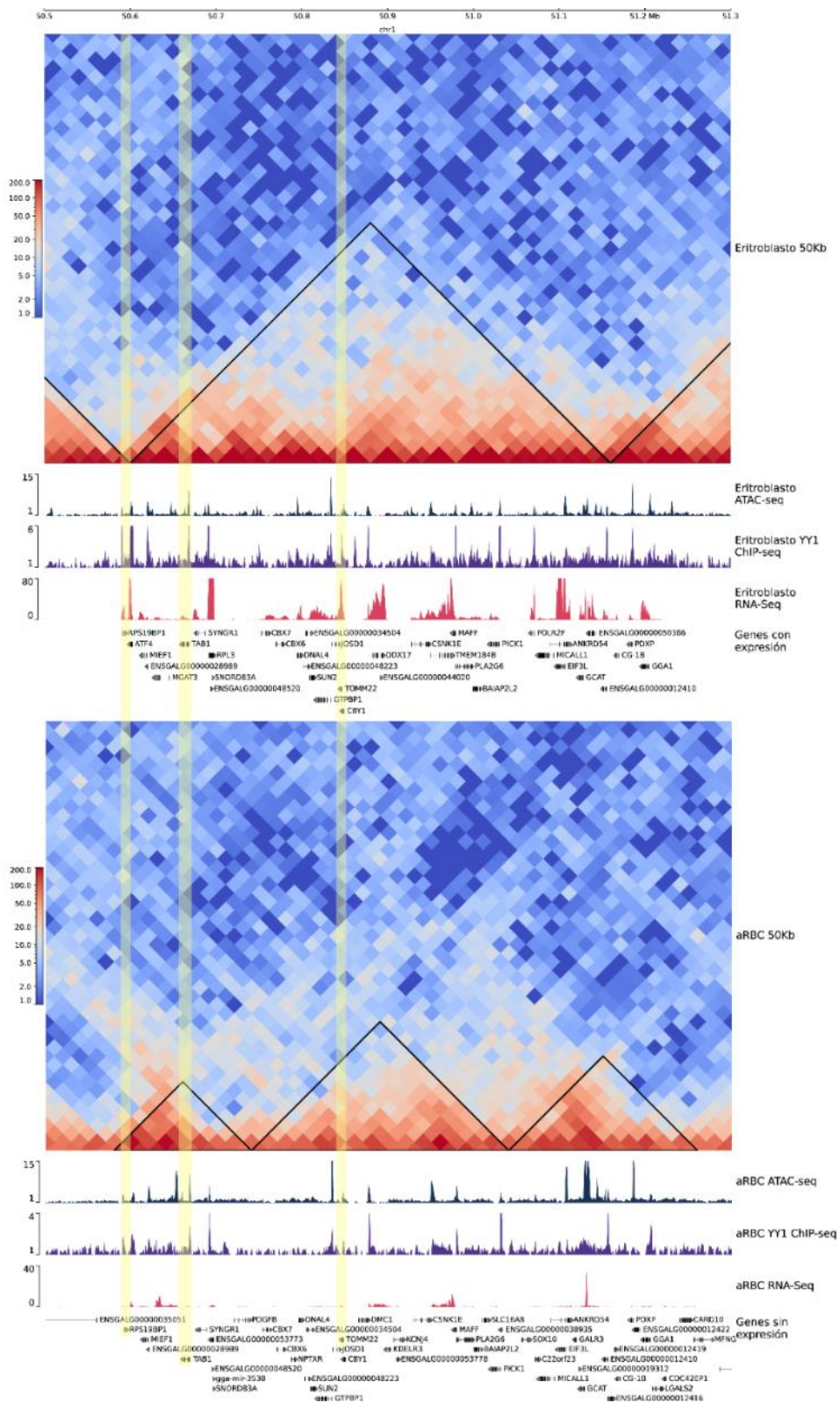


Figura 16. Posicionamiento de YY1 durante la reestructuración de TADs a lo largo de la diferenciación terminal eritroide. Integración de datos de Hi-C con los paneles de ATAC-seq, YY1 ChIP-seq y RNA-seq en eritroblastos y aRBC. En esta región genómica los TADs de eritroblastos se reestructuran a minidominios en aRBC. Dicha reestructuración está acompañada del silenciamiento de genes, como RPS19BP1, TAB1 y TOMM22, que cuentan con picos de YY1 y continúan estructurados en minidominios.

Subsecuentemente, exploramos ejemplos de genes que pertenecían a la categoría de genes pausados estructurados como TAF1A (**Figura 17**) y POLR3H (**Figura 18**). En el gen TAF1A, codificante para una subunidad de la RNA polimerasa I⁶², fue posible observar que, en eritroblastos, este gen se encuentra estructurado en un TAD de mayor tamaño. A través de la diferenciación, observamos que TAF1A se mantiene dentro de un minidominio, pierde expresión y conserva accesibilidad en aRBC. El caso del gen POLR3H, que codifica para la subunidad H de la RNA polimerasa III⁶³, es análogo al de TAF1A. De forma interesante, observamos que los patrones de señal del panel de CHIP-seq YY1 de estos genes que pertenecen a promotores de genes pausados estructurados de aRBC, comparados contra los perfiles de otros genes estructurados sin expresión (ver Figura 16), son más intensos posiblemente por la presencia otros factores como la RNA polimerasa II.

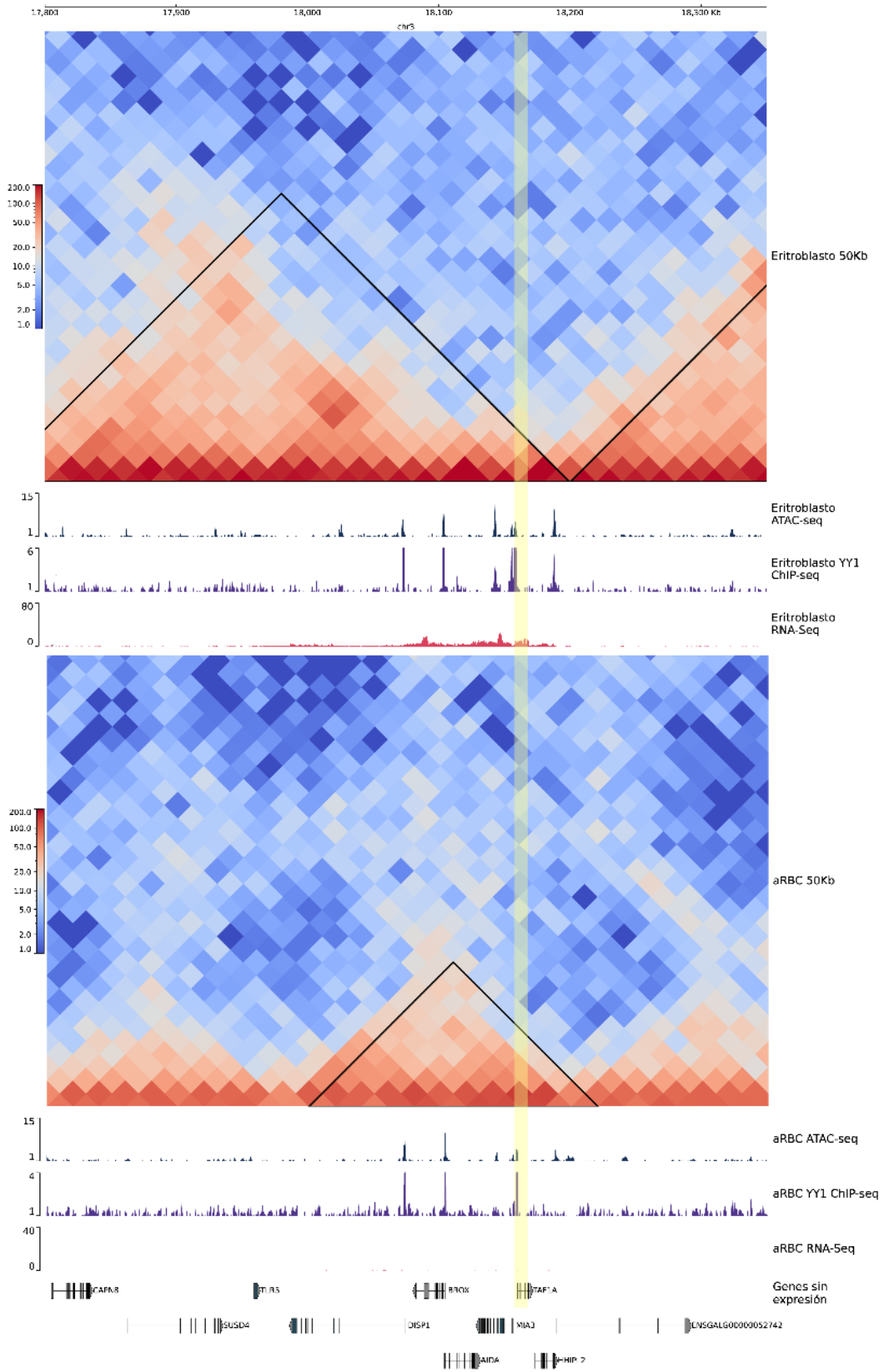


Figura 17. Visualización de genes pausados estructurados: TAF1A. TAF1A es un gen que se inactiva en el estado eritroide maduro y se mantiene estructurado. El análisis bioinformático reveló que TAF1A cuenta con un pico evidente de YY1.

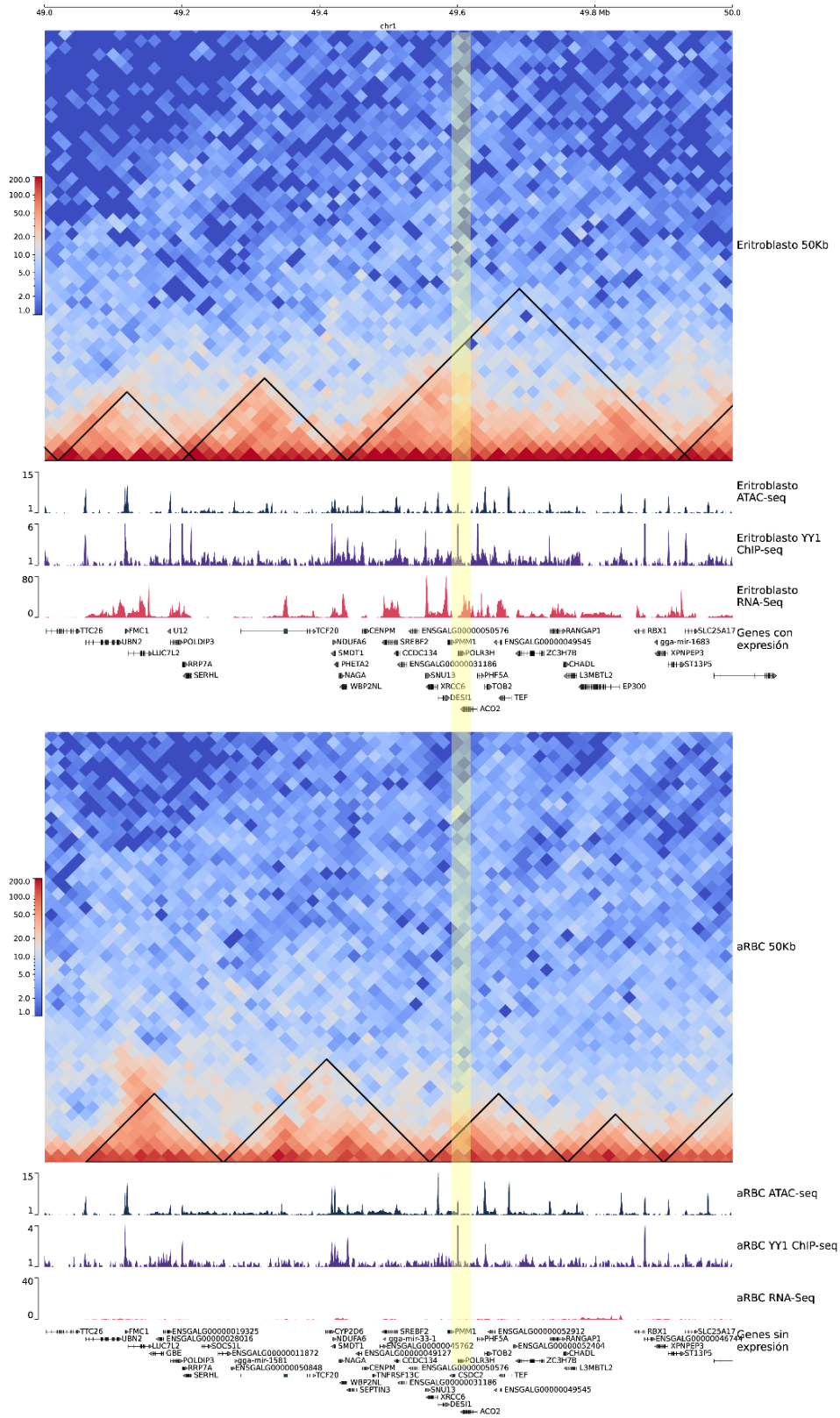


Figura 18. Visualización de genes pausados estructurados: POLR3H. POLR3H es un gen que se inactiva en el estadio aRBC y se mantiene estructurado. El análisis bioinformático reveló que POLR3H cuenta con un pico evidente de YY1.

Finalmente, decidimos explorar el panorama donde los picos de YY1 cayeran fuera de dominios. En este escenario, encontramos ambos panoramas transcripcionales, con expresión y sin expresión. Un ejemplo de un gen con expresión fuera de dominios es RPL18A, codificante para la proteína ribosomal L18a que forma parte de la subunidad 60s del ribosoma. En eritroblastos, este gen se encuentra dentro de un dominio que no puede ser detectado en estadios maduros. Si bien es evidente que la señal de RNA-seq para este gen en aRBC es menor a la encontrada en eritroblastos, la presencia de actividad transcripcional remanente es notoria en aRBC. Respecto a los paneles de señal de ChIP-seq y ATAC-seq, las señales para los picos de YY1 y accesibilidad son consistentes en los tres estadios de diferenciación estudiados (**Figura 19**). Por otro lado, AVPR1 es un gen que codifica al receptor uno de vasopresina⁶⁴. En el estadio de eritroblasto, AVPR1 se encuentra inactivo pero estructurado. A lo largo de la diferenciación terminal permanece inactivo, la estructura se pierde quedando fuera de su dominio, pero la unión de YY1 a este sitio se conserva (**Figura 20**). Este panorama podría indicar un posible papel represor de YY1 durante el establecimiento del linaje eritroide y su importancia en el silenciamiento de genes no específicos del linaje.

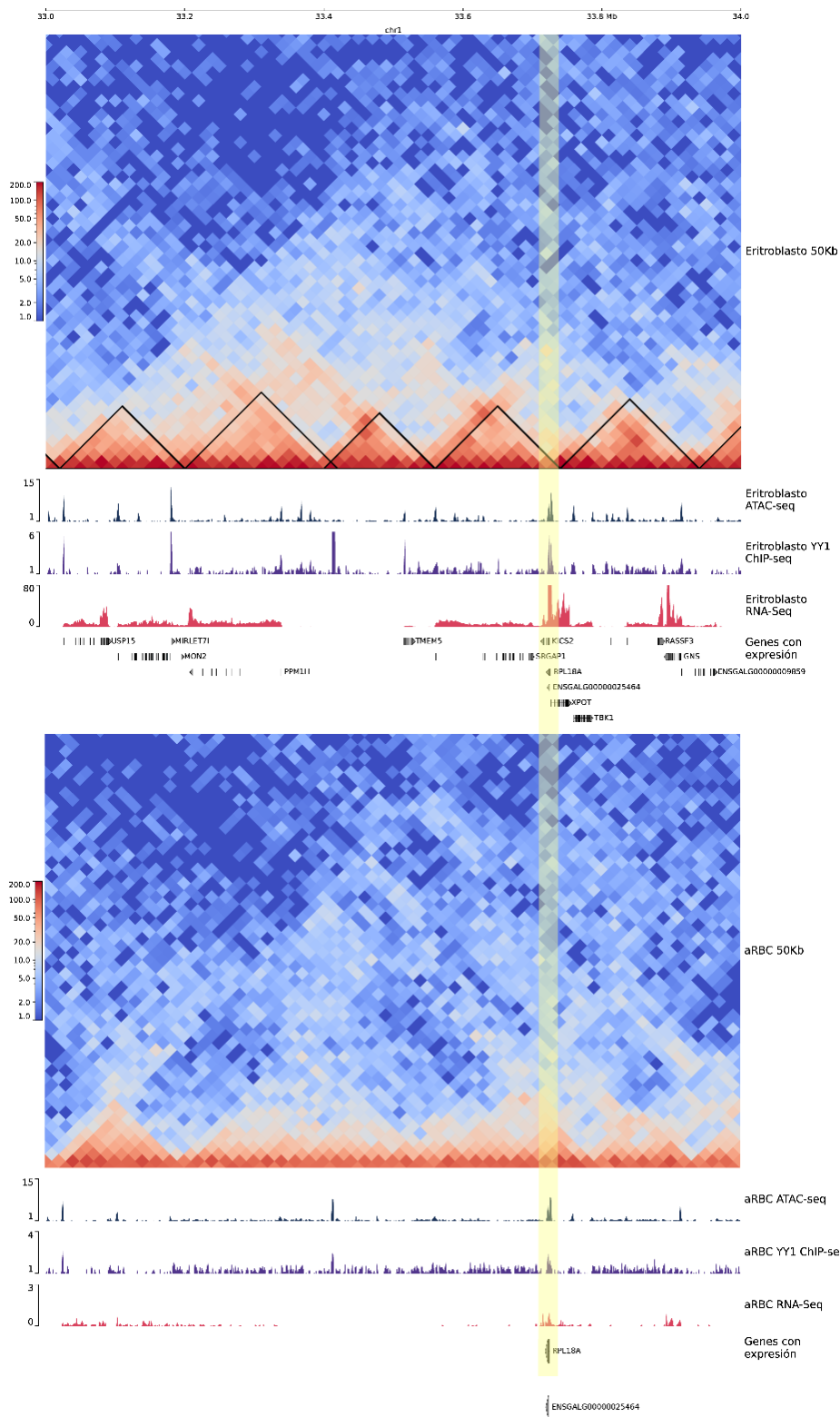


Figura 19. Visualización de picos de YY1 de genes activos fuera de dominios. RPL18A es un gen que presenta elevada actividad transcripcional en el estadio de eritroblasto, el cual se inactiva subsecuentemente durante la diferenciación y maduración eritroide. El enriquecimiento de señal de YY1 se conserva en aRBC. Esta zona muestra además una región donde un gen conserva su expresión incluso después de la pérdida de estructura del resultado de la diferenciación y maduración eritroide.

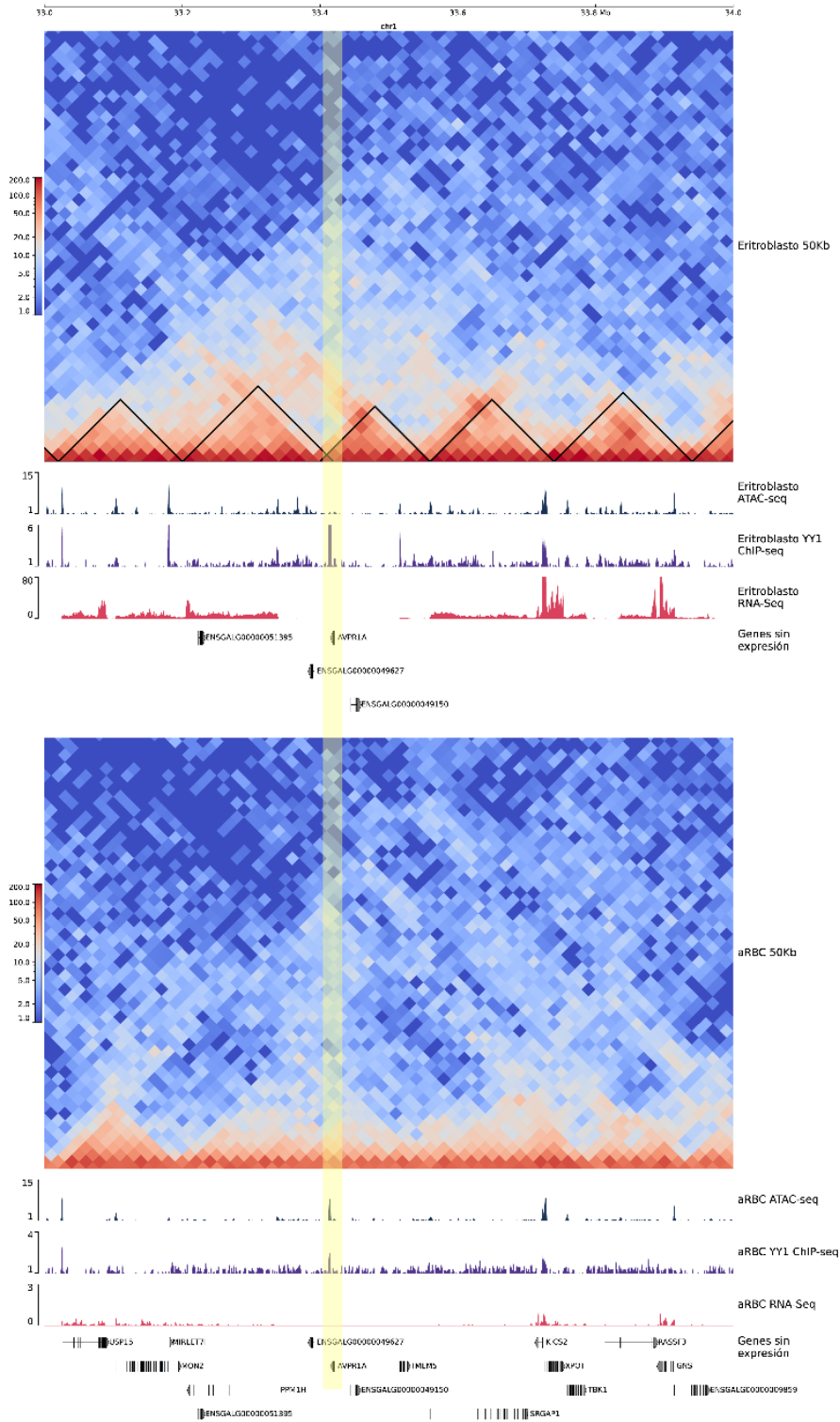


Figura 20. Visualización de picos de YY1 de genes inactivos fuera de dominios. AVPR1A es un gen que presenta nula actividad transcripcional desde el estadio de eritroblasto y hasta el eritrocito maduro. No obstante, es posible observar enriquecimiento de señal de YY1 a lo largo de la diferenciación. Esta zona corresponde a una región donde un gen está inactivo y estructurado, y continua de esta forma después de perderse la estructura topológica en aRBC.

9. Discusión y conclusiones

En el presente trabajo se estudió a YY1 como un potencial estructurador del genoma eritroide en eritrocitos maduros de pollo. Debido a que previos estudios del laboratorio detectaron que durante el proceso de diferenciación y maduración eritroide ocurría una pérdida generalizada de TADs acompañada de la disminución en expresión de proteínas estructurales, decidimos elucidar el posicionamiento de YY1 en el genoma eritroide para interrogar si este factor pudiera ser responsable de la estructura genómica remanente en minidominios en eritrocitos terminalmente diferenciados. Los resultados muestran que YY1, a nivel de proteína, disminuye significativamente a lo largo de la diferenciación terminal eritroide. No obstante, incluso en el estadio maduro, aRBC, YY1 aún es detectable por western blot e inmunofluorescencia. A diferencia de la marca H3K9me3 o la señal de ATAC-seq, YY1 no muestra una distribución preferencial dentro del núcleo por lo que YY1 parece encontrarse tanto en el centro nuclear como en la periferia. Al estudiar este factor transcripcional mediante ChIP-seq, fue posible determinar que YY1 se posiciona preferencialmente en regiones promotoras e intergénicas. Además, se encontró que YY1 presenta un posicionamiento preferencial en promotores de genes inactivos en aRBC y aun cuando YY1 no se localizan en fronteras de minidominios, más de la mitad de los picos se ubican dentro de los mismos en aRBC. Esto podría ser indicativo de un presunto papel de pausador transcripcional o de una configuración espacial necesaria para reactivar loci en los que se posiciona dentro de minidominios.

9.1. La proteína YY1 se preserva en los eritrocitos terminalmente diferenciados

Los experimentos de western blot e inmunofluorescencia evidenciaron la presencia de YY1 en todos los diferentes estadios de diferenciación y maduración eritroide estudiados. Aunque la abundancia de esta proteína disminuye conforme el estado de maduración avanza, la presencia en eritrocitos terminalmente diferenciados podría tener una relevancia mayor dado que YY1 ha sido descrito como una proteína con capacidad estructural para formar asas de cromatina de tipo promotor-*enhancer*. En el contexto de nuestro modelo, donde el eritroblasto compacta su núcleo al diferenciarse y los niveles de expresión de las principales proteínas estructurales disminuyen, la presencia de YY1

cobra relevancia como potencial estructurador de la porción del genoma que permanece estructurada en eRBC y aRBC.

No obstante, otro posible panorama de la presencia de esta proteína en núcleo que considera su naturaleza dual de activador y represor transcripcional podría estar relacionado con auxiliar a activar/reprimir la expresión de genes importantes durante la diferenciación. Modelos alternativos de diferenciación terminal de músculo han demostrado que YY1 actúa como un mecanismo de rescate para permitir la expresión de set de genes relacionados al desarrollo muscular cuando los complejos de CTCF-cohesina son alterados. La alteración de estos complejos está dada por cambios en los componentes de la matriz nuclear, como Matrin3²³. Trabajos previos del laboratorio demostraron que las aRBC pierden a la lámina A/C, provocando el desprendimiento de la heterocromatina de la periferia nuclear y la inversión de la eucromatina⁴¹. Existen reportes que hablan sobre la interacción de Matrin3 con la lámina A/C, donde se describe que el knockout de la lámina A/C en la envoltura nuclear resulta en la disminución en los niveles de Matrin3⁴². Aunque no se ha estudiado la ausencia de la proteína de Matrin3 en eritrocitos de pollo, su expresión se pierde en aRBC de acuerdo con datos de RNA-seq obtenidos en el laboratorio (no publicados)⁴⁵. De existir estas modificaciones en la composición de la matriz nuclear, los cambios observados en CTCF y cohesina en eritrocitos terminalmente diferenciados podrían justificar un modelo de rescate facilitado a través del reclutamiento de YY1 a minidominios para asegurar la expresión o participar en reestablecer la expresión de genes pausados previamente reportados.

9.2. YY1 no presenta patrones de distribución periférica nuclear en eritrocitos terminalmente diferenciados

La reorientación de la arquitectura nuclear dada por el posicionamiento de cromatina accesible hacia la periferia nuclear y H3K9me3 hacia el centro, en aRBC, dieron pauta a explorar si YY1 pudiera exhibir un patrón similar. Para eritroblastos y eRBC, la distribución de la señal de YY1 es similar a los fibroblastos donde la señal de DAPI es más fuerte en la periferia y la de YY1 se incrementa hacia el centro. Congruente con una arquitectura nuclear convencional. No obstante, en aRBC la señal de YY1 es ligeramente mayor que la de DAPI en la periferia, pero conserva una distribución uniforme a lo largo del núcleo. El ligero incremento de la señal de YY1 en la periferia nuclear podría indicar algún papel

relevante favoreciendo la accesibilidad de la cromatina. Sin embargo, no podría explicarla en su totalidad ya que la mayoría de la señal no está posicionada ahí.

Sabemos que YY1 tiene la capacidad dual de fungir como activador o represor transcripcional. Los dominios correspondientes al extremo N-terminal de GK y REPO presentan la capacidad de asociarse a proteínas remodeladoras de la cromatina. Específicamente, GK y REPO pueden asociarse a HDACs y proteínas del grupo Polycomb^{26,65-67}. Mientras que el dominio N-terminal con acetiltransferasas de histonas (HATs) como p300 o CBP. Esta dualidad respecto a su función justificaría su distribución generalizada dentro del núcleo al encontrarse en zonas de cromatina laxa y densa, según la señal de DAPI. De hecho, algunos reportes han identificado mediante inmunofISH y microscopía electrónica la distribución nuclear de YY1 a regiones pertenecientes a heterocromatina, eucromatina y las fronteras entre ambos tipos de cromatina⁶⁸.

9.3. Las distribuciones nucleares de las marcas H3K27Ac y H3K27me3 son similares a la de YY1 en eritrocitos maduros

En el contexto de la diferenciación eritroide, sabemos que los eritrocitos terminalmente diferenciados presentan estructuración remanente de su genoma en minidominios. Además, estas estructuras contienen regiones de cromatina con estado abierto y genes pausados por el pausado proximal de la RNA polimerasa II. De acuerdo con la naturaleza bivalente de YY1 de modificar la transcripción, exploramos las marcas de activación y represión H3K27Ac y H3K27me3. H3K27me3 es una marca asociada a heterocromatina facultativa^{31,60}, cromatina condensada y transcripcionalmente inactiva que tienen la capacidad de descondensarse y permitir la transcripción de genes en función de espacio y tiempo⁶⁹. Por el contrario, H3K27Ac es una marca antagónica y mutuamente excluyente de H3K27me3⁸. H3K27Ac ha sido establecida como un sello de *enhancers* activos. En el estudio que describe a YY1 como estructurador de asas de tipo promotor-*enhancer*, se resalta el posicionamiento preferencial de YY1 a sitios *enhancer*⁴⁹. Los resultados de la exploración de distribución nuclear de la marca de activación H3K27Ac y la de represión H3K27me3 revelaron un patrón similar al de YY1 en aRBC. Aunque por mucho tiempo se consideraron a los eritrocitos de pollo como células transcripcionalmente inertes, hoy sabemos que retienen niveles bajos de transcripción y un set de genes pausados los cuales podrían estar preparados para responder a estímulos. Por este motivo, hace

sentido pensar que YY1 podría estar posicionándose en regiones que permiten reprimir regiones del genoma o mantener activadas a otras.

H3K27me3 es una marca que el complejo represivo de Polycomb 2 (PRC2) coloca⁷⁰. El interactoma de YY1 está constituido por un vasto repertorio de proteínas entre las cuales se encuentran aquellas que constituyen a PRC2. Algunos reportes han demostrado que YY1 tiene la capacidad de reclutar a PRC2 a sitios del genoma que necesitan ser silenciados. Por ejemplo, YY1 funciona como un puente interactuando con DNA y el RNA *Xist* para reclutar a PRC2 al sitio de nucleación de *Xist* facilitando la inactivación de un cromosoma X⁷¹. Por otro lado, YY1 interactúa con EZH2, una metiltransferasa de lisinas contenida en PRC2, permitiendo su reclutamiento a sitios de modificación H3K27me3⁷²⁻⁷⁴. La relevancia de una posible interacción entre YY1 y PRC2 en eritrocitos terminalmente diferenciados radica en que YY1 podría fungir como enlace entre regiones que necesitan ser silenciadas durante la diferenciación terminal eritroide mediante la marca H3K27me3.

También, YY1 podría representar un factor importante que, junto con la interacción con otras proteínas como las que componen a la maquinaria transcripcional o modificadores de cromatina, eviten el silenciamiento total de locus importantes, mediante la transcripción *per se* o depositando la marca H3K27Ac. Evidencia proveniente de modelos de células troncales pluripotentes extendidas muestran que la pérdida de YY1 provoca la reducción de las marcas H3K4me3 y H3K27Ac de promotores de genes de pluripotencia⁷⁵. Adicionalmente, durante el desarrollo embrionario del ratón también se ha observado la importancia de YY1 en el enriquecimiento de la marca de activación⁷⁶. En los estadios de ocho células y la mórula se ha visto que las zonas de unión de YY1 ganan la marca de H3K27Ac, mientras que la delección de YY1 de estos sitios evita la transición de mórula a blastocisto⁷⁶. Esto hace evidente que YY1 es un factor indispensable en el remodelado de histonas permitiendo la activación de programas transcripcionales indispensables en procesos que involucran diferenciación celular y desarrollo.

Un posible dinamismo entre ambas marcas mediado por YY1 durante la diferenciación y maduración eritroide podría ser relevante en el establecimiento de un ambiente nuclear que permita a la transcripción remanente ejecutarse de forma adecuada y como un

mecanismo de silenciamiento. En otros organismos como el arroz, *Oryza sativa L. ssp. indica*, se ha estudiado cómo el dinamismo entre las marcas H3K27me3 y H3K27Ac permiten a los cultivos de arroz adaptarse a cambios ambientales de temperatura⁷⁷. El frío induce el enriquecimiento de la marca de acetilación sobre H3K27 permitiendo un rearrreglo en los nucleosomas brindando mayor accesibilidad a la maquinaria transcripcional en loci de respuesta a estrés favoreciendo la transcripción. Por otro lado, la respuesta al frío en el arroz también mostró efecto en la disminución de metiltransferasas impactando directamente a los niveles de H3K27me3. En conjunto, ambas modificaciones establecen un entorno transcripcional favorable y dinámico para responder a un estímulo externo al activar programas transcripcionales específicos⁷⁷. Este panorama podría ocurrir en el núcleo de los eritrocitos terminalmente diferenciados. Una importante primera aproximación a este escenario sería explorar la distribución de las marcas H3K27me3 y H3K27Ac respecto a YY1 mediante inmunofluorescencias. Al generar inmunofluorescencias dobles, H3K27me3-YY1 y H3K27Ac-YY1, podríamos aproximar una posible colocalización para cada una de las marcas con YY1 y establecer experimentos que nos permitan conocer la dinámica de estas modificaciones en H3K27 que potencialmente podrían ayudar a regular panorama transcripcional en eritrocitos de pollo.

9.4. YY1 se enriquece en promotores y regiones intergénicas, pero no en fronteras de minidominios en eritrocitos maduros

Exploramos el posicionamiento de YY1 en la hebra de cromatina en alta resolución mediante ChIP-seq. Debido a la dificultad inherente de trabajar con células con un núcleo altamente compacto, los experimentos de la réplica 1 de los diferentes estadios de diferenciación de interés fueron deficientes. Adicionalmente, las lecturas en eRBC mostraban mucho ruido, lo que dificultó el llamado de picos. Por tal motivo, decidimos excluir a eRBC de la mayor parte de la discusión. Los experimentos con eRBC serán realizados nuevamente para integrar todos sus resultados.

Tanto en eritroblastos como en aRBC encontramos que YY1 se posiciona preferencialmente en regiones promotoras e intergénicas. Estos resultados son congruentes con el reporte inicial⁴⁹ de YY1 como proteína arquitectónica en células de ratón, donde se encontró que YY1 se enriquece con la marca de *enhancers* activos H3K27Ac, permitiendo formar contactos con moléculas de YY1 posicionadas en

promotores a través de asas de cromatina. Esto evidencia que el posicionamiento de YY1 está altamente conservado en vertebrados. El mecanismo de homodimerización es análogo a la forma en la que CTCF forma fronteras de TADs. Ensayos *in vitro* revelaron la capacidad de YY1 de promover la circularización de moléculas de DNA que contienen los motivos de YY1⁴⁹. Esta capacidad de potenciar interacción de DNA podría fungir como mecanismo organizador de los minidominios si YY1 se encontrara en las fronteras. Para responder esto, analizamos el enriquecimiento de YY1 en fronteras en los tres estadios de diferenciación eritroide. Nuestros resultados evidenciaron que solo el 6% de picos de YY1 se encuentran en fronteras de minidominios en aRBC. De esta manera, es evidente que YY1 no tienen como función principal estructurar a los minidominios de una forma análoga a CTCF en eritrocitos maduros.

Sin embargo, otros mecanismos indirectos controlados por YY1 podrían estar involucrados en la estructuración. Considerando que YY1 tiene la capacidad de interactuar con RNAs y que la mayoría de los transcritos presentes en eritrocitos son de RNA pequeños no codificantes^{41,45}, el mecanismo podría involucrar a estos RNAs con YY1 como guía. Por tal, resulta relevante investigar a futuro que especies de RNAs presentes en los eritrocitos pueden interactuar con YY1, pues se ha discutido que estos RNAs podrían formar un papel importante en la formación de heterocromatina⁴⁵, un componente importante en la topología genómica⁷⁸. Será interesante analizar si YY1 se encuentra enriquecido formando asas de cromatina entre promotores de genes activos y *enhancers* en eritroblastos que posteriormente pierden estructuración y expresión en eritrocitos, pero preservan a YY1.

Al analizar a los picos de YY1 desde la perspectiva de compartimentos A y B encontramos que la mayoría de ellos se encuentran en el compartimento A, asociado a eucromatina. También observamos que la proporción de picos de YY1 en el compartimento B crece en eritrocitos maduros. Es decir, en eritrocitos maduros existe un incremento de picos de YY1 que se encuentran en regiones heterocromatinizadas del genoma. La ganancia de esta proporción es indicativa de que algunas regiones donde YY1 se une designadas como A o A/B en eritroblastos, fueron silenciadas pasando a B en aRBC; un fenómeno esperable para este modelo. Estas observaciones refuerzan la idea de que YY1 participa en el silenciamiento de algunas regiones, como se discutió en la sección anterior. Algunos reportes han descrito la influencia de YY1 en el silenciamiento mediante la formación de

heterocromatina al reclutar HDACs que favorecen una conformación cerrada de la cromatina^{68,79}. Por ejemplo, un reporte en cáncer de mama triple negativo mostró que YY1 recluta a HDACs a promotores de genes para controlar su expresión⁸⁰, además de ser capaz de regular la expresión *per se* de HDAC1⁷⁹. Independiente a las deacetilasas, YY1 ha sido descrito como un regulador positivo de HP1 α , una proteína indispensable en el silenciamiento génico⁸¹. A esta serie de candidatos debemos agregar al complejo PRC2, por los motivos previamente descritos en el silenciamiento genómico.

Un ejemplo proveniente de este estudio que relaciona YY1 con el silenciamiento proviene del gen AVPR1A. AVPR1 es uno de los tres receptores de vasopresina presentes en el cerebro, hígado, riñón y vasculatura⁸². La vasopresina es una hormona que tiene la capacidad de estimular la proliferación y diferenciación de precursores hematopoyéticos favoreciendo la producción de eritrocitos⁶⁴. En células troncales hematopoyéticas y progenitoras, todos los receptores de vasopresina están expresados y disminuyen en estadios tempranos de la diferenciación terminal eritroide⁶⁴. En el pollo, es claro que AVPR1A están inactivo en el estadio de eritroblasto cuyo locus está acompañado por la presencia de YY1 (ver **Figura 20**). Esta configuración se mantiene hasta el eritrocito maduro, lo que sugiere que YY1 podría ser responsable de mantener su silenciamiento desde su establecimiento y a lo largo de la diferenciación terminal eritroide. En conjunto, todas las observaciones anteriores hacen evidente la posible relación compleja entre el posicionamiento de YY1 y el silenciamiento de las regiones mediante la formación de heterocromatina.

9.5. YY1 se posiciona dentro y fuera de minidominios

Continuando con la integración de los datos topológicos y los picos de YY1 encontramos que, tanto en eritroblastos como aRBC, más del 50% de los dominios contiene picos de YY1. Sin embargo, aunque el 80% de los picos de YY1 se encuentran en TADs en eritroblastos, en aRBC esta proporción se reduce a solo el 47%. La observación que el 53% de los picos de YY1 se encuentren fuera en minidominios en aRBC podría parecer confusa, ya que las estructuras como los minidominios están asociados a regiones donde la transcripción ocurre al mantener regiones eucromáticas estructuradas⁴¹. No obstante, la mayoría de los picos de YY1 se encuentran en el compartimento A. Al realizar una inspección visual a regiones con picos de YY1 donde no hay de minidominios, es posible identificar estructuras similares.

Este panorama podría ser consecuencia directa de que los métodos bioinformáticos para realizar la identificación de TADs no son infalibles, y pueden omitir el llamado de algunos minidominios o regiones que presentan estructuración remanente pero que son dependientes a otros mecanismos. Por ejemplo, de forma reciente fue descrita una nueva estructura conocida como “fuente” que representa el sitio de inicio de la extrusión para el plegado de la cromatina mediadas por la cohesina durante el desarrollo⁸³. Las fuentes están altamente relacionadas con elementos *enhancer* y se encontró que su fuerza está relacionada con la marca H3K27Ac⁸³. De hecho, algunos minidominios encontrados en eRBC y aRBC, visualmente se asemejan a estructuras fuente. Por el momento, estas estructuras han sido descritas durante el desarrollo temprano del pez cebra y la rana, por lo que futuras investigaciones determinaran la presencia de estas estructuras en la diferenciación eritroide del pollo. Finalmente, la mera presencia de YY1 con factores, como la maquinaria de transcripción^{41,84} o RNAs asociados a la cromatina²⁸, podrían indirectamente proveer de estructura a regiones identificadas como compartimento A, que están fuera de minidominios en aRBC.

9.6. La unión de YY1 a sus blancos está acompañada por cambios en la accesibilidad del genoma

Métodos como ATAC-seq han permitido mapear los cambios de accesibilidad a lo largo del genoma. En el pollo, el número de picos de accesibilidad detectados es similar en todos los tres estadios de diferenciación estudiados. Los eritrocitos exhiben dos clases de picos de accesibilidad: picos de regiones constitutivamente abiertas o aquellas que se comparten con eritroblastos y mantienen su accesibilidad a lo largo de la diferenciación terminal⁴¹. Al intersecar los picos de YY1 con los picos de ATAC-seq descubrimos dos categorías de picos principales. La primera corresponde a aquellos picos que colocalizan con picos de accesibilidad. La segunda con los que no tienen accesibilidad. Los cambios entre los eritroblastos y los eritrocitos son drásticos, pues la proporción de picos sin accesibilidad en eritroblastos es mucho mayor que los eRBC y aRBC.

En un inicio, esta noción podría parecer contradictoria pues la unión de factores de transcripción está generalmente asociada a cambios en la accesibilidad de la cromatina y por ende a cambios en la expresión génica⁸⁵. No obstante, estos resultados podrían explicarse por la sensibilidad de los métodos bioinformáticos de llamado de picos de ATAC-seq. En eritroblastos pueden existir picos de ATAC-seq fuertes y otros más débiles.

Al realizar el llamado de picos, el umbral de detección establecido puede provocar la exclusión de picos débiles. Por tanto, algunos picos de YY1 en el genoma podrían coincidir con regiones que fueron inicialmente excluidas por no ser picos fuertes en el pipeline de análisis de ATAC-seq.

Por otro lado, otro escenario posible sería la discordancia entre la unión del factor de transcripción y la generación de picos de accesibilidad. Un ejemplo de este tipo en nuestro estudio vuelve al gen AVPR1A en eritroblastos, donde observamos un pico evidente de YY1 sin accesibilidad. Algunos estudios sugieren que el eje “unión-accesibilidad-cambio en la expresión” no siempre es concordante⁸⁶⁻⁸⁸. Se ha sugerido que pueden existir grandes cambios en la expresión génica sin que existan cambios en la accesibilidad de la cromatina^{86,87}. De la misma forma, la unión de factores de transcripción no necesariamente implica accesibilidad detectable en interfase⁸⁸.

9.7. YY1 podría ser importante para el mantenimiento del pausado de genes estructurados y la estructuración local de la cromatina

Los resultados de integrar los datos de RNA-seq de genes expresados o no expresados con los picos de YY1 revelaron que, en aRBC, YY1 se posiciona de forma preferencial en promotores de genes inactivos. De forma interesante, en eritroblastos observamos un enriquecimiento en la misma zona, pero de genes activos. Lo anterior es indicativo que YY1 se conserva unido a regiones que perdieron expresión a lo largo de la diferenciación terminal y maduración eritroide.

Además, una proporción importante de los picos de YY1 que no tiene expresión se encuentra accesible y estructurada en minidominios. Esta configuración genómica corresponde a la previamente reportada donde los minidominios eran encontrados alrededor de zonas con cromatina abierta y que no presentaban actividad transcripcional⁴¹. En este reporte, los minidominios fueron clasificados en función de su estado transcripcional y estados cromatínicos. Se reportó que 80% de los minidominios contienen promotores de genes con accesibilidad y sin expresiones referidas como “promotores pausados”, o una combinación de promotores pausados y activos⁴¹. Esta evidencia apunta a que los promotores no expresados, accesibles y estructurados con YY1 representan un subconjunto de los promotores pausados inicialmente reportados.

En el grupo de investigación se encontró evidencia que relaciona al pausado proximal de la RNA polimerasa II con el control transcripcional de genes pausados y su relevancia en el mantenimiento de la organización local del genoma^{41,45}. En ese trabajo se hipotetizó que el mecanismo de pausado proximal podría preparar a los eritrocitos de pollo para restaurar la transcripción de elementos importantes de la maquinaria transcripcional y de procesamientos de RNA con el objetivo de montar una respuesta ante un estímulo⁴¹. La existencia de los picos de YY1 en promotores accesibles, no expresados y estructurados podría implicar que, dentro de minidominios, YY1 podría ser relevante en el pausado de estos loci génicos. De hecho, al explorar nuestra lista de genes de promotores, encontramos algunos como TAF1A y POLR3H que corresponden a promotores pausados (ver Figuras 17 y 18).

Por otro lado, la presencia de RNA polimerasa II pausada y YY1 en algunos loci génicos podría ser necesaria para la reactivación de estos en respuesta a un estímulo. En este escenario, además de actuar sinérgicamente con la RNA polimerasa II promoviendo la estructuración local del genoma, YY1 potenciaría la expresión de un gen tan pronto como la RNA polimerasa II pausada se libera. Existen evidencias separadas que apoyan esta noción. Por ejemplo, YY1 potencia la expresión de HDAC1 al posicionarse directamente en su promotor⁷⁹ y se ha reportado que este factor transcripcional es un potente regulador de la transcripción dependiente de la RNA polimerasa II⁸⁹. Además, un estudio acerca de las funciones de la eritropoyetina durante la diferenciación eritroide de ratones encontró que YY1 se redistribuye en el genoma a genes responsivos a eritropoyetina. Estos genes contienen a la RNA polimerasa II en estado pausado, posiblemente para montar una respuesta transcripcional rápida ante el estímulo de eritropoyetina⁹⁰. De esta manera, la combinación de ambos componentes permite la activación genes ante el estímulo de eritropoyetina. En el pollo, existen reportes de estimulación inmunológica a eritrocitos que resulta en la expresión genes relacionados a respuesta inmune como interleucinas, interferones y TLRs⁹¹⁻⁹³. La configuración YY1-RNA polimerasa II pausada podría contribuir en la activación de genes pausados para responder a estímulo inmunes.

Los resultados de las GOs de genes accesibles, no expresados y estructurados están involucrados en el metabolismo de compuestos nitrogenados y RNAs, control de la expresión génica, traducción y organización o modificación de complejos DNA-proteína. Algunos ejemplos de genes interesantes dentro de estas categorías son TAF1A y

POLR3H. TAF1A codifica para una subunidad de la RNA polimerasa I que permite la formación del complejo de preiniciación para la transcripción de RNAs ribosomales⁶². Por otro lado, el producto del gen POLR3H es la subunidad H de la RNA polimerasa III, la cual está involucrada en la transcripción de RNAs pequeños no codificantes y RNAs ribosomales como U6 y tRNAs⁶³. De forma interesante, uno de los RNAs pequeños no codificantes encontrados con expresión elevada en eritrocitos es 7SK⁴⁵, quien es transcrito por la RNA polimerasa III⁹⁴. También, es importante destacar la detección mediante GO del proceso biológico de transcripción por la RNA polimerasa III en genes inactivos con YY1 en su promotor. En conjunto, la reactivación de este subconjunto de genes pausados permitiría montar una respuesta al reactivar genes necesarios para la transcripción y traducción, además de proveer de retroalimentación negativa al pausado de la polimerasa para mantener la organización de la topología local.

En conclusión, la naturaleza dual de YY1 como activador y represor transcripcional se ve reflejada en los escenarios genómicos encontrados durante la diferenciación terminal eritroide del pollo. Estos efectos estarían mediados directamente por los complejos proteicos con los que interactúe dependiente del contexto genómico. Los resultados presentados sugieren que YY1 podría tener un papel importante al contribuir indirectamente a la estructuración local del genoma de la mano del pausado proximal de la RNA polimerasa II. De esta manera, YY1 podría contribuir a fomentar un entorno idóneo para la reactivación transcripcional ante estímulos externos, como desafíos inmunológicos. Además, la presencia de YY1 en regiones de heterocromatina y loci silenciados a lo largo de la diferenciación, tiene implicaciones importantes en el estudio de los mecanismos de silenciamiento global en este modelo (**Figura 21**).

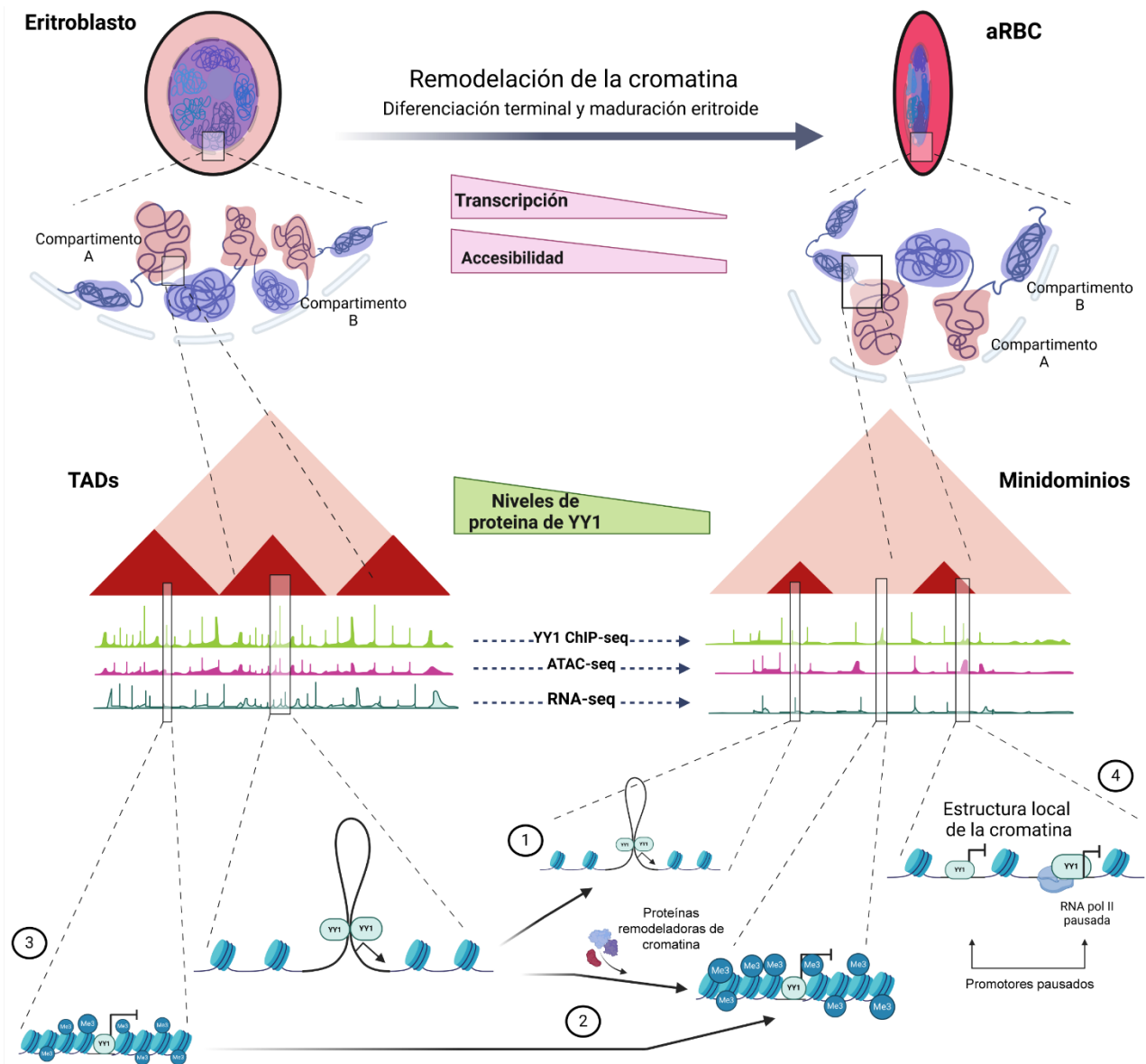


Figura 21. Modelo de la participación de YY1 en la diferenciación y maduración eritroide. A través del proceso de diferenciación terminal y maduración, YY1 se posiciona dentro y fuera de minidominios principalmente en regiones asociadas a eucromatina. Los sitios donde YY1 se posiciona en aRBC son principalmente regiones que conservan accesibilidad, promotores sin expresión pero que permanecen estructurados dentro de minidominios. Enumerados de 1-4 encontramos los escenarios genómicos donde YY1 fue encontrado y se describen las diferentes hipótesis de los mecanismos en los que YY1 podría estar involucrado. 1. Dentro de minidominios, YY1 permite mantener la expresión de algunos genes promoviendo la accesibilidad de la región probablemente mediante el reclutamiento de proteínas remodeladoras, como HATs, a zonas cercanas a sus promotores. 2. YY1 participa en el silenciamiento génico al reclutar a remodeladores de la cromatina como HDACs o PRC2 durante la pérdida de dominios y transición de regiones A a regiones B, que acompañan la diferenciación terminal eritroide. 3. YY1 permite mantener reprimidos a genes no específicos del linaje eritroide que fueron silenciados previo al estadio eritroide, como AVPR1A. 4. YY1 contribuye al pausado de genes estructurados, posiblemente cooperando con la RNA polimerasa II y otros factores para mantener la estructura local de la cromatina y permitir la reactivación de estos loci génicos en respuesta a estímulos específicos.

10. Perspectivas

Este trabajo representa una pieza más del rompecabezas sobre la caracterización de los cambios en el escenario genómico durante la diferenciación terminal eritroide. La posible participación de YY1 con la RNA polimerasa II en el pausado proximal y mantenimiento en la estructuración local del genoma abren un nuevo panorama de preguntas interesantes. En un inicio, realizar réplicas adicionales de ChIP-seq YY1 en nuestras células resulta necesario para robustecer nuestros resultados y poder integrar aquellos relacionados con eRBC. Además, la integración de los datos de ChIP-seq YY1 con ChIP-seq de la RNA polimerasa II pausada (fosforilada en la serina 5) permitiría conocer loci donde ambas proteínas estén posicionadas. Subsecuentemente, explorar los cambios en la distribución de YY1 ante estímulos inmunes que posibiliten la activación transcripcional de genes pausados en los eritrocitos.

De la misma manera, explorar los mecanismos de silenciamiento mediados por YY1 en eritrocitos maduros es un componente importante. La interacción de YY1 con PRC2 y HATs podría ser un mecanismo importante en el establecimiento de ambientes nucleares específicos que repriman o activen genes. Por este motivo, resultaría interesante estudiar la colocalización de las marcas de histonas H3K27Ac y H3K27me3 mediante experimentos de inmunofluorescencia y ChIP-seq.

Finalmente, los RNAs codificantes están cobrando relevancia en el campo de la topología genómica. YY1 ha sido descrito como un factor de transcripción capaz de interactuar con RNA. Por tal motivo, estudiar las interacciones de este factor con RNAs no codificantes podría iluminar el camino hacia mecanismos estructurales mediados por RNA aún no descritos.

11. Anexos

11.1. Comparación de las proteínas homólogas YY1 en humano y pollo

Debido a la falta de anticuerpos producidos específicamente para este modelo, se decidió investigar cual anticuerpo podría ser un buen candidato para utilizarse en los experimentos. Por tal motivo, decidimos comparar el grado de similitud de las proteínas humana y de pollo. Nuestros anticuerpos candidato anti-YY1 (Cell Signaling (D5D9Z) y Diagenode (C15410345)), reaccionaban contra epítomos mapeados a la región C-terminal de YY1. Los alineamientos realizados demostraron que la proteína YY1 humana y de pollo son exactamente iguales en dicha región, y presentan mayor variabilidad en la región N-terminal y central. Por tal motivo, decidimos utilizar ambos anticuerpos y probar su especificidad mediante experimentos de western blot e inmunofluorescencia. Se concluyó que el mejor candidato, dada su especificidad, es el anticuerpo anti-YY1 de Cell Signaling.

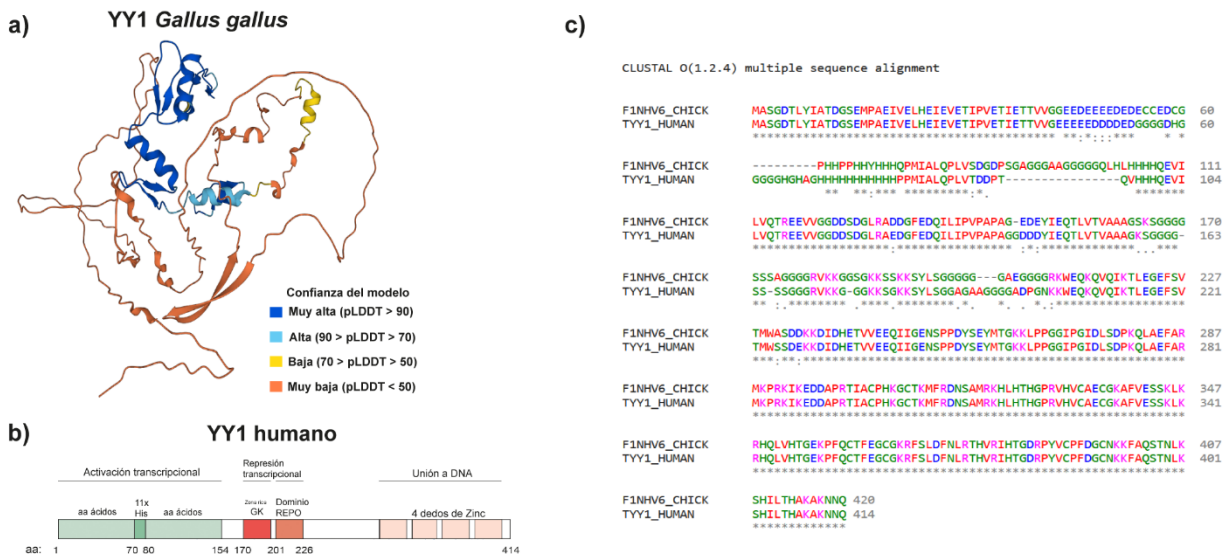


Figura Suplementa 1. Estructura de YY1. a) Modelo de la estructura terciaria de YY1 a través de AlphaFold²⁵². Puntajes altos de pLDDT (>80) indican alta confianza en la predicción de la estructura. Por el contrario, valores bajos (<50) son indicadores de regiones donde los residuos de la proteína podrían estar intrínsecamente desordenados (IDRs). b) Diagrama de los dominios que conforman a YY1. El dominio REPO y GK están asociados a represión transcripcional. El dominio N-terminal está ligado a funciones de activación transcripcional, mientras que el C-terminal permite la unión a DNA. Figura modificada de: Verheul, T. C. J. et al. (2020)²⁶ c) Alineamiento de la secuencia de aminoácidos de las proteínas YY1 humana y de pollo. Se observa gran similitud entre las secuencias en la región de unión a DNA.

11.2. Western blot e inmunofluorescencia contra YY1

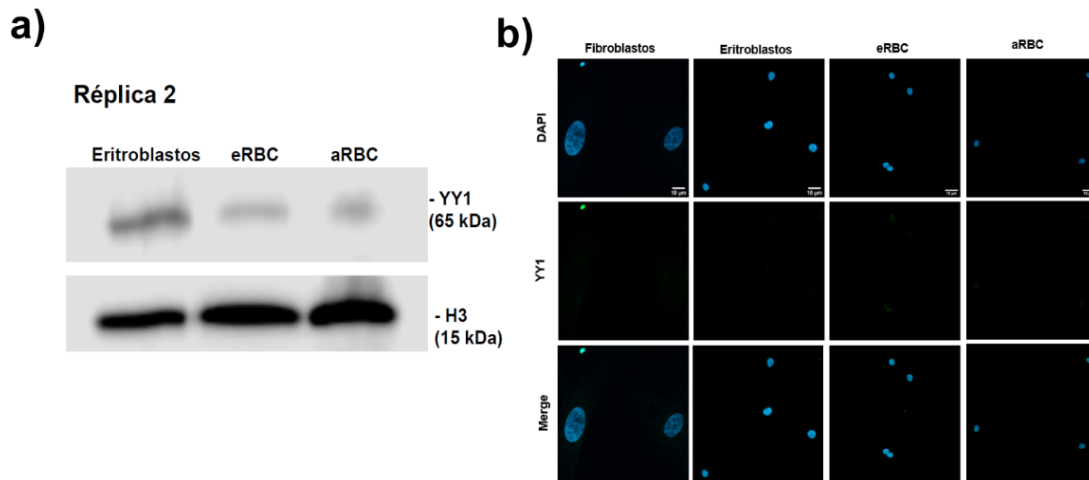


Figura Suplementaria 2. a) Resultado de la réplica 2 de western blot contra YY1 utilizada en la cuantificación. b) Controles negativos para cada experimento de inmunofluorescencia contra YY1 realizado.

11.3. Establecimiento de ChIP contra YY1 en eritrocitos embrionarios

Se fijaron 150 millones de eRBC con formaldehído al 1%. El material fue sonificado para obtener fragmentos entre 200-500 pb (**Figura Suplementaria 3a**). Subsecuentemente, el material fue cuantificado y se dividió en tres alícuotas de 50 ug de cromatina, una para el experimento mock (sin anticuerpo) y dos para cada anticuerpo a probar (Cell Signaling o Diagenode) (**Figura Suplementaria 3b**). Se diseñaron diferentes oligonucleótidos candidatos para regiones control y aquellos usados en el experimento ChIP de YY1⁵³. En este caso, el par de oligonucleótidos denominado A120P, contiene un sitio de unión de YY1 situado en una porción del *enhancer* de la región 3'- del locus de α -globina. Por otro lado, el otro par denominado $\beta^{A/\epsilon}$, contiene una porción del *enhancer* β -globina, el cual fue utilizado como control negativo pues en este sitio no se detectó la unión de YY1. Subsecuentemente, los oligonucleótidos fueron estandarizados mediante PCR punto final y qPCR utilizando DNA genómico de pollo para determinar la Tm adecuada para cada par.

Una vez estandarizados, se realizó la inmunoprecipitación (IP) de la cromatina para cada anticuerpo y el control mock. Subsecuentemente, el DNA obtenido de las IPs fue purificado y se evaluó la calidad del experimento mediante qPCR. Los resultados de la qPCR con el material del ChIP muestran que el realizado con el anticuerpo de Cell Signaling fue el más eficiente, en contraste a Diagenode (**Figura Suplementaria 3c-d**).

El enriquecimiento observado con el anticuerpo de Cell Signaling si bien es mayor, no se observaron cambios drásticos en tal enriquecimiento al compararlo con el experimento mock. Utilizamos un input genómico y primers para IFN α como control interno de la qPCR. Se piensa que la dificultad para realizar ChIP en células con cromatina altamente compacta representa un desafío metodológico responsable de estos resultados. El desarrollo de técnicas como CUT&RUN podrían ser de utilidad para afrontar estos problemas a futuro. Sin embargo, considerando los resultados previos de western blot e inmunofluorescencia, se decidió realizar el set completo de experimentos en las células HD3, eRBC, aRBC e hígado de ratón únicamente con Cell Signaling.

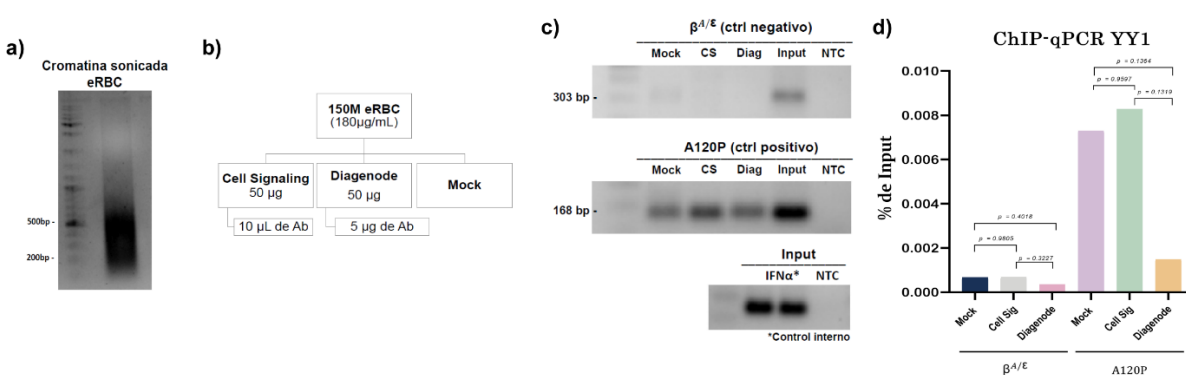


Figura Suplementaria 3. Establecimiento de ChIP-qPCR contra YY1 en eRBC *Gallus gallus*. **a)** Cromatina de eRBC sonicada. Se muestra un gel de agarosa donde la mayoría de las poblaciones de moléculas de DNA se distribuyen en los rangos de entre 200-500 pb. **b)** Distribución de material utilizado para el experimento de ChIP en eRBC. **c)** Productos amplificados utilizando los oligonucleótidos controles, positivo A120P y negativo $\beta^{A/\epsilon}$, en el material de ChIP purificado. Se observa una mayor intensidad en el amplicón del control positivo donde se utilizó el anticuerpo de Cell Signaling. Se utilizó un control interno contra IFN alfa con DNA genómico. NTC: control sin templado. **d)** Cuantificación de YY1 ChIP-qPCR y análisis estadístico mediante una ANOVA de un factor, seguido de una prueba post hoc de Tukey.

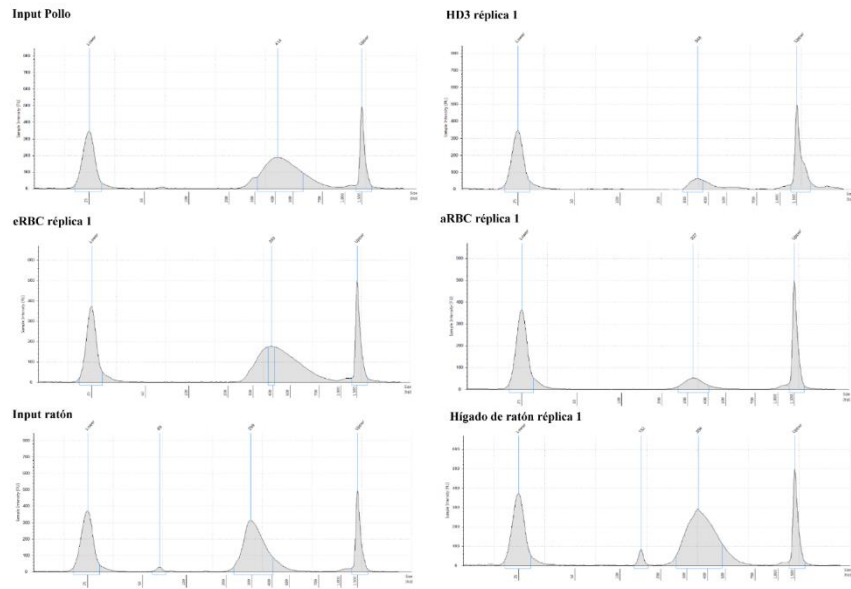


Figura Suplementaria 4. Electroforesis capilar de bibliotecas de ChIP-seq de YY1 réplica 1. Un análisis de electroforesis capilar se realizó mediante el equipo 4200 TapeStation System (Agilent) para determinar si los tamaños de la biblioteca construida para secuenciación eran adecuados. Se muestran las réplicas 1 e inputs pertenecientes a células eritroides de pollo e hígado de ratón.

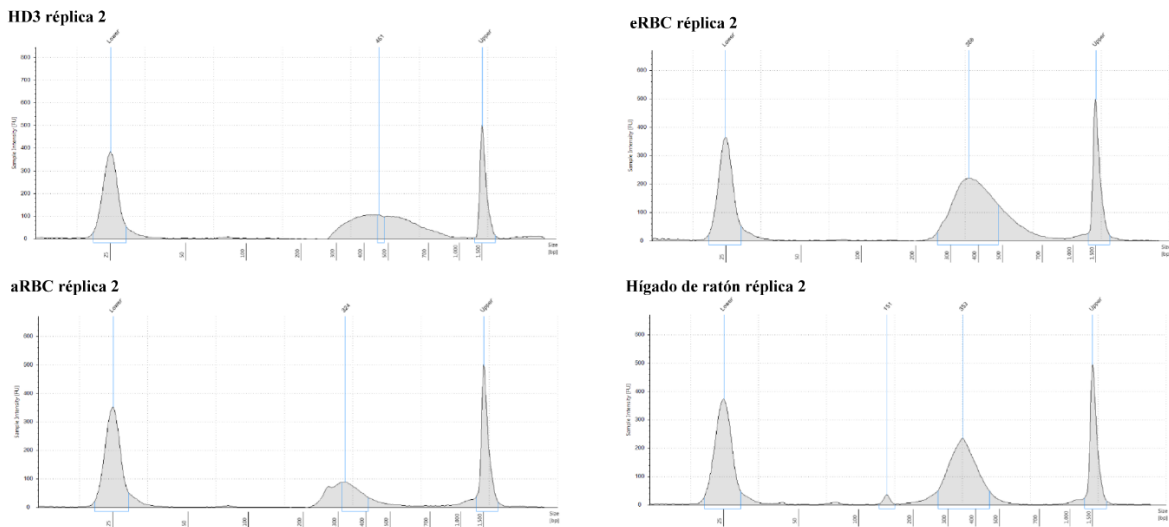


Figura Suplementaria 5. Electroforesis capilar de bibliotecas de ChIP-seq de YY1 réplica 2. Un análisis de electroforesis capilar se realizó mediante el equipo 4200 TapeStation System (Agilent) para determinar si los tamaños de la biblioteca construida para secuenciación eran adecuados. Se muestran las réplicas 2 pertenecientes a células eritroides de pollo e hígado de ratón.

11.4. Análisis de calidad de secuenciación y mapeo

Una vez que recibimos las secuencias de nuestras muestras, un análisis de calidad de secuenciación se realizó con el fin de esclarecer el tipo de procesamiento de los datos. Mediante FASTQC se determinó que las todas las muestras secuenciadas exhiben un

Phred Score superior a 30, permitiéndonos continuar con nuestro análisis (**Figura Suplementaria 6**).

Una vez comprobada la calidad de las secuencias, se realizó el mapeo de las lecturas al genoma de referencia correspondiente mediante MACS2 y se removieron las dúplicas de PCR. Los resúmenes de calidad de alineamiento muestran, a excepción de la réplica dos de la muestra de hígado de ratón (57%), porcentajes de alineamiento superior a 86% (**Figura Suplementaria 7**). Para los análisis posteriores únicamente fueron consideradas aquellas lecturas que mapeaban de forma única a su genoma de referencia.

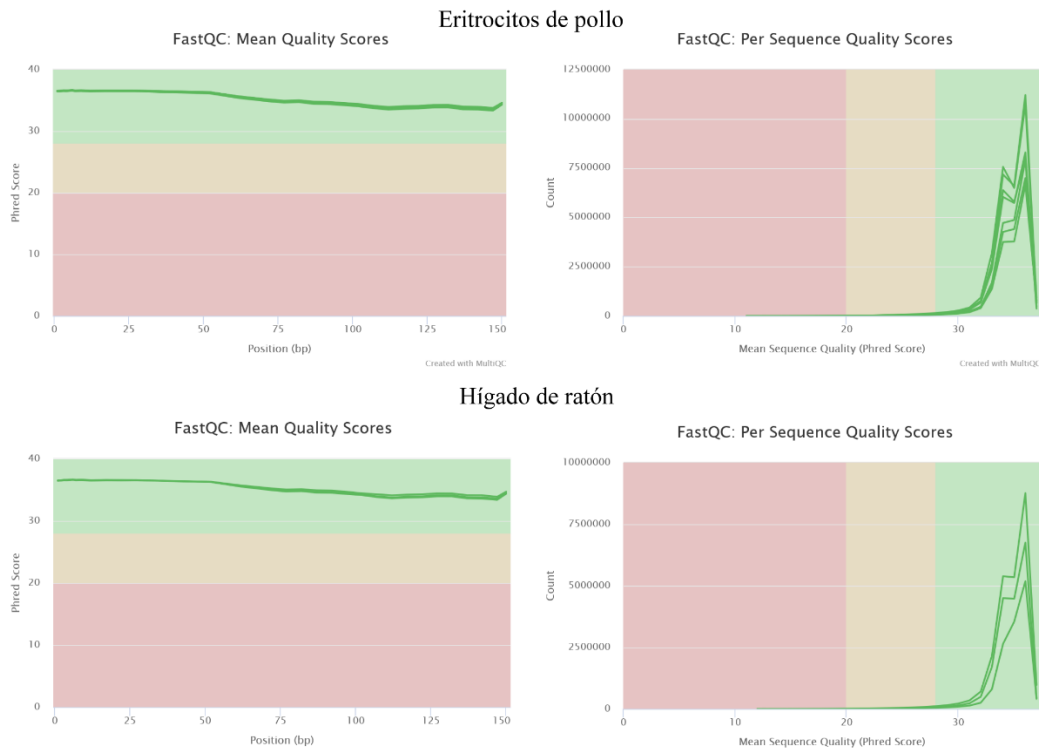


Figura Suplementaria 6. Análisis de calidad de secuenciación. Se muestran los resúmenes de los análisis de calidad de secuenciación realizados mediante FASTQC, y compactados con MultiQC. Panel superior: muestras de células eritroides de pollo. Panel inferior: muestras de hígado de ratón.

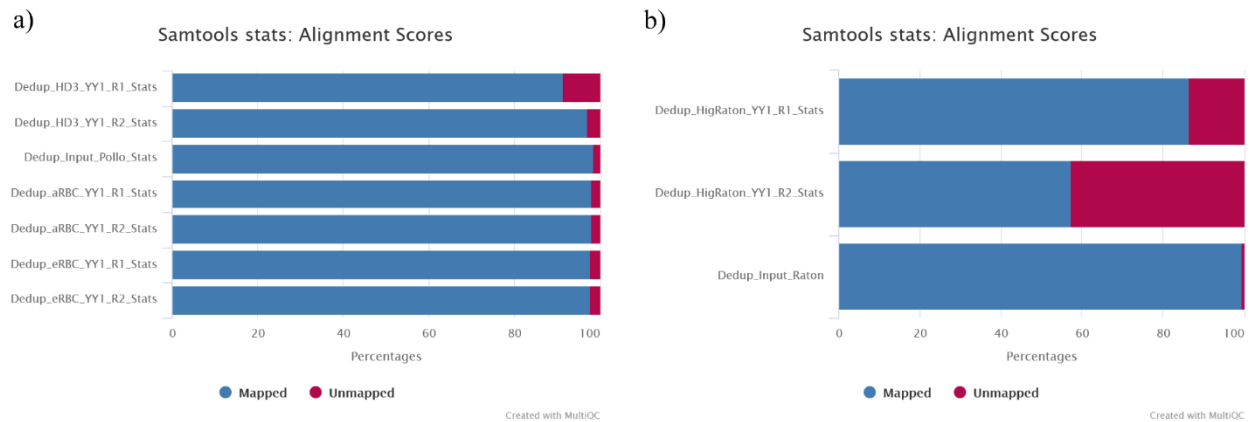


Figura Suplementaria 7. Estadística de mapeo. a) Resúmenes de mapeo de las muestras eritroides de pollo. b) Resúmenes de mapeo de las muestras hígado de ratón.

Tabla Suplementaria 1. Número de lecturas mapeadas al genoma de referencia y dúplicas.

| Muestra | Lecturas mapeadas (millones) | Lecturas totales (millones) | Dúplica de PCR (millones) |
|-----------------------|------------------------------|-----------------------------|---------------------------|
| Dedup_HD3_YY1_R1 | 17.5 | 19.1 | 1.7 |
| Dedup_HD3_YY1_R2 | 16.8 | 17.3 | 1.5 |
| Dedup_Input_Pollo | 29.9 | 30.3 | 1.3 |
| Dedup_aRBC_YY1_R1 | 30.4 | 31 | 1.4 |
| Dedup_aRBC_YY1_R2 | 20.7 | 21.1 | 0.7 |
| Dedup_eRBC_YY1_R1 | 24.6 | 25.1 | 2.03 |
| Dedup_eRBC_YY1_R2 | 24.1 | 24.6 | 2.1 |
| Dedup_HigRaton_YY1_R1 | 16.5 | 19.1 | 1.5 |
| Dedup_HigRaton_YY1_R2 | 7.6 | 13.3 | 1.9 |
| Dedup_Input_Raton | 24.3 | 24.4 | 1.5 |

11.5. Cálculo de factores de normalización Spike-In para ChIP-seq

Para tener un mayor control del experimento de ChIP-seq y la detección de picos de YY1, se empleó DNA exógeno de *Drosophila melanogaster* como spike-in. Los factores para la normalización para cada IP fueron calculados mediante la siguiente ecuación⁵⁴:

$$\text{Factor de normalización} = \frac{\text{Muestra con el menor \# de lecturas alienadas de forma única al genoma exógeno}}{\text{\# de lectura alineadas de forma única al genoma exógeno de cada IP}}$$

En este caso, aRBC_R2 fue la muestra con menor número de lecturas que alinearon de forma única al genoma de la mosca (**Tabla Suplementaria 2**). Únicamente se realizó esta normalización para las muestras del pollo.

Tabla Suplementaria 2. Factores de normalización por Spike-In

| Muestras | Lecturas mapeadas de forma única al genoma de <i>D. melanogaster</i> | Factor |
|-------------|----------------------------------------------------------------------|------------|
| Input Pollo | 112122 | 1.22826029 |
| HD3_R1 | 162790 | 0.8459672 |
| HD3_R2 | 161598 | 0.85220733 |
| eRBC_R1 | 241124 | 0.57113767 |
| eRBC_R2 | 213952 | 0.64367241 |
| aRBC_R1 | 297174 | 0.46341537 |
| aRBC_R2 | 137715 | 1 |

11.6. Resultados de llamado de picos de YY1

Se muestra un resumen de la cantidad de picos filtrados llamados a través de los métodos *BroadPeaks* y *NarrowPeaks*.

Tabla Suplementaria 3. Número de picos de YY1 identificados

| Muestras | Número de picos |
|------------------------------|-----------------|
| aRBC_YY1_R1_BPeaks | 743 |
| aRBC_YY1_R2_Bpeaks | 3396 |
| aRBC_YY1_R1_NPeaks.bed | 399 |
| aRBC_YY1_R2_NPeaks.bed | 1921 |
| eRBC_YY1_R1_BPeaks.bed | 231 |
| eRBC_YY1_R2_BPeaks.bed | 829 |
| eRBC_YY1_R1_NPeaks.bed | 104 |
| eRBC_YY1_R2_NPeaks.bed | 451 |
| HD3_YY1_R1_BPeaks.bed | 4698 |
| HD3_YY1_R2_BPeaks.bed | 15667 |
| HD3_YY1_R1_NPeaks.bed | 3915 |
| HD3_YY1_R2_NPeaks.bed | 8036 |
| HigRaton_YY1_R1_BPeaks_peaks | 3109 |
| HigRaton_YY1_R1_NPeaks_peaks | 2889 |
| HigRaton_YY1_R2_BPeaks_peaks | 8498 |
| HigRaton_YY1_R2_NPeaks_peaks | 4714 |

11.7. Análisis de picos llamados por BroadPeaks

Se muestra el anotado de picos de los picos de YY1 llamados por *BroadPeaks*, los picos comunes entre réplicas y el análisis de búsqueda de motivos. Además, se realizaron análisis motivos a los picos no comunes en ambos métodos de llamado.

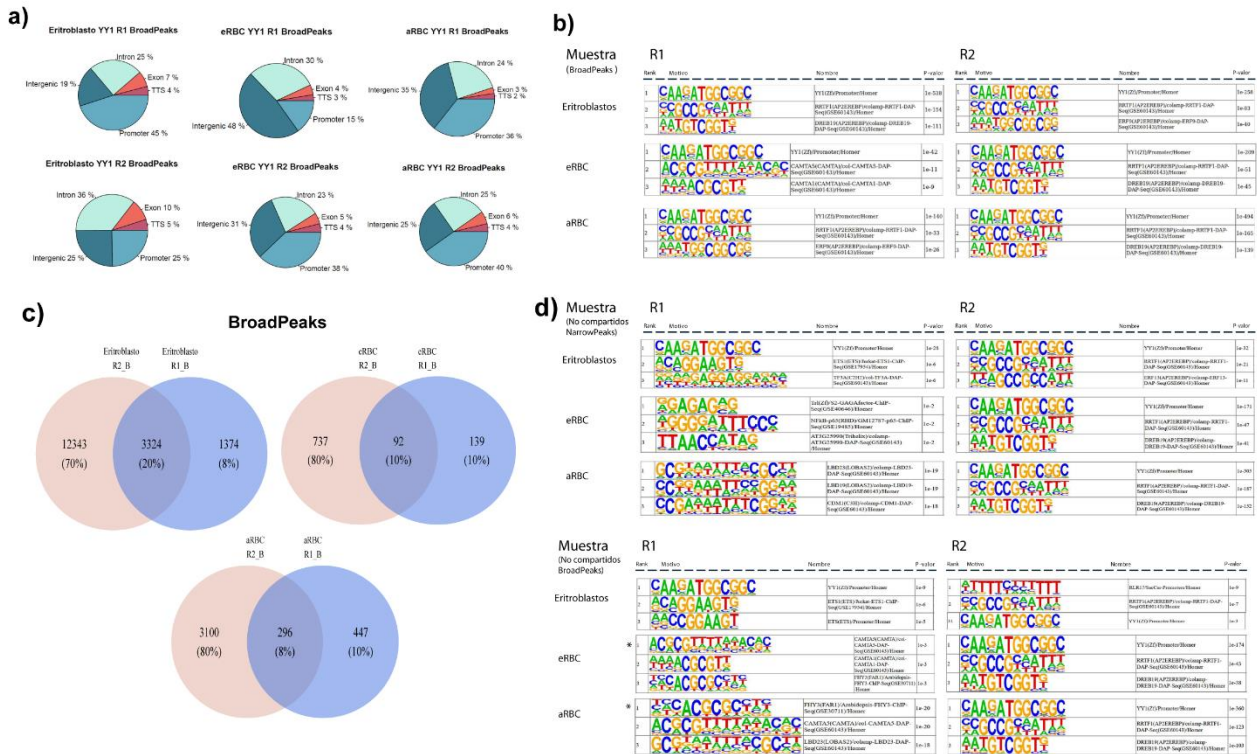


Figura Suplementaria 8. Análisis de picos de YY1. a) Anotado de picos de YY1 llamados mediante el método BroadPeaks. b) Resultados del análisis de búsqueda de motivos de los BroadPeaks. Tanto R1 como R2 muestran a YY1 como el motivo más enriquecido. c) Diagramas de Venn de picos compartidos entre réplicas. d) Análisis de búsqueda de motivos de los picos no compartidos entre réplicas experimentales. Se observa que R2 preserva a YY1 como motivo principal. Sin embargo, en R1 el motivo de YY1 se pierde en aRBC y eRBC (asteriscos).

11.8. YY1 se preserva dentro de TADs, pero no se enriquece en fronteras (picos compartidos entre réplicas)

Se muestran los resultados asociados a las secciones 8.4.3. y 8.4.4, pero con los picos compartidos entre las réplicas.

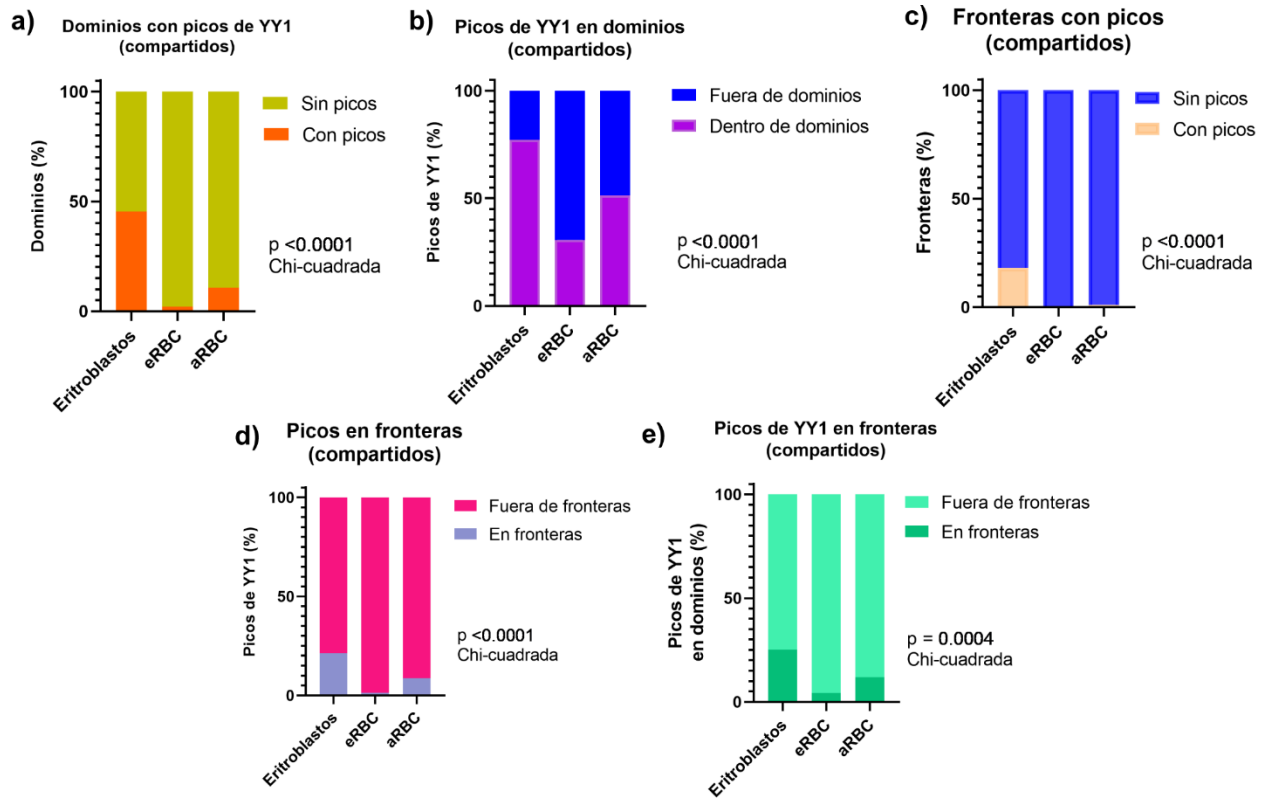


Figura Suplementaria 9. YY1 (compartidos) se preserva en minidominios, pero no está enriquecido en sus fronteras. **a)** Dominios que contienen picos de YY1. **b)** Picos de YY1 que se encuentran dentro de dominios. Los picos dentro de fronteras disminuyen en aRBC respecto a los eritroblastos. **c)** Porcentaje de las fronteras que contiene picos de YY1. **d)** Porcentaje de picos contenidos en las fronteras. **e)** Picos que están contenidos dentro de dominios y además están en sus fronteras. Las diferencias encontradas entre los estadios diferentes estadios de eritroblasto, eRBC y aRBC en los análisis a), b), c), d) ($p < 0.0001$, prueba de Chi-cuadrada), y e) ($p = 0.0004$, prueba de Chi-cuadrada), son estadísticamente significativos.

11.9. Los picos compartidos de YY1 están enriquecidos en el compartimento A

Resultados asociados a la sección 8.4.5. Se muestran los resultados de la intersección entre los picos de YY1 que se encuentran dentro y fuera de los dominios cromatínicos, y su distribución respecto a los compartimentos.

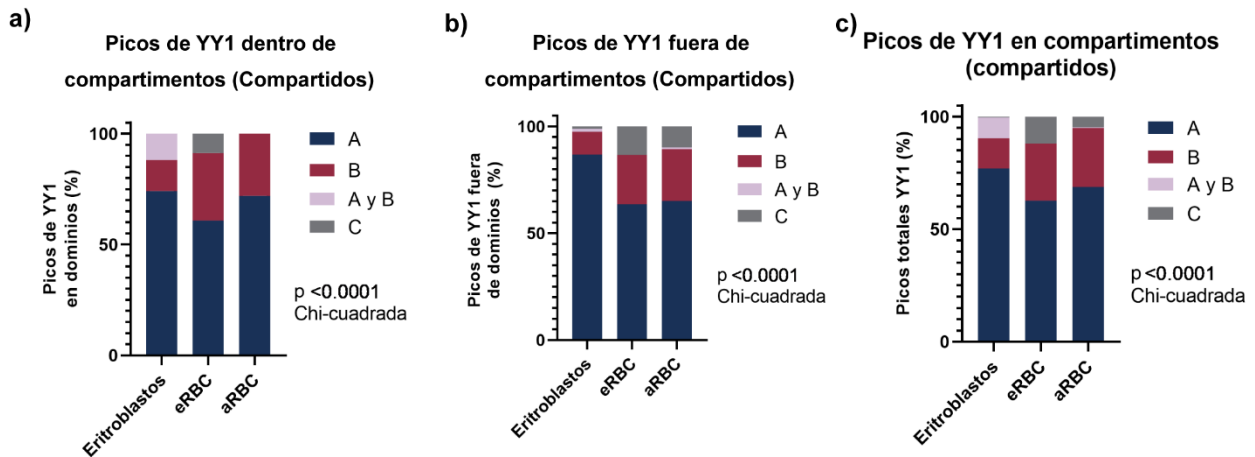


Figura Suplementaria 10. Análisis de picos compartidos de YY1 en los compartimentos. Los picos de YY1 compartidos entre réplicas dentro (a) y fuera (b) de dominios exhiben un posicionamiento preferencial en el compartimento A en los tres estadios de diferenciación. c) Picos compartidos totales. Las diferencias entre los estadios de eritroblasto, eRBC y aRBC en a), b) y c) son estadísticamente significativas ($p < 0.0001$, prueba de Chi-cuadrada).

11.10. Los picos compartidos de YY1 preservan accesibilidad y se enriquecen en promotores de genes no expresados en aRBC

Se muestran los resultados asociados a las secciones 8.4.6. y 8.4.7 con los análisis realizados en el conjunto de picos compartidos entre ambas réplicas.

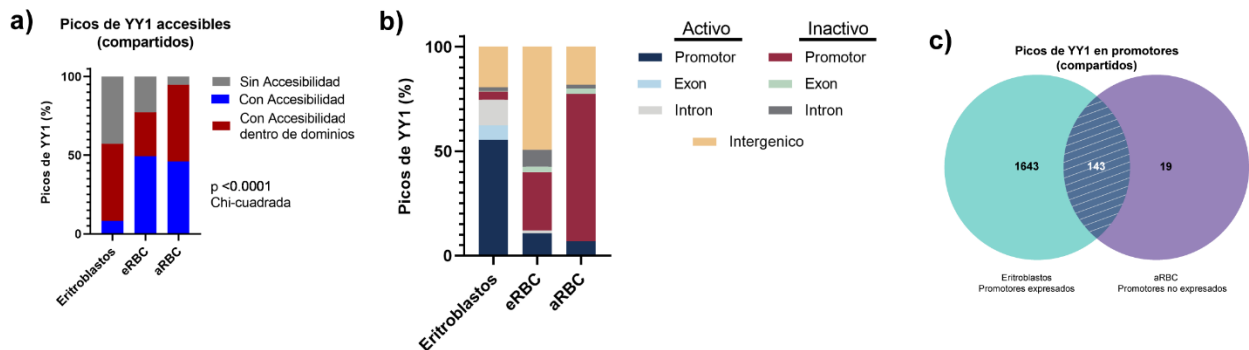


Figura Suplementaria 11. Los picos compartidos de YY1 preservan accesibilidad y se enriquecen en promotores de genes inactivos. a) Porcentaje de picos de YY1 compartidos entre réplicas que preservan accesibilidad. La mayoría de los picos en aRBC tienen accesibilidad ($p < 0.0001$, prueba de Chi-cuadrada). b) Anotado de picos YY1 ChIP-seq de promotores, exones, intrones y regiones intergénicas respecto a su estado transcripcional. YY1 se enriquece en promotores de genes inactivos en aRBC. c) Picos de YY1 en promotores que se pierden expresión en eritrocitos maduros.

11.11. Picos de YY1 posicionados en promotores de genes

Se despliega el número de genes con picos de YY1 en sus promotores y se subdividen respecto a su estado transcripcional y su posicionamiento respecto a los dominios topológicos. Para cada tipo celular se analizaron los picos correspondientes a la réplica 2 (R2) y los picos compartidos entre réplicas.

Tabla Suplementaria 4. Número de picos de YY1 en promotores de genes.

| | | Estado | Dentro dominios | Fuera de dominios | Total |
|------|-------------|---------------|-----------------|-------------------|-------|
| HD3 | R2 | Expresado | 2367 | 316 | 2683 |
| | | Sin expresión | 272 | 61 | 333 |
| | Compartidos | Expresado | 1416 | 190 | 1606 |
| | | Sin expresión | 89 | 20 | 109 |
| eRBC | R2 | Expresado | 18 | 19 | 37 |
| | | Sin expresión | 91 | 107 | 198 |
| | Compartidos | Expresado | 1 | 7 | 8 |
| | | Sin expresión | 9 | 12 | 21 |
| aRBC | R2 | Expresado | 94 | 56 | 150 |
| | | Sin expresión | 541 | 497 | 1038 |
| | Compartidos | Expresado | 11 | 5 | 16 |
| | | Sin expresión | 88 | 74 | 162 |

Subsecuentemente, relacionado con la sección 8.4.8., se realizó la GO sin considerar si los genes estaban dentro de dominios o no. De esta forma pudimos observar que las categorías de términos de los procesos biológicos enriquecidos eran altamente congruentes con aquellas realizadas únicamente con los genes estructurados. Esto hace sentido pues el subconjunto de picos en promotores de genes que caen dentro de dominios es un subconjunto del conjunto estudiado en esta sección (**Figura Suplementaria 12a-b**). Adicionalmente, para la selección de ejemplos en la integración visual de los datos de Hi-C, ATAC-seq, RNA-seq y ChIP-seq realizamos el mismo refinamiento relacionado con la Figura 14c (**Figura Suplementaria 12c**).

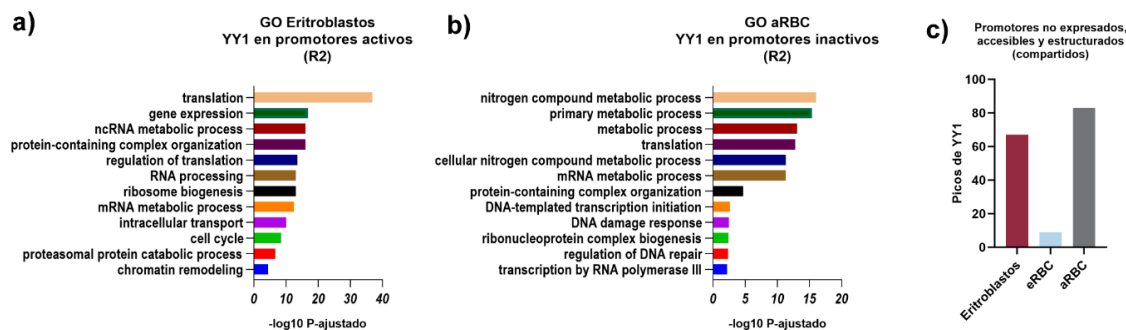


Figura Suplementaria 12. Ontologías de genes (GO) con YY1 en su promotor. a) GO en eritroblastos de genes activos (R2). Los procesos biológicos más enriquecidos están relacionados con la traducción, expresión génica y procesos metabólicos. b) GO en aRBC (R2) de genes inactivos. Los procesos biológicos más enriquecidos se relación con procesos metabólicos, modificación postraducciona de proteínas y expresión génica. c) Número de picos compartidos en promotores génicos no expresados, accesibles y estructurados.

12. Referencias bibliográficas

1. Misteli, T. The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* **183**, 28–45 (2020).
2. Jerkovic ´, I. & Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nat Rev Mol Cell Biol* **22**, 511–528 (2021).
3. Lanctôt, C., Cheutin, T., Cremer, M., Cavalli, G. & Cremer, T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**, 104–115 (2007).
4. Denker, A. & de Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes & development* **30**, 1357—1382 (2016).
5. Nakato, R. & Sakata, T. Methods for ChIP-seq analysis: A practical workflow and advanced applications. *Methods* **187**, 44–53 (2021).
6. Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930 (2019).
7. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, e21856 (2017).
8. Chen, T. & Dent, S. Y. R. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet* **15**, 93–106 (2014).
9. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21.29.1–21.29.9 (2015).
10. Withanage, M. H. H., Liang, H. & Zeng, E. RNA-Seq Experiment and Data Analysis Data analysis. in *Estrogen Receptors: Methods and Protocols* (ed. Eyster, K. M.) 405–424 (Springer US, New York, NY, 2022). doi:10.1007/978-1-0716-1920-9_22.
11. Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**, 292–301 (2001).
12. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat Rev Genet* **19**, 789–800 (2018).
13. Szalaj, P. & Plewczynski, D. Three-dimensional organization and dynamics of the genome. *Cell Biol Toxicol* **34**, 381–404 (2018).
14. Phillips, J. E. & Corces, V. G. CTCF: Master Weaver of the Genome. *Cell* **137**, 1194–1211 (2009).
15. Merckenschlager, M. & Nora, E. P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu Rev Genomics Hum Genet* **17**, 17–43 (2016).
16. Li, M. *et al.* Architectural proteins for the formation and maintenance of the 3D genome. *Sci China Life Sci* **63**, 795–810 (2020).
17. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930–944.e22 (2017).

18. Gonzalez-Sandoval, A. & Gasser, S. M. On TADs and LADs: Spatial Control Over Gene Expression. *Trends Genet* **32**, 485–495 (2016).
19. Peters, J.-M., Tedeschi, A. & Schmitz, J. The cohesin complex and its roles in chromosome biology. *Genes Dev* **22**, 3089–3114 (2008).
20. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305–320.e24 (2017).
21. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
22. Gassler, J. *et al.* A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J* **36**, 3600–3618 (2017).
23. Liu, T. *et al.* Matrin3 mediates differentiation through stabilizing chromatin accessibility and chromatin loop-domain interactions, and YY1 mediated enhancer-promoter interactions. *bioRxiv* 2023.03.30.534959 (2023) doi:10.1101/2023.03.30.534959.
24. Lomvardas, S. *et al.* Interchromosomal Interactions and Olfactory Receptor Choice. *Cell* **126**, 403–413 (2006).
25. Monahan, K., Horta, A. & Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **565**, 448–453 (2019).
26. Verheul, T. C. J., van Hijfte, L., Perenthaler, E. & Barakat, T. S. The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1. *Front Cell Dev Biol* **8**, (2020).
27. Liu, Z., Lee, D.-S., Liang, Y., Zheng, Y. & Dixon, J. R. Foxp3 orchestrates reorganization of chromatin architecture to establish regulatory T cell identity. *Nat Commun* **14**, 6943 (2023).
28. Khosraviani, N., Ostrowski, L. A. & Mekhail, K. Roles for non-coding RNAs in spatial genome organization. *Front Cell Dev Biol* **7**, 336 (2019).
29. Rippe, K. Liquid–Liquid Phase Separation in Chromatin. *Cold Spring Harb Perspect Biol* **14**, (2022).
30. Hirose, T., Ninomiya, K., Nakagawa, S. & Yamazaki, T. A guide to membraneless organelles and their various roles in gene regulation. *Nat Rev Mol Cell Biol* **24**, 288–304 (2023).
31. Ahn, J. H. *et al.* Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**, 591–595 (2021).
32. Ayala-Guerrero, L., Claudio-Galeana, S., Furlan-Magaril, M. & Castro-Obregón, S. Chromatin Structure from Development to Ageing. in *Biochemistry and Cell Biology of Ageing: Part III Biomedical Science* (eds. Harris, J. R. & Korolchuk, V. I.) 7–51 (Springer International Publishing, Cham, 2023). doi:10.1007/978-3-031-21410-3_2.
33. Cha, H. J. *et al.* Inner nuclear protein Matrin-3 coordinates cell differentiation by stabilizing chromatin architecture. *Nat Commun* **12**, 6241 (2021).
34. Zhao, X., Zhu, S., Peng, W. & Xue, H.-H. The Interplay of Transcription and Genome Topology Programs T Cell Development and Differentiation. *The Journal of Immunology* **209**, 2269–2278 (2022).

35. Ji, P. Chapter Four - New Insights into the Mechanisms of Mammalian Erythroid Chromatin Condensation and Enucleation. in *International Review of Cell and Molecular Biology* (ed. Jeon, K. W.) vol. 316 159–182 (Academic Press, 2015).
36. Zhao, B., Yang, J. & Ji, P. Chromatin condensation during terminal erythropoiesis. *Nucleus* **7**, 425–429 (2016).
37. Kowalski, A. & Palyga, J. Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non-histone protein MENT. *Chromosome Research* **19**, 579–590 (2011).
38. Geerts, W. J. C., Lamers, W. H. & Moorman, A. F. M. Differences in erythropoiesis in normal chicken and quail embryos. *Histochem J* **25**, 280–290 (1993).
39. Sheng, G. Primitive and definitive erythropoiesis in the yolk sac: a bird's eye view. *International Journal of Developmental Biology* **54**, 1033–1043 (2010).
40. Campbell, T. W. & Ellis, C. K. *Avian and Exotic Animal Hematology and Cytology*. (Wiley, 2013).
41. Penagos-Puig, A. *et al.* RNA polymerase II pausing regulates chromatin organization in erythrocytes. *Nat Struct Mol Biol* **30**, 1092–1104 (2023).
42. Rajgor, D. *et al.* Lamin A/C controls nuclear matrin-3 levels and localization, but not alternative splicing of cassette exons. *bioRxiv* 378240 (2018) doi:10.1101/378240.
43. CHEN, H. Y., SUN, J.-M., HENDZEL, M. J., RATTNER, J. B. & DAVIE, J. R. Changes in the nuclear matrix of chicken erythrocytes that accompany maturation. *Biochemical Journal* **320**, 257–265 (1996).
44. Fishman, V. *et al.* 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin. *Nucleic Acids Res* **47**, 648–665 (2019).
45. Claudio Galeana, G. S. & Furlan Magaril, M. Estudio del transcriptoma y la abundancia de proteínas estructurales durante la diferenciación y maduración eritroides en *Gallus gallus*. (Universidad Nacional Autónoma de México, Ciudad de México, 2020).
46. Wang, W. *et al.* A histidine cluster determines YY1-compartmentalized coactivators and chromatin elements in phase-separated enhancer clusters. *Nucleic Acids Res* **50**, 4917–4937 (2022).
47. Consortium, T. U. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523–D531 (2023).
48. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
49. Weintraub, A. S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573–1588.e28 (2017).
50. Li, L. *et al.* YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat Chem Biol* **17**, 161–168 (2021).

51. Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).
52. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444 (2022).
53. Rincón-Arano, H., Valadez-Graham, V., Guerrero, G., Escamilla-Del-Arenal, M. & Recillas-Targa, F. YY1 and GATA-1 Interaction Modulate the Chicken 3'-Side α -Globin Enhancer Activity. *J Mol Biol* **349**, 961–975 (2005).
54. Egan, B. *et al.* An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition. *PLoS One* **11**, e0166438- (2016).
55. Lopez-Delisle, L. *et al.* pyGenomeTracks: reproducible plots for multivariate genomic datasets . *Bioinformatics* **37**, 422–423 (2021).
56. BOLTE, S. & CORDELIÈRES, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* **224**, 213–232 (2006).
57. Kolberg, L. *et al.* g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res* **51**, W207–W212 (2023).
58. Gavrilov, A. A. & Razin, S. V. Spatial configuration of the chicken α -globin gene domain: immature and active chromatin hubs. *Nucleic Acids Res* **36**, 4629–4640 (2008).
59. Singh, S., Vanden Broeck, A., Miller, L., Chaker-Margot, M. & Klinge, S. Nucleolar maturation of the human small subunit processome. *Science (1979)* **373**, eabj5338 (2024).
60. Yano, M., Hoogenraad, N., Terada, K. & Mori, M. Identification and Functional Analysis of Human Tom22 for Protein Import into Mitochondria. *Mol Cell Biol* **20**, 7205–7213 (2000).
61. Pathak, S. *et al.* O⁶-Methylguanine Cytosine Acylation of TAB1 modulates TAK1-mediated cytokine release. *EMBO J* **31**, 1394-1404–1404 (2012).
62. Rudloff, U., Eberhard, D., Tora, L., Stunnenberg, H. & Grummt, I. TBP-associated factors interact with DNA and govern species specificity of RNA polymerase I transcription. *EMBO J* **13**, 2611-2616–2616 (1994).
63. Han, Y., Yan, C., Fishbain, S., Ivanov, I. & He, Y. Structural visualization of RNA polymerase III transcription machineries. *Cell Discov* **4**, 40 (2018).
64. Mayer, B. *et al.* Vasopressin stimulates the proliferation and differentiation of red blood cell precursors and improves recovery from anemia. *Sci Transl Med* **9**, eaao1632 (2017).
65. Lee, J.-S. *et al.* Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev* **9**, 1188–1198 (1995).
66. Austen, M., Lüscher, B. & Lüscher-Firzlaff, J. M. Characterization of the transcriptional regulator YY1: The bipartite transactivation domain is independent of interaction with the TATA box-binding protein, transcription factor IIB, TAFII55, or cAMP-responsive element-

- binding protein (CBP)-binding protein. *Journal of Biological Chemistry* **272**, 1709–1717 (1997).
67. Yant, S. R. *et al.* High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human β globin cluster. *Nucleic Acids Res* **23**, 4353–4362 (1995).
 68. Shestakova, E. A., Mansuroglu, Z., Mokrani, H., Ghinea, N. & Bonnefoy, E. Transcription factor YY1 associates with pericentromeric γ -satellite DNA in cycling but not in quiescent (G0) cells. *Nucleic Acids Res* **32**, 4390–4399 (2004).
 69. Trojer, P. & Reinberg, D. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell* **28**, 1–13 (2007).
 70. Laugesen, A., Højfeldt, J. W. & Helin, K. Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Mol Cell* **74**, 8–18 (2019).
 71. Thorvaldsen, J. L., Weaver, J. R. & Bartolomei, M. S. A YY1 Bridge for X Inactivation. *Cell* **146**, 11–13 (2011).
 72. Basu, A., Wilkinson, F. H., Colavita, K., Fennelly, C. & Atchison, M. L. YY1 DNA binding and interaction with YAF2 is essential for Polycomb recruitment. *Nucleic Acids Res* **42**, 2208–2223 (2014).
 73. Hoxha, S. *et al.* YAP-Mediated Recruitment of YY1 and EZH2 Represses Transcription of Key Cell-Cycle Regulators. *Cancer Res* **80**, 2512–2522 (2020).
 74. Hao, W. *et al.* PGC7 regulates H3K27me3 modification by inhibiting the interaction of YY1 with PRC2. *American Journal of Physiology-Cell Physiology* **325**, C286–C299 (2023).
 75. Dong, X. *et al.* YY1 safeguard multidimensional epigenetic landscape associated with extended pluripotency. *Nucleic Acids Res* **50**, 12019–12038 (2022).
 76. Sakamoto, M. *et al.* Dynamic nucleosome remodeling mediated by YY1 underlies early mouse development. *Genes Dev* (2023) doi:10.1101/gad.350376.122.
 77. Dasgupta, P., Prasad, P., Bag, S. K. & Chaudhuri, S. Dynamicity of histone H3K27ac and H3K27me3 modifications regulate the cold-responsive gene expression in *Oryza sativa* L. ssp. indica. *Genomics* **114**, 110433 (2022).
 78. Penagos-Puig, A. & Furlan-Magaril, M. Heterochromatin as an Important Driver of Genome Organization. *Front Cell Dev Biol* **8**, (2020).
 79. Dong, S. *et al.* YY1 promotes HDAC1 expression and decreases sensitivity of hepatocellular carcinoma cells to HDAC inhibitor. *Oncotarget* **8**, 40583 (2017).
 80. Wang, Z.-T. *et al.* Histone deacetylase inhibitors suppress mutant p53 transcription via HDAC8/YY1 signals in triple negative breast cancer cells. *Cell Signal* **28**, 506–515 (2016).
 81. Lieberthal, J. G., Kaminsky, M., Parkhurst, C. N. & Tanese, N. The role of YY1 in reduced HP1 α gene expression in invasive human breast cancer cells. *Breast Cancer Research* **11**, R42 (2009).

82. Charles, R. *et al.* Introduction of the human AVPR1A gene substantially alters brain receptor expression patterns and enhances aspects of social behavior in transgenic mice. *Dis Model Mech* **7**, 1013–1022 (2014).
83. Galitsyna, A. *et al.* Extrusion fountains are hallmarks of chromosome organization emerging upon zygotic genome activation. *bioRxiv* 2023.07.15.549120 (2023)
doi:10.1101/2023.07.15.549120.
84. Zhang, S. *et al.* RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis. *Sci Adv* **7**, eabg8205 (2024).
85. Tsompana, M. & Buck, M. J. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* **7**, 33 (2014).
86. Chereji, R. V, Eriksson, P. R., Ocampo, J., Prajapati, H. K. & Clark, D. J. Accessibility of promoter DNA is not the primary determinant of chromatin-mediated gene regulation. *Genome Res* **29**, 1985–1995 (2019).
87. Kiani, K., Sanford, E. M., Goyal, Y. & Raj, A. Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations. *Mol Syst Biol* **18**, e10979 (2022).
88. Coux, R.-X., Owens, N. D. L. & Navarro, P. Chromatin accessibility and transcription factor binding through the perspective of mitosis. *Transcription* **11**, 236–240 (2020).
89. Oei, S. L. *et al.* Interaction of the Transcription Factor YY1 with Human Poly(ADP-Ribosyl) Transferase. *Biochem Biophys Res Commun* **240**, 108–111 (1997).
90. Perreault, A. A., Brown, J. D. & Venters, B. J. Erythropoietin Regulates Transcription and YY1 Dynamics in a Pre-established Chromatin Architecture. *iScience* **23**, 101583 (2020).
91. Jahan, S., Beacon, T. H., Xu, W. & Davie, J. R. Atypical chromatin structure of immune-related genes expressed in chicken erythrocytes. *Biochemistry and Cell Biology* **98**, 171–177 (2019).
92. Morera, D. *et al.* RNA-Seq reveals an integrated immune response in nucleated erythrocytes. *PLoS One* **6**, e26998 (2011).
93. St Paul, M., Paolucci, S., Barjesteh, N., Wood, R. D. & Sharif, S. Chicken erythrocytes respond to Toll-like receptor ligands by up-regulating cytokine transcripts. *Res Vet Sci* **95**, 87–91 (2013).
94. Egloff, S., Studniarek, C. & Kiss, T. 7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features. *Transcription* **9**, 95–101 (2018).