



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE FISIOLÓGÍA CELULAR

Estudio de los mecanismos de regulación epigenética mediados por la
proteína ATRX

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
LIC. EN CIENCIAS GENÓMICAS
DAVID VALLE GARCÍA

TUTOR PRINCIPAL: DR. FÉLIX RECILLAS TARGA
INSTITUTO DE FISIOLÓGÍA CELULAR

TUTORA ADJUNTA: DRA. EMILY BERNSTEIN
*ICAHN SCHOOL OF MEDICINE AT
MOUNT SINAI*

COMITÉ TUTORAL: DR. MARIO ENRIQUE ZURITA ORTEGA
INSTITUTO DE BIOTECNOLOGÍA
DRA. LEONOR PÉREZ MARTÍNEZ
INSTITUTO DE BIOTECNOLOGÍA



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Este trabajo se realizó para cumplir el compromiso adquirido con el Programa de Doctorado en Ciencias Biomédicas en el Laboratorio de Epigenética y Biología de la Cromatina del Departamento de Genética Molecular del Instituto de Fisiología Celular de la Universidad Nacional Autónoma de México (UNAM) y en el Laboratorio de Epigenética del Cáncer del Departamento de Ciencias Oncológicas de la Icahn School of Medicine at Mount Sinai bajo la tutoría del Dr. Félix Recillas Targa y la Dra. Emily Bernstein. Se recibió financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACyT) (donativos 42653-Q, 128464 y 220503, beca doctoral 239663), la Dirección General de Asuntos del Personal Académico (DGAPA)-PAPIIT de la UNAM (donativos IN209403, IN203811 e IN201114) y el National Cancer Institute of the National Institutes of Health (EUA, donativo R01CA154683).

“Books permit us to voyage through time, to tap the wisdom of our ancestors [...] I think the health of our civilization, the depth of our awareness about the underpinnings of our culture and our concern for the future can all be tested by how well we support our libraries.”

- Carl Sagan, *Cosmos*

“The best life is the one in which the creative impulses play the largest part and the possessive impulses the smallest.”

- Bertrand Russell, *Political Ideals*

“Hay hombres que luchan un día y son buenos. Hay otros que luchan un año y son mejores. Hay quienes luchan muchos años y son muy buenos. Pero hay los que luchan toda la vida, esos son imprescindibles.”

- Bertolt Brecht

A mi mamá, mi papá y mi hermana.
A mi familia, mis amigos y mis compañeros.
A mis maestros.
A la Universidad.
A los imprescindibles.

Agradecimientos

A la UNAM por darme una educación de excelencia de forma prácticamente gratuita, por haberme formado académica y personalmente, y por ser el motor académico y científico del país.

Al Doctorado en Ciencias Biomédicas, por haberme formado y darme todas las facilidades para hacer un doctorado a mi medida, cuidando siempre la excelencia y velando por el bienestar de sus alumnos.

A mi tutor, el Dr. Félix Recillas Targa, por enseñarme todo lo que sé de epigenética, por enseñarme a buscar la excelencia ante todo y por siempre estar ahí para guiarme a pesar de los momentos difíciles.

To my other tutor, Emily Bernstein, for being the best PI one could ask for, for always believing in me, for giving me the chance of being part of her extraordinary lab and for guiding me and sincerely look for my personal and scientific growth.

A los miembros de mi comité tutorial: la Dra. Leonor Pérez Martínez y el Dr. Mario Zurita Ortega. Por su tiempo, guía, dedicación y pasión por mi proyecto. Por siempre apoyarme aún en los peores momentos.

A los miembros de mi jurado de examen: el Dr. David Romero Camarena, la Dra. Viviana Valadez Graham, el Dr. José Luis Reyes Taboada y el Dr. Enrique Merino Pérez. Por haber sido mis maestros ahora y desde siempre.

Al Dr. Federico Sánchez Rodríguez. Por haber sido siempre uno de los imprescindibles. Por haber sido mi Maestro, con M mayúscula. Por haberme mostrado que ser buena persona es una parte esencial de ser buen científico. Porque aunque nos haces falta, tu ejemplo siempre nos ilumina.

A todos mis compañeros del lab Recillas, porque de todos ustedes aprendí muchísimo y disfruté de su compañía en cada momento. Porque hicieron del laboratorio una experiencia inolvidable y por todos los momentos en los que gritamos que no pasa nada. Porque el doctorado no hubiera sido lo mismo sin ustedes: Rodrigo, Estela, Geito, Giane, Charis, Erandi, Andrea, Yael, Mariana, Carla, Paco, Edgar, Christian V., Cristian A.,

Ricardo y todos los alumnos nuevos con los que no tuve la oportunidad de convivir pero estoy seguro que mantienen vivo el espíritu del lab Recillas.

To all my wonderful labbies from the Bernstein lab, past and present. For showing me what passion for science means. For always being the best lab one can ask for. For making me feel like home. Thank you for all the teachings, the laughs and the night whiskeys: Flávia, Zu, Chiara, Maddy, Taniya, Asif, Dan H, Jay, Alex, Kajan, Dom, Mooki, Zhen, Thomas, Luis, Saúl, Nick and Dan F.

To the Icahn School of Medicine, for being my second home and for always supporting me in every way, both personal and academic.

A todos mis maestros, presentes y pasados. Sobre todo a los que se tomaron un tiempo para enseñarme y guiarme durante mis primeros meses en el laboratorio: Christian Valdes, Kajan Ratnakumar y Dan Hasson.

A todas las secretarías que me han ayudado con mis trámites y me han hecho la vida burocrática mucho más sencilla: Sara Méndez, Zenaida Martínez, Genevieve Joseph, Donna Ann Chiodi, Kazuko Robinson y todas las que me faltan por culpa de mi mala memoria.

A los amigos que hice durante el doctorado o los que ya tenía pero han seguido conmigo a pesar del tiempo y la distancia: Rodrigo, Estela, Yael, David Plata, Paco Jiménez.

A Karla Meza Sosa, por ser mi compañera de viaje, por estar conmigo en todos los momentos, por guiarme siempre y por seguir acompañándome a través de los años, las ciudades, las distancias, los enojos, las risas, los desencantos y los triunfos. Por siempre apoyarme y hacerme sacar lo mejor de mí en todos los aspectos.

A mis papás, Fermín Valle García y Elsa Rosaura García Salgado, porque es gracias a ustedes que ahora soy lo que soy. Por enseñarme que las calificaciones no son importantes, que lo importante es aprender. Por estar ahí y comprenderme y apoyarme siempre en todo, aunque no estuvieran siempre de acuerdo con el camino que tomé. Por darme alas, y luego dejarme volar.

A mi hermana, Claudia Valle García, por ser siempre mi mejor amiga y confidente. Por nuestras eternas discusiones, por siempre poner en perspectiva mi visión de la ciencia y la vida y por estar ahí siempre, a pesar de todo.

Índice

Resumen	7
Abstract.....	9
Lista de abreviaturas	11
Introducción.....	13
Mecanismos de regulación epigenética	14
ATRX: Un remodelador peculiar de la cromatina.....	19
Parte I: ATRX y la regulación de los genes ZNF	25
I.I - Introducción.....	26
I.I.I - Entendiendo la función de ATRX: hipótesis e incógnitas	26
II.I.II - K562 como un modelo para entender el papel de ATRX en la regulación epigenómica	28
I.II - Justificación y planteamiento del problema	29
I.III - Hipótesis	29
I. IV - Objetivos.....	29
I.V - Resultados.....	30
I.V.I - ATRX se une al exón 3' de los genes de tipo <i>zinc finger</i>	31
I.V.II - Los genes ZNF a los que ATRX se une poseen una cromatina atípica y características genéticas particulares	32
I.V.III ATRX forma un complejo con SETDB1, TRIM28 y ZNF274 que se une a los exones 3' de los genes ZNF	33
I.V.IV - La unión de ATRX a los genes ZNF está conservada en distintos linajes celulares humanos	34
I.V.V - La deficiencia de ATRX causa una pérdida de H3K9me3 en los exones 3' de los ZNFs	35
I.V.VI - Los niveles de H3K9me3 en dominios de cromatina atípica son sensibles a la pérdida de ATRX.....	36
I.V.VII - ZNF274 recluta a ATRX a un subgrupo de genes ZNF para regular sus niveles de H3K9me3 y la estabilidad genómica.....	37
I.VI - Discusión	38
I.VII - Perspectivas.....	41
I.VIII - Materiales y métodos	42
Parte II: ATRX y la regulación del dominio α-globina.....	44
II.I - Introducción.....	45
II.I.I - El dominio α -globina como un caso de estudio de la regulación epigenética.....	45

II.I.II - ATRX como regulador del dominio α -globina: mecanismos y suposiciones	46
II.II - Justificación y planteamiento del problema	48
II.III - Hipótesis	48
II.IV - Objetivos.....	48
II.V - Resultados.....	49
II.V.I - Caracterizar la Expresión de ATRX y las α -globinas a lo largo del desarrollo del embrión de pollo.....	49
II.V.II - Análisis de la unión global de ATRX en el genoma de pollo y su relación con el locus α -globina	50
II.V.III - La sobre-expresión de ATRX aumenta la expresión de los genes α -globina adultos en pollo.....	52
II.V.IV - ATRX no se une al locus α -globina en células humanas.....	54
II.IV.V La pérdida de ATRX afecta la diferenciación de células eritroides humanas a través de un mecanismo de regulación indirecta.....	55
II.VI – Discusión y perspectivas	57
II.VII - Materiales y métodos	59
Parte III: Dificultades técnicas al trabajar con ATRX.....	62
III.I – Dificultades en el uso de plásmidos de sobre-expresión que contienen el cDNA de ATRX.....	63
III.I.I - Los vectores que contienen el cDNA de ATRX tienen inserciones frecuentes cuando se expanden en bacterias	63
III.I.II - Problemas para transfectar el cDNA de ATRX en la línea K562	65
III.II - ChIP-seq de ATRX: problemas de los protocolos comunes y optimización de un ChIP-seq eficiente para ATRX.....	67
III.II.I - La sonicación degrada a ATRX	68
III.II.II - Estandarización de un protocolo de XL-MNasa para realizar experimentos de ChIP-seq de ATRX en humano.....	69
Conclusión.....	77
Bibliografía.....	80
Apéndice I - Publicaciones.....	87

Resumen

La regulación génica en las células eucariontes requiere de un complejo conjunto de actores epigenéticos que permiten el establecimiento y el mantenimiento de patrones espacio-temporales exquisitamente refinados. Cuando los factores epigenéticos son alterados, los patrones de expresión pueden ser alterados de forma tal que pueden causar el desarrollo de enfermedades complejas. Un ejemplo notable de esos factores epigenéticos es ATRX, un remodelador de la cromatina de tipo SWI/SNF. Las mutaciones del gen *ATRX* en la línea germinal son la causa de una compleja enfermedad humana llamada *Alpha thalassemia X-linked intellectual disability syndrome* (síndrome de discapacidad intelectual y α -talasemia ligada al cromosoma X). Por otro lado, las mutaciones somáticas en *ATRX* están frecuentemente ligadas a distintos tipos de cáncer.

Debido a su relación con el desarrollo de α -talasemia, se ha sugerido que ATRX es un regulador importante del locus α -globina. Adicionalmente, estudios recientes han sugerido que ATRX promueve la estabilidad genómica a través de la regulación de secuencias repetitivas tales como retrotransposones y repetidas teloméricas y pericentroméricas. Sin embargo, a pesar de su importante rol en varias enfermedades humanas, se han descrito pocos genes blanco directos de ATRX y se entiende muy poco acerca de su impacto global en la regulación del epigenoma.

En este trabajo, utilizamos dos aproximaciones complementarias para entender la regulación mediada por ATRX. Por un lado, realizamos un análisis epigenómico imparcial para estudiar la relevancia de ATRX a nivel de todo el genoma. Por el otro, usamos al locus α -globina como un modelo paradigmático para analizar la función de ATRX en los contextos del desarrollo y la diferenciación celular.

Nuestro análisis global identificó una nueva clase de genes blancos de ATRX: la familia de factores de transcripción de tipo *zinc finger* (ZNFs). De forma inesperada, observamos que ATRX se une a los exones 3' de los genes ZNF que poseen una estructura de la cromatina atípica: son genes transcripcionalmente activos, contienen altos niveles de H3K36me₃, una marca de cromatina asociada con la elongación de la polimerasa, y están paradójicamente enriquecidos en H3K9me₃, una marca típica de heterocromatina y regiones transcripcionalmente inactivas. Mostramos, además, que

ATRX se une a estas regiones de cromatina atípica en un complejo conformado también por las proteínas SETDB1, TRIM28 y ZNF274.

La eliminación genética de ATRX o ZNF274 causa la reducción de H3K9me3 en los exones 3' de los genes ZNF que tienen una marca de cromatina atípica. Más aún, la inactivación de ATRX genera una pérdida a nivel genómico de H3K9me3 específicamente en las regiones de cromatina atípica. Finalmente, la disminución de H3K9me3 en los genes ZNF se correlaciona con un aumento al daño al DNA y alteraciones en el ciclo celular, características de un incremento en la inestabilidad genómica.

En conjunto, nuestros análisis sugieren que ATRX regula el enriquecimiento de H3K9me3 en los genes ZNF para establecer un dominio de cromatina atípica que favorece la estabilidad genómica.

Adicionalmente, nuestros estudios en el locus α -globina han demostrado que ATRX es un regulador importante de la expresión de los genes α -globina. Usando al pollo como organismo modelo, probamos que durante el desarrollo eritroide ATRX se une directamente a los genes α -globina y forma un gran dominio que abarca todo el locus. Usando modelos de pollo y humanos, mostramos que ATRX promueve la correcta sobre-expresión de los genes α -globina durante la diferenciación eritroide. Se necesitan de más análisis para determinar los mecanismos moleculares precisos a través de los cuales ATRX regula al locus.

Durante el curso de esta investigación, nos enfrentamos a diversos retos técnicos y desarrollamos nuevas herramientas para superarlos, que son discutidas en este escrito. De forma importante, logramos establecer un protocolo de ChIP-seq para ATRX en humano, eficiente y reproducible, que puede ser útil a la comunidad.

En resumen, descubrimos un rol inesperado de ATRX como regulador de los genes ZNF y las regiones de cromatina atípica. También, proveemos evidencia de su importancia en la regulación del locus α -globina. Esperamos que estos conocimientos puedan ser importantes para el entendimiento a futuro de las enfermedades en las que ATRX está alterado.

Abstract

The regulation of gene expression in eukaryote cells requires a complex gathering of epigenetic actors that allow the establishment and maintenance of exquisitely refined spatiotemporal patterns. When epigenetic factors are altered, the expression patterns can be perturbed in such a way that they may cause the development of complex disease. One notable example of such epigenetic factors is the SWI/SNF chromatin remodeler ATRX. Germ-line mutations of the *ATRX* gene cause a complex human disease called Alpha thalassemia X-linked intellectual disability syndrome. In contrast, somatic mutations of *ATRX* are frequently linked to several types of cancers.

Because of its involvement with alpha thalassemia, it has been suggested that ATRX is an important regulator of the α -globin locus. Furthermore, recent studies have suggested that ATRX promotes genome stability through the regulation of repetitive sequences such as retrotransposons, and telomeric and pericentromeric repeats. However, despite its important role in human disease, very few direct ATRX gene targets have been identified so far, and its genome-wide impact in the regulation of the epigenome is poorly understood.

Here, we took two complementary approaches to understand ATRX-mediated regulation. On the one hand, we performed an unbiased epigenomic analysis to understand the genome-wide relevance of ATRX. On the other, we used the α -globin locus as a paradigmatic model to analyze ATRX function in developmental and cell-differentiation contexts.

Our genome-wide analysis found a new class of ATRX targets: the zinc finger gene family of transcription factors (ZNFs). Unexpectedly, ATRX binds to the 3' exons of ZNFs displaying an atypical chromatin signature: they are transcriptionally active, contain high levels of H3K36me₃, a chromatin mark associated with polymerase elongation, and are paradoxically enriched in H3K9me₃, a mark associated with gene silencing and heterochromatin formation. We further show that ATRX binds to these atypical chromatin regions in a complex formed with the proteins SETDB1, TRIM28 and ZNF274.

The genetic deletion of either ATRX or ZNF274 causes the decrease of H3K9me₃ at the 3' exons of ZNFs that have an atypical chromatin signature. Furthermore, we show that

ATRX inactivation causes a genome-wide decrease of H3K9me3 specifically at atypical chromatin regions. Finally, the decrease of H3K9me3 at ZNFs is correlated with higher levels of DNA damage and an altered cell cycle, which are characteristics of increased genome instability.

Taken together, our analyses suggest that ATRX regulates the enrichment of H3K9me3 at ZNFs to establish an atypical chromatin signature that favors genomic stability.

Moreover, our studies in the α -globin locus have shown that ATRX is an important regulator of α -globin expression. Using chicken as a model organism, we show that during erythroid development ATRX binds directly to the α -globin genes and forms a large domain spanning the entire locus. Using chicken and human models, we show that ATRX promotes the correct over-expression of the α -globin genes during erythroid differentiation. Further analyses are needed to determine the precise molecular mechanisms by which ATRX regulates the locus.

During the course of this investigation, we confronted several technical hurdles and developed new tools to overcome them, which are discussed here. Importantly, we were able to establish an efficient and reproducible protocol for ATRX ChIP-seq in human cells that may be useful for the community.

In summary, we discovered an unexpected role of ATRX as a regulator of ZNFs and atypical chromatin. Additionally, we provide evidence of its importance for the regulation of the α -globin locus. We envision that these findings may be important for the future understanding of diseases in which ATRX is altered.

Lista de abreviaturas

5mC – Metilación de la Citosina en el carbono 5.
ALT – Alternative Lengthening of Telomeres
ATRX – Alpha-Thalassemia, Mental Retardation X-linked protein
BrdU – 5-bromo-2-desoxiuridina
DNA – Ácido Desoxirribonucleico
ChIP-qPCR – Chromatin Immunoprecipitation followed by quantitative PCR
ChIP-seq – Chromatin Immunoprecipitation followed by massive parallel sequencing
CpG –di-nucleótidos C-G
CTCF – CCCTC-binding factor
H3K4me3 – Tri-metilación de la lisina 4 en la histona H3
H3K9me3 – Tri-metilación de la lisina 9 en la histona H3
H3K36me3 – Tri-metilación de la lisina 39 en la histona H3
H4K20me3 – Tri-metilación de la lisina 20 en la histona H4
hESC – human Embryonic Stem Cells
HR – Homologous Recombination
KD – Knock Down
KO – Knock Out
KRAB – Krüppel Associated Box, dominio represor típico de algunos ZNFs
lncRNA – long non-coding RNA
µg – microgramo
µl – microlitro
µM – micromolar
MEF – Mouse Embryonic Fibroblasts
mESC – Mouse Embryonic Stem Cells
mH2A – Variante de histona macroH2A
MNasa – Micrococcal Nuclease
PI – Propidium iodide
PTMs – Post-Transductional Modifications
RBC - Red Blood Cell (eritrocitos)
RNA – Ácido Ribonucleico
shRNAs – small hairpin RNAs
TF – Transcription Factor
WB – Western Blot
WT – Wild Type
ZNFs – Genes de la familia de factores de transcripción de tipo C2H2 Zinc Einger

Introducción

"[...] nos esse quasi nanos, gigantium humeris incidentes [...]"

- Atribuida a Bernardus Carnotensis en el *Metalogicon* de Juan de Salisbury

Mecanismos de regulación epigenética

Una de las grandes áreas de interés en la biología moderna consiste en entender cómo programas específicos de expresión génica son capaces de regular y llevar a cabo procesos complejos como el desarrollo de un organismo y la diferenciación celular. Los organismos pluricelulares, en particular los vertebrados, poseen una gran cantidad de células especializadas que cumplen funciones específicas y que poseen patrones de expresión únicos. Aunque el establecimiento de dichos patrones de expresión involucran mecanismos variados, investigaciones realizadas en los últimos años han dejado claro que la regulación epigenética es central para el correcto funcionamiento de los mismos (Goldberg et al. 2007).

A grandes rasgos, podemos definir a la regulación epigenética como el conjunto de mecanismos que regulan la formación, el acceso, la organización y la herencia de la cromatina (Goldberg et al. 2007).

La cromatina es el complejo nucleoproteico conformado principalmente por el DNA y las histonas que lo compactan y lo protegen. Las histonas son proteínas básicas, globulares, capaces de dimerizarse entre sí para conformar un nucleosoma, la unidad básica de la cromatina. Existen 4 tipos de histonas canónicas que forman parte del nucleosoma: H2A, H2B, H3 y H4. Los nucleosomas son octámeros conformados por 2 pares de cada una de esas histonas y tienen la capacidad de enrollar aproximadamente 146 nucleótidos de DNA a su alrededor. Este enrollamiento permite compactar espacialmente al DNA dentro del núcleo, pero también impide en buena medida el acceso de otras proteínas a la secuencia nucleotídica. Es por ello que la célula desarrolló un complejo sistema de mecanismos que regulan la formación, la compactación y el acceso a la cromatina (Jenuwein & Allis 2001; Felsenfeld & Groudine 2003).

Existen esencialmente dos tipos de cromatina: 1) la heterocromatina: una cromatina compacta, transcripcionalmente inactiva, e inaccesible para la mayoría de las proteínas que requieren de la secuencia del DNA como templado 2) la eucromatina: una cromatina poco compacta, transcripcionalmente activa y que permite el fácil acceso a la secuencia de DNA. Estos tipos de cromatina suelen formar dominios excluyentes entre sí, y su localización espacial a lo largo del genoma permite establecer los patrones de expresión que generan la identidad celular (Jenuwein & Allis 2001; Felsenfeld & Groudine 2003).

La distinción entre eucromatina y heterocromatina se establece esencialmente a base de marcas post-transduccionales (PTMs) específicas, primordialmente en la región N

terminal de las histonas, un extremo muy flexible que no forma parte de la estructura globular del nucleosoma y que posee varios residuos que pueden ser modificados post-transduccionalmente. A estas modificaciones se les conoce comúnmente como “marcas de la cromatina” o “marcas de histonas” y ambos términos se usarán de forma intercambiable a lo largo del texto. La lista de PTMs es larga y se está expandiendo continuamente conforme se van refinando los métodos bioquímicos para identificarlas. Sin embargo, las más estudiadas son la metilación (me), la acetilación (ac), la fosforilación (ph) y la ubiquitinación (ub) (Jenuwein & Allis 2001).

Lo que hace especial a las marcas de histonas es que existe todo un conjunto de enzimas que son capaces de generarlas (escritores o *writers*), de removerlas (borradores o *erasers*) y de unirse a ellas de forma específica (lectores o *readers*). Estas características llevaron a David Allis a proponer que las marcas de cromatina conforman un “código de histonas”, un lenguaje adicional al código genético (es decir, un lenguaje epigenético) que regula el estado transcripcional del genoma, entre otras funciones. Diversos estudios posteriores comprobarían la validez de dicho modelo (Jenuwein & Allis 2001).

A través de una extensa red de investigadores a lo largo del mundo y utilizando diversos organismos modelo, se fue descifrando poco a poco el código epigenético de las marcas de histonas, y se estableció de manera formal el campo de estudio que hoy conocemos como epigenética (Jenuwein & Allis 2001).

Las marcas de histonas tienen una notación específica. Usualmente se abrevian mencionando al inicio la histona modificada, luego el tipo y número de residuo (por convención los residuos se numeran a partir del primer aminoácido del extremo amino terminal), y finalmente el tipo de modificación. Algunas modificaciones como la metilación pueden darse en forma de monómeros, dímeros o trímeros. En ese caso, se denota el tipo de modificación con un número. De esta forma la modificación “H3K4me3” corresponde a la tri-metilación de la histona H3 en la cuarta lisina a partir del extremo N terminal.

A continuación se presenta una muy breve lista de las modificaciones más comunes y su relación con diferentes elementos reguladores y genes:

<i>Modificación</i>	<i>Tipo de cromatina</i>	<i>Región asociada</i>
H3K4me1	Eucromatina	Enhancers
H3K4me3	Eucromatina	Promotores
H3K9ac	Eucromatina	Enhancers/promotores

H3K27ac	Eucromatina	Enhancers
H3K36me3	Eucromatina	Genes transcripcionalmente activos
H3K9me3	Heterocromatina	Dominios de heterocromatina constitutiva
H3K27me3	Heterocromatina	Dominios de heterocromatina facultativa

Por regla general, las marcas de eucromatina son mutuamente excluyentes con las marcas de heterocromatina, aunque existen algunas excepciones reportadas siendo el ejemplo más estudiado las marcas bivalentes presentes en las células madre embrionarias (Voigt et al. 2012; Bernstein et al. 2006; Shema et al. 2016). En algunos casos ésta es una exclusión estérica; por ejemplo las lisinas 9 y 27 de la histona H3 pueden ser metiladas o acetiladas, pero usualmente no pueden presentar ambas modificaciones. En otros casos se debe a una exclusión funcional llevada a cabo por los complejos de lectura, escritura y borrado de las marcas; por ejemplo, las regiones enriquecidas en H3K9me3 no se sobreponen a las regiones de H3K27ac, debido a que los complejos que “escriben” la H3K9me3 suelen incluir proteínas que eliminan las marcas de acetilación (Beisel & Paro 2011).

Es importante mencionar que la cromatina tiene un rol regulatorio no sólo en la transcripción, sino en prácticamente todos los procesos que usan al DNA como templado o requieren del acceso a la secuencia nucleotídica. A su vez, el código epigenético no es solamente un código transcripcional: existen marcas de la cromatina específicas para los procesos de replicación y reparación del DNA, por poner un par de ejemplos (Ransom et al. 2010).

Además de las modificaciones post-transduccionales de las histonas, se han descubierto otros mecanismos relevantes para la regulación epigenética. Uno de ellos es la metilación del DNA. La metilación del DNA es un fenómeno conservado tanto en eucariontes como en procariontes y en éstos últimos conforma una forma de lenguaje epigenético que se ha desarrollado de forma independiente del lenguaje de los eucariontes (Casadesús & Low 2006). En los eucariontes vertebrados la metilación de la Citosina (5mC) en el DNA es la modificación al DNA más abundante y mejor estudiada. La 5mC ocurre casi exclusivamente en Citosinas seguidas de una Guanina, lo que se conoce como di nucleótidos CpG. Los di nucleótidos CpG no se distribuyen de forma aleatoria en el genoma, sino que tienden a agruparse en regiones conocidas como islas CpG. Los primeros análisis del genoma de vertebrados encontraron que las islas CpG se encuentran en una frecuencia desproporcionadamente alta en promotores de genes. Estudios posteriores encontraron que los genes que contienen islas CpGs en sus promotores pueden ser silenciados transcripcionalmente si dichas islas son metiladas.

Por el contrario, los genes que se expresan activamente tienden a mantener sus islas CpGs desmetiladas. Esta correlación no es exclusiva de los genes. Diversas regiones transcripcionalmente inactivas, como los transposones y otros tipos de regiones con secuencias repetitivas, suelen tener niveles de metilación del DNA altos (Smith & Meissner 2013).

Además de las modificaciones a las histonas y la metilación del DNA, existe otra capa de regulación epigenética importante: las variantes de histonas. Las variantes de histonas son histonas alternativas a las histonas canónicas que poseen diversos cambios en su secuencia peptídica. A diferencia de las histonas canónicas, que son expresadas exclusivamente durante la replicación del DNA, las variantes de histonas suelen expresarse a lo largo del ciclo celular. Algunas variantes de histonas han sido asociadas con funciones específicas y a diferencia de las histonas canónicas, las variantes suelen depositarse en regiones particulares del genoma (Maze et al. 2014).

La variante de histona más estudiada es H3.3, una variante de la histona H3 que difiere de su contraparte canónica H3.1 en sólo 5 residuos pero que posee funciones particulares. La variante H3.3 suele encontrarse enriquecida en regiones de eucromatina, específicamente en los promotores de genes activos y otras regiones regulatorias como *enhancers* e *insulators*. Diversos análisis genómicos han encontrado que H3.3 también es depositada en algunas regiones de heterocromatina, como los telómeros y las regiones repetitivas pericentroméricas. Aunque reportes recientes han demostrado que esta variante es la histona que sufre la modificación H3K9me3 necesaria para el silenciamiento de dichas regiones (Udugama et al. 2015), su función en dichos contextos no es bien entendida todavía. La importancia de las variantes de histonas ha sido puesta en relieve en los últimos años pues diversos estudios han encontrado que la desregulación en su función o mutaciones en su secuencia suelen asociarse con el desarrollo de cáncer y otras patologías (Maze et al. 2014; Vardabasso et al. 2014).

Algunas PTMs de las histonas tienen una consecuencia fisico-química en la forma en la que el nucleosoma interactúa con el DNA. La acetilación, por ejemplo, reduce la carga total positiva de las histonas lo cual reduce a su vez su capacidad de interactuar con las cargas negativas de la cadena de fosfatos del DNA. Otras modificaciones como la metilación, por ejemplo, no tienen un efecto físico. Las metilaciones en sí mismas son inertes y deben ser leídas por una proteína lectora para tener un efecto funcional en la conformación de la cromatina. Muchas veces, sin embargo, no son las proteínas lectoras las encargadas de realizar dicho efecto. Existe un cuarto tipo de proteína involucrado en

la interpretación del código epigenético: las proteínas efectoras (*effectors*) (Jenuwein & Allis 2001; Goldberg et al. 2007). Al grupo de efectores que son capaces de modificar físicamente las interacciones DNA-histona se les conoce como remodeladores de la cromatina ATP-dependientes. Como su nombre lo indica, sus funciones dependen de un gasto energético a partir de la hidrólisis del ATP. Los remodeladores de la cromatina son entonces los traductores últimos del código epigenético. Además de su función como efectores, los remodeladores de la cromatina también están involucrados en la función de las variantes de histonas, pues son ellos los encargados de regular la deposición y el recambio de las histonas canónicas y sus variantes en la cromatina. Es por ello que no es de sorprender que las mutaciones en esta clase de proteínas sean particularmente deletéreas y están involucradas en el desarrollo de cáncer y otras enfermedades (Lusser & Kadonaga 2003; Clapier & Cairns 2009; Hargreaves & Crabtree 2011).

Otro mecanismo que se ha empezado a investigar en los últimos años es la regulación de la cromatina mediada por RNAs no codificantes. A pesar de que los RNAs no codificantes forman parte esencial del proceso de traducción (RNA de transferencia, RNA ribosomal, etc.), no ha sido sino hasta años recientes que se ha empezado a apreciar su influencia en el proceso transcripcional por medio de la regulación del templado de la cromatina. Distintos tipos de RNAs no codificantes participan en la regulación de la cromatina, pero en particular los RNA codificantes largos (lncRNAs, por sus siglas en inglés) parecen ser los que están mayormente involucrados en el proceso epigenético. Se sabe muy poco acerca de los mecanismos por los cuales los lncRNAs regulan a la cromatina, sin embargo diversos estudios han mostrado que suelen unirse a complejos lectores, escritores y efectores para modular su función, ya sea atenuándola o aumentándola. Adicionalmente, se ha propuesto que varios lncRNAs funcionan como andamios que guían a los distintos complejos a sus sitios blanco en el genoma (Bonasio & Shiekhataar 2014).

Finalmente, existe un sexto mecanismo de regulación epigenética: la regulación de la arquitectura tridimensional de la cromatina. La cromatina no se encuentra distribuida aleatoriamente dentro del núcleo. A pesar de que la secuencia del DNA es lineal, su distribución en el núcleo no lo es. Existe una compleja distribución tridimensional que aproxima físicamente ciertas regiones mientras que aleja otras. Esta distribución, lejos de ser aleatoria, parece estar finamente regulada pues es precisamente a través de estas interacciones de largo alcance que el genoma se puede dividir en "dominios" que pueden ser regulados de forma distinta entre sí. La distribución espacial del genoma, y las consecuencias que ésta tiene en la regulación transcripcional apenas comienzan a ser estudiadas. Un ejemplo paradigmático de esta regulación es la formación de bucles de

cromatina entre las regiones *enhancer* y los promotores de sus genes regulados. Cuando los bucles aproximan físicamente a la región *enhancer* a la región promotora, el proceso de transcripción se lleva a cabo de forma eficiente. Si la formación del bucle se inhibe, la transcripción se interrumpe. Sin embargo, las interacciones *enhancer*-promotor son apenas el primer nivel topológico de organización de la cromatina. Existen estructuras todavía más grandes que permiten establecer grandes dominios que generalmente separan a la eucromatina de la heterocromatina. Aunque la existencia de los dominios es bien aceptada, los mecanismos a través de los cuáles se forman y se regulan son en gran parte un misterio (Bickmore & Van Steensel 2013; Furlan-Magaril et al. 2015).

En resumen, existen seis grandes mecanismos de regulación epigenética: las marcas de histonas, la metilación del DNA, las variantes de histona, los remodeladores de la cromatina, los lncRNAs y la arquitectura tridimensional de la cromatina. Sin embargo, estas divisiones son esencialmente conceptuales, ya que todos estos fenómenos convergen entre sí para regular de una manera fina, precisa y dinámica la estructura de la cromatina y la transcripción génica.

Entender cómo es que los diversos mecanismos interactúan entre sí para generar patrones específicos de transcripción es una de las áreas que más interés está despertando en el campo de la epigenética actualmente. Sin embargo, existen pocos ejemplos de proteínas que estén involucradas en todos los mecanismos antes descritos. Uno de esos pocos ejemplos es la proteína multifuncional ATRX, la cual describiremos a detalle en la siguiente sección.

ATRX: Un remodelador peculiar de la cromatina

ATRX es un remodelador de la cromatina ATP-dependiente perteneciente a la familia de los remodeladores de tipo SWI/SNF con una gran relevancia médica. La mayoría de los remodeladores de la cromatina de la familia SWI/SNF forman parte de grandes complejos remodeladores que han sido ampliamente caracterizados (Hargreaves & Crabtree 2011). Sin embargo, ATRX es un remodelador atípico que aunque ha sido encontrado en complejos de alto peso molecular no ha sido formalmente identificado en ninguno de los complejos típicos de la familia SWI/SNF (Xue et al. 2003; Ratnakumar & Bernstein 2013). En cambio, ha sido ligado directamente a cada uno de los distintos mecanismos de regulación epigenética. A continuación, repasaremos uno a uno dichos vínculos, no sin antes analizar la relevancia médica de ATRX.

Relevancia médica de ATRX

ATRX comenzó a ser estudiado originalmente debido a que mutaciones en la línea germinal en el gen que codifica para la proteína son responsables de una compleja enfermedad llamada *Alpha thalassemia X-linked intellectual disability syndrome* (síndrome de discapacidad intelectual y α -talasemia ligada al cromosoma X, Gibbons et al. 1991; Gibbons et al. 1995; Gibbons 2006). Los pacientes que presentan dicha enfermedad, como su nombre lo indica, tienen grados variables de α -talasemia (la deficiencia en la producción de las subunidades de la hemoglobina conocidas como α -globinas) y una severa discapacidad intelectual. Debido a que el síndrome es una enfermedad rara, por muchos años el estudio de ATRX se vio limitado al estudio del síndrome y a su relación con el desarrollo neuronal y la regulación de los genes α -globina (Bérubé 2011; Ratnakumar & Bernstein 2013). Se encontró posteriormente que algunos pacientes con tipos raros de leucemia desarrollaban α -talasemia al adquirir mutaciones *de novo* en el gen ATRX (Gibbons et al. 2003). Esta fue la primera evidencia que ligó a ATRX con el desarrollo de cáncer, pero al tratarse de casos relativamente raros y aislados permaneció como un fenómeno interesante para los grupos dedicados a investigar el síndrome.

Recientemente y gracias a las nuevas tecnologías de secuenciación masiva del DNA, se encontró que las mutaciones o grandes alteraciones somáticas en el gen ATRX son también una característica frecuente de varios tipos de cáncer entre los que se encuentran: glioblastoma, neuroblastoma, osteosarcoma, tumores pancreáticos neuroendocrinos, por mencionar sólo algunos ejemplos (Jiao et al. 2011; Jiao et al. 2012; Cheung et al. 2012; Chen et al. 2014; Watson et al. 2015). Adicionalmente se encontró que algunos otros cánceres, como melanoma, silencian la expresión de ATRX durante la progresión de la enfermedad, sin necesidad de acumular mutaciones en el gen (Qadeer et al. 2014). Todas estas evidencias han hecho que ATRX sea reconocido en la actualidad como un regulador onco-epigenético (Watson et al. 2015). Sin embargo, quedan poco claros los mecanismos a través de los cuáles la inactivación genética o transcripcional de ATRX promueve la tumorigénesis.

ATRX y las marcas de histonas

ATRX posee dos dominios altamente conservados (Picketts et al. 1998). Uno de ellos es un dominio de tipo PHD capaz de unirse a la H3K9me3 y reconocer la falta de metilación de la H3K4 (HeK4me0). Esto ubicaría a ATRX en la categoría de un lector de la cromatina (Iwase et al. 2011; Eustermann et al. 2011; Dhayalan et al. 2011). El otro

dominio de ATRX es un dominio de helicasa dependiente de ATP lo que le da la característica de remodelador de la cromatina. Aunque la función precisa de este dominio se desconoce, se ha observado que ATRX presenta actividad de translocasa de DNA y que requiere de ATP para la incorporación de H3.3 (ver más abajo). Adicionalmente, la mayoría de las mutaciones en el síndrome de ATRX se presentan en alguno de estos dos dominios (Gibbons et al. 2008; Mitson et al. 2011; Lewis et al. 2010).

Estudios recientes han encontrado que ATRX es un regulador de la deposición de H3K9me3 en regiones repetitivas y loci improntados (ver más adelante). La eliminación de ATRX reduce la acumulación de la marca en dichas regiones y parece tener efectos negativos en la estabilidad genómica de la célula (Elsässer et al. 2015; Sadic et al. 2015; He et al. 2015; Voon et al. 2015; Voon & Wong 2016).

ATRX y la metilación del DNA

El primer vínculo entre ATRX y la metilación del DNA fue encontrado en un estudio en pacientes con el síndrome de ATRX donde se encontró que los pacientes con dicha enfermedad presentan una disminución en la metilación del DNA en algunas regiones repetitivas como las repetidas de DNA ribosomal, y algunas repetidas satelitales y sub-teloméricas (Gibbons et al. 2000).

Estudios más recientes encontraron que ATRX es un regulador importante de genes improntados. Los genes improntados son genes expresados exclusivamente del alelo de uno de los padres. La copia que no es expresada es silenciada por métodos epigenéticos, frecuentemente a través de la metilación del DNA. Varios estudios han demostrado que ATRX es un regulador global de la impronta. Cuando ATRX está mutado, la metilación del DNA de la copia silenciada se pierde y ésta comienza a expresarse de forma aberrante (Kernohan et al. 2010; Kernohan et al. 2014; Voon et al. 2015). Por ejemplo, en el caso del dominio improntado *H19/Igf2*, en el que típicamente la copia paterna de H19 está silenciada, muestra un incremento en la expresión de H19 en la ausencia de ATRX (Kernohan et al. 2010).

Finalmente, se ha demostrado que ATRX interactúa con Mecp2, una importante proteína de unión al DNA metilado, aunque se desconocen las implicaciones de esta interacción más allá de los loci improntados que ya se mencionaron (Nan et al. 2007; Kernohan et al. 2010; Kernohan et al. 2014).

ATRX y las variantes de histona

ATRX, en conjunto con la chaperona de histonas DAXX, está encargada de la deposición de la variante de histonas H3.3 en regiones de heterocromatina tales como telómeros, regiones pericentroméricas, retrotransposones, etc. (Wong et al. 2010; Goldberg et al. 2010; Drané et al. 2010; Lewis et al. 2010; Elsässer et al. 2015; Sadic et al. 2015; He et al. 2015). La función de la deposición de H3.3 en dichas regiones todavía no es bien entendida, aunque parece estar relacionada con el mantenimiento de la H3K9me3 y la represión transcripcional.

Adicionalmente, ATRX también regula negativamente la deposición de la variante de histona macroH2A (mH2A) (Ratnakumar et al. 2012). mH2A es una variante que se ha relacionado principalmente a la formación de heterocromatina, particularmente en el contexto del silenciamiento del cromosoma X. Esta función parece ser relevante para la regulación del locus α -globina, ya que la inactivación de ATRX promueve la deposición de mH2A en dicho locus y el silenciamiento de los genes α -globina (Ratnakumar et al. 2012).

ATRX y la remodelación de la cromatina

A pesar de que ATRX ha sido definido como un remodelador de la cromatina de tipo SWI/SNF por el dominio de helicasa que posee, sólo existen hasta la fecha un par de reportes que muestran que ATRX es una ATPasa que es estimulada por la unión al DNA y mononucleosomas (Xue et al. 2003; Tang et al. 2004) y que tiene la capacidad de desplazar al DNA en presencia de ATP (Xue et al. 2003; Mitson et al. 2011). Una de las funciones de los remodeladores de la cromatina SWI/SNF: su capacidad de alterar la unión entre los nucleosomas y el DNA, no ha sido demostrada aún para ATRX. El hecho de que ATRX participa en la deposición de H3.3 sugiere que en efecto posee esta capacidad, pero aún no existe evidencia experimental clara que lo demuestre.

ATRX y los lncRNAs

Un interesante estudio que usó el modelo de la inactivación del cromosoma X describió que ATRX se requiere para mediar la interacción entre el lncRNA *Xist*, un lncRNA esencial para el establecimiento y la extensión del silenciamiento en el cromosoma X inactivo, y el complejo PRC2, un complejo especializado en la "escritura" de la H3K27me3 y el establecimiento de heterocromatina. Si se disminuye la expresión de ATRX, la interacción entre *Xist* y PRC2 se reduce, la H3K27me3 no se establece y el cromosoma X no se puede inactivar adecuadamente. El artículo demostró además que

ATRX tiene una fuerte afinidad por el RNA y que es capaz de unirse de forma directa a una secuencia repetitiva de Xist (Sarma et al. 2014). Éste es hasta la fecha el único reporte que ha demostrado la interacción de ATRX con lncRNAs, pero es posible que en el futuro se descubran más lncRNAs regulados por ATRX.

ATRX y la arquitectura tridimensional de la cromatina

En el contexto de la regulación de genes improntados, se ha demostrado que ATRX también está involucrado con la regulación de la arquitectura tridimensional de la cromatina. Los loci improntados son notables porque suelen requerir de complejas estructuras tridimensionales que aseguran que sólo uno de los alelos se mantendrá activo. Esto implica que, de forma alelo específica, se promueven o se impiden las interacciones con elementos regulatorios. Una de las proteínas que regula dicha función es la proteína *insulator* y regulador tridimensional de la cromatina CTCF (Phillips & Corces 2009). La eliminación de ATRX no sólo provoca la desmetilación del DNA del locus silenciado, también altera las interacciones tridimensionales de los loci improntados al modificar la unión de CTCF a sus sitios blancos (Kernohan et al. 2014).

Estudiando la función de ATRX a través de dos aproximaciones distintas

Como el lector podrá apreciar, existen algunos ejemplos en los roles de ATRX en los que se puede apreciar la interacción entre distintos mecanismos epigenéticos. Por ejemplo, ATRX interactúa con lncRNAs para regular la deposición de marcas de cromatina. En el caso de las regiones repetitivas, deposita H3.3 para regular el enriquecimiento de H3.3 y consecuentemente el de la H3K9me3. En el ejemplo de los loci improntados, regula la metilación del DNA y la deposición de H3.3 para mantener la arquitectura tridimensional de la cromatina.

Sin embargo, muchos de los mecanismos en los que está involucrado ATRX parecen ser dispares, e inconexos. Cabe preguntarse, por ejemplo si la regulación de los genes improntados está relacionada con la capacidad de ATRX de depositar H3.3 en los telómeros. O si el enriquecimiento de H3K9me3 en las regiones repetitivas está ligado a su rol como regulador de la arquitectura tridimensional de la cromatina.

Para poder responder a dichas preguntas consideramos que es viable tomar dos aproximaciones complementarias entre sí. Por un lado, es necesario tomar una aproximación imparcial y global para tratar de determinar el rol funcional de ATRX en un contexto genómico. Esta clase de aproximación, conocida como epigenómica, ya ha

dado frutos importantes en nuestro entendimiento de los diversos mecanismos epigenéticos y su relación entre ellos (Higgs et al. 2007). Por otro lado, el estudio cuidadoso y detallado de casos paradigmáticos que involucran la interacción de diversos mecanismos genéticos y epigenéticos ha sido una fuente invaluable de conocimiento (Guerrero et al. 2007). Proponemos que el análisis de la regulación de ATRX en el locus α -globina es uno de esos casos paradigmáticos.

A continuación describiremos con detalle ambas aproximaciones y los resultados obtenidas con ellas. La tesis está dividida en tres secciones. La primera detalla el análisis epigenómico y el descubrimiento de una nueva familia de genes regulados por ATRX. La segunda describe los análisis que realizamos usando el locus α -globina como modelo de estudio. La tercera es un recuento breve de algunos problemas y soluciones técnicas que podrían ser valiosas para la comunidad.

Parte I: ATRX y la regulación de los genes ZNF

“The study of epigenetics is founded on trying to understand unexpected observations”

- Daniel Gottschling

I.I - Introducción

A pesar de que se han hecho descubrimientos considerables en los últimos años, aún quedan muchas incógnitas respecto a las funciones de ATRX y existen muy pocos estudios que hayan abordado el tema de forma global e imparcial. A continuación se describirán los hallazgos más importantes a nivel de la función molecular de ATRX, se discutirán sus limitaciones, se abordarán algunas de las incógnitas que aún quedan por resolver y se analizarán algunos modelos de estudio que nos permitirán comprender mejor el rol funcional de ATRX .

I.I.I - Entendiendo la función de ATRX: hipótesis e incógnitas

ATRX posee dos dominios altamente conservados: el dominio ADD (ATRX-Dnmt3-DNMT3L) y el dominio helicasa SWI/SNF (Park et al. 2004). El dominio ADD contiene un dedo de tipo PHD que es capaz de reconocer la combinación de marcas de histona H3K9me3/H3K4me0 (Eustermann et al. 2011; Iwase et al. 2011; Dhayalan et al. 2011), mientras que el dominio SWI/SNF es una helicasa dependiente de ATP responsable de la posible capacidad de remodelación de la cromatina de ATRX (Xue et al. 2003; Park et al. 2004). A pesar de que se une a H3K9me3/H3K4me0 *in vitro*, ATRX se une sólo a un subconjunto de regiones que contienen H3K9me3 *in vivo* (McDowell et al. 1999; Law et al. 2010; Eustermann et al. 2011; Iwase et al. 2011) por razones que se desconocen. En particular, ATRX está enriquecido en ciertas regiones repetitivas que contienen H3K9me3, tales como los telómeros y las regiones pericentroméricas, así como algunas familias de retrotransposones (Drané et al. 2010; Goldberg et al. 2010; Wong et al. 2010; Lewis et al. 2010; Sadic et al. 2015; Elsässer et al. 2015; He et al. 2015). Además, ATRX interactúa físicamente con otras proteínas de unión a H3K9me3 como HP1 α (Bérubé et al. 2000; Lechner et al. 2005). En conjunto, estas evidencias sugieren que ATRX está implicado en la regulación de la cromatina, en particular en las regiones ricas en H3K9me3. Sin embargo, se desconoce por qué ATRX se une sólo a un subgrupo específico de regiones ricas en H3K9me3, y si existen características comunes entre dichas regiones más allá de que poseen secuencias repetitivas con la capacidad de formar estructuras G4.

Un papel bien caracterizado de ATRX es el control de la deposición de variantes de histonas en la cromatina. Las variantes de histonas son histonas no canónicas que se transcriben de forma independiente al ciclo celular y que poseen características funcionales que las diferencian de las histonas canónicas (Maze et al. 2014). Por ejemplo,

ATRX y DAXX (Death-Domain Associated Protein) actúan juntos como un complejo de chaperonas de histona para la variante H3.3, una variante que se ha encontrado asociada a regiones de eucromatina, en particular promotores de genes transcripcionalmente activos (Maze et al. 2014). ATRX se requiere para la localización de H3.3 en telómeros y repetidas pericentroméricas (Goldberg et al. 2010; Drané et al. 2010; Wong et al. 2010; Lewis et al. 2010), retrotransposones (Sadik et al. 2015; Elsässer et al. 2015; He et al. 2015) y loci improntados (Voon et al. 2015). Todas las regiones en las que ATRX/DAXX deposita H3.3 contienen H3K9me3 y suelen ser transcripcionalmente inactivas. Esta capacidad parece ser única para ATRX, ya que otro de los complejos capaz de depositar H3.3, HIRA, solamente la deposita en regiones de eucromatina (Goldberg et al. 2010; Ray-Gallet et al. 2011; Schneiderman et al. 2012; Filipescu et al. 2013). Esta evidencia sugeriría que el rol primario de ATRX es el silenciamiento transcripcional. Sin embargo, nuestro grupo demostró que, además de promover la deposición H3.3, ATRX regula negativamente la deposición de variante de la histona macroH2A (mH2A) en el locus α -globina (Ratnakumar et al. 2012). Dado que mH2A es un regulador negativo de la transcripción, el efecto neto de ATRX al impedir su deposición es la activación de los genes α -globina (ver parte I, Ratnakumar et al. 2012). Esto pone de relieve la complejidad de los mecanismos de regulación mediados por ATRX y sugiere que ATRX posee diversas funciones que pueden ser sitio o tejido específicas.

ATRX también se ha implicado en la resolución de estructuras secundarias aberrantes de DNA, llamados G-quadruplex o estructuras G4, que se forman en regiones ricas en guanina durante la replicación y la transcripción (Law et al. 2010; Clynes et al. 2013; Levy et al. 2015). Las estructuras G4 son una característica común de algunas familias de secuencias repetitivas y repetidas en tándem, tales como las que se encuentran en los telómeros. Curiosamente, las mutaciones de ATRX en células cancerosas se han relacionado con el fenómeno de alargamiento alternativo de los telómeros (ALT, por sus siglas en inglés), un mecanismo alternativo a la reactivación de la telomerasa por el cual las células cancerosas pueden mantener la longitud de sus telómeros por medio de un mecanismo de recombinación telomérica y así dividirse indefinidamente (Heaphy et al. 2011; Schwartzenuber et al. 2012; Lovejoy et al. 2012; Bower et al. 2012). Aunque la función precisa de ATRX en el mantenimiento del ALT sigue siendo poco clara, se ha sugerido que en condiciones normales, ATRX evita la recombinación homóloga (HR) entre secuencias teloméricas (Clynes et al. 2015; Ramamoorthy & Smith 2015; Napier et al. 2015). Una posible hipótesis del mecanismo a través del cual ATRX regula las regiones teloméricas, es precisamente la resolución de las estructuras G4, las cuáles podrían servir como un impedimento para la duplicación de las repetidas teloméricas y una posible fuente de los cortes de DNA en doble cadena requeridos para iniciar el

proceso de recombinación necesario para el ALT (Clynes & Gibbons 2013; Clynes et al. 2015). Sin embargo, este modelo aún es especulativo.

A nivel genómico, se ha sugerido que ATRX es un regulador de la estabilidad del genoma. Esto en base a varios reportes que han demostrado en diferentes modelos *in vitro* e *in vivo* que la eliminación de ATRX causa disfunciones en los telómeros, aumento en el número de horquillas de replicación estancadas cuando las células se someten a estrés replicativo, mayores tasas de daño al DNA, y defectos en la segregación mitótica (Wong et al. 2010; De La Fuente et al. 2011; Huh et al. 2012; Watson et al. 2013; Leung et al. 2013; Clynes et al. 2014; He et al. 2015).

II.I.II - K562 como un modelo para entender el papel de ATRX en la regulación epigenómica

A pesar de que ATRX ha sido involucrado en distintas funciones como la deposición de variantes de histonas, la regulación de H3K9me3, la resolución de estructuras G4 y la estabilidad genómica, se han identificado sorprendentemente pocos genes blanco de ATRX. Más aún, se desconoce si todos estos mecanismos están ligados entre sí o son independientes.

Aunque ciertas regiones modelo, tales como los telómeros, se han usado para intentar entender algunas de las funciones de ATRX, se necesita hacer un análisis global e imparcial que intente analizar la función molecular de ATRX en un contexto genómico.

Para ello proponemos que la línea celular eritroleucémica K562 es un buen modelo de estudio por diversas razones: 1) La relevancia funcional de ATRX en dicha línea y en el modelo eritroide ha sido establecida por diversos grupos, incluido el nuestro (Law et al. 2010; Ratnakumar et al. 2012; Ratnakumar & Bernstein 2013). 2) La línea K562 es una línea tier1 (el más alto) del proyecto ENCODE (Dunham et al. 2012), lo que significa que todos los experimentos relacionados al consorcio se deben hacer en esa línea antes que todas las demás. Por lo tanto, existen, literalmente, cientos de sets de datos de diferentes experimentos que pueden ser usados como una rica fuente de información. 3) La línea K562 se puede crecer en grandes cantidades lo que la hace un buen modelo para realizar experimentos bioquímicos.

A pesar de dichas características, también planteamos la necesidad de diseñar un estudio que pueda ser extendido a otros modelos de estudio. En este sentido, consideramos que

una buena aproximación global es hacer uso de la tecnología de ChIP-seq para analizar el patrón de unión global y realizar un análisis global y exploratorio agnóstico, integrando estos datos con ChIP-seqs de otras marcas de cromatina, factores de transcripción (TFs por sus siglas en inglés), etc. para generar hipótesis a través de un análisis imparcial de los datos.

I.II - Justificación y planteamiento del problema

Un creciente número de reportes han mostrado que distintos tipos de cáncer poseen mutaciones somáticas en el gen que codifica para ATRX (Watson et al. 2015). A su vez, las mutaciones en línea germinal de *ATRX* generan un complejo síndrome humano (Gibbons 2006). A pesar de que *ATRX* se ha involucrado en diversas funciones celulares, sus mecanismos moleculares permanecen poco claros. Debido a su cada vez más extendida relevancia en la salud humana, es imperante entender las funciones precisas de *ATRX*. A pesar del interés creciente que el campo ha recibido en los últimos años, existen muy pocos análisis sistemáticos que intenten estudiar de forma global los mecanismos a través de los cuáles *ATRX* regula el epigenoma (Bérubé 2011; Ratnakumar & Bernstein 2013; Watson et al. 2015). No obstante se han realizado algunos estudios de ChIP-seq de *ATRX* en ratón que han arrojado resultados interesantes (Sarma et al. 2014; Levy et al. 2015; Voon & Wong 2016), hasta la fecha sólo se ha realizado un análisis de ChIP-seq en células humanas cuya calidad estuvo limitada por el tipo de muestra que se usó (Law et al. 2010). Es por ello que planteamos la necesidad del desarrollo de técnicas y modelos que permitan un análisis epigenómico imparcial e integrativo del patrón de unión de *ATRX* en células humanas y que aproveche la enorme cantidad de datos generados por consorcios internacionales como ENCODE para estudiar su relación con otras marcas, actores epigenéticos y procesos celulares.

I.III - Hipótesis

A través del análisis del patrón de unión global de *ATRX* al genoma se pueden establecer nuevos blancos de *ATRX* y determinar algunas de las funciones regulatorias del mismo

I. IV - Objetivos

I.IV.I - Objetivo general

Realizar un ChIP-seq de ATRX en la línea celular K562 y hacer un análisis epigenómico para encontrar nuevos blancos de ATRX y entender su rol como regulador global de la cromatina. Corroborar experimentalmente la importancia de ATRX en la regulación de los nuevos blancos y su función global en la cromatina.

I.IV.II - Objetivos particulares

- 1) Estandarizar un protocolo de ChIP-seq para ATRX en la línea celular humana
- 2) Analizar los datos de ChIP-seq y realizar un análisis integrativo con otros datos epigenómicos de la misma línea celular obtenidos por el consorcio ENCODE.
- 3) Buscar nuevos blancos de ATRX. Estudiar su contexto genómico y epigenómico. Analizar su correlación con otros reguladores epigenéticos para los que existan datos de ChIP-seq
- 4) Generar células KO de ATRX por medio de CRISPR y analizar los cambios globales en la cromatina.
- 5) Analizar el efecto de la eliminación de ATRX en los nuevos blancos encontrados en el punto 3.

I.V - Resultados

Los resultados del objetivo 1) se describen a detalle en la sección III.II. En breve, se logró establecer un método de ChIP-seq eficiente y reproducible que nos permitió obtener datos de muy alta calidad en la línea celular K562. Por comunicaciones personales con diversos laboratorios en el mundo que están trabajando con ATRX sabemos que realizar experimentos de ChIP-seq en humano es desafiante. De hecho, además de nuestro análisis, sólo hay otro artículo que ha publicado un ChIP-seq de ATRX en humano (Law et al. 2010), a pesar del intenso interés que ATRX ha generado a partir de su involucramiento con diversos tipos de cáncer (Ratnakumar & Bernstein 2013; Watson et al. 2015).

El análisis de nuestros datos de ChIP-seq generó una publicación en la revista *Epigenetics* titulada “ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment”. Dicha publicación se adjunta en el Apéndice I. A continuación se hará un resumen de los principales resultados de dicho artículo. Se hará referencia a las figuras del artículo de la siguiente forma: ZNF-Fig1, ZNF-Fig2A, etc.

I.V.I - ATRX se une al exón 3' de los genes de tipo *zinc finger*

Tras realizarse un análisis de ChIP-seq de ATRX usando células K562 como modelo se comparó el perfil de unión de ATRX con dominios funcionales de cromatina previamente identificados en K562 por el consorcio ENCODE (Ernst et al. 2011). Se encontró que, como ya había sido reportado previamente (Law et al. 2010; Eustermann et al. 2011; Ratnakumar & Bernstein 2013), ATRX se une a regiones repetidas y regiones de heterocromatina.

Sin embargo, de forma inesperada, se observó que ATRX también se une a regiones que son transcripcionalmente activas, lo que contrasta con su rol como regulador de la heterocromatina (ZNF-Fig1A). Al analizar los genes a los que se une ATRX, se encontró de forma sorpresiva que una cantidad significativa pertenece a la familia de TFs de tipo Zinc Finger (ZNFs) (ZNF-Fig1B-E). La familia ZNF es la familia de TFs más grande del genoma humano con aproximadamente 700 miembros. Es además la familia de genes con una tasa de evolución y duplicación más alta del genoma, lo que sugiere que recombinan de forma frecuente (Tadepally et al. 2008).

Los genes ZNFs se encuentran en grandes clústeres contiguos en el genoma humano. Inesperadamente, ATRX demarca estos clústeres (ZNF-Fig1D-E). Al analizar de forma detallada la manera en la que ATRX se une a los ZNFs encontramos que se enriquece en el cuerpo de los mismos, específicamente en su exón 3' (ZNF-FigF,G). Para entender el contexto epigenético de los ZNFs unidos por ATRX, se re-analizaron todas las marcas de histonas con datos de ChIP-seq disponibles en ENCODE y se estudió su patrón de enriquecimiento en los genes ZNF. En acuerdo con su papel como regulador de la heterocromatina, se vio que H3K9me3 tiene un patrón muy similar al de ATRX en los genes ZNFs, demarca los clústeres y está enriquecido en las regiones 3' (ZNFs-FigE,G). Sin embargo, también se observó que H3K36me3, una marca asociada a la elongación de la polimerasa se encontraba enriquecida a la par de H3K9me3. Dado que las regiones de H3K9me3 son casi siempre regiones de heterocromatina silenciadas, es muy atípico

encontrarla asociada con marcas de cromatina abierta, incluida la H3K36me3. Es por ello que denominamos a estas regiones de los exones 3' de los genes ZNF como dominios de cromatina atípica.

I.V.II - Los genes ZNF a los que ATRX se une poseen una cromatina atípica y características genéticas particulares

Para estudiar más a fondo la unión de ATRX a los ZNFs, los dividimos en tres categorías basados en su enriquecimiento de ATRX por medio de un algoritmo de *k-means clustering*: La clase I, contiene a los genes altamente enriquecidos para ATRX. La clase II contiene a genes medianamente o débilmente enriquecidos en ATRX. La clase III contiene a genes sin enriquecimiento de ATRX (ZNF-Fig2A). Al analizar el enriquecimiento de H3K9me3 y H3K36me3 en estas categorías, descubrimos que el patrón de H3K9me3 mimetiza al de ATRX, en tanto que H3K36me3 está altamente enriquecido en la clase I y medianamente enriquecido en las clases II y III (ZNF-Fig2A).

A pesar de que en promedio, el enriquecimiento de H3K9me3 y H3K36me3 es más alto en los ZNFs enriquecidos en ATRX de la clase I, cabe la posibilidad de que sean marcas independientes (es decir, que no se encuentren enriquecidas al mismo tiempo en los genes). Para descartar esa posibilidad comparamos el nivel de correlación entre el enriquecimiento de ATRX, H3K9me3 y H3K36me3 en los ZNFs de clase I y clase III. Nuestros resultados prueban que el enriquecimiento de H3K9me3, H3K36me3 y ATRX se encuentra correlacionado en los genes de clase I, mientras que esta correlación se pierde en los genes de clase III (ZNF-Fig2B). Estos resultados corroboran que ATRX se une específicamente a genes ZNF que presentan una cromatina atípica.

Posteriormente, analizamos algunas características genéticas de las distintas clases de ZNFs. Dado que encontramos que los ZNFs que contienen al dominio represor KRAB están también enriquecidos entre los blancos de ATRX (ZNF-Fig1B), estudiamos si los ZNFs de la clase I y II poseen, en promedio, más dominios KRAB que los genes de clase III. Nuestros resultados corroboraron que efectivamente, los ZNFs de clase I y II casi siempre poseen dominios KRAB en tanto los de clase III lo hacen con una frecuencia significativamente menor (ZNF-Fig1B).

La mayoría de los genes ZNF poseen varias copias del motivo (o dominio) zinc finger. Un ZNF promedio posee 7 copias del motivo, y existen genes que pueden llegar a poseer hasta 40. Curiosamente, estos motivos están frecuentemente codificados en el exón 3', el

sitio de unión de ATRX. Por esa razón nos preguntamos si el número de motivos zinc finger que los genes poseen estaba relacionado con la unión de ATRX. Efectivamente, los genes de clase I y II poseen más motivos zinc finger en promedio que los de la clase III (ZNF-Fig1B).

Dado que se ha sugerido que ATRX regula a las estructuras G4 que se forman en regiones ricas en G, nos preguntamos si los exones 3' de las clases I y II tenían alto contenido en G y presentan alta probabilidad de formar estructuras G4. Contrario a lo esperado, los genes de las clases I y II poseen un bajo contenido en G, y una baja probabilidad de formar estructuras G4. Los genes de la clase III, en cambio, son más ricos en G y tienen mucha mayor probabilidad de formar estructuras G4. Esto sugiere que los ATRX reconoce a los ZNFs por un mecanismo alternativo a su afinidad por las estructuras G4 (ZNF-Fig1B).

Finalmente, analizamos los niveles de expresión de los genes ZNFs, re-analizando los datos de RNA-seq de ENCODE. Sorprendentemente, los ZNFs presentan niveles de expresión muy heterogéneos y sin ninguna correlación clara en su unión por ATRX. Tanto en la clase I como en la clase III, los genes poseen el mismo nivel de expresión promedio, y se pueden encontrar varios ejemplos de genes sobre- y sub-expresados. Esto sugiere que la unión de ATRX a los exones 3' de los genes ZNF no tiene un efecto claro sobre su nivel de expresión (ZNF-Fig1B).

I.V.III ATRX forma un complejo con SETDB1, TRIM28 y ZNF274 que se une a los exones 3' de los genes ZNF

Para intentar descubrir otras proteínas que pudieran estar regulando a los genes ZNF en conjunto con ATRX, re-analizamos los 98 sets de datos de ChIP-seq disponibles para K562 en el proyecto ENCODE. De todos los distintos TFs, marcas, reguladores de la cromatina, etc., tres de ellos tienen un patrón de unión llamativamente similar al de ATRX en los genes ZNF: SETDB1, TRIM28, y ZNF274. Estos datos sugieren que ATRX forma parte del complejo SETDB1/TRIM28/ZNF274.

El complejo SETDB1/TRIM28/ZNF274 ha sido identificado previamente como un complejo que se une específicamente a las regiones de cromatina atípica en los exones 3' de los genes ZNF (Frietze et al. 2010). ZNF274 es un factor de transcripción del tipo KRAB-ZNF que se une específicamente a la región 3' de los genes ZNF. TRIM28 es un cofactor que reconoce el dominio KRAB y se une a ZNF274. TRIM28 tiene la capacidad

de reclutar metiltransferasas de H3K9 para silenciar sus genes blanco. En el caso de los genes ZNF recluta a la metiltransferasa SETDB1 que es a su vez responsable del establecimiento de la H3K9me3 en dichas regiones (Frietze et al. 2010; Iyengar et al. 2011).

Tras confirmar que el enriquecimiento por CHIP-seq de estos factores está correlacionado entre sí específicamente en los genes ZNFs (ZNF-Fig2E), se hicieron experimentos de co-inmunoprecipitación con un anticuerpo contra ATRX en extractos de cromatina de células K562. Encontramos que al inmunoprecipitar a ATRX, pudimos co-inmunoprecipitar a SETDB1, TRIM28 y a ZNF274 (ZNF-Fig2F). Estos datos sugieren que ATRX interactúa físicamente con SETDB1/TRIM28/ZNF274 y que este nuevo complejo regula a los genes ZNF.

I.V.IV - La unión de ATRX a los genes ZNF está conservada en distintos linajes celulares humanos

Para demostrar que la unión de ATRX a los genes ZNF no es una característica única de las células K562, se usaron otras líneas celulares humanas con datos en ENCODE que incluyen células normales y células derivadas de cáncer de distintos linajes (H1 hESC, HeLa y HepG2) así como dos líneas celulares humanas de neuroblastoma (LAN6 y SKNFI). Se eligieron 11 ZNFs al azar pertenecientes a la clase I (que poseen alto enriquecimiento de ATRX en K562) que están distribuidos en distintos cromosomas. A su vez se eligieron dos genes al azar de la clase III (que no están unidos por ATRX en K562) como control. Se realizaron análisis de CHIP-qPCR en todas las líneas celulares elegidas, incluida K562, para el panel de ZNFs elegidos. Como se muestra en la figura ZNF-Fig3A, todas las líneas celulares probadas muestran altos enriquecimientos de ATRX en los ZNFs de clase I y baja en los ZNFs de clase III. Estos resultados corroboran los datos obtenidos por CHIP-seq y muestran que ATRX no se asocia con ZNFs exclusivamente en K562, sino que esta parece ser una característica general de las líneas celulares humanas.

Para analizar si la asociación de ATRX con H3K9me3 y con ZNF274 se mantiene en otras líneas celulares en los ZNFs, se re-analizaron los datos de CHIP-seq disponibles en ENCODE para K562, H1 hESC, HeLa y HepG2. Efectivamente, tanto H3K9me3 y ZNF274 están enriquecidos en el panel de ZNFs de clase I analizados por CHIP-qPCR y su enriquecimiento es menor en los ZNFs de clase III (ZNF-Fig2B). Desafortunadamente no hay datos de H3K36me3, TRIM28 ni SETDB1 disponibles para analizar si estos ZNFs

poseen dominios de cromatina atípica y si los otros miembros del complejo también se unen a los ZNFs de clase I.

Adicionalmente, se re-analizaron los datos de RNA-seq disponibles para K562, H1 hESC, HeLa y HepG2 y se estudió el perfil de expresión del panel de ZNFs. Tal como se había observado previamente en K562 (ZNF-Fig1C), el patrón de expresión es heterogéneo, y variable entre líneas celulares. No se observa una diferencia evidente entre los ZNFs de clase I y los de clase III, lo que sugiere que los niveles transcripcionales de los ZNFs son independientes de la unión de ATRX (ZNF-Fig2C).

Con la finalidad de corroborar a nivel genómico que ATRX se une a los ZNFs en otras líneas celulares además de K562, e investigar su asociación con el enriquecimiento de H3K9me3, se realizaron ChIP-seqs de ATRX y H3K9me3 en la línea celular de neuroblastoma LAN6. Elegimos esta línea porque es la que muestra el mayor enriquecimiento de ATRX en los ZNFs por ChIP-qPCR (ZNF-Fig2A) además de que puede ser un modelo relevante a futuro para estudiar las funciones de ATRX en el desarrollo de neuroblastoma (un cáncer en el que se han encontrado mutaciones frecuentes en ATRX, Cheung et al. 2012; Molenaar et al. 2012). Los análisis del enriquecimiento de ATRX y H3K9me3 en los genes ZNF muestran que efectivamente ambos están enriquecidos en la región 3' de los genes ZNF pertenecientes a la clase I y II, y están ausentes en los ZNFs de clase III (ZNF-Fig2D-G).

I.V.V - La deficiencia de ATRX causa una pérdida de H3K9me3 en los exones 3' de los ZNFs

Con el objetivo de analizar si la unión de ATRX en los exones 3' de los ZNFs es relevante, se generaron dos líneas clonales KO para ATRX en células K562 usando el sistema CRISPR/Cas9. Se validó la eliminación de ATRX por WB y ChIP-qPCR (ZNF-Fig4A,B). Estos análisis, aunados a un detallado estudio de la secuencia genómica de las clonas, muestran que una de ellas tiene una mutación hipomórfica (ATRX KO1, uno de los alelos presenta una eliminación que preserva el marco de lectura y genera un producto proteico detectable por Western Blot) y la otra es una eliminación completa (ATRX KO2, ZNF-Fig4A,B). Ambas mutantes de ATRX fueron viables, aunque presentan fenotipos que se describirán más abajo.

Con estas dos clonas KO para ATRX y una clona control que sobre-expresa a Cas9 sola (V2), se realizaron experimentos de ChIP-qPCR de H3K9me3 en el panel de ZNFs. Se

observó que la mutante hipomórfica presenta una pérdida ligera (10-20%) de H3K9me3 en los ZNFs de clase I y la mutante completa tiene una pérdida mayor (30-50%) (ZNF-Fig4C). Estos datos sugieren fuertemente que la unión de ATRX sirve para establecer o preservar los niveles de H3K9me3 en los genes ZNF. Por otro lado, la marca de elongación H3K36me3, así como los niveles de SETDB1, TRIM28, ZNF274 y los niveles de expresión de los ZNFs no se ven afectados en las mutantes de ATRX (ZNF-SuppFigS7).

Para corroborar si este fenómeno sucede a nivel global, se realizaron experimentos de ChIP-seq de la marca H3K9me3 en células control y el KO completo de ATRX. Como se esperaba, los ZNFs pertenecientes a la clase I y clase II pierden considerablemente su enriquecimiento de H3K9me3, mientras que los ZNFs de la clase III se mantienen en niveles similares (ZNF-Fig4D,E).

I.V.VI - Los niveles de H3K9me3 en dominios de cromatina atípica son sensibles a la pérdida de ATRX

Durante los análisis del enriquecimiento de H3K9me3, se notó que los ZNFs no eran la única región que pierde la marca en el KO de ATRX. Para entender mejor el rol de ATRX como un regulador global de H3K9me3, analizamos todos los sitios enriquecidos para H3K9me3 y los clasificamos como "reducidos" si tenían una pérdida significativa de H3K9me3 en las mutantes de ATRX o "inalterados" si sus niveles de H3K9me3 se mantenían estables en la mutante. Debido a que según nuestros análisis previos ya se sabía que ATRX se une principalmente a las regiones de cromatina atípica, se dividieron a su vez las regiones de H3K9me3 en dos categorías "cromatina atípica" (aquellas regiones que tienen H3K9me3 que coinciden con enriquecimientos en H3K36me3) y regiones de "sólo H3K9me3" (aquellas que únicamente están enriquecidas en la marca). Posteriormente se analizó la proporción de regiones reducidas y estables que se encuentran en la cromatina atípica y en las regiones de sólo H3K9me3. Sorprendentemente, el 80% de los dominios de cromatina atípica son regiones que pierden H3K9me3. En cambio, sólo el ~42% de las regiones que tienen únicamente H3K9me3 pierden la marca y el resto permanece estable (ZNF-Fig4F,G). Estos resultados implican que ATRX es un regulador global de los niveles de H3K9me3 en las regiones de cromatina atípica.

Con la finalidad de entender de forma global qué tipo de regiones del genoma son afectadas por la pérdida de H3K9me3 en las mutantes de ATRX, analizamos la

distribución de las regiones de cromatina atípica. Consistente con su rol como un regulador de regiones repetitivas (Clynes et al. 2013), encontramos que la mayor parte de regiones de cromatina atípica que pierden H3K9me3 se encuentra en regiones repetitivas. En segunda posición se encontraron las regiones génicas, entre las que se encuentran los genes ZNFs (ZNF-Fig4H).

I.V.VII - ZNF274 recluta a ATRX a un subgrupo de genes ZNF para regular sus niveles de H3K9me3 y la estabilidad genómica

Debido a que el ZNF274 es la única proteína del complejo con la capacidad de reconocer un motivo de unión al DNA, hipotetizamos que ZNF274 reconoce una secuencia particular en la región 3' de los genes ZNF, y recluta al complejo ATRX/SETDB1/TRIM28 para establecer y/o mantener el enriquecimiento de H3K9me3. Con el fin de corroborar este modelo, se generaron clonas KO para ZNF274 y clonas doble KO para ZNF274 y ATRX usando el sistema CRISPR/Cas9. Para validar dichas clonas se realizaron experimentos de ChIP-qPCR de ZNF274 en el panel de ZNFs. Como se esperaba, el enriquecimiento de ZNF274 se pierde por completo tanto en las mutantes simples (ZNF274 KO) como en las doble mutantes (ZNF274/ATRX KO) (ZNF-Fig5A).

Posteriormente, se analizó por ChIP-seq el enriquecimiento de ATRX en estas clonas. De forma interesante, el nivel de ATRX se ve reducido en varios ZNFs en las células ZNF274 KO, pero no en todos (ZNF-Fig5B). Más aún, sólo algunos ZNFs presentan reducciones comparables a las células ZNF274/ATRX KO (ZNF-Fig5B), lo que sugiere que ZNF274 recluta a ATRX sólo a un subgrupo de genes ZNF y que en el resto ATRX puede ser reclutado por algún método alternativo, posiblemente por algún otro TF.

Después se analizó el enriquecimiento de H3K9me3 por ChIP-qPCR en el panel de ZNFs. La clona ZNF274 KO presenta una reducción del enriquecimiento de la H3K9me3 en los ZNFs de clase I, comparada con la clona control (random). Curiosamente, los ZNFs que presentan una pérdida completa de ATRX en estas clonas, según los resultados previos (ZNF-Fig5B), presentan una pérdida total de H3K9me3 (p. ej. ZNF267, ZNF721, ZNF891, ZNF-Fig5C). En cambio, los ZNFs que tienen una pérdida media de ATRX en las células ZNF274 KO, también tienen pérdidas moderadas de H3K9me3. Estos resultados sugieren fuertemente que el reclutamiento de ATRX y ZNF274 es esencial para el mantenimiento de la H3K9me3 en los genes ZNF. Apoyando esta idea, la clona ZNF274/ATRX KO presenta una mayor pérdida de H3K9me3 que los KO individuales (ZNF-Fig5C).

Reportes previos han sugerido que ATRX es un importante regulador de la estabilidad genómica, sin embargo se ignoran los mecanismos moleculares precisos a través de los cuáles lleva a cabo esta función (De La Fuente et al. 2011; Lovejoy et al. 2012; Huh et al. 2012; Marinoni et al. 2014). Para analizar el efecto funcional de la pérdida de H3K9me3 en la región 3' de los genes ZNF, se analizó por WB a la marca de histona γ H2A.X, un marcador que se enriquece cuando el DNA sufre daños, como un indicador para analizar la estabilidad genómica en las células ZNF274 KO y las dobles mutantes ZNF274/ATRX KO. Los WB muestran una acumulación mayor de γ H2A.X en las células ZNF274 KO con respecto al control. La acumulación es aún mayor en las mutantes ZNF274/ATRX KO (ZNF-Fig5D). Esto sugiere que, en efecto, la pérdida de H3K9me3 en los ZNFs explica, al menos en parte, la mayor acumulación de daño al DNA y el posible aumento de la inestabilidad genómica.

Para analizar si la pérdida de ATRX y ZNF274 tiene otros efectos además de incrementar la acumulación del daño al DNA, se analizó el patrón del ciclo celular por medio de experimentos de tinción con PI y BrdU en las mutantes sencillas ATRX KO2, ZNF274 KO y las dobles mutantes ATRX/ZNF274 KO. Notablemente, todas las mutantes, tanto las sencillas como las dobles, presentan un ligero defecto en la transición de la fase G1 a la fase S, con una mayor acumulación de células en la fase G1 en comparación con las células control (ZNF-Fig5E,F).

I.VI - Discusión

A través de un análisis epigenómico integrativo e imparcial, se logró identificar una nueva familia de blancos para la proteína ATRX: los TF de tipo zinc finger o ZNFs. Los resultados de esta investigación muestran que ATRX se une al exón 3' de aquellos ZNFs que poseen una cromatina atípica: están enriquecidos en H3K9me3, una marca de silenciamiento transcripcional, y a la vez presentan altos niveles de H3K36me3, una marca de elongación (ver modelo en la Fig. 1). La unión de ATRX es independiente del estatus transcripcional de dichos ZNFs y éstos presentan algunas características genéticas que los hacen únicos: suelen poseer un dominio represor KRAB, tienen un alto número de dominios zinc finger y tienen poca probabilidad de formar estructuras G4. ATRX interactúa físicamente con SETDB1/TRIM28/ZNF274 y las cuatro proteínas forman un complejo que se une a los genes ZNF. La eliminación de ATRX o ZNF274 tiene como consecuencia la reducción de la H3K9me3 en los exones 3' de los genes ZNFs. A nivel global, los niveles de H3K9me3 en las mayoría de regiones de cromatina

atípica son particularmente sensibles a la eliminación de ATRX. Finalmente, la eliminación de ATRX, ZNF274 o ambas proteínas tiene como consecuencia una mayor acumulación de daño a DNA, y una ligera desregulación del ciclo celular.

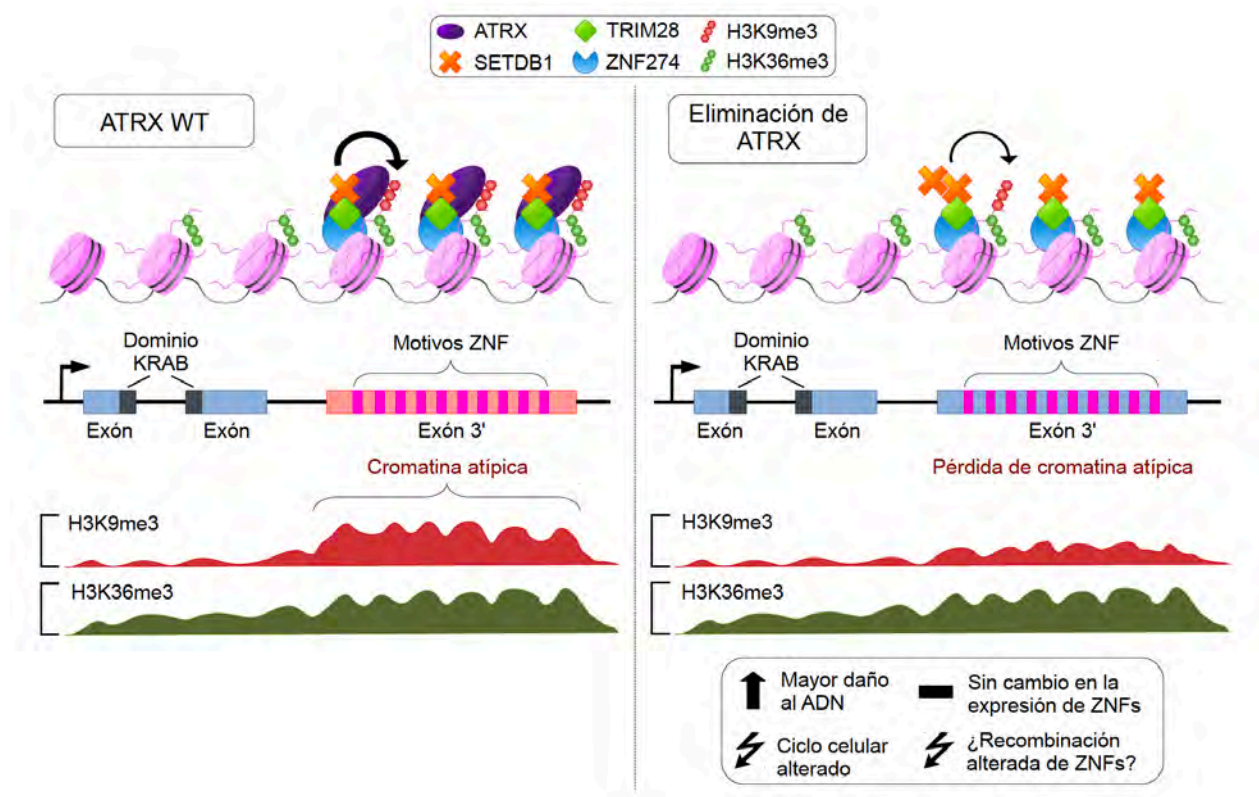


Figura 1. Modelo de regulación de los genes ZNF mediado por ATRX.

Izquierda: ATRX establece un dominio de cromatina atípica en el exón 3' de los genes ZNF, el cual contiene los motivos zinc finger. Derecha: Cuando ATRX se pierde, la acumulación de H3K9me3 se reduce y la cromatina atípica se pierde lo cual puede generar consecuencias en la estabilidad genómica de los ZNFs.

Aunque nuestros resultados muestran claramente que ATRX es importante para mantener los niveles de H3K9me3, todavía queda poco claro cómo es que ATRX mantiene dichos niveles. Basados en nuestras observaciones proponemos tres escenarios no excluyentes entre sí: 1) ATRX genera una estructura de la cromatina que facilita la función de metilación de SETDB1 2) ATRX se une a la H3K9me3 y bloquea la actividad de las de-metilinas de histonas 3) ATRX ayuda al re-establecimiento de H3K9me3 después de la transcripción o la replicación.

Funcionalmente, la pérdida de H3K9me3 en las células KO, podría afectar la estabilidad genómica de los ZNFs. Los genes ZNF son la familia de genes que más rápidamente se ha

expandido en el linaje de los primates y se sugiere que estas expansiones se dan gracias a duplicaciones génicas derivadas de eventos de recombinación homóloga (HR por sus siglas en inglés) (Tadepally et al. 2008). Esto sugiere que los eventos de recombinación entre ZNFs son frecuentes. Debido a que el exón 3' es el que suele contener los motivos de tipo zinc finger, estas regiones suelen tener una gran similitud de secuencia entre diversos ZNFs, haciéndolas regiones particularmente propensas a recombinar entre sí. Adicionalmente, reportes recientes han mostrado que las regiones enriquecidas en H3K36me3, como los exones 3' de los genes ZNF, tienden a reparar los cortes de DNA de doble cadena por medio de HR, en lugar del sistema de *Non-Homologous End Joining* que es el comúnmente utilizado (Pfister et al. 2014; Aymard et al. 2014). Estas evidencias sugieren que los exones 3' de los genes ZNF son potencialmente regiones altamente recombinogénicas. Debido a que se ha demostrado que el enriquecimiento de H3K9me3 puede actuar como una marca que impide la HR (Murray et al. 2012; Kalousi et al. 2015), sugerimos que ATRX se une a los ZNFs para evitar su recombinación y contribuir a la estabilidad genómica. Sin embargo, dicha hipótesis debe ser probada experimentalmente.

Como evidencia de la plausibilidad de la hipótesis propuesta, varios estudios recientes han demostrado que ATRX puede actuar como un represor de la recombinación en las regiones teloméricas (Clynes et al. 2015; Ramamoorthy & Smith 2015; Napier et al. 2015; He et al. 2015).

A pesar de la evidencia circunstancial que apoya la teoría de ATRX como un regulador de la recombinación de genes ZNF, pueden existir explicaciones alternativas que encajan con los resultados obtenidos. Por ejemplo, ATRX podría facilitar la replicación de los ZNFs. Estudios previos han mostrado que cuando se somete a células mutantes en ATRX a estrés replicativo, éstas tienen una mayor acumulación de horquillas de replicación estancadas que células WT (Leung et al. 2013; Clynes et al. 2014). Un fenotipo similar podría explicar la acumulación de daño a DNA y los defectos en el ciclo celular observados en las células ATRX y ZNF274 KO.

Estudios recientes han demostrado que, junto con DAXX, ATRX regula la deposición de H3.3 y el enriquecimiento de H3K9me3 en varias regiones repetidas en el genoma de ratón (Elsässer et al. 2015; Sadic et al. 2015; He et al. 2015; Voon & Wong 2016). Todavía queda por definir si estas regiones poseen las mismas características que las regiones de cromatina atípica de los genes ZNF o si se trata de dos mecanismos de regulación independientes.

En resumen, demostramos que ATRX regula los niveles de H3K9me3 en los exones 3' de los genes ZNF y otros loci que poseen una cromatina atípica. Esta función inesperada pone de relieve la compleja red de regulación génica orquestada por ATRX y podría ser importante para entender en el futuro algunas enfermedades que presentan alteraciones en ATRX.

I.VII - Perspectivas

Aún quedan varias preguntas abiertas que necesitan ser exploradas para confirmar el modelo propuesto. Por ejemplo, a pesar de que es claro que ZNF274 interactúa con ATRX y es necesario para que éste último sea reclutado a un subgrupo de ZNFs de clase I, ATRX es capaz de unirse, aunque con menor afinidad, al grupo restante de ZNFs que fueron analizados. Adicionalmente, las células doble mutantes de ZNF274 y ATRX presentan una mayor reducción de H3K9me3 en los ZNFs que las mutantes individuales, lo que sugiere que sus blancos no son del todo redundantes. Una posible explicación a estos fenómenos es que existe un mecanismo compensatorio, posiblemente otros TF de tipo ZNF, que pueden reclutar a ATRX a sus regiones blanco. Esta idea se refuerza por el hecho de que existen regiones de cromatina atípica que muestran enriquecimiento de ATRX/SETDB1/TRIM28 pero no de ZNF274 (datos no mostrados). Esto abre la posibilidad de que distintos ZNFs guíen al complejo ATRX/SETDB1/TRIM28 a distintas regiones blanco, de una forma que puede ser específica al tipo o estadio celular. En los ejemplos recientes en la literatura en los que se ha demostrado que ATRX se une a distintas regiones repetitivas en ratón no se ha podido identificar el TF que logra que ATRX reconozca dichas regiones, aunque se ha sugerido que se puede tratar de ZNFs (Elsässer et al. 2015; Sadic et al. 2015; He et al. 2015; Voon & Wong 2016). Identificar los TFs que reclutan a ATRX a regiones específicas de la cromatina seguramente será un área de intensa investigación en el futuro.

Una de las perspectivas más atractivas es explorar las implicaciones que los hallazgos de esta investigación pueda tener en entender la etiología de las diversas enfermedades en las que está mutado ATRX, en particular en cáncer. Aunque diversos estudios han puesto de manifiesto la relevancia funcional entre las mutaciones de ATRX en diferentes tipos de cáncer y el fenómeno de ALT (alargamiento alternativo de los telómeros) (Heaphy et al. 2011; de Wilde et al. 2012; Lovejoy et al. 2012; Episkopou et al. 2014; Napier et al. 2015; Clynes et al. 2015; Ramamoorthy & Smith 2015), se desconoce si otros blancos son también afectados. Es por ello que consideramos que analizar la estabilidad genómica de los genes ZNF en distintos contextos en los que ATRX se ve afectado podría ayudar a

entender algunos de los mecanismos que contribuyen a la oncogénesis mediada por ATRX.

Finalmente, será importante determinar en el futuro si ATRX efectivamente actúa como un supresor de HR en los genes ZNF. Aunque existe evidencia creciente que sugiere que ese es precisamente el rol de ATRX en la regulación de las regiones teloméricas a las cuales se une (Napier et al. 2015; Clynes et al. 2015; Ramamoorthy & Smith 2015), debe comprobarse experimentalmente que éste también sea el caso para los genes ZNF. Desafortunadamente, a diferencia de los telómeros, no hay estrategias experimentales diseñadas a la fecha que nos permitan medir de forma directa y específica la tasa de HR en los genes ZNF. Existen, sin embargo, sistemas que son capaces de medir las tasas globales de HR y que podrían ser adaptados para medir los niveles de recombinación de ZNFs específicos. Por ejemplo, se podría inducir un corte de doble cadena específicamente en un ZNF mediante CRISPR y luego analizar si éste es reparado por recombinación usando un plásmido con secuencias homólogas al ZNF y un gen reportero. Este sistema permitiría cuantificar la frecuencia de HR entre células WT y mutantes de ATRX. También se han desarrollado algunos sistemas basados en secuenciación para analizar *hotspots* de recombinación (Khil et al. 2012). Si la tasa de recombinación en los ZNFs cambia drásticamente con la mutación de ATRX, debería poder identificarse por medio de este sistema. Adicionalmente, aprovechando la creciente capacidad y el decremento en los costos de secuenciación masiva de DNA, se podría secuenciar el genoma con una gran cobertura para identificar las posibles amplificaciones, fusiones o eliminaciones resultado de eventos de recombinación, aún si éstos se presentan sólo en una pequeña proporción en la población. La ventaja de esta técnica es que podría ser usada para investigar la contribución de ATRX a la estabilidad genómica en distintos modelos y contextos, incluidos distintos modelos de cáncer.

En resumen, la investigación aquí presentada ha abierto un nuevo panorama en el entendimiento de los mecanismos de regulación mediados por ATRX. El entendimiento cada vez más profundo de las funciones de ATRX podría permitirnos encontrar métodos terapéuticos más eficientes y específicos para tratar algunos de los tipos de cáncer en los que ATRX se encuentra mutado.

I.VIII - Materiales y métodos

La descripción detallada de todos los métodos usados para los análisis bioinformáticos y experimentales de esta sección pueden ser encontrados en el Apéndice I en la sección de

"Materials and methods" del artículo "ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment". La versión on-line del artículo (PMID: 27029610) contiene además información detallada y relevante para los métodos, incluyendo secuencias de primers, cantidades y tipos de anticuerpos usados en los distintos ensayos, parámetros para los análisis de CHIP-seq y los resultados a todas las pruebas estadísticas realizadas en el artículo. Los datos que se usaron para generar la mayor parte de las gráficas del artículo pueden además ser descargados en el material suplementario del artículo y en el caso de los datos crudos de secuenciación, éstos pueden ser consultados en el portal de *Gene Expression Omnibus* con el número de acceso GSE70920.

Parte II: ATRX y la regulación del dominio α -globina

“It seems to me what is called for is an exquisite balance between two conflicting needs: the most skeptical scrutiny of all hypotheses that are served up to us and at the same time a great openness to new ideas”

- Carl Sagan, *The Burden of Skepticism*

II.I - Introducción

Debido a su relevancia médica, se han estudiado los roles de ATRX como regulador del desarrollo neuronal y como regulador de los genes α -globina. Diversas investigaciones han puesto de manifiesto varios mecanismos a través de los cuáles ATRX regula la estructura de la cromatina en el contexto neuronal (Bérubé et al. 2002; Bérubé et al. 2005; Seah et al. 2008; Kernohan et al. 2010; Watson et al. 2013; Kernohan et al. 2014; Levy et al. 2015). Sin embargo, el rol que ATRX juega en la regulación de los genes α -globina no se ha estudiado a fondo y los reportes disponibles (Law et al. 2010; Ratnakumar et al. 2012) no ofrecen un mecanismo incluyente y coherente entre sí (ver discusión más abajo). Es por ello que consideramos necesario hacer uso de diferentes modelos para entender la función de ATRX en el contexto de la regulación eritroide.

II.I.I - El dominio α -globina como un caso de estudio de la regulación epigenética

Los loci β - y α -globina han sido estudiados como modelos paradigmáticos de la regulación epigenética. Debido a que los genes globina presentan una expresión diferencial a lo largo del desarrollo, y a que son expresados de forma específica en células eritroides maduras, han constituido un modelo invaluable para entender la regulación transcripcional durante la diferenciación y el desarrollo eritroide (Guerrero et al. 2007; Kiefer et al. 2008). Varios de los elementos claves en la regulación génica fueron descubiertos o estudiados inicialmente en los locus globina. Algunos de estos incluyen a los enhancers, los *insulators*, los locus control regions, conocidos hoy en día como "*super enhancers*", y los dominios génicos (Guerrero et al. 2007; Kiefer et al. 2008).

Diversos organismos modelo han sido utilizados en el estudio de la expresión diferencial de los genes globina. Entre ellos destaca el uso de *Gallus gallus*, el pollo doméstico, debido, entre otras razones, a su facilidad de crianza, su versatilidad, su desarrollo exógeno, y a que los eritrocitos son nucleados lo que permite el estudio de los mecanismos de regulación epigenética aún en etapas muy tardías del desarrollo lo que es imposible en modelos mamíferos.

Nuestro laboratorio ha trabajado durante muchos años en el estudio de los diversos mecanismos que regulan la expresión diferencial del locus α -globina en pollo.

El locus α -globina en pollo se compone principalmente de tres genes que presentan patrones de expresión diferenciales: el gen temprano π , y los genes tardíos αD y αA

(Guerrero et al. 2007). Dichos patrones son regulados, establecidos y mantenidos por diversos mecanismos epigenéticos durante el desarrollo del pollo y la diferenciación eritroide (Guerrero et al. 2007; Rincón-Arano et al. 2009; Valdes-Quezada et al. 2013; Arriaga-Canon et al. 2014). El laboratorio ha establecido y estudiado principalmente dos modelos para entender la expresión diferencial de los genes α -globina. Estos incluyen un modelo *in vivo* de desarrollo usando eritrocitos de pollo en distintas etapas de desarrollo, y un modelo *in vitro* de diferenciación en el que usamos líneas celulares de precursores eritroides de humano o pollo que se pueden diferenciar a células más maduras. Ambos modelos requieren de la expresión diferencial de los genes α -globina y nos permiten estudiar diferentes mecanismos regulatorios: en el modelo de diferenciación podemos analizar el cambio de expresión del gen π , que se expresa sólo durante etapas tempranas en el desarrollo, a la expresión de los genes α^D y α^A , que se expresan durante etapas más tardías (Rincón-Arano et al. 2009). En el modelo de diferenciación celular en cultivo, estudiamos la sobre-expresión de los genes tardíos tras inducir las células a diferenciarse (Valdes-Quezada et al. 2013). Nuestros estudios han mostrado que cuando los diferentes mecanismos epigenéticos que regulan la expresión del dominio α -globina son alterados, los genes α -globina no pueden expresarse (Rincón-Arano et al. 2009; Valdes-Quezada et al. 2013; Arriaga-Canon et al. 2014). Por lo tanto, planteamos que en un contexto normal, la expresión diferencial de los genes α -globina está regulada por ATRX. En contextos en los que ATRX está alterado, se genera un entorno epigenético anormal que altera la expresión de los genes α -globina.

II.I.II - ATRX como regulador del dominio α -globina: mecanismos y suposiciones

A pesar de que el vínculo entre las mutaciones de ATRX y la des-regulación de los genes α -globina se conoce desde 1991, hasta la fecha sólo se han publicado dos reportes que intentan explicar los mecanismos por los cuáles ATRX regula al locus α -globina (Law et al. 2010; Ratnakumar et al. 2012). Por un lado, el grupo de Richard Gibbons realizó un análisis de ChIP-seq de ATRX en proeritroblastos (precursores eritroides) humanos (el primer y único ChIP-seq de ATRX en humano hasta antes de nuestros estudios) con el fin de determinar la localización de ATRX en el genoma y en particular en el locus α -globina (Law et al. 2010). A través de diversos análisis, Gibbons y su equipo encontraron que ATRX se une principalmente a repetidas en tándem ricas en G que son propensas a formar estructuras secundarias conocidas como G-quadruplexes o estructuras G4. Una de esas repetidas se encuentra cerca del promotor de los genes globina adultos, sitio al que ATRX se une con gran afinidad según los resultados de ChIP-seq. Debido a que ya se sabe que dichas estructuras son refractarias a la transcripción, el grupo sugirió que

ATRX regula indirectamente la expresión de los genes α -globina al evitar la formación de estructuras G4 cerca de su promotor. Como prueba de ello, reportan que los pacientes con síndrome ATR-X que poseen mayor número de copias de la repetida en el locus α -globina presentan fenotipos más severos de α -talasemia (Law et al. 2010). Sin embargo dicha hipótesis no fue corroborada directamente.

El segundo reporte, a cargo del grupo de Emily Bernstein, encontró a través de un estudio bioquímico global, que ATRX se une con gran afinidad a la variante de histona macroH2A (mH2A) (Ratnakumar et al. 2012). mH2A es una variante de histona única en vertebrados que ha sido correlacionada con represión transcripcional y la generación de heterocromatina. La variante está particularmente enriquecida en el cromosoma X inactivo en mamíferos, aunque también está distribuida en otras regiones en los autosomas (Gamble & Kraus 2010). El grupo de la doctora Bernstein mostró que ATRX interactúa directamente con mH2A. Más aún, demostró que ATRX es un regulador negativo de la incorporación de mH2A a la cromatina. De forma inesperada, cuando ATRX fue silenciado por medio de shRNAs en la línea humana eritroleucémica K562 existe una sobreacumulación de mH2A a lo largo de todo el dominio α -globina, causando el silenciamiento de los genes α -globina. Estos resultados sugerirían que ATRX regula de forma global la deposición de mH2A en todo el locus α -globina, evitando así la formación de heterocromatina en todo el locus (Ratnakumar et al. 2012).

Los resultados del grupo de Bernstein contrastan con el mecanismo propuesto por el grupo de Gibbons, pues en tanto Gibbons sugiere que la regulación de ATRX en el locus es puntual y los genes globina se ven afectados de forma circunstancial, Bernstein encuentra un mecanismo global y dirigido específicamente al locus. Desafortunadamente, los modelos usados en ambos reportes son muy distintos entre sí y podrían no ser comparables. En tanto Gibbons utilizó precursores eritroides humanos primarios, Bernstein optó por el uso de una línea celular derivada de una paciente con leucemia (Lozzio & Lozzio 1975). Aunque esta línea celular en particular es un buen modelo de diferenciación eritroide (Cioe et al. 1981; Villeval et al. 1983), no se estudió el papel de ATRX en dicho proceso. Una de las principales debilidades del reporte de Bernstein es que no demostró si ATRX se une directamente al locus o si afecta indirectamente a la deposición de mH2A. Por otro lado el mecanismo propuesto por Gibbons no puede ser extendido a otras especies, pues algunas de ellas, como el pollo, no presentan las repeticiones en tándem que son características del locus humano. Sin embargo, el porcentaje de identidad entre la proteína humana y la de pollo (~81%) sugiere que muy probablemente su función está conservada entre los organismos. Es por ello que se requieren de estudios adicionales que determinen los mecanismos por los

cuáles ATRX regula a los genes α -globina, particularmente en el contexto de la diferenciación eritroide y el desarrollo embrionario.

II.II - Justificación y planteamiento del problema

La presencia de α -talasemia, la desregulación de la expresión de los genes α -globina, es uno de los principales síntomas de los pacientes que presentan el síndrome de ATR-X. Por ello, se ha sugerido que ATRX está relacionado con la regulación de los genes α -globina, sin embargo los mecanismos a través de los cuáles se da esta regulación quedan poco claros (Steensma et al. 2005; Higgs et al. 2005; Gibbons 2006). A pesar de que se sabe que las mutaciones somáticas del gen ATRX son comunes en varios tipos de cáncer (Watson et al. 2015), los pacientes que presentan el síndrome de ATR-X no tienen una mayor probabilidad de desarrollar tumores malignos (Gibbons 2006). Esto sugiere que ATRX regula de forma única la expresión de los genes α -globina, posiblemente a lo largo del desarrollo o la diferenciación eritroide. A pesar de que se han hecho algunas investigaciones en modelos eritroides humanos que han mostrado que ATRX puede regular al dominio α -globina (Law et al. 2010; Ratnakumar et al. 2012), no se ha estudiado sistemáticamente la contribución de ATRX en la regulación de la expresión diferencial de los genes α -globina a lo largo del desarrollo y la diferenciación. Es por ello que planteamos la necesidad del uso de modelos humanos y de pollo para entender los mecanismos a través de los cuáles ATRX regula los genes α -globina en el desarrollo y la diferenciación.

II.III - Hipótesis

ATRX regula directamente la estructura de la cromatina del dominio α -globina a lo largo del desarrollo y la diferenciación.

II.IV - Objetivos

II.IV.I - Objetivo general

Estudiar los mecanismos a nivel epigenético a través de los cuales ATRX participa en la estructuración de la cromatina y la regulación de la expresión de los genes α -globina usando modelos de pollo y humano como herramientas de estudio.

II.IV.II - Objetivos particulares

- 1) Caracterizar la expresión de ATRX y su relación con los genes α -globina a lo largo del desarrollo embrionario de pollo
- 2) Analizar la unión de ATRX en el locus α -globina a lo largo del desarrollo embrionario de pollo por medio de ChIP y ChIP-seq
- 3) Alterar la expresión de ATRX en pollo para analizar sus consecuencias en los genes α -globina
- 4) Analizar la unión de ATRX en el dominio α -globina de células humanas.
- 5) Alterar la expresión de ATRX en humano para analizar sus consecuencias en los genes α -globina

II.V - Resultados

II.V.I - Caracterizar la Expresión de ATRX y las α -globinas a lo largo del desarrollo del embrión de pollo

Para determinar el patrón de expresión de ATRX, se obtuvo RNA de eritrocitos primarios de pollo en distintas etapas del desarrollo (ver métodos, sección II.VII). Se analizó por RT-qPCR el patrón de expresión de las α -globina, así como el de ATRX. Debido a que CTCF es también un importante regulador del dominio según resultados previos de nuestro laboratorio (Valdes-Quezada et al. 2013) también se midieron sus niveles de expresión.

Los resultados muestran que la expresión de las α -globinas sigue el patrón establecido previamente (Guerrero et al. 2007; Rincón-Arano et al. 2009): Mientras que el gen temprano π se expresa en la etapa más temprana del desarrollo (5 días), su expresión disminuye a los 7 días, que es cuando se genera el *switch* de expresión entre la globina temprana y las globinas tardías. En cambio los genes tempranos αD y αA poseen tasas relativamente bajas de expresión en las etapas tempranas (5 y 7 días), y su expresión aumenta en etapas tardías (10 y 14 días) una vez que el gen π ha dejado de expresarse (Fig. 2A). Por otro lado, CTCF, una proteína multifuncional con capacidad de actuar como *insulator* y que es crítico para determinar los dominios de cromatina en el locus α -

globina (Valdes-Quezada et al. 2013) tiene una expresión estable a lo largo del desarrollo (Fig. 2B). Curiosamente, ATRX tiene un patrón de expresión diferencial a lo largo del desarrollo de los eritrocitos de pollo. En tanto su expresión se mantiene alta en la etapa más temprana (5 días) y en las etapas tardías (10 días), su expresión disminuye en la etapa del *switch* de expresión. Esto podría sugerir que ATRX puede participar en el proceso de establecer los estados transcripcionales del locus α -globina (tanto el estado temprano en el que π es expresado, como el estado tardío en el que αD y αA son expresados). Sin embargo es necesario realizar experimentos más directos para confirmar si la expresión de ATRX correlaciona con su capacidad de unión al locus α -globina o si éstos son dos eventos independientes.

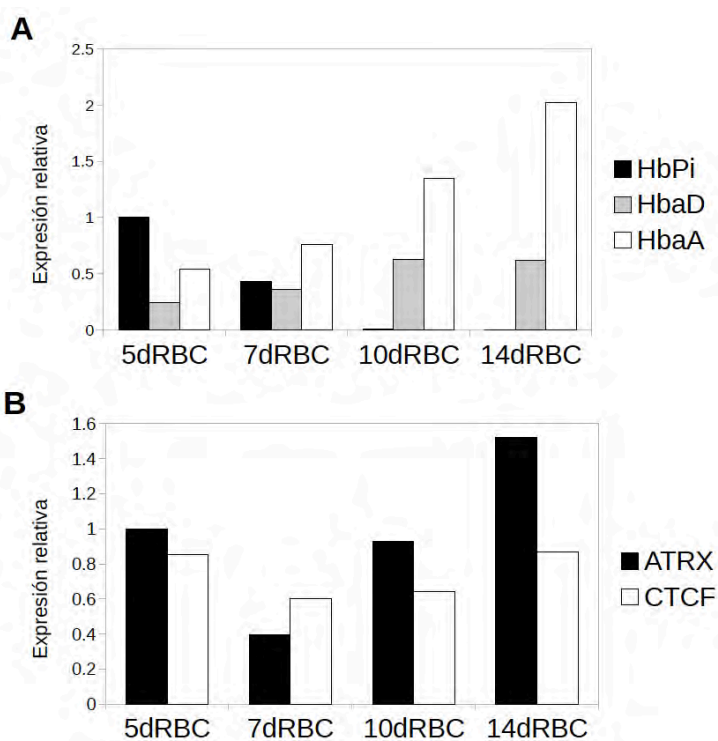


Figura 2. Expresión de ATRX, CTCF y las α -globina en el desarrollo de eritrocitos de pollo.

(A) Cuantificación por RT-qPCR de los tres genes α -globina: π (HbPi), αD (HbaD) y αA (HbaA) en diferentes etapas del desarrollo de eritrocitos de pollo: 5 días (5dRBC), 7 días (7dRBC), 10 días (10dRBC) y 14 días (14dRBC) La expresión se normalizó con respecto a la expresión de β -actina.

(B) Expresión de ATRX y CTCF en diferentes etapas del desarrollo de eritrocitos de pollo. La expresión se normalizó con respecto a la expresión de β -actina.

II.V.II - Análisis de la unión global de ATRX en el genoma de pollo y su relación con el locus α -globina

Como un primer acercamiento para entender si ATRX se une de forma directa al locus α -globina, se realizaron ensayos de ChIP-qPCR de ATRX en eritrocitos de 10 días de desarrollo (10dRBC) utilizando un protocolo de ChIP estandarizado en el laboratorio especialmente para poder analizar la unión de ATRX (ver sección III.II). Debido a que no hay anticuerpos diseñados específicamente para reconocer a ATRX en pollo y a que la proteína de pollo es muy similar a la de humano (comparten el 81% de identidad) se usó

un anticuerpo contra humano que ya había sido validado en células humanas (ver sección I).

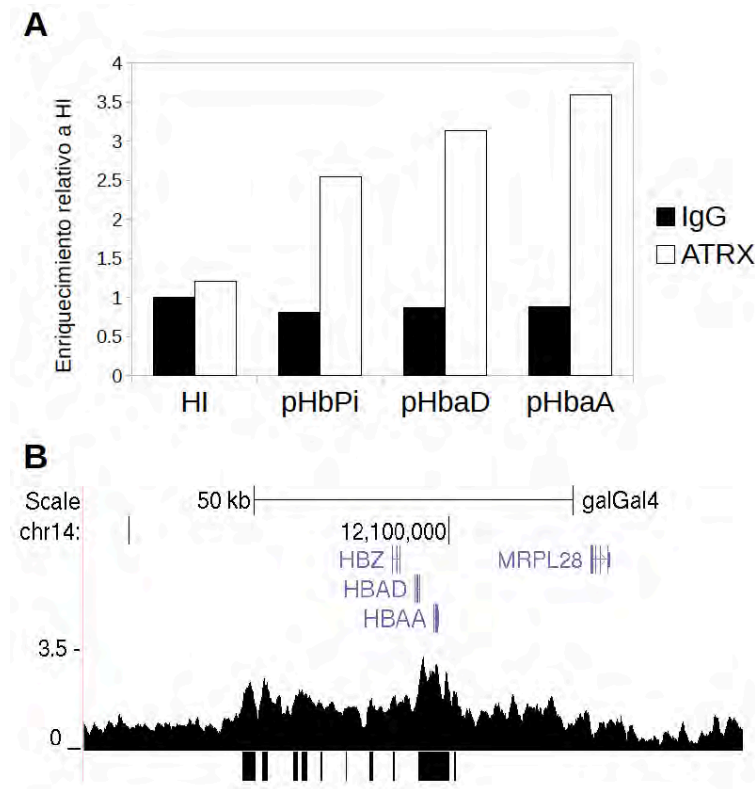


Figura 3. Enriquecimiento de ATRX en el locus α -globina en eritrocitos de 10 días de desarrollo de pollo.

(A) ChIP-qPCR de ATRX en eritrocitos de 10 días de pollo. Se utilizaron primers de los promotores de los tres genes α -globina. Como control negativo se usó un anticuerpo inespecífico y primers contra una región de heterocromatina del locus β -globina.

(B) Enriquecimiento de ATRX en el locus α -globina en eritrocitos de 10 días por ChIP-seq. En la parte superior se muestran los genes α -globina en gris. En negro se muestra el enriquecimiento de ATRX. Las regiones que están significativamente enriquecidas (p -value < 0.01) se muestran como cajas negras debajo del enriquecimiento.

Como se puede observar en la figura 3A, ATRX está enriquecido en los promotores de los tres genes α -globina en eritrocitos de 10 días. Estos datos sugieren que efectivamente ATRX regula de forma directa a los genes α -globina. Sin embargo, fue inesperado observar que ATRX se une a todos los genes, tanto el gen π , que permanece silenciado en esta etapa (Fig. 2A), como a los genes α D y α A que son expresados abundantemente. Este resultado implica que ATRX no regula de forma directa la expresión de los genes globina y que su presencia en los promotores de éstos tiene un rol adicional a la expresión o la represión. Adicionalmente, ATRX podría estar promoviendo la incorporación de H3.3 a lo largo de todo el dominio para establecerlo como un dominio de cromatina abierta o dinámica.

Con la finalidad de analizar el patrón preciso de ATRX a lo largo de todo el locus, se realizaron ensayos de ChIP-seq en eritrocitos de 10 días siguiendo el protocolo que se estableció y validó en el laboratorio. Sorprendentemente, se encontró que ATRX está enriquecido a lo largo de todo el locus α -globina y forma un gran dominio que abarca los tres genes y una gran zona de aproximadamente 30Kb río arriba de los mismos (Fig.

3B). Estos datos sugieren que ATRX podría jugar un papel estructural en el establecimiento del dominio del locus α -globina.

Desafortunadamente, prácticamente no existen datos de ChIP-seq para eritrocitos de pollo. El único set de datos que se ha realizado en eritrocitos corresponde a un ChIP-seq de CTCF realizado en nuestro laboratorio (Martin et al. 2011). A diferencia de ATRX, CTCF tiene sitios de unión muy puntuales por lo que es difícil hacer una comparación válida entre ambos. Al menos en el locus α -globina, no existe una evidente correlación entre los dos factores (datos no mostrados).

Otro problema deriva de la dificultad técnica de estudiar ATRX sin un anticuerpo específico que reconozca la proteína de pollo. A pesar de que el experimento de ChIP-seq en eritrocitos de pollo fue exitoso, análisis bioinformáticos revelan que alrededor del 99.98% de las lecturas del ChIP son ruido de fondo. Eso hace que la identificación de sitios de unión sutiles sea complicada. De hecho, el locus α -globina es la región más enriquecida en ATRX del genoma de pollo en estas células. Análisis de ontología de genes con los genes que tienen sitios de unión a ATRX muestran que la categoría más enriquecida son genes involucrados en la producción de hemoglobina, es decir, los genes α -globina.

Una limitación del modelo de eritrocitos de pollo es que es complicado realizar manipulaciones genéticas en los mismos. Con la finalidad de entender la importancia de ATRX en la expresión de los genes α -globina en el contexto eritroide, se intentaron hacer KD con shRNAs contra ATRX en eritrocitos primarios de 10 días, sin embargo los experimentos no fueron exitosos y la expresión de ATRX se mantuvo estable en las células transfectadas con el shRNAs vs las células control (datos no mostrados).

II.V.III - La sobre-expresión de ATRX aumenta la expresión de los genes α -globina adultos en pollo

Con el fin de estudiar de forma funcional el rol de ATRX en la expresión de los genes α -globina se utilizó un modelo alternativo en cultivo. El laboratorio posee una línea celular de eritroblastos de pollo llamada HD3. Esta línea celular es un buen modelo para estudiar la diferenciación eritroide ya que se derivó de precursores arrestados por un virus sensible a temperatura y a ciertos agentes químicos antes de la diferenciación terminal de los eritrocitos. Por lo tanto, la línea se puede inducir a un estadio de diferenciación maduro.

En su contexto normal (sin inducir) el locus α -globina permanece silenciado. Cuando las células son inducidas a diferenciarse, los genes adultos del locus se re-expresan.

Debido a que la transfección de ATRX es muy ineficiente (ver apartado III.I), se seleccionaron clonas estables (ver métodos, sección II.VII) de HD3 que sobre-expresan un plásmido que contiene la versión completa del cDNA de ATRX de humano. Tras confirmar la sobre-expresión de ATRX en una de ellas por medio de western blot, comparada con clonas control, se realizaron ensayos de RT-qPCR de los genes α -globina (Fig. 4A). De forma interesante, la clona que sobre-expresa ATRX (A5) también muestra la sobre-expresión de los genes α -globina adultos (Fig. 4B)

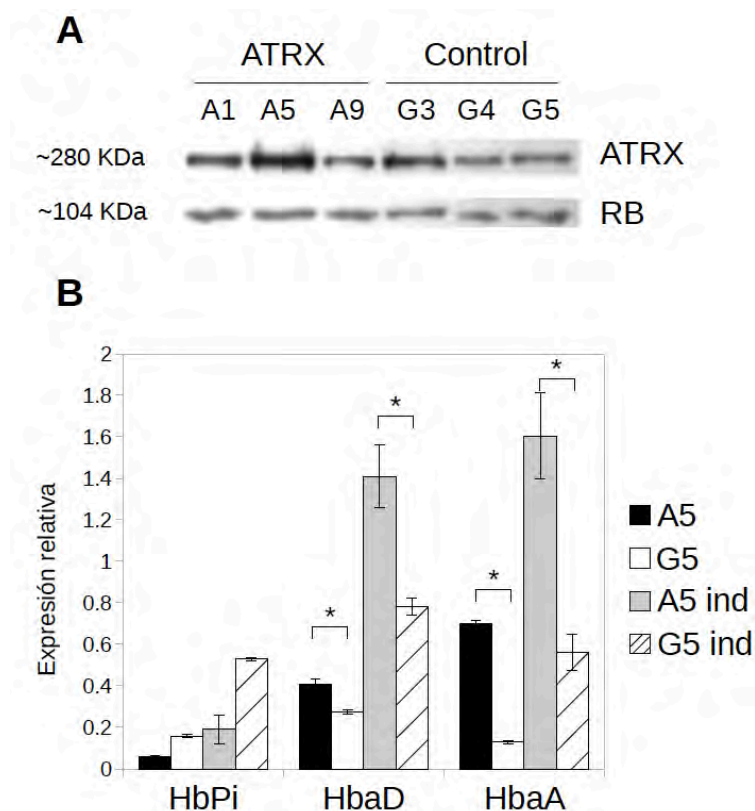


Figura 4. Efecto de la sobre-expresión de ATRX en los genes α -globina en pollo.

(A) Western blot de las clonas de HD3 que sobre-expresan el cDNA humano de ATRX y clonas control. RB se utilizó como control de carga. La clona A5 muestra una clara sobre-expresión de ATRX

(B) RT-qPCR de los genes α -globina en la clona A5 y la clona control G5 en condiciones normales y tras ser inducida a diferenciarse (A5 ind y G5 ind). Los asteriscos marcan las comparaciones estadísticamente significativas entre la clona A5 y la clona control (p-value < 0.05).

Para determinar si ATRX afecta también la inducción de los genes α -globina en el proceso de diferenciación, tanto la clona A5 como la clona control fueron inducidas a diferenciarse por medio de calor (incubación a 42°C). Tal como se esperaba, todas las clonas aumentaron los niveles de expresión de los genes globina adulto α -globina. Sin embargo la clona que sobre-expresa a ATRX logra mayores niveles de sobre-expresión de los genes globina adulto comparada con la clona control (Fig. 4B). De forma curiosa, sucede el efecto contrario con la globina temprana π : la clona que sobre-expresa ATRX tiene menores niveles de inducción de la globina temprana comparada con la clona

control. Estos resultados sugieren que ATRX, en el contexto de inducción de las α -globina adultas, es un regulador positivo de su expresión. Queda pendiente determinar si este efecto es directo, o indirecto. Sin embargo estos resultados sugieren que ATRX es un actor importante en el control de la expresión de los genes α -globina en este modelo de pollo.

II.V.IV - ATRX no se une al locus α -globina en células humanas

Para complementar los estudios realizados en pollo, se utilizó la célula eritroleucémica humana K562 como modelo de estudio para entender la regulación de ATRX en el locus α -globina. Previamente el laboratorio de nuestra colaboradora, la Dra. Bernstein demostró que ATRX regula negativamente la deposición de la variante de histona mH2A en el locus α -globina. Eliminar ATRX a través de un shRNA provoca la disminución de la expresión de las α -globina en K562, y aumenta la deposición de mH2A en todo el locus (Ratnakumar et al. 2012). Por ello se decidió analizar el patrón de unión de ATRX en el locus α -globina de humano, usando las células K562 como modelo.

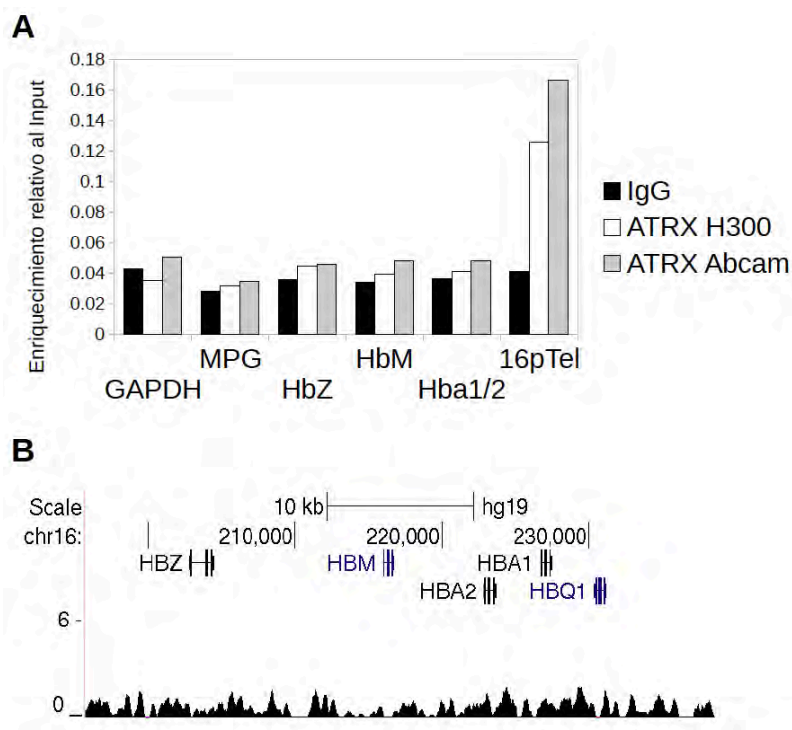


Figura 5. ATRX no está enriquecido en el locus α -globina de K562.

(A) ChIP-qPCR de ATRX en K562 usando distintos primers del locus α -globina en K562. IgG se usó como control negativo. La región de GAPDH y MPG se usaron como controles negativos de la qPCR. 16pTel es una región telomérica que se usa como control positivo. Se usaron dos anticuerpos de ATRX distintos para el experimento

(B) Enriquecimiento de ATRX en el locus α -globina por ChIP-seq. No existe ninguna región significativamente enriquecida en el locus.

Al igual que los análisis en pollo, se comenzó con una aproximación basada en ChIP-qPCR de ATRX. Para este análisis se utilizaron primers que ya habían sido usados para demostrar que ATRX se une a algunas regiones del locus en eritroblastos primarios

humanos (Law et al. 2010). Sin embargo, a pesar de que se utilizó un protocolo de ChIP-qPCR optimizado para ATRX (ver sección III.II) no se pudo observar ningún enriquecimiento significativo de ATRX en el locus α -globina con ninguno de los primers usados (Fig. 5A).

Debido a que la técnica de ChIP-qPCR está limitada sólo a las regiones amplificadas por los primers, se analizaron los datos de ChIP-seq en células K562 que se habían generado previamente (ver sección I). Al estudiar el enriquecimiento de ATRX en la región del locus α -globina, se comprobó que efectivamente ATRX no se une a ninguna región del locus (Fig. 5B). Estos resultados fueron corroborados nuevamente por ChIP-qPCR usando distintos primers en diversas regiones del locus (datos no mostrados). Debido a que el experimento fue realizado en múltiples ocasiones, con dos anticuerpos independientes, y con controles positivos y negativos adecuados, concluimos que ATRX no se une al locus α -globina en la línea K562. Estos resultados sugieren que la regulación que ATRX ejerce a través de la deposición de mH2A es un mecanismo indirecto. Todavía queda por definir precisamente cómo es que ATRX puede influenciar la deposición de la variante de histona en el locus sin unirse a él.

II.IV.V La pérdida de ATRX afecta la diferenciación de células eritroides humanas a través de un mecanismo de regulación indirecta.

Las células K562, al igual que las células HD3, se pueden diferenciar con el uso de diversos fármacos como la hemina (Cioe et al. 1981). Debido a que ATRX no se une al locus α -globina en condiciones normales, decidimos investigar si estaba involucrado en el proceso de diferenciación tal como lo observamos en las células de pollo. Inicialmente se intentó sobre-expresar a ATRX para ver si éste era capaz de inducir la expresión de los genes globina como ya lo se había observado en las células HD3 de pollo (Fig. 4B). Sin embargo, estos experimentos fueron infructuosos ya que no se logró sobre-expresar a ATRX con ningún método de transformación probado (ver detalles en sección III.I). Como alternativa, se usaron shRNAs para disminuir la expresión de ATRX y analizar su efecto en los genes α -globina durante el proceso de inducción. Posteriormente, se analizó el patrón de expresión de todos los genes α -globina en las células KD y las células control con y sin hemina.

El locus de humano, a diferencia del de pollo, cuenta con 5 genes α -globina. HbZ es un gen con expresión temprana que hasta hace poco tiempo se consideraba un pseudogen pero que recientemente se encontró que es un gen funcional. HbM es el homólogo a la

globina temprana π . Hba 1 y Hba 2 son homólogos a αA . La secuencia entre ellos es tan similar que es imposible distinguirlos por métodos basados en PCR. HbQ es un parálogo de reciente evolución que se asemeja a αD pero del cual no se tiene evidencia clara si es funcional o no (Higgs et al. 2005; Garrick et al. 2008). De todos los genes, sólo HbZ y Hba1/2 tuvieron niveles de expresión detectables. Tal como se había demostrado previamente, el KD de ATRX reduce la expresión de los genes α -globina en células K562 sin inducir (Fig. 6A). Sorprendentemente, el KD de ATRX tuvo un efecto severo en la capacidad de diferenciación de las células. Lejos de aumentar su expresión, los genes α -globina redujeron sus niveles de transcrito en las células KD para ATRX tras el tratamiento de hemina en comparación con las células no tratadas y las células control. Estos datos sugieren que ATRX es indispensable para la sobre-expresión de los genes α -globina durante el proceso de diferenciación.

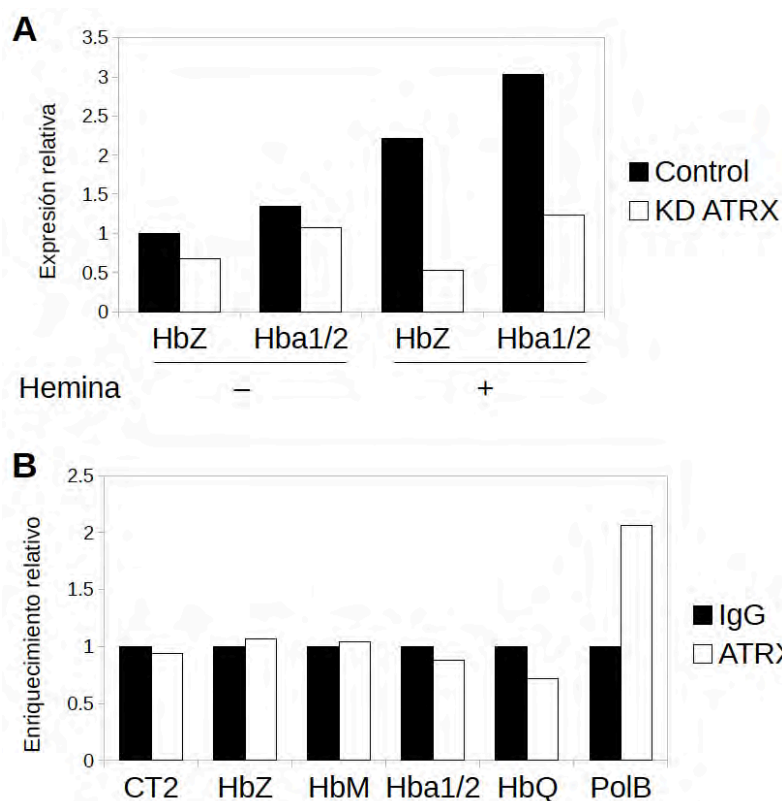


Figura 6. ATRX regula a los genes α -globina en la diferenciación de K562 por un mecanismo indirecto.

(A) RT-qPCR de los genes α -globina en líneas de células K562 KD para ATRX y líneas con un shRNA control. Las células fueron tratadas con 30 μ M de hemina por 4 días para diferenciarlas y aumentar la expresión del locus α -globina. Como control se analizaron células sin tratamiento.

(B) ChIP-qPCR de ATRX en células K562 tratadas con hemina. No se observa enriquecimiento de ATRX en ninguno de los genes α -globina a pesar del tratamiento. CT2 es una región control negativa para el enriquecimiento de ATRX. PolB es una región control positiva enriquecida en ATRX.

Para corroborar esta hipótesis, se realizaron experimentos de ChIP-qPCR de ATRX en el locus α -globina en células K562 tratadas con hemina. Sorprendentemente, no se encontró enriquecimiento de ATRX en ninguno de los genes α -globina (Fig. 6B). Estos datos sugieren fuertemente que, a pesar de que ATRX es necesario para la sobre-expresión de los genes α -globina en el proceso de inducción, la regulación se da a través de un mecanismo indirecto que no involucra la unión de ATRX al promotor de los genes.

II.VI – Discusión y perspectivas

A través de diversos experimentos en modelos de desarrollo y diferenciación, tanto en humano como en ratón, hemos aportado suficiente evidencia para concluir que ATRX está involucrado en la regulación de los genes α -globina. Sin embargo, dicha regulación se da a través de mecanismos más complejos de lo que se había previsto previamente y nuestros datos han abierto numerosas incógnitas y posibles líneas de investigación.

Una de las posibilidades más atractivas es determinar la función del dominio de unión a ATRX que se forma en los eritrocitos de pollo de 10 días. Debido a la poca información existente en este modelo, sería conveniente realizar una caracterización epigenómica completa a través de ChIP-seq de diversas marcas de cromatina en los principales estadios del desarrollo. A pesar de que extensos análisis funcionales realizados en diversos laboratorios, incluido el nuestro, han logrado caracterizar a detalle los diversos elementos funcionales del locus α -globina de pollo (Guerrero et al. 2007), aún no existe ningún análisis que haya estudiado las marcas de cromatina a un nivel epigenómico. Determinar la dinámica de la cromatina a lo largo del locus en el desarrollo sería una importante avenida de desarrollo futuro. Debido a que diversos estudios, incluido el nuestro (ver sección II), muestran que ATRX es un importante regulador de la deposición de la H3K9me3, se podría comenzar mapeando dicha marca en eritrocitos de 5, 7, 10 y 14 días. Adicionalmente sería relevante realizar ChIP-seqs de ATRX en eritrocitos de 5, 7 y 14 días con la finalidad de analizar la dinámica de unión de ATRX a lo largo del desarrollo.

Otra avenida de desarrollo importante sería implementar las diversas tecnologías basadas en CRISPR/Cas9 en el modelo de células HD3. Por ejemplo, existen sistemas basados en versiones inactivas de Cas9 que permiten sobre-expresar genes de forma endógena. Dichos sistemas podrían ser utilizados para sobre-expresar el gen de ATRX endógeno en células HD3. Esta técnica no sólo evitaría los problemas asociados a la transformación de plásmidos que contienen la secuencia de ATRX (ver sección III.I), sino que además sobre-expresaría el gen de pollo lo que tendría potencialmente un efecto más dramático que el sistema heterólogo que se usó en nuestros experimentos previos. Asimismo, se puede utilizar la tecnología para inactivar al gen endógeno y analizar el efecto de esta inactivación en la diferenciación.

Los estudios en humano, resultan extremadamente valiosos porque no solamente son complementarios a los análisis en pollo, sino que pueden mostrar mecanismos de regulación radicalmente diferentes. En tanto que en los eritrocitos de pollo se pudo de mostrar claramente que ATRX se une al locus, fue imposible encontrar cualquier clase de unión de ATRX en su contraparte humano. Irónicamente, ha sido demostrado que la eliminación de ATRX afecta la estructura de la cromatina en el locus, pues propicia la deposición de mH2A. Asimismo, nuestros datos muestran que tiene un efecto dramático en la capacidad de sobre-expresar los genes del locus durante la diferenciación. Descubrir el mecanismo a través del cual ATRX ejerce su efecto regulatorio en el locus α -globina en las células humanas queda como una de las grandes incógnitas a resolver en el futuro cercano. Una posibilidad muy atractiva es que ATRX esté regulando al locus α -globina a través de RNAs no codificantes largos (lncRNAs). Un estudio reciente del laboratorio de la Dra. Lee demostró que ATRX participa en la inactivación del cromosoma X a través de su unión a lncRNA. Curiosamente, si ATRX es afectado, la función del lncRNA, el establecimiento de la marca H3K27me3 y la deposición de mH2A en el cromosoma X inactivo se ven alterados (Sarma et al. 2014). Es posible que un mecanismo similar esté ocurriendo en el locus α -globina. Como un posible apoyo para esta hipótesis, nuestro grupo descubrió recientemente que el locus α -globina de pollo contiene un RNA no codificante largo que es esencial para la expresión de los genes α -globina en estadios adultos (Arriaga-Canon et al. 2014). Sería interesante explorar si ATRX se puede unir a dicho lncRNA y si el locus α -globina de humano también posee un transcrito similar.

Otra posibilidad interesante sería analizar los datos de ChIP-seq de ATRX para determinar si ATRX se une a los promotores de factores de transcripción esenciales para el desarrollo eritroide, tales como GATA-1 (Garrick et al. 2008) o realizar ensayos de ChIP-seq de ATRX en células K562 tratadas con Hemina y analizar los cambios de unión de ATRX, si los hubiera.

En resumen, nuestros datos muestran que ATRX es un regulador importante de la expresión de los genes α -globina pero que dicha regulación se da por mecanismos que todavía no comprendemos. Estudiar dichos mecanismos será una fuente importante para entender la función global de ATRX. Es posible que el locus α -globina sirva como un modelo excepcional para entender algunos de los mecanismos de regulación mediados por ATRX más complejos, lo cual brindará información invaluable para entender la función de ATRX en otros contextos.

II.VII - Materiales y métodos

II.VII.I Obtención y cuantificación de eritrocitos primarios de pollo

Los huevos pre-incubados de pollo son obtenidos a través de la compañía Charles River SPAFAS (SPF Premium pre-incubated eggs). Se obtuvieron huevos pre-incubados por 5, 7, 10 y 14 días. La sangre de los embriones es recolectada y resuspendida en 10ml de PBS + 50mM de EDTA como anticoagulante. El número relativo de eritrocitos se cuantifica lisando una muestra de 10 μ l en 1.5 ml de buffer de lisis (200mM NaCl, 0.5% triton X-100), clarificando el lisado por centrifugación (13,000 rpm por 2 min.) y leyendo el sobrenadante en un espectrómetro a una longitud de onda de $\lambda=412$ nm. Las ODs obtenidas se multiplican por el factor de dilución (150), el número de μ l tomados de la muestra (10) y el factor de conversión 8.56×10^6 .

II.VII.II Cultivo celular

Línea HD3

La línea celular de pollo, derivada de eritroblastos, HD3 es cultivada en medio DMEM suplementado con 8% de suero fetal bovino y 2% de suero de pollo y antibióticos (pen/strep Gibco). La línea crece en suspensión en condiciones estándar (37°C, atmósfera con 5% de CO₂).

Línea K562

La línea celular humana, derivada de células eritroleucémicas, K562 es cultivada en medio RPMI suplementado con 10% de suero fetal bovino y antibióticos-antimicóticos (anti-anti, Gibco). La línea crece en suspensión en condiciones estándar (37°C, atmósfera con 5% de CO₂).

II.VII.III RT-qPCR

Todas las extracciones de RNA se realizaron usando el kit RNeasy® mini de qiagen siguiendo las instrucciones del fabricante. El RNA fue cuantificado en un nanodrop y 1 μ g fue usado para preparar cDNA usando el kit de retro transcripción SuperScript II de invitrogen con oligo dT (# de cat. 18064014) según las instrucciones del fabricante. El cDNA fue diluido en un volumen final de 200 μ l. La qPCR se realizó mezclando 1 μ l de cDNA con primers específicos (conc. final 0.5 μ M) y el FastStart Universal SYBR Green

Master de Roche (# de cat. 04913914001) en placas de 96 pozos en una máquina de real time PCR de Bio-Rad.

II.VII.IV Generación de líneas estables de HD3.

Las células HD3 se mantuvieron en cultivo como ha sido descrito anteriormente. Aproximadamente 1×10^6 células fueron transfectadas con 1 μ g de plásmido GFP-ATRX lineal usando lipofectamina de acuerdo a las instrucciones del fabricante. 2 días post-transfección, 5×10^5 células fueron transferidas a una matriz de celulosa (Methocel®, Fluka) con 800 μ g/ml de geneticina (G418) como agente de selección. Las células fueron incubadas por alrededor de 20 días hasta que clonas aisladas fueran visibles a simple vista. Clonas individuales fueron recuperadas y crecidas en medio HD3 con selección (800 μ g/ml de geneticina). Posteriormente las clonas fueron evaluadas por citometría de flujo para analizar la expresión de GFP como un marcador de la expresión de ATRX y aquellas que presentaban niveles bajos fueron descartadas.

II.VII.V Primers

A continuación se lista la secuencia de los primers usados en esta sección de la tesis:

Nombre	Organismo	Aplicación	Blanco	Secuencia 5'-3'
RT-ATRX-1-gga-F	<i>G. gallus</i>	RT-qPCR	ATRX	TCTGTCAGCGATGAGCAATC
RT-ATRX-1-gga-R	<i>G. gallus</i>	RT-qPCR	ATRX	TGTCTTCAAGGGGCTGAATC
RT-CTCF-1-gga-F	<i>G. gallus</i>	RT-qPCR	CTCF	GCACCACCTCCCAGTAAG
RT-CTCF-1-gga-R	<i>G. gallus</i>	RT-qPCR	CTCF	GGTCTTCAACCTGAATGATTGC
RT-Pi-1-gga-F	<i>G. gallus</i>	RT-qPCR	π	CCATTGGGCTGGAATCACTG
RT-Pi-1-gga-R	<i>G. gallus</i>	RT-qPCR	π	CGAAGCTGAACTGAGCCTTG
RT-aD-1-gga-F	<i>G. gallus</i>	RT-qPCR	α D	CAACCTGCATGCCTACAAC
RT-aD-1-gga-R	<i>G. gallus</i>	RT-qPCR	α D	CCATGTGTACAGCCAGCA
RT-aA-1-gga-F	<i>G. gallus</i>	RT-qPCR	α A	GGCCATGCTGAGGAGTATG
RT-aA-1-gga-R	<i>G. gallus</i>	RT-qPCR	α A	CGTGTGACAGATCGAAGTGG
RT-BActin-gga-F	<i>G. gallus</i>	RT-qPCR	β -actina	CCCAGACATCAGGGTGTGATGG
RT-BActin-gga-R	<i>G. gallus</i>	RT-qPCR	β -actina	CCAGTTGGTGACAATACCGTGTTT
H1-1-gga-F	<i>G. gallus</i>	ChIP-qPCR	H1	GGAAGTAGTTGGCAAGGTCCTCT
H1-1-gga-R	<i>G. gallus</i>	ChIP-qPCR	H1	TCTTCTGCCCTGCCCGTAT
pPi-1-gga-F	<i>G. gallus</i>	ChIP-qPCR	π	GCACTGACCCTTAGCGAAATG
pPi-1-gga-R	<i>G. gallus</i>	ChIP-qPCR	π	CAGAGCAGTTGTACTGAGTG
paD-1-gga-F	<i>G. gallus</i>	ChIP-qPCR	α D	CGTGCGGATAAGATAAGGC
paD-1-gga-R	<i>G. gallus</i>	ChIP-qPCR	α D	GGCAATGTGTGATGGAGG
paA-1-gga-F	<i>G. gallus</i>	ChIP-qPCR	α A	GGCTGGGGTCCAGAATCTATG

paA-1-gga-R	<i>G. gallus</i>	ChIP-qPCR	α A	CTGAAGTCTGCTGTCAGCCTTAT
GAPDH-hs-ChIP-F	<i>H. sapiens</i>	ChIP-qPCR	GAPDH	CGGCTACTAGCGGTTTTACG
GAPDH-hs-ChIP-R	<i>H. sapiens</i>	ChIP-qPCR	GAPDH	GCTGCGGGCTCAATTTATAG
MPG-F	<i>H. sapiens</i>	ChIP-qPCR	MPG	TCATGCATTTAGCGGCTTTG
MPG-R	<i>H. sapiens</i>	ChIP-qPCR	MPG	TCATGCATTTAGCGGCTTTG
CT2-hs-F	<i>H. sapiens</i>	ChIP-qPCR	Intergenic	CCGTGCCTGAGGTTTCTCC
CT2-hs-R	<i>H. sapiens</i>	ChIP-qPCR	Intergenic	GGGAAGAGGGCACAGACG
16pTelomere-F	<i>H. sapiens</i>	ChIP-qPCR	Telómero	CCTCGCCTTGCCTTGGGAG
16pTelomere-R	<i>H. sapiens</i>	ChIP-qPCR	Telómero	CGGTTTCAGTGTGGAAATGGGAAAC
PolB-1-hs-F	<i>H. sapiens</i>	ChIP-qPCR	PolB	CCTTCTTAGTAACGGAAGTAATAGATTC
PolB-1-hs-R	<i>H. sapiens</i>	ChIP-qPCR	PolB	GGAAATTTTCAGATTTCTAGAATTACTGAG
pHbZ-F	<i>H. sapiens</i>	ChIP-qPCR	HbZ	CCACAGTTCCTGCCCTGACTCC
pHbZ-R	<i>H. sapiens</i>	ChIP-qPCR	HbZ	TCTCCCTCCTTCCCTCCAATAC
pHbM-F	<i>H. sapiens</i>	ChIP-qPCR	HbM	GCGGAGCTGCTGCTCAGGTC
pHbM-R	<i>H. sapiens</i>	ChIP-qPCR	HbM	AAACCCTCATCTCCACCTCCCTG
pHba1/2-F	<i>H. sapiens</i>	ChIP-qPCR	Hba1/2	CCCACAGACTCAGAGAGAACCCACCATG
pHba1/2-R	<i>H. sapiens</i>	ChIP-qPCR	Hba1/2	GGCCTTGACGTTGGTCTTGT
pHbQ-F	<i>H. sapiens</i>	ChIP-qPCR	HbQ	GCCACTGCACTACCCG
pHbQ-R	<i>H. sapiens</i>	ChIP-qPCR	HbQ	CCTCTGAATCGCCCTGG
RT-GAPDH-hs-F	<i>H. sapiens</i>	RT-qPCR	GAPDH	TTTGTCAAGCTCATTTCCTGG
RT-GAPDH-hs-R	<i>H. sapiens</i>	RT-qPCR	GAPDH	TGATGGTACATGACAAGGTGC
RT-HbZ-hs-F	<i>H. sapiens</i>	RT-qPCR	HbZ	GTGGGCCAAGATCTCCAC
RT-HbZ-hs-R	<i>H. sapiens</i>	RT-qPCR	HbZ	GAAGTGCGGGAAGTAGGTC
RT-HbM-hs-F	<i>H. sapiens</i>	RT-qPCR	HbM	CCAAATCGCGCAGGTCTG
RT-HbM-hs-R	<i>H. sapiens</i>	RT-qPCR	HbM	GGTACACCGTGAAGAGCCT
RT-Hba1/2-hs-F	<i>H. sapiens</i>	RT-qPCR	Hba1/2	CGACAAGACCAACGTC AAGG
RT-Hba1/2-hs-R	<i>H. sapiens</i>	RT-qPCR	Hba1/2	GGGAAGGACAGGAACATCCT
RT-HbQ-hs-F	<i>H. sapiens</i>	RT-qPCR	HbQ	CCCTGTGGAAGAAGCTGG
RT-HbQ-hs-R	<i>H. sapiens</i>	RT-qPCR	HbQ	GGTCCAGGTGGGAGAAGTAG

Parte III: Dificultades técnicas al trabajar con ATRX

“Humans are good, she knew, at discerning subtle patterns that are really there, but equally so at imagining them when they are altogether absent.”

- Carl Sagan, *Contact*

III.I – Dificultades en el uso de plásmidos de sobre-expresión que contienen el cDNA de ATRX

ATRX es una proteína de gran tamaño: Posee casi 2,500 aminoácidos y tiene una masa de alrededor de 280 KDa. Su cDNA mide aproximadamente 7.5 Kb y su gen se extiende a lo largo de más de 281 Kb. Estas son sólo algunas de las características que dificultan el trabajo con dicha proteína. Durante la realización de los experimentos de esta tesis, en múltiples ocasiones intentamos realizar experimentos de sobre-expresión de ATRX con resultados adversos. Es por ello que en esta sección se discutirán algunos de los problemas encontrados, sus soluciones, y sugerencias para facilitar esta clase de experimentos en el futuro.

III.I.I - Los vectores que contienen el cDNA de ATRX tienen inserciones frecuentes cuando se expanden en bacterias

Debido al gran tamaño de su cDNA, no resulta sencillo clonar a ATRX. Sin embargo, el grupo del Dr. Picketts logró clonar el cDNA de ATRX humano e insertarlo en un plásmido de sobre-expresión asociado al promotor CMV (Bérubé et al. 2008). Este plásmido se ha utilizado para otras publicaciones y a partir de él se han derivado otras versiones. Cuando nuestro laboratorio adquirió una de esas versiones (Newhart et al. 2012), como un regalo generoso del laboratorio de la Dra. Janicki, lo analizamos por medio de enzimas de restricción, como es rutina. Tras encontrar que el patrón observado no correspondía con el esperado, contactamos al grupo de la Dra. para hacérselo saber. Al mismo tiempo, el grupo de uno de nuestros colaboradores, el Dr. Dyer, nos informó que el plásmido que habían adquirido de parte del grupo del Dr. Picketts también tenía un patrón inesperado. Al secuenciar y analizar ambos plásmidos, nos percatamos que ambos contenían inserciones del mismo transposón bacteriano, el elemento IS10. Curiosamente, el sitio de inserción difería en ambos plásmidos, lo que sugería que la inserción de los elementos había ocurrido de forma independiente. Debido a que nuestro análisis reveló que el cDNA de ATRX contenía la secuencia reconocida por el elemento IS10 como sitio de inserción, decidimos investigar de forma metódica si este era un fenómeno aislado o un problema generalizado.

Nuestras investigaciones nos llevaron a publicar un artículo original en la revista SpringerPlus titulado: "The ATRX cDNA is prone to bacterial IS10 element insertions that alter its structure" que puede ser consultado en el Apéndice I de esta tesis. A

continuación se hace un breve resumen de las conclusiones principales de dicho artículo. Se hará referencia a las figuras de dicho artículo como IS10-Fig1, etc.

Como ya se había comentado, corroboramos tanto por PCR como por un análisis de restricción que los plásmidos obtenidos del laboratorio del Dr. Picketts y la Dra. Janicki contienen la inserción de un elemento IS10 en el exón 8 del cDNA de ATRX (IS10-Fig1A,B). Adicionalmente, se sub-clonó una versión libre del transposón que fue puesta a disposición de la comunidad a través de la plataforma addgene (addgene id: 45444) y que al momento de hoy ha sido requerida más de 50 veces.

Para analizar si la inserción es un evento aislado o algo frecuente, se transformó el cDNA de ATRX libre del transposón en dos distintas cepas de *E. coli*: DH5 α , una cepa usada rutinariamente en prácticamente todos los laboratorios de biología molecular del mundo, y Stb14, una cepa modificada genéticamente para inhibir los eventos de recombinación entre el genoma de la bacteria y los plásmidos que contiene. Sorprendentemente, el plásmido crecido en la cepa DH5 α tuvo una nueva inserción del elemento IS10, mientras que la el plásmido crecido en la cepa Stb14 se mantuvo libre de inserciones (IS10-Fig1C). Esto demuestra que la inserción *de novo* del elemento IS10 en el cDNA de ATRX, cuando éste se crece en cepas estándar de laboratorio, es un evento frecuente.

Según nuestros análisis, la inserción del elemento Is10 posiblemente generaría un codón de paro temprano que truncaría a ATRX. Para comprobar nuestra predicción, se sobre-expresaron las versiones con y sin transposón en células HEK293T. Dado que el plásmido contiene un reportero GFP, pudimos analizar la expresión del plásmido por medio de microscopía. La versión completa de ATRX tiene un patrón puntuado muy similar al que ya se ha reportado previamente para la proteína de forma endógena (IS10-Fig2A). En cambio la versión con la inserción IS10 presenta un patrón mucho más difuso, pero no totalmente aleatorio. Al analizar la acumulación de la proteína en el citoplasma y la cromatina a través de western blot, comprobamos que efectivamente el plásmido con transposón genera una versión de ATRX trunca que se acumula mayormente en el citoplasma, ha diferencia del ATRX intacto que se acumula mayormente en la cromatina. Sin embargo, todavía existe una pequeña fracción de ATRX truncado que es capaz de unirse a la cromatina y que potencialmente podría tener un efecto, presumiblemente distinto al de la proteína completa. Esta versión truncada no está presente en las células que sobre-expresan la versión completa de ATRX (IS10-Fig2B,C).

Este reporte técnico tiene como finalidad advertir a la comunidad de las posibles inserciones frecuentes del elemento IS10 cuando plásmidos que contienen el cDNA de ATRX crecen en bacterias. Debido a que el reportero GFP se encuentra antes del cDNA de ATRX, un simple análisis de sobre-expresión por microscopía no evidenciaría de forma inmediata la presencia del elemento IS10. Es por ello que en el artículo incluimos un protocolo para crecer plásmidos que contengan ATRX. El protocolo se puede resumir en los siguientes puntos clave: 1) Usar cepas que inhiban la recombinación como StbI4 2) crecer las bacterias a 30°C 3) analizar frecuentemente por secuenciación o análisis de restricción la integridad del plásmido.

Las descripciones detalladas de los métodos usados para esta publicación pueden ser encontradas en el texto del artículo en el Apéndice I. Ligas a datos suplementarios relevantes para los métodos pueden ser encontrados de forma gratuita en la versión on-line del artículo (PMID: 24834375).

III.I.II - Problemas para transfectar el cDNA de ATRX en la línea K562

Como parte de nuestros análisis para entender la función de ATRX en el contexto de los genes α -globina y los genes ZNF, y haciendo uso del plásmido libre de inserciones que desarrollamos, intentamos sobre-expresar ATRX por diferentes métodos en la línea eritroleucémica K562. Se intentaron dos métodos basados en lípidos: transfección por lipofectamina® y transfección por FuGene®. Se intentó un método por electroporación: El nucleofector® que ya ha sido optimizado para distintas líneas celulares, entre ellas K562. Se intentó un método químico: transfección por Fosfato de Calcio (CaCl). En todos los casos, y aprovechando que el plásmido de ATRX contiene un reportero GFP, se analizó, la eficiencia de transfección utilizando Citometría de flujo. Se utilizó un plásmido que contiene sólo GFP como control. Adicionalmente se usaron versiones del plásmido de ATRX que sólo contienen la región amino terminal (ATRX NT, aa 1-841), la región intermedia (ATRX Int, aa 800-1670) o la región carboxilo terminal (ATRX CT, aa 1670-2492). Para el caso de los métodos de transfección comerciales se siguieron las instrucciones de los fabricantes respectivos, usando sus recomendaciones de estandarización cuando fuera posible. En el caso de la transfección por fosfato de calcio se usó el protocolo "Calcium phosphate-mediated transfection of eukaryotic cells with plasmid DNAs" del libro *Molecular cloning: a laboratory Manual*, 3^{ra} edición (Sambrook & Russell 2001).

Se probaron diversas condiciones para los métodos que permitían alguna estandarización, con el fin de optimizar la eficiencia de transformación. En la Figura 8 se puede ver un resumen de las mejores eficiencias obtenidas con las condiciones óptimas de los diferentes métodos usados.

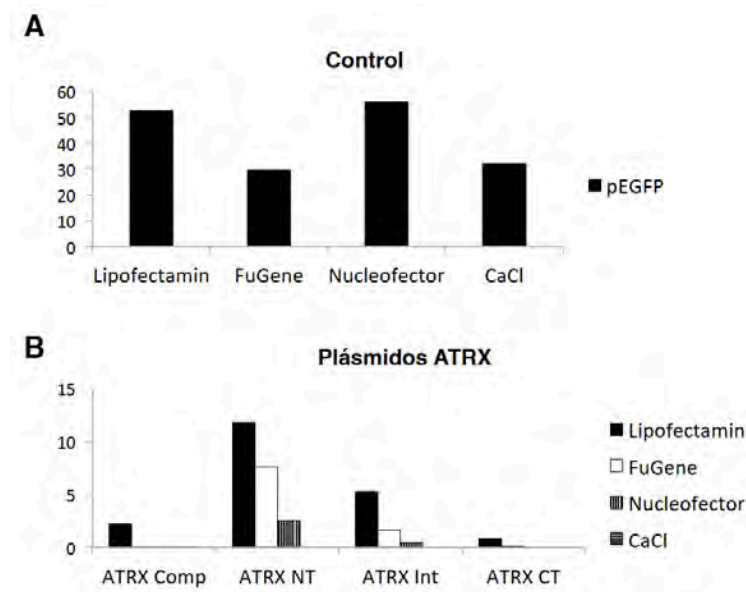


Figura 8. Eficiencias de transfección de plásmidos de ATRX en K562.

(A) Eficiencia de transfección (% de células GFP positivas) de un plásmido control (pEGFP) con las condiciones óptimas para los distintos métodos de transfección.

(B) Eficiencia de transfección de plásmidos que contienen la versión completa de ATRX (ATRX Comp), la región amino terminal (ATRX NT), la región intermedia (ATRX Int) o la región carboxilo terminal (ATRX CT) con los distintos métodos de transfección.

Como se puede observar, los métodos más eficientes para el plásmido control fueron la transfección por lipofectamina y el uso del Nucleofector, con eficiencias superiores al 50%. Los otros dos métodos (FuGene y CaCl) aunque tuvieron menores eficiencias (de alrededor del 30%) todavía muestran niveles aceptables para el control (Fig. 8A). Sin embargo, todas las versiones que contienen a ATRX tienen eficiencias drásticamente inferiores. El método más eficiente para transfectar la versión completa del plásmido fue lipofectamina con apenas un 3% de eficiencia. El resto de los métodos tuvieron eficiencias inferiores al 0.5%. En principio, se asumió que estas bajas eficiencias eran un problema del tamaño del plásmido (el plásmido mide en total más de 12Kb). Sin embargo, los plásmidos que contienen fragmentos más pequeños tampoco se pueden transfectar de forma eficiente (Fig. 8B). Notablemente, a pesar de que los tres fragmentos tienen tamaños similares entre sí, mostraron un comportamiento distinto en cuanto a su eficiencia de transfección. Mientras que el fragmento NT pudo ser transfectado con una eficiencia de hasta ~12% con el mejor método, el CT tiene eficiencias muy inferiores, similares a las del plásmido completo. Aún quedan por explorarse las razones de estos cambios y de la baja eficiencia del plásmido. Este no parece ser un problema técnico ya que el plásmido control se pudo transfectar de forma adecuada. Una posible explicación es que, al menos en las células K562, la sobre-

expresión de ATRX es tóxica, en particular la región CT. Eso podría explicar los bajos niveles de sobre-expresión del plásmido y la diferencia de eficiencias entre los distintos fragmentos.

Las condiciones optimizadas para la realización de la Figura 8. Se detallan a continuación.

Método	Tipo	Concentración del plásmido	Concentración del reactivo	Número de células
Lipofectamina	Lípido	2 µg	1:4 (plásmido:lipo)	1 x 10 ⁶
FuGene	Lípido	3 µg	1:3 (plásmido:Fu)	1 x 10 ⁶
Nucleofector	Electroporación	1 µg	NA. Se usó el protocolo específico para K562	1 x 10 ⁶
CaCl	Químico	5 µg	NA.	1 x 10 ⁶

Actualmente no contamos con un protocolo establecido para superar este problema. Comunicaciones personales con diversos laboratorios alrededor del mundo que están trabajando con ATRX y han intentado sobre-expresarlo en distintos modelos nos han informado que estos experimentos son particularmente desafiantes y frecuentemente imposibles de realizar. Si, como sospechamos, la sobre-expresión de ATRX es tóxica para la célula, se podría desarrollar un vector inducible para sobrellevar este problema. Otra alternativa es el uso de un método basado en CRISPR para sobre-expresar genes de forma endógena (Koneremann et al. 2015). La ventaja de esta aproximación es que la sobre-activación del propio plásmido de ATRX generaría niveles “naturales” que posiblemente sean tolerables para la célula. La desventaja es que este tipo de aproximación sólo se puede aplicar a células que posean una copia WT del gen, por lo que no se pueden realizar experimentos de rescate en células ATRX KO con este sistema.

III.II - CHIP-seq de ATRX: problemas de los protocolos comunes y optimización de un CHIP-seq eficiente para ATRX

Desarrollar un protocolo eficiente de ChIP-seq para ATRX fue esencial para realizar los estudios presentados en esta tesis. Este esfuerzo técnico tomó un tiempo considerable de pruebas y estandarizaciones, debido a que los protocolos estándar de ChIP-seq no se pueden aplicar directamente a líneas celulares humanas para el caso de ATRX. A

continuación se describirán los principales problemas que se tuvieron en la implementación del ChIP-seq y se detalla el protocolo final que se siguió.

III.II.I - La sonicación degradada a ATRX

En todos los protocolos de inmunoprecipitación de la cromatina se realiza un paso de fragmentación de la cromatina. Este paso es determinante para la resolución del ensayo y la mayoría de los protocolos de ChIP-seq recomiendan fragmentar la cromatina a fragmentos de menos de 300 pares de bases (Park 2009). Actualmente, el protocolo estándar de ChIP-seq recomienda usar un sonicador, un aparato que es capaz de fragmentar la cromatina con ondas de ultrasonido. Nuestro laboratorio ha realizado exitosamente experimentos de ChIP-seq de otras proteínas usando esta técnica (Martin et al. 2011).

Sin embargo, en el caso de ATRX, el protocolo estándar no estaba dando resultados para la línea celular humana K562. El primer problema es que a pesar de utilizar largos tiempos de sonicación usando un sonicador típico de la marca Bioruptor Twin, la cromatina de las células K562 no se fragmentaba de forma adecuada (datos no mostrados). Por ello se recurrió a usar un sonicador especializado de la compañía covaris, una compañía que se especializa en crear sonicadores para aplicaciones biológicas. La sonicación por covaris logró fragmentar la cromatina en tiempos razonables (Fig. 9A), sin embargo éramos incapaces de observar algún enriquecimiento en nuestros experimentos piloto de ChIP-qPCR (datos no mostrados). Cuando realizamos ensayos de WB para corroborar la integridad de ATRX tras la sonicación descubrimos que en el tiempo en el que estábamos obteniendo una fragmentación adecuada (12 mins), ~80% de la cantidad original de ATRX ya se había degradado. Como control realizamos el mismo blot para otra proteína (Suz12) y encontramos el mismo efecto, aunque ligeramente atenuado (Fig. 9B). A través de comunicaciones personales con la compañía Covaris, descubrimos que las proteínas de gran tamaño suelen ser particularmente sensibles a los tiempos de sonicación. Dado que después de muchas pruebas fuimos incapaces de encontrar una condición en la que pudiéramos obtener fragmentos de tamaño adecuado sin degradar a ATRX, decidimos cambiar la estrategia a un protocolo basado en la fragmentación de la cromatina por un método enzimático.

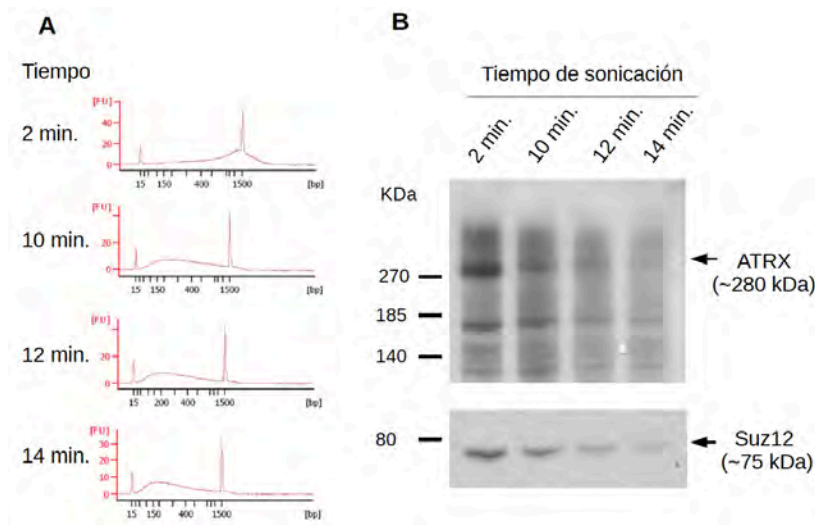


Figura 9. La sonicación degrada a ATRX en K562.

(A) Fragmentación de la cromatina de K562 tras diferentes tiempos de sonicación en un sonicador Covaris. Se muestran los perfiles de bioanalyzer (eje de las X, tamaño en pares de bases, eje de las Y, abundancia).

(B) WB de ATRX (arriba) y Suz12 (abajo) en cromatina sonicada por distintos tiempos en el sonicador Covaris.

Cabe mencionar que el protocolo de sonicación se probó también en otras líneas celulares humanas (las líneas de neuroblastoma LAN6 y SKNF1) con resultados negativos, por lo que es posible que este efecto no sea particular de la línea K562. El laboratorio de nuestro colaborador, el Dr. Dyer intentó también realizar ChIP-seq de ATRX en las líneas de neuroblastoma y otras líneas celulares siguiendo un protocolo basado en sonicación sin obtener resultados positivos. El laboratorio del Dr. Dyer encargó también el ChIP-seq de ATRX en dichas líneas celulares a la compañía Active Motif, que ofrece un servicio de ChIP-seq por pedido. La compañía declaró que no le fue posible realizar el ChIP-seq aunque no ofreció detalles técnicos acerca de los métodos empleados.

III.II.II - Estandarización de un protocolo de XL-MNasa para realizar experimentos de ChIP-seq de ATRX en humano

Debido a que demostramos claramente que la sonicación no era un método factible para realizar el ChIP-seq de ATRX en células humanas, decidimos utilizar una estrategia basada en la fragmentación de la cromatina por un método enzimático. Debido a que nuestros colaboradores en el laboratorio de la Dra. Bernstein tienen amplia experiencia realizando experimentos de ChIP basados en MNasa, probamos este protocolo. Comúnmente, el protocolo de fragmentación por MNasa se usa como un método para estudiar la distribución de marcas o variantes de histonas en condiciones nativas, es decir, sin la necesidad de fijar las interacciones proteína-DNA por medio de tratamiento con formaldehído o algún otro método similar (Zentner & Henikoff 2014). El método no

es comúnmente usado para estudiar factores transcripcionales aunque existen algunos reportes de CHIP-seq en condiciones nativas para TFs (Zentner & Henikoff 2014).

Es por ello que en primera instancia probamos un protocolo de Native CHIP que ya había sido estandarizado en la línea K562 en el laboratorio de la Dra. Bernstein (Ratnakumar et al. 2012). Desafortunadamente, los resultados fueron negativos (datos no mostrados). Debido a que la interacción de los factores de transcripción es mucho más lábil que la interacción entre el DNA y las histonas, y a que posiblemente ATRX no interactúa de forma directa con el DNA, sino a través de un complejo multiproteico, razonamos que fijar la muestra con formaldehído era un paso necesario para preservar intacta las interacciones de ATRX con sus sitios de unión. Es por ello que adaptamos un protocolo de *crosslink* seguido de MNasa de un kit de Cell Signaling (SimpleChip® Enzymatic Chromatin IP kit, número de catálogo #9003).

Lograr las condiciones óptimas para la generación de ChIPs de forma eficiente y reproducible requirió de la estandarización cuidadosa y la modificación del protocolo original del kit. A continuación se describe de forma detallada dicho protocolo. Esta es una versión extendida del protocolo publicado en la sección de Métodos del artículo “ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment” que se puede consultar en el Apéndice I de esta tesis. El protocolo ha sido aplicado de forma exitosa en varias líneas celulares humanas (K562, LAN6, SKNFI, ver sección II) y en eritrocitos de pollo de 10 días de desarrollo (ver sección I.V.II)

Protocolo de XL-MNase ChIP-seq para ATRX

Consideraciones previas:

- 1) Es necesario estandarizar cuidadosamente la cantidad de MNasa necesaria para obtener la fragmentación deseada. Cada línea celular requiere de su propia estandarización. Hemos visto diferencias notables usando la misma concentración de MNasa entre viales de distintos lotes. Así que es recomendable realizar el proceso de estandarización cuando se compran nuevos viales de MNasa.
- 2) En las distintas pruebas que hemos realizado, hemos observado que el protocolo es escalable a distintas cantidades de células. La descripción dada aquí asume que el experimento se realizará con 4×10^7 células, pero hemos realizado experimentos exitosos en el rango de $1-8 \times 10^7$ células escalando las cantidades de acuerdo a la cantidad de células.

Parte 1. Crosslink de las células.

Preparar:

- Solución de formaldehído 11x (ver 1.2)

1.1 - Para cultivos en suspensión: contar las células. Es recomendable que la concentración no sea mayor a 1×10^6 por mililitro. Lavar las células una vez con PBS a temperatura ambiente, y resuspender 4×10^7 en 10 ml de PBS a temperatura ambiente.

Para células adherentes: Contar una caja de células. Es recomendable realizar todos los experimentos con células a la misma confluencia y que ésta no sobrepase el 80%.

1.2 - Preparar la solución de formaldehído 11x. Para 1 ml mezclar:

- 100 μ l de buffer de *crosslink* 100X.

- 300 μ l de Formaldehído 37% (es recomendable que el formaldehído sea fresco)

- 600 μ l de agua destilada.

Volumen final: 1ml.

Esta solución se debe preparar fresca cada vez. La concentración de formaldehído de esta solución es de 11%.

1.3 - Mezclar la solución de formaldehído 11x con las células a temperatura ambiente para obtener una concentración final de 1% de formaldehído. Por cada 10ml de células en suspensión se debe agregar 1ml de solución 11x. Para células adherentes, agregar 1 ml por cada 10ml de medio. Mezclar gentilmente y contar exactamente 10 minutos.

1.4 - Agregar glicina a una concentración final de 124 mM (500 μ l de stock 2.5 M por cada 10ml de medio o células en suspensión). Rotar suavemente por 5 minutos.

1.5 - Para células en suspensión: centrifugar gentilmente (~1,200 rpm) por 5 minutos a 4°C y lavar una vez con PBS frío y desechar el sobrenadante. Para células adherentes: desechar el medio y lavar una vez con PBS frío. Despegar las células con un *scraper* y colectarlas en PBS frío, usar la cantidad necesaria para tener 4×10^7 de acuerdo a la cuenta del paso 1.1. Centrifugar gentilmente y desechar el sobrenadante. NOTA: El medio contiene formaldehído y debe ser desechado adecuadamente.

1.6 - Opcional: Congelar el pellet de células en nitrógeno líquido o hielo seco. A partir de este momento el pellet puede ser guardado a -80°C por largos periodos de tiempo. Este

paso no es necesario y se puede seguir con el protocolo sin necesidad de congelar el pellet.

Parte 2: Digestión con MNasa y preparación de la cromatina.

Preparar:

- Buffer A (kit) 1x con inhibidores de proteasas y DTT.
- Buffer B (kit) 1x con DTT.
- CHIP buffer (kit) 1x con inhibidores de proteasas.

Nota: Esta parte se realiza siguiendo básicamente las instrucciones del kit. Sólo se cambió la cantidad usada de buffer A y B. A partir de este momento todo el procedimiento se realiza en hielo y los buffers se deben mantener a 4°C a menos que se indique lo contrario.

2.7 - En caso de que el pellet haya sido congelado, descongelarlo en hielo. Preparar suficiente buffer A con inhibidores de proteasas y DTT. Resuspender el pellet de 4×10^7 de células en 4 ml de buffer A frío. Incubar en hielo por 10 minutos y mezclar gentilmente cada tres minutos aproximadamente. El buffer A lizará las células pero debe mantener los núcleos intactos. Es importante que la primera vez que se realiza este protocolo se tome una muestra, se tiña con azul de metileno y se observe al microscopio para analizar si se están obteniendo núcleos y si éstos están íntegros. El tiempo de incubación puede extenderse o acortarse para obtener una mejor lisis.

2.8 - Centrifugar a 4°C a 3,000 rpm por 5 minutos. Descartar el sobrenadante y resuspender cuidadosamente los núcleos en 3 ml de buffer B. Centrifugar nuevamente y descartar sobrenadante.

2.9 - Resuspender los núcleos en 1 ml de buffer B. Añadir la cantidad adecuada de MNasa, mezclar suavemente por inversión e incubar por 20 minutos exactamente a 37°C. Nota: la cantidad precisa de MNasa debe ser titulada cuidadosamente. La primera vez que se realiza este experimento se deben usar distintas concentraciones de MNasa y el patrón de corte debe ser analizado para determinar la cantidad adecuada.

2.10 - Pasar el tubo inmediatamente al hielo y agregar 10µl de EDTA 0.5M para detener la digestión. Centrifugar por 1 minuto a máxima velocidad y resuspender los núcleos en 1ml de CHIP buffer frío. Incubar en hielo por 10 minutos.

2.11 - Dividir los núcleos en dos tubos (500µl cada uno) y sonicar ligeramente para lisarlos. Usualmente 3-4 ciclos de 20 segundos en el bioruptor en poder alto son suficientes. La primera vez que se realiza el protocolo se debe determinar la cantidad adecuada de ciclos. Se debe dar la menor cantidad de ciclos posible que lise la mayor cantidad de núcleos posible.

2.12 - Centrifugar a 10,000 rpm a 4°C por 10 minutos para eliminar el *debris* del lisado. Pasar el sobrenadante a un nuevo tubo, teniendo cuidado de no perturbar el pellet. En este paso se pueden juntar las muestras que previamente se habían separado. La cromatina en este estado se puede guardar a 4°C por aproximadamente una semana. La cromatina no debe ser congelada.

Parte 3. Análisis de la digestión y cuantificación de la proteína.

Preparar:

- CHIP Elution buffer (kit) 1x

3.14 - Tomar 20µl de cromatina y mezclarla con 80µl de CHIP Elution buffer 1x del kit. Añadir 4µl de NaCl 5M y 1µl de RNasa. Incubar a 37°C por 30 minutos.

3.15 - Agregar 2µl de proteinasa K e incubar las muestras a 56°C por 1 hora. Incubar posteriormente a 65°C por 2-3 horas. Esta incubación se puede extender *overnight*.

3.16 - Purificar el DNA. En el laboratorio usamos las columnas del MinElute PCR Purification kit de qiagen, pero es posible usar las columnas del kit, algún otro kit similar o purificar el DNA por fenol cloroformo y precipitación en etanol. Eluir o resuspender el DNA en 20µl de agua o elution buffer de qiagen.

3.17 - Analizar la concentración del DNA por nanodrop. Analizar la digestión de la MNasa por medio de un gel de agarosa o preferentemente en un bionalyzer usando un chip de *high-sensitivity*.

3.18 - Con base en la concentración medida, determinar la cantidad de cromatina que se va a utilizar para la inmunoprecipitación. Para ATRX se deben usar como mínimo 30µg de material (la concentración de la cromatina es equivalente a la concentración de la muestra purificada) para CHIP-qPCR. Si se va a realizar CHIP-seq es preferente usar al menos 50µg. En general, entre más material se usa, más fuerte es la señal (la señal parece alcanzar un tope alrededor de 60-70µg, al menos para K562).

Parte 4. Inmunoprecipitación.

Día 1:

Preparar:

- CHIP buffer (kit) 1X con inhibidores de proteasas.

4.19 - Diluir la cantidad de cromatina con CHIP buffer 1x frío con inhibidores de proteasas a una concentración final de 100ng/μl. Si se van a realizar experimentos de qPCR, tomar 1% de la muestra diluida como Input y guardarla a -20°C. Añadir los anticuerpos e incubar *overnight* (~12 horas) en rotación a 4°C. Hemos probado 3 anticuerpos distintos de ATRX y todos han dado señales positivas aunque los resultados varían mucho de lote a lote. Es necesario determinar de forma empírica si un lote en particular sirve para CHIP-seq o no. Se debe usar una cantidad adecuada de anticuerpo. Para ATRX usamos al menos 2μg de anticuerpo para 30μg de cromatina para CHIP-qPCR. Para CHIP-seq usamos como mínimo 5μg para 50μg de cromatina.

Día 2:

Preparar:

- CHIP buffer (kit) 1x
- CHIP buffer (kit) 1x high salt
- LiCl wash buffer
- TE buffer 1X
- CHIP Elution buffer 1x, mantener a 37°

4.20 - Añadir a la cromatina las perlas magnéticas del kit. Se deben añadir aproximadamente 10μl de perlas por cada μg de anticuerpo usado aunque añadir ligeramente más o menos no afecta los resultados. Asegurarse de que las perlas están bien resuspendidas. Incubar de 2 a 4 horas en rotación a 4°C.

4. 21 - Colocar los tubos en el magneto para perlas y esperar entre 30 y 60 segundos a que el lisado se aclare. Descartar el sobrenadante y añadir 1ml de CHIP buffer 1x frío. Incubar por 5 minutos en rotación a 4°C.

4.22 - Repetir el lavado del paso 4.21 otras dos veces para un total de 3 lavados con CHIP buffer.

4.23 - Lavar las perlas una vez con ChIP buffer high salt frío como se describe en el paso 4.21

4.24 - Lavar las perlas una vez con LiCl wash buffer frío como se describe en el paso 4.21

4.25 - Lavar las perlas una vez con TE frío como se describe en el paso 4.21

4.26 - Descartar el TE y resuspender las perlas en 100µl de ChIP Elution buffer. Si se tiene una muestra de Input 1%, sacarla de -20° y agregarle también 100µl de ChIP Elution buffer.

4.27 - Incubar todos los tubos en un thermomixer a 1,200 rpm a 65°C por una hora.

4.28 - Poner las muestras de ChIP en el magneto y esperar 60 segundos a que la muestra se aclare. Cuidadosamente pasar el sobrenadante a un nuevo tubo. Descartar las perlas. NO descartar el sobrenadante.

4.29 - A las muestras del input y a los sobrenadantes del paso anterior agregar 4µl de NaCl 5M y 1µl de RNasa. Incubar a 37°C por 30 minutos.

4.30 - Agregar 2µl de proteinasa K. Incubar por una hora a 56°C

4.31 - Incubar por 2-3 horas a 65°C. Esta incubación puede extenderse *overnight*

4.32 - Purificar el DNA usando las el MinElute PCR Purification kit de qiagen. Es posible usar las columnas del kit, o purificar el DNA por fenol cloroformo y precipitación en etanol. Eluir o resuspender el DNA en 30-40µl de elution buffer de qiagen.

4.33 - Opcional: Analizar los ChIPs por bioanalyzer (*high-sensitivity chip*).

4.34 - Realizar experimentos de qPCR. Usar 1µl de ChIP por PCR.

4.35 - Opcional: Preparar librerías para secuenciación.

Buffers no incluidos en el kit:

Buffer de *crosslink* 100X:

- 1M NaCl

- 500 mM HEPES pH 8.0
- 10mM EDTA pH 8.0

LiCl wash buffer:

- 10mM Tris-HCl pH 8.0
- 1mM EDTA pH 8.0
- 1% sodium deoxycholate
- 1% Igepal
- 250mM LiCl

Conclusión

En los últimos años ha quedado de manifiesto la importancia del entendimiento de los diversos mecanismos epigenéticos que regulan la expresión génica. La epigenética ha ido evolucionando lentamente como un campo de estudio propio y su combinación con las nuevas tecnologías de análisis genómico y manipulación genética ha llevado a una explosión de nuevos descubrimientos. Dichos hallazgos han puesto en relieve la importancia de diversos reguladores epigenéticos en el desarrollo de enfermedades humanas. Un ejemplo de ello es el remodelador de la cromatina ATRX, el cual ha sido identificado como uno de los genes frecuentemente mutados en diversos tipos de cáncer gracias a las nuevas tecnologías de secuenciación masiva. Adicionalmente, mutaciones en la línea germinal en el gen ATRX, son responsables de un complejo síndrome humano conocido como *Alpha thalassemia X-linked intellectual disability syndrome*.

A pesar de la importancia de ATRX en la salud humana, se conoce relativamente poco acerca de sus mecanismos moleculares de acción y existen muy pocos análisis epigenómicos que intenten entender de forma global su función regulatoria. Es por ello que utilizando un conjunto de herramientas genéticas, genómicas, y bioinformáticas en diversos sistemas modelo buscamos entender las funciones moleculares de ATRX a un nivel epigenómico. Desprendidas de estas investigaciones, hemos podido llegar a las siguientes conclusiones generales:

1) ATRX se une a dominios de cromatina atípica en los exones 3' de los genes de tipo zinc finger. Realizando un extensivo análisis epigenómico, logramos identificar a la familia de factores transcripcionales de tipo zinc finger (ZNF) como una extensa familia de blancos de ATRX que no había sido reportada previamente. Con una combinación de análisis bioinformáticos, bioquímicos y moleculares hemos podido demostrar que ATRX forma un complejo con las proteínas SETDB1, TRIM28 y ZNF274. Dicho complejo se une a los exones 3' de los genes ZNF que presentan una configuración de cromatina atípica, caracterizada por la presencia de dos marcas de histonas que suelen ser mutuamente excluyentes: H3K9me3 y H3K36me3.

2) ATRX es necesario para el enriquecimiento de la marca H3K9me3 en los dominios de cromatina atípica. La eliminación de ATRX por medio de métodos genéticos tiene como consecuencia la disminución de la marca H3K9me3 en los genes ZNF que poseen una cromatina atípica. Esta disminución no es exclusiva de los ZNFs, la mayoría de las

regiones del genoma que presentan una configuración de cromatina atípica, tienen una disminución en sus niveles de H3K9me3 tras la eliminación de ATRX. Por el contrario, las regiones que están enriquecidas sólo en H3K9me3 y que no pertenecen a regiones atípicas, suelen ser resistentes a la eliminación de ATRX. Esto sugiere que un papel fundamental de ATRX, desconocido previamente, es la regulación de la marca H3K9me3 en regiones de cromatina atípica.

3) El establecimiento de la cromatina atípica en los genes ZNF podría contribuir a la estabilidad genómica de la célula. La eliminación tanto de ATRX como de ZNF274 tiene como consecuencia la pérdida del enriquecimiento de H3K9me3 en las regiones de cromatina atípica de los genes ZNF. Este decrecimiento está correlacionado con una mayor inestabilidad genómica. Las células que carecen de ATRX o de ZNF274 presentan mayores niveles de daño al DNA y tienen alteraciones en su ciclo celular consistentes con un entorno genómico más inestable.

4) ATRX es un regulador importante de la expresión de los genes α -globina. Combinando distintos modelos que permiten estudiar el desarrollo y la diferenciación eritroide tanto en células humanas como en organismos modelo, hemos logrado determinar que ATRX es necesario para la diferenciación terminal eritroide y la expresión de los genes α -globina. Asimismo, nuestros datos sugieren que ATRX participa también en la regulación transcripcional del locus α -globina durante el desarrollo.

5) Trabajar con la proteína ATRX presenta diversos retos técnicos. Es posible que la sobre-expresión de ATRX sea tóxica para ciertas líneas celulares. Adicionalmente, identificamos que los plásmidos que contienen el cDNA de ATRX son propensos a sufrir inserciones de transposones bacterianos cuando son amplificados en cepas estándar de *E. coli*. Finalmente, se requiere de un protocolo especial para realizar experimentos de CHIP-seq de ATRX de forma exitosa en células humanas.

En conjunto, nuestros resultados revelan una función inesperada de ATRX que hasta antes de esta investigación permanecía desconocida. A nivel global estos datos, junto con las investigaciones de varios grupos alrededor del mundo, demuestran que ATRX es un importante regulador de la deposición de H3K9me3 en distintas regiones del genoma, particularmente zonas de secuencias repetitivas que suelen ser una fuente de inestabilidad genómica cuando no son reguladas de forma adecuada. Esto sugiere que ATRX actúa como un centinela epigenético de la estabilidad genómica. Esta función podría explicar, al menos en parte, la razón por la cual las mutaciones en el gen ATRX son comunes en distintos tipos de cáncer. Esperamos que, en el futuro, el entendimiento

cada vez más profundo de los mecanismos a través de los cuáles ATRX regula al epigenoma nos permitirá el desarrollo de mejores estrategias para el tratamiento de las diversas enfermedades en las que ATRX está alterado.

Bibliografía

- Arriaga-Canon, C. et al., 2014. A long non-coding RNA promotes full activation of adult gene expression in the chicken α -globin domain. *Epigenetics*, 9(1), pp.173–181.
- Aymard, F. et al., 2014. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nature structural & molecular biology*, 21(4), pp.366–74.
- Beisel, C. & Paro, R., 2011. Silencing chromatin: comparing modes and mechanisms. *Nature reviews. Genetics*, 12(2), pp.123–35.
- Bernstein, B.E. et al., 2006. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*, 125(2), pp.315–326.
- Bérubé, N.G., 2011. ATRX in chromatin assembly and genome architecture during development and disease. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 89(5), pp.435–44.
- Bérubé, N.G. et al., 2002. Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Human molecular genetics*, 11(3), pp.253–61.
- Bérubé, N.G. et al., 2008. Patient mutations alter ATRX targeting to PML nuclear bodies. *European journal of human genetics : EJHG*, 16(2), pp.192–201.
- Bérubé, N.G. et al., 2005. The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *The journal of clinical investigation*, 115(2).
- Bérubé, N.G., Smeenk, C. a & Picketts, D.J., 2000. Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Human molecular genetics*, 9(4), pp.539–547.
- Bickmore, W.A. & Van Steensel, B., 2013. Genome architecture: Domain organization of interphase chromosomes. *Cell*, 152(6), pp.1270–1284.
- Bonasio, R. & Shiekhattar, R., 2014. Regulation of transcription by long noncoding RNAs. *Annual review of genetics*, 48(176), pp.433–55.
- Bower, K. et al., 2012. Loss of wild-type ATRX expression in somatic cell hybrids segregates with activation of Alternative Lengthening of Telomeres. *PLoS one*, 7(11), p.e50062.
- Casadesús, J. & Low, D., 2006. Epigenetic gene regulation in the bacterial world. *Microbiology and molecular biology reviews : MMBR*, 70(3), pp.830–56.
- Chen, X. et al., 2014. Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. *Cell Reports*, 7(1), pp.104–112.
- Cheung, N.-K. V et al., 2012. Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA: the journal of the American Medical Association*, 307(10), pp.1062–71.
- Cioe, L. et al., 1981. Differential expression of the globin genes in human leukemia K562(S) cells induced to differentiate by hemin or butyric acid. *Cancer research*, 41(1), pp.237–43.
- Clapier, C.R. & Cairns, B.R., 2009. The biology of chromatin remodeling complexes. *Annual review of biochemistry*, 78, pp.273–304.
- Clynes, D. et al., 2014. ATRX Dysfunction Induces Replication Defects in Primary Mouse

- Cells. *PloS one*, 9(3), p.e92915.
- Clynes, D. et al., 2015. Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nature Communications*, 6(May), p.7538.
- Clynes, D. & Gibbons, R.J., 2013. ATRX and the replication of structured DNA. *Current Opinion in Genetics and Development*, 23(3), pp.289–294.
- Clynes, D., Higgs, D.R. & Gibbons, R.J., 2013. The chromatin remodeler ATRX: a repeat offender in human disease. *Trends in biochemical sciences*, 38(9), pp.461–6.
- Dhayalan, A. et al., 2011. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Human molecular genetics*, 20(11), pp.2195–203.
- Drané, P. et al., 2010. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes & development*, 24(12), pp.1253–1265.
- Dunham, I. et al., 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), pp.57–74.
- Elsässer, S.J. et al., 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, 522(7555), pp.240–4.
- Episkopou, H. et al., 2014. Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. *Nucleic acids research*, 42(7), pp.4391–405.
- Ernst, J. et al., 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, 473(7345), pp.43–9.
- Eustermann, S. et al., 2011. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nature structural & molecular biology*, 18(7), pp.777–82.
- Felsenfeld, G. & Groudine, M., 2003. Controlling the double helix. *Nature*, 421(6921), pp.448–53.
- Filipescu, D., Szenker, E. & Almouzni, G., 2013. Developmental roles of histone H3 variants and their chaperones. *Trends in Genetics*, 29(11), pp.630–640.
- Frietze, S. et al., 2010. ZNF274 recruits the histone methyltransferase SETDB1 to the 3' ends of ZNF genes. *PloS one*, 5(12), p.e15082.
- Furlan-Magaril, M. et al., 2015. 3D genome architecture from populations to single cells. *Current Opinion in Genetics & Development*, 31, pp.36–41.
- Gamble, M.J. & Kraus, W.L., 2010. Multiple facets of the unique histone variant macroH2A: From genomics to cell biology. *Cell Cycle*, 9(13), pp.2568–2574.
- Garrick, D. et al., 2008. Switching genes on and off in haemopoiesis. *Biochemical Society transactions*, 36(Pt 4), pp.613–8.
- Gibbons, R., 2006. Alpha thalassaemia-mental retardation, X linked. *Orphanet journal of rare diseases*, 1, p.15.
- Gibbons, R.J. et al., 1991. A newly defined X linked mental retardation syndrome associated with alpha thalassaemia. *Journal of medical genetics*, 28(11), pp.729–33.
- Gibbons, R.J. et al., 2003. Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). *Nature genetics*, 34(4), pp.446–9.
- Gibbons, R.J. et al., 1995. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell*, 80(6), pp.837–45.

- Gibbons, R.J. et al., 2000. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nature genetics*, 24(4), pp.368–71.
- Gibbons, R.J. et al., 2008. Mutations in the chromatin-associated protein ATRX. *Human mutation*, 29(6), pp.796–802.
- Goldberg, A.D. et al., 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell*, 140(5), pp.678–91.
- Goldberg, A.D., Allis, C.D. & Bernstein, E., 2007. Epigenetics: a landscape takes shape. *Cell*, 128(4), pp.635–8.
- Guerrero, G. et al., 2007. Globin genes transcriptional switching, chromatin structure and linked lessons to epigenetics in cancer: a comparative overview. *Comparative biochemistry and physiology. Part A, Molecular & Integral physiology*, 147(3), pp.750–60.
- Hargreaves, D.C. & Crabtree, G.R., 2011. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell research*, 21(3), pp.396–420.
- He, Q. et al., 2015. The Daxx/Atrx Complex Protects Tandem Repetitive Elements during DNA Hypomethylation by Promoting H3K9 Trimethylation. *Cell stem cell*, 17(3), pp.273–86.
- Heaphy, C.M. et al., 2011. Altered telomeres in tumors with ATRX and DAXX mutations. *Science (New York, N.Y.)*, 333(6041), p.425.
- Higgs, D.R. et al., 2005. Understanding alpha-globin gene regulation: Aiming to improve the management of thalassemia. *Annals of the New York Academy of Sciences*, 1054, pp.92–102.
- Higgs, D.R. et al., 2007. Using genomics to study how chromatin influences gene expression. *Annual review of genomics and human genetics*, 8, pp.299–325.
- Huh, M.S. et al., 2012. Compromised genomic integrity impedes muscle growth after Atrx inactivation. *The Journal of clinical investigation*, 122(12), pp.4412–23.
- Iwase, S. et al., 2011. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nature structural & molecular biology*, 18(7), pp.769–776.
- Iyengar, S. et al., 2011. Functional analysis of KAP1 genomic recruitment. *Molecular and cellular biology*, 31(9), pp.1833–1847.
- Jenuwein, T. & Allis, C.D., 2001. Translating the histone code. *Science (New York, N.Y.)*, 293(5532), pp.1074–80.
- Jiao, Y. et al., 2011. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science (New York, N.Y.)*, 331(6021), pp.1199–203.
- Jiao, Y., Killela, P. & Reitman, Z., 2012. Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas. *Oncotarget*, 3(7), pp.709–722.
- Kalouisi, A. et al., 2015. The Nuclear Oncogene SET Controls DNA Repair by KAP1 and HP1 Retention to Chromatin. *Cell Reports*, 11(1), pp.149–163.
- Kernohan, K.D. et al., 2014. Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. *Nucleic Acids Research*, 42(13), pp.8356–8368.
- Kernohan, K.D. et al., 2010. ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Developmental cell*, 18(2),

- pp.191–202.
- Khil, P.P. et al., 2012. Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. *Genome Research*, 22(5), pp.957–965.
- Kiefer, C.M. et al., 2008. Epigenetics of beta-globin gene regulation. *Mutation research*, 647(1-2), pp.68–76.
- Konermann, S. et al., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536), pp.583–8.
- De La Fuente, R., Baumann, C. & Viveiros, M.M., 2011. Role of ATRX in chromatin structure and function: implications for chromosome instability and human disease. *Reproduction (Cambridge, England)*, 142(2), pp.221–234.
- Law, M.J. et al., 2010. ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. *Cell*, 143(3), pp.367–378.
- Lechner, M.S. et al., 2005. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochemical and Biophysical Research Communications*, 331(4), pp.929–937.
- Leung, J.W.-C.C. et al., 2013. Alpha Thalassemia/Mental Retardation Syndrome X-linked Gene Product ATRX Is Required for Proper Replication Restart and Cellular Resistance to Replication Stress. *The Journal of biological chemistry*, 288(9), pp.6342–50.
- Levy, M. a. et al., 2015. ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions. *Human molecular genetics*, 24(7), pp.1824–35.
- Lewis, P.W. et al., 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America*, 107(32), pp.14075–14080.
- Lovejoy, C. a et al., 2012. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS genetics*, 8(7), p.e1002772.
- Lozzio, C.B. & Lozzio, B.B., 1975. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*, 45(3), pp.321–34.
- Lusser, A. & Kadonaga, J.T., 2003. Chromatin remodeling by ATP-dependent molecular machines. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 25(12), pp.1192–200.
- Marinoni, I. et al., 2014. Loss of DAXX and ATRX are associated with chromosome instability and reduced survival of patients with pancreatic neuroendocrine tumors. *Gastroenterology*, 146(2), pp.453–460.e5.
- Martin, D. et al., 2011. Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nature structural & molecular biology*, 18(6), pp.708–14.
- Maze, I. et al., 2014. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nature Reviews Genetics*, (March).
- McDowell, T.L. et al., 1999. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proceedings of the National Academy of Sciences of the United*

- States of America*, 96(24), pp.13983–8.
- Mitson, M. et al., 2011. Functional significance of mutations in the Snf2 domain of ATRX. *Human molecular genetics*, 20(13), pp.2603–10.
- Molenaar, J.J. et al., 2012. Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature*, 483(7391), pp.589–93.
- Murray, J.M., Stiff, T. & Jeggo, P. a, 2012. DNA double-strand break repair within heterochromatic regions. *Biochemical Society transactions*, 40(1), pp.173–8.
- Nan, X. et al., 2007. Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(8), pp.2709–14.
- Napier, C.E. et al., 2015. ATRX represses alternative lengthening of telomeres. *Oncotarget*, 6(18), pp.16543–58.
- Newhart, A. et al., 2012. Single-cell analysis of Daxx and ATRX-dependent transcriptional repression. *Journal of cell science*, 125(Pt 22), pp.5489–501.
- Park, D.J. et al., 2004. Comparative analysis of ATRX, a chromatin remodeling protein. *Gene*, 339(1-2), pp.39–48.
- Park, P.J., 2009. ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics*, 10(10), pp.669–80.
- Pfister, S.X. et al., 2014. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell reports*, 7(6), pp.2006–18.
- Phillips, J.E. & Corces, V.G., 2009. CTCF: master weaver of the genome. *Cell*, 137(7), pp.1194–211.
- Picketts, D.J. et al., 1998. Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. *Mammalian Genome*, 9, pp.400–403.
- Qadeer, Z. a et al., 2014. Decreased Expression of the Chromatin Remodeler ATRX Associates with Melanoma Progression. *The Journal of investigative dermatology*, (January).
- Ramamoorthy, M. & Smith, S., 2015. Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. *Cancer cell*, 28(3), pp.357–69.
- Ransom, M., Dennehey, B.K. & Tyler, J.K., 2010. Chaperoning histones during DNA replication and repair. *Cell*, 140(2), pp.183–95.
- Ratnakumar, K. et al., 2012. ATRX-mediated chromatin association of histone variant macroH2A1 regulates α -globin expression. *Genes & development*, 26(5), pp.433–8.
- Ratnakumar, K. & Bernstein, E., 2013. ATRX: the case of a peculiar chromatin remodeler. *Epigenetics: official journal of the DNA Methylation Society*, 8(1), pp.3–9.
- Ray-Gallet, D. et al., 2011. Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3.3 to Maintain Chromatin Integrity. *Molecular Cell*, 44(6), pp.928–941.
- Rincón-Arano, H. et al., 2009. Chicken alpha-globin switching depends on autonomous silencing of the embryonic pi globin gene by epigenetics mechanisms. *Journal of cellular biochemistry*, 108(3), pp.675–87.
- Sadic, D. et al., 2015. Atrx promotes heterochromatin formation at retrotransposons. *EMBO reports*, 16(7), pp.836–50.

- Sambrook, J. & Russell, D., 2001. *Molecular Cloning: A Laboratory Manual* 3rd ed., Cold Spring Harbor Laboratory Press.
- Sarma, K. et al., 2014. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell*, 159(4), pp.869–83.
- Schneiderman, J.I. et al., 2012. Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. *Proceedings of the National Academy of Sciences of the United States of America*, 109(48), pp.19721–6.
- Schwartzentruber, J. et al., 2012. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature*, 482(7384), pp.226–31.
- Seah, C. et al., 2008. Neuronal death resulting from targeted disruption of the Snf2 protein ATRX is mediated by p53. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 28(47), pp.12570–80.
- Shema, E. et al., 2016. Single-molecule decoding of combinatorially modified nucleosomes. *Science*, 352(6286), pp.717–721.
- Smith, Z.D. & Meissner, A., 2013. DNA methylation: roles in mammalian development. *Nature reviews. Genetics*, 14(3), pp.204–20.
- Steensma, D.P., Gibbons, R.J. & Higgs, D.R., 2005. Acquired alpha-thalassemia in association with myelodysplastic syndrome and other hematologic malignancies. *Blood*, 105(2), pp.443–52.
- Tadepally, H.D., Burger, G. & Aubry, M., 2008. Evolution of C2H2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains. *BMC evolutionary biology*, 8, p.176.
- Tang, J. et al., 2004. A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. *The Journal of biological chemistry*, 279(19), pp.20369–77.
- Udugama, M. et al., 2015. Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres. *Nucleic acids research*, 43(21), pp.10227–37.
- Valdes-Quezada, C. et al., 2013. CTCF demarcates chicken embryonic α -globin gene autonomous silencing and contributes to adult stage-specific gene expression. *Epigenetics*, 8(8), pp.827–838.
- Vardabasso, C. et al., 2014. Histone variants: emerging players in cancer biology. *Cellular and molecular life sciences: CMLS*, 71(3), pp.379–404.
- Villevall, J.L. et al., 1983. Erythroid properties of K562 cells. Effect of hemin, butyrate and TPA induction. *Experimental cell research*, 146(2), pp.428–35.
- Voigt, P. et al., 2012. Asymmetrically modified nucleosomes. *Cell*, 151(1), pp.181–93.
- Voon, H.P.J. et al., 2015. ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. *Cell reports*, 11(3), pp.405–18.
- Voon, H.P.J. & Wong, L.H., 2016. New players in heterochromatin silencing: histone variant H3.3 and the ATRX/DAXX chaperone. *Nucleic acids research*, 44(4), pp.1496–501.
- Watson, L.A. et al., 2013. Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *Journal of Clinical Investigation*, 123(5), pp.2049–2063.
- Watson, L.A., Goldberg, H. & Bérubé, N.G., 2015. Emerging roles of ATRX in cancer. *Epigenomics*, 7, pp.1365–1378.
- de Wilde, R.F. et al., 2012. Loss of ATRX or DAXX expression and concomitant acquisition of the alternative lengthening of telomeres phenotype are late events in a

- small subset of MEN-1 syndrome pancreatic neuroendocrine tumors. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*, 25(7), pp.1033–9.
- Wong, L.H. et al., 2010. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome research*, 20(3), pp.351–60.
- Xue, Y. et al., 2003. The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19), pp.10635–40.
- Zentner, G.E. & Henikoff, S., 2014. High-resolution digital profiling of the epigenome. *Nature Reviews Genetics*, 15(12), pp.814–827.

Apéndice I - Publicaciones

A continuación, se anexan las publicaciones de primer autor que se derivaron del trabajo de investigación de esta tesis. Ambos artículos fueron publicados en revistas internacionales indexadas y arbitradas.

- David Valle-García, Zulekha A. Qadeer, Domhnall S. McHugh, Flávia G. Ghiraldini, Asif H. Chowdhury, Dan Hasson, Michael A. Dyer, Félix Recillas-Targa, Emily Bernstein (2016) ATRX binds to atypical chromatin domains at the 3' exons of ZNF genes to preserve H3K9me3 enrichment. *Epigenetics* [Epub ahead of print]. PMID: 27029610
- David Valle-García*, Lyra Griffiths*, Michael A. Dyer, Emily Bernstein, Félix Recillas-Targa (2014) The ATRX cDNA is prone to bacterial IS10 element insertions that alter its structure. *SpringerPlus* 3:222. PMID: 24834375
* equal contribution

Adicionalmente, durante la realización de este doctorado se contribuyó a los siguientes artículos originales publicados en revistas internacionales indexadas y arbitradas. Estos textos no se anexan a la tesis.

- Chiara Vardabasso, Alexandre Gaspar-Maia*, Dan Hasson*, Sebastian Pünzeler*, David Valle-García*, Tobias Straub, Eva C. Keilhauer, Thomas Strub, Joanna Dong, Taniya Panda, Chi-Yeh Chung, Jonathan L. Yao, Rajendra Singh, Miguel F. Segura, Barbara Fontanals-Cirera, Amit Verma, Matthias Mann, Eva Hernando, Sandra B. Hake, Emily Bernstein (2015) Histone variant H2A.Z.2 mediates proliferation and drug sensitivity of malignant melanoma. *Molecular Cell* 59(1):75-88. PMID: 26051178
* equal contribution
- Zulekha A. Qadeer, Sara Harcharik, David Valle-García, Chen Chen, Miriam B. Birge, Chiara Vardabasso, Luis F. Duarte, Emily Bernstein (2014) Decreased expression of the chromatin remodeler ATRX associates with melanoma progression. *Journal of Investigative Dermatology* 134(6):1768-72. PMID: 24468746
- Alexandre Gaspar-Maia, Zulekha A Qadeer, Dan Hasson, Kajan Ratnakumar, N Adrian Leu, Gary Leroy, Shichong Liu, Carl Costanzi, David Valle-García, Christoph Schaniel, Ihor Lemischka, Benjamin Garcia, John R Pehrson, Emily Bernstein (2013) MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nature Communications* 4:1565. PMID: 23463008

- Karla F Meza-Sosa, David Valle-García, Gustavo Pedraza-Alva, Leonor Pérez-Martínez (2012) Role of microRNAs in Central Nervous System Development and Pathology. *Journal of Neuroscience Research* 90(1):1-12. PMID: 21922512
- Rodrigo G Arzate-Mejía, David Valle-García, Félix Recillas-Targa (2011) Signaling epigenetics: Novel insights on cell signaling and epigenetic regulation. *IUBMB Life* 63(10):881-95. PMID: 21901818



ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment

David Valle-García, Zulekha A. Qadeer, Domhnall S. McHugh, Flávia G. Ghiraldini, Asif H. Chowdhury, Dan Hasson, Michael A. Dyer, Félix Recillas-Targa & Emily Bernstein

To cite this article: David Valle-García, Zulekha A. Qadeer, Domhnall S. McHugh, Flávia G. Ghiraldini, Asif H. Chowdhury, Dan Hasson, Michael A. Dyer, Félix Recillas-Targa & Emily Bernstein (2016): ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment, Epigenetics, DOI: [10.1080/15592294.2016.1169351](https://doi.org/10.1080/15592294.2016.1169351)

To link to this article: <http://dx.doi.org/10.1080/15592294.2016.1169351>



© 2016 The Author(s). Published with license by Taylor & Francis Group, LLC



View supplementary material



Accepted author version posted online: 30 Mar 2016.
Published online: 30 Mar 2016.



Submit your article to this journal



Article views: 440



View related articles









View Crossmark data

RESEARCH PAPER

 OPEN ACCESS

ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment

David Valle-García ^{a,b}, Zulekha A. Qadeer ^{a,c}, Domhnall S. McHugh ^a, Flávia G. Ghiraldini ^a, Asif H. Chowdhury^a, Dan Hasson^a, Michael A. Dyer^d, Félix Recillas-Targa ^b, and Emily Bernstein ^{a,c,d}

^aDepartments of Oncological Sciences and Dermatology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^bInstituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Ciudad de México, México; ^cGraduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^dDepartment of Developmental Neurobiology, St Jude Children's Research Hospital, Memphis, Tennessee, USA

ABSTRACT

ATRX is a SWI/SNF chromatin remodeler proposed to govern genomic stability through the regulation of repetitive sequences, such as rDNA, retrotransposons, and pericentromeric and telomeric repeats. However, few direct ATRX target genes have been identified and high-throughput genomic approaches are currently lacking for ATRX. Here we present a comprehensive ChIP-sequencing study of ATRX in multiple human cell lines, in which we identify the 3' exons of zinc finger genes (ZNFs) as a new class of ATRX targets. These 3' exonic regions encode the zinc finger motifs, which can range from 1–40 copies per ZNF gene and share large stretches of sequence similarity. These regions often contain an atypical chromatin signature: they are transcriptionally active, contain high levels of H3K36me₃, and are paradoxically enriched in H3K9me₃. We find that these ZNF 3' exons are co-occupied by SETDB1, TRIM28, and ZNF274, which form a complex with ATRX. CRISPR/Cas9-mediated loss-of-function studies demonstrate (i) a reduction of H3K9me₃ at the ZNF 3' exons in the absence of ATRX and ZNF274 and, (ii) H3K9me₃ levels at atypical chromatin regions are particularly sensitive to ATRX loss compared to other H3K9me₃-occupied regions. As a consequence of ATRX or ZNF274 depletion, cells with reduced levels of H3K9me₃ show increased levels of DNA damage, suggesting that ATRX binds to the 3' exons of ZNFs to maintain their genomic stability through preservation of H3K9me₃.

ARTICLE HISTORY

Received 21 October 2015
Revised 14 March 2016
Accepted 17 March 2016

KEYWORDS

Atypical chromatin; ATRX; ESET; H3K9me₃; KRAB-ZNFs; KAP1; SETDB1; TRIM28; zinc finger genes; ZNF274

Introduction


Chromatin remodeling proteins act through shifting, sliding, deposition, and eviction of nucleosomes and histones. Members of the SWI/SNF family of chromatin remodelers are fundamental in many cellular processes, such as transcription, replication, DNA repair, and recombination.^{1–5} One notable chromatin remodeler involved in all of these processes is ATRX. Increasing evidence supports that ATRX acts as a sentinel of genome integrity by maintaining heterochromatin at repetitive sequences.^{6,7} Interestingly, ATRX germline mutations are responsible for a complex genetic disorder called α thalassemia, mental retardation X-linked (ATR-X) syndrome, while somatic mutations, deletions, and altered ATRX expression levels are highly prevalent in a wide variety of cancers.^{8,9}

ATRX contains 2 highly conserved domains: the ADD (ATRX-DNMT3-DNMT3L) and the SWI/SNF helicase domain.¹⁰ The ADD domain contains a PHD finger that binds H3K9me₃/H3K4me₀,^{11,12} whereas the SWI/SNF domain is an ATP-dependent helicase responsible for the chromatin remodeling capacity of

ATRX.^{10,13} Despite the fact that ATRX binds H3K9me₃/H3K4me₀ *in vitro*, ATRX binds to only a subset of H3K9me₃-containing regions *in vivo*.^{11,12,14,15} In particular, ATRX is highly enriched at certain H3K9me₃-containing repetitive regions, such as telomeric and pericentromeric repeats, as well as some retrotransposon families.^{16–22} Furthermore, ATRX physically interacts with other H3K9me₃ binding proteins, such as HP1 α .^{23,24} All together, these pieces of evidence suggest that ATRX is involved in the regulation of particular H3K9me₃-modified chromatin.

A well-characterized role of ATRX is deposition of histone variants into the chromatin template. For example, ATRX and Death-domain associated protein (DAXX) act together as a histone chaperone complex for the H3 variant H3.3. ATRX is required for the localization of H3.3 at telomeres and pericentromeric repeats,^{16–19} retrotransposons,^{20–22} and imprinted loci,²⁵ which all contain H3K9me₃. This ability appears to be unique for ATRX, as the HIRA complex deposits H3.3 only at euchromatic regions.^{17,26–28} In addition to promoting H3.3 deposition, our group showed that ATRX negatively regulates

CONTACT Emily Bernstein  Emily.Bernstein@mssm.edu; Félix Recillas-Targa  frecilla@ifc.unam.mx.

 Supplemental data for this article can be accessed on the publisher's website.

Published with license by Taylor & Francis Group, LLC © DavidValle-García, Zulekha A. Qadeer, Domhnall S. McHugh, Flávia G. Ghiraldini, Asif H. Chowdhury, Dan Hasson, Michael A. Dyer, Félix Recillas-Targa, and Emily Bernstein.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

the deposition of histone variant macroH2A at the α -globin locus.²⁹

ATRX has also been implicated in resolving aberrant secondary DNA structures, called G-quadruplexes, which form in guanine-rich regions during replication and transcription.^{15,30,31} G-quadruplexes are a common feature of some families of repetitive sequences and tandem repeats, such as those found in telomeres. Intriguingly, ATRX mutations in cancer have been linked to the alternative lengthening of telomeres (ALT) pathway.^{32–35} Although the precise role of ATRX in ALT remains unclear, it has been suggested that ATRX prevents homologous recombination (HR) between telomeric sequences through the resolution of stalled replication forks in G-rich regions.^{6,36} In accordance with its role as a regulator of genome stability, several reports demonstrated that ATRX depletion causes telomere dysfunction, increased replication fork stalling, and increased sensitivity to replicative stress across different cellular and *in vivo* models.^{18,22,37–40}

Despite these important functions, surprisingly few direct ATRX target genes have been identified. To address this, we utilized an unbiased approach using the ENCODE Tier 1 human erythroleukemic cell line K562 as a model system to analyze ATRX genomic occupancy. Through comprehensive ChIP-seq analyses, we identified an unexpected binding pattern of ATRX at the 3' exons of zinc finger genes (ZNFs). ZNFs represent the largest family of putative transcription factors in the human genome with more than 700 identified members.^{41–43} This enrichment of ATRX at ZNF 3' exons was further confirmed in additional human cell lines of both normal or cancer origin.

The 3' exons of ZNFs are enriched in chromatin that is permissive to transcription yet contains high levels of H3K9me3 and H3K36me3.⁴⁴ These atypical chromatin regions do not possess the characteristics of any known regulatory region (i.e., promoter, enhancer, insulator) and their functional significance remains unclear.⁴⁴ Here we show that ATRX co-occupies 3' ZNF exons containing an H3K9me3/H3K36me3 chromatin signature, together with the H3K9 methyltransferase SETDB1 (also known as ESET), the co-repressor TRIM28 (also known as KAP1), and the transcription factor ZNF274. Deletion of ATRX or ZNF274 leads to a reduction of H3K9me3, particularly at 3' ZNF exons and other H3K9me3/H3K36me3-containing regions, as well as increased DNA damage, and defects in the cell cycle. Taken together, our studies suggest that ATRX binds the 3' exons of ZNFs to maintain genomic stability by regulating H3K9me3 levels.

Results

ATRX binds to the 3' exons of ZNF genes in K562 cells

In order to perform an unbiased search for novel direct ATRX target genes, we examined its genomic distribution by ChIP-seq analysis in the human erythroleukemic cell line K562 using 2 independent antibodies (see Methods for details). We chose K562 as a model system for 2 reasons: first, it has been established that ATRX has important roles in the regulation of the erythroid lineage;^{15,29} second, K562 is a Tier 1 ENCODE cell line that has been extensively analyzed using a wide array of genomic and epigenomic methodologies that are publicly available.⁴⁵

To determine the global ATRX binding pattern in relation to other chromatin modifications, we re-analyzed the available ENCODE ChIP-seq datasets for K562 (see Methods) and performed a correlation analysis of their binding profiles. The only data sets that show a positive correlation with ATRX are H3K9me3 ($r = 0.46$, Spearman correlation) and macroH2A ($r = 0.19$, Spearman correlation) (Fig. S1A), consistent with its role as an H3K9me3 binding protein^{11,12} and macroH2A regulator, respectively.^{29,46,47} Furthermore, we examined the genomic distribution of ATRX significant peaks and found that, consistent with previous reports,^{15,20–22} ATRX is bound mainly to repetitive sequences (~56% of ATRX peaks overlap with repeats) (Fig. S1B). In order to understand ATRX distribution at a functional level, we analyzed its distribution across Hidden Markov Model-derived chromatin states.⁴⁸ While ATRX binds to repressed and repetitive regions (Fig. 1A), it is significantly enriched in transcribed regions as well (Fig. 1A). In order to further investigate the functional significance of ATRX occupancy at these transcriptionally active regions, we performed Gene Ontology (GO) analysis with ATRX-bound genes ($n = 374$, Table S1). Strikingly, C2H2 ZNFs were the most overrepresented gene family and comprised one quarter of the ATRX-bound genes, many of which contain the repressive KRAB domain (Fig. 1B). We next analyzed ATRX binding at promoters and gene bodies and found that enrichment of ATRX at gene bodies of ZNFs was highly significant as compared to non-ZNF genes, but that promoter regions had minimal binding in either group of genes (Fig. 1C).

ZNFs represent the fastest expanding gene family in the primate lineage. Frequent gene duplications and rapid divergence of paralogs are characteristic of ZNFs.^{41–43} Because of this, ZNFs are often arranged in large continuous clusters in the human genome and share stretches of highly similar DNA sequence, particularly at their 3' exons where the DNA sequence encoding the zinc finger motifs is contained.^{41,42} Chromosome 19 contains the majority of ZNF clusters in the human genome.^{41–43} By examining ATRX enrichment on chromosome 19, we found that ZNF clusters are demarcated by ATRX occupancy (Fig. 1D, E). We next analyzed the binding pattern of ATRX over individual ZNF genes and found ATRX to be preferentially enriched at the 3' exons of ZNFs (Fig. 1F, G). These results were confirmed by ChIP-seq with a second ATRX antibody, which showed nearly identical enrichment patterns at ZNF genes (Fig. S2). Overall, our ChIP-seq studies demonstrate that ZNFs are a novel set of ATRX targets and that ATRX is highly enriched at their 3' exons.

ATRX is enriched at ZNF genes harboring an atypical chromatin signature and distinctive epigenetic and genomic features

A large proportion of ZNFs contain an atypical chromatin signature at their 3' exons.⁴⁴ This includes high levels of H3K9me3, permissibility to transcription, and enrichment of H3K36me3, a mark associated with transcriptional elongation. Our analysis of ENCODE ChIP-seq data in K562 cells corroborated these observations (Fig. 1E–G).

To investigate the epigenetic and genetic characteristics of the ATRX-enriched ZNFs and their relationship with the above atypical chromatin signature, we categorized all ZNF genes into

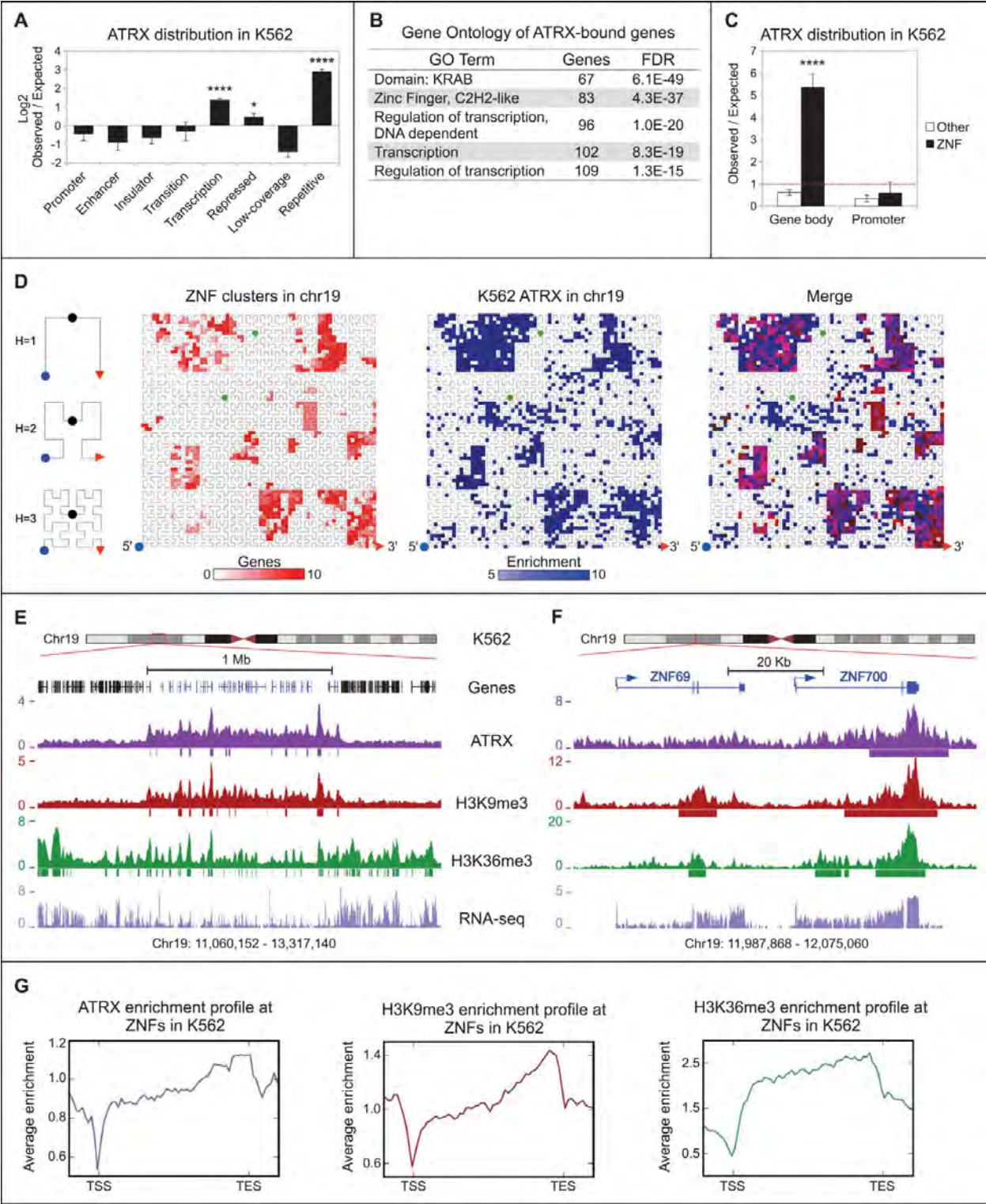


Figure 1. ATRX binds to the 3' regions of ZNF genes in K562 cells. (A) Observed over expected random distribution of significant ATRX peaks within HMM chromatin categories in K562. ⁴⁸ In (A) and (C) error bars represent standard deviation. Asterisks represent significantly overrepresented regions (* P -value < 0.05; **** P -value < 1×10^{-4}) assessed by the hypergeometric test (see Tables S2 and S3 for details of statistical tests). (B) Gene Ontology analysis of genes that overlap with significant ATRX peaks in their gene bodies. (C) Observed over expected random distribution of significant ATRX peaks in ZNFs and non-ZNF (other) promoters and gene bodies in K562. The red line represents the expected value of a random distribution. (D) Hilbert curve plot of chromosome 19 showing the ZNF clusters (red, left), highly enriched ATRX regions in K562 (fold enrichment over input > 5, blue, middle), and overlap (right). The blue dot and the red arrow mark the start (5') and end (3') of the chromosome, respectively (see diagram on left). The green dots indicate the start and end of centromere, which is excluded from the analysis. (E) UCSC Genome browser screenshot of a typical ZNF cluster (genes in blue) on chromosome 19. The enrichment over input signal in K562 for ATRX, H3K9me3 and H3K36me3 ChIP-seq is shown. Significant peaks represented as bars below enrichment tracks. Normalized RPKM signal for RNA-seq is also shown. (F) Zoomed in snapshot of 2 ZNFs contained within the ZNF cluster shown in (E). (G) Average K562 enrichment ChIP-seq profiles of ATRX, H3K9me3 and H3K36me3 over all ZNF gene bodies ± 1 kb ($n = 736$).

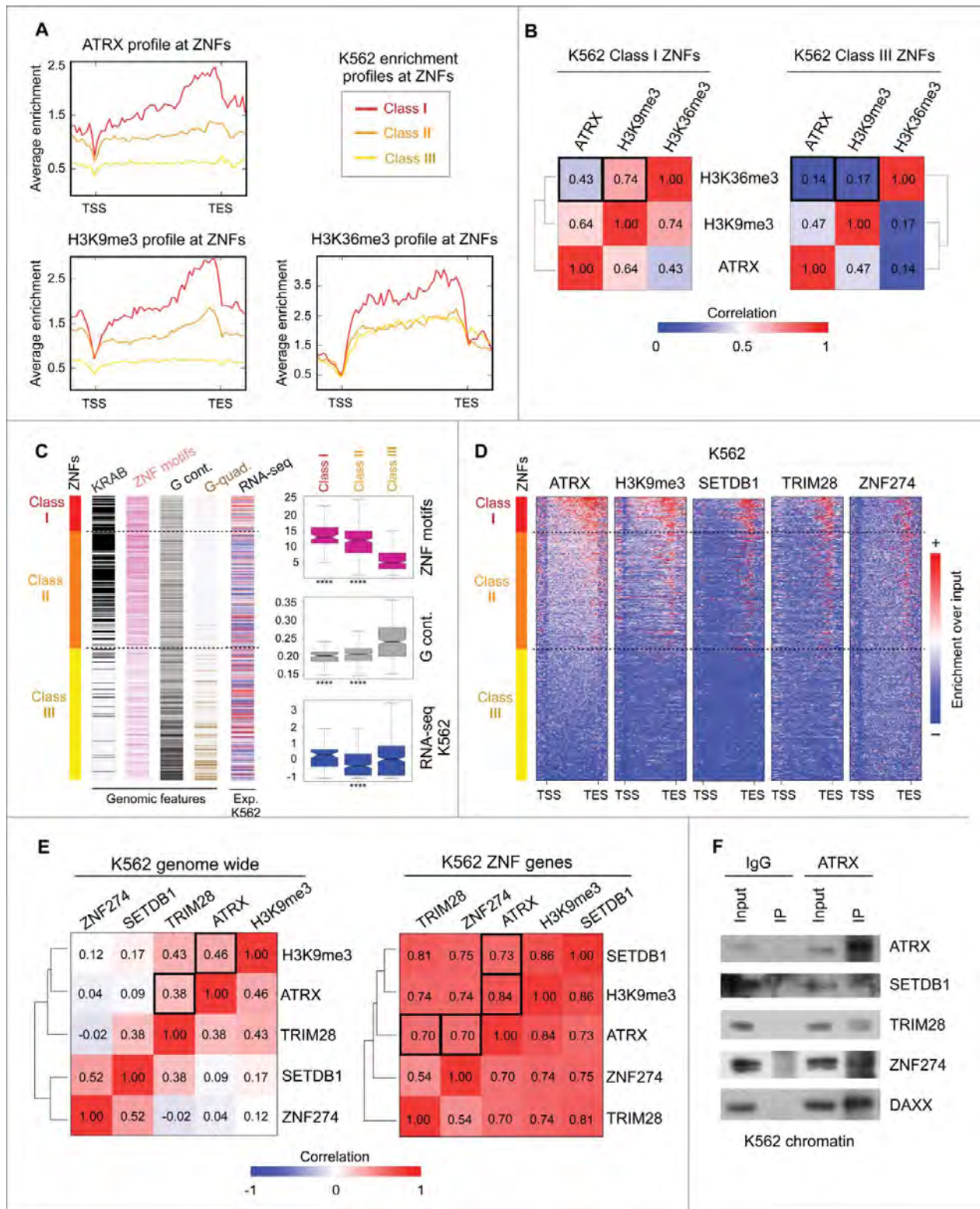


Figure 2. ATRX and the ZNF274/TRIM28/SETDB1 complex bind to ZNF genes with an atypical chromatin signature and distinctive genomic and epigenetic features. (A) Average K562 enrichment ChIP-seq profiles of ATRX, H3K9me3 and H3K36me3 at ZNFs classified by their ATRX content. Class I contains high levels of ATRX ($n = 91$), Class II contains medium to low levels of ATRX ($n = 303$) and Class III is devoid of ATRX enrichment ($n = 342$). (B) Spearman correlation heatmap of K562 ChIP-seq signal at ZNF Class I genes (left) and ZNF Class III genes (right). (C) Left: Distribution of genetic features among ZNF classes (sorted from high to low ATRX enrichment from top to bottom). Dashed lines show separation of the 3 classes. Colors represent presence of KRAB domains (black), number of zinc finger motifs (pink), G content at the C-terminal ZNF region (last 3 kb of the gene) (gray) and presence of sequences predicted to form G-quadruplexes (brown). RNA-seq bar shows the Z score of the normalized RPKM signal ($\log_2(\text{RPKM}+1)$) in K562; red: high expression signal and blue = low expression signal. For statistical tests between the classes see Table S3. Right: Box plots displaying the number of ZNF motifs, G-content at the ZNF region and RNA-seq values in K562 per ZNF Class. Asterisks show significant differences ($P\text{-value} < 1 \times 10^{-4}$). (D) Metagene analysis of ChIP-seq enrichment over input profiles at ZNF gene bodies ± 1 kb. (E) Spearman correlation heatmaps between ChIP-seq profiles genome-wide (left) and at ZNF genes (right). Black boxes indicate the significant correlations. (F) Immunoblots for endogenous ATRX Co-IP of chromatin bound proteins in K562 cells after pulldown with IgG or ATRX antibody. DAXX used as a positive control for the ATRX IP.

3 classes based on their ATRX enrichment levels: Class I represents ZNFs highly enriched for ATRX, Class II contains ZNFs moderately enriched, and Class III for ZNFs depleted of ATRX enrichment (Fig. 2A, top). We next quantified the average ChIP-seq signals of ATRX, H3K9me3, and H3K36me3 over the gene bodies of the ZNF classes. As shown in Figure 2A, Class I ZNFs show high levels of both H3K9me3 and H3K36me3. In contrast, Class III genes are largely depleted of H3K9me3 and show less enrichment of H3K36me3. To analyze if ATRX enrichment is correlated with the presence of H3K9me3 and H3K36me3 at the same loci, we calculated the Spearman correlation of these marks in Class I and Class III ZNFs. H3K9me3 is correlated with ATRX and moderately correlated with H3K36me3 in Class I ZNFs, whereas a poor correlation for both marks was observed in Class III ZNFs (Fig. 2B). These data suggest that ATRX is specifically enriched at those ZNFs displaying an atypical chromatin signature.

To analyze the relation between ATRX and atypical chromatin genome-wide, we re-analyzed ENCODE H3K36me3 and H3K9me3 ChIP-seq data for K562 cells and plotted ATRX enrichment levels at atypical chromatin peaks (overlapping H3K9me3/H3K36me3 peaks), H3K9me3-only peaks, and H3K36me3-only peaks. ATRX is enriched at a subset of atypical and H3K9me3-only regions, and largely absent from H3K36me3-only regions, suggesting that ATRX recruitment is independent of the H3K36me3 mark (Fig. S3A,B). Interestingly, ATRX enrichment is more significant at atypical chromatin regions than at H3K9me3-only regions (Fig. S3C). These data suggest that ATRX is a *bona fide* binder of atypical chromatin genome-wide.

To further understand ATRX recruitment and function at ZNFs, we analyzed the ZNF-related genomic features in the 3 defined classes. ZNF genes can be classified as transcriptional activators or KRAB-containing repressors. KRAB is a potent repressor domain that is contained in about half of the ZNFs and is generally encoded in 2 exons distinct from the 3' exon containing the zinc finger motifs.⁴² As KRAB-containing genes were enriched in our GO analysis of ATRX-bound genes (Fig. 1B), we plotted the number of KRAB domains contained by each ZNF ordered by class. Most of the ATRX-enriched Class I and Class II ZNFs contained KRAB domains, while very few of Class III ZNF genes contain this feature (Fig. 2C, left).

Because the DNA sequence encoding C2H2 zinc finger motifs is similar between ZNFs, it has been proposed that ZNF genes are prone to homologous recombination (HR), particularly those with more zinc finger motifs.^{41,42,44} Therefore, the presence of H3K9me3 at the ZNF 3' exons has been suggested to protect against HR.^{44,49} To support this idea, we plotted the number of predicted C2H2 zinc finger motifs per ZNF gene. On average, human ZNF genes contained ~9 zinc finger motifs per gene. In contrast, Class I ZNF genes contained significantly more motifs with an average of ~14, while ATRX depleted Class III genes contained only ~6 domains per gene (Fig. 2C, left and top right). These results suggest that ATRX enrichment at the ZNF 3' exons positively correlates with the number of C2H2 zinc finger motifs. This is in accordance with a study that reported H3K9me3 enrichment at 3' ZNF exons positively associated with the number of zinc finger motifs.⁴⁴

A genomic feature proposed to be important for ATRX binding is the Guanine DNA content (G-content). ATRX binds to G-quadruplexes with high affinity *in vitro* and facilitates polymerase elongation through deposition of H3.3 specifically in G-rich regions that have a tendency to form these structures.^{15,31} Based on these observations, we measured the G-content and predicted the potential of G-quadruplex formation at the 3' ends of the ZNF genes. Surprisingly, ATRX enrichment levels negatively correlated with both G-content and the potential to form G-quadruplexes (Fig. 2C, left and middle right). This strongly indicates that ATRX recruitment to ZNF 3' exons is not mediated by its ability to recognize G-quadruplexes, but by an alternative mechanism(s).

We then investigated whether ATRX enrichment and the presence of the atypical chromatin signature correlates with ZNF transcriptional levels. Thus, we analyzed the ENCODE RNA-seq datasets for K562 and plotted the normalized RPKM signal for the 3 ZNF classes. Intriguingly, there was no evident association between RNA-seq expression levels and ATRX enrichment. This suggests that neither ATRX binding nor the formation of the atypical chromatin signature have a direct effect on ZNF expression levels (Fig. 2C, left and bottom right).

As ATRX regulates late stalled replication forks and H3K9me3-marked chromatin is often late replicating,^{37,39,40,50} we queried whether ATRX binds to late-replicating ZNFs. To address this, we analyzed the K562 Repli-seq data from ENCODE and quantified the signal for the ZNF classes throughout S phase. Interestingly, we found that ZNF Classes I and II tend to be late replicating while Class III ZNFs replicate early (Fig. S4A).

In summary, we have established that ATRX levels positively correlate with H3K9me3 at atypical chromatin found at the 3' of ZNF genes. ATRX enrichment at ZNFs is independent of transcriptional levels. Moreover, ATRX-enriched ZNFs tend to be late replicating, contain KRAB domains, and have a large number of zinc finger motifs. Such ZNFs also contain low levels of G-content and low potential for G-quadruplex formation. These trends are all statistically significant (Tables S2-S3).

SETDB1, TRIM28 and ZNF274 co-localize at 3' exons of ATRX-bound ZNF genes and form a complex with ATRX

In order to find additional chromatin factors that bind the ZNF 3' exons, we performed metagene analyses at the ZNF gene bodies with all available K562 ChIP-seq data sets from the ENCODE project.⁵¹ From the 98 datasets we analyzed, we found that the enrichment levels of the H3K9me3 methyltransferase SETDB1 (also known as ESET) and the SETDB1-interacting protein TRIM28 (also known as KAP1) correlated appreciably with ATRX and H3K9me3 at the 3' exons of ZNF genes (Fig. 2D). Interestingly, TRIM28 is a co-repressor that interacts with KRAB-containing ZNF transcription factors and recruits HDACs and H3K9 methyltransferases to enforce silencing.⁵² Moreover, we found a striking co-localization of ATRX and H3K9me3 with the KRAB-containing transcription factor ZNF274 (Fig. 2D). This is in agreement with a study that reported ZNF274 to bind the 3' region of ZNF genes and recruit SETDB1 through its interaction with TRIM28.⁵³

We then performed a correlation analysis of the ChIP-seq signals of ATRX, ZNF274, TRIM28, SETDB1, and H3K9me3 to determine if they preferentially co-localize at ZNFs. Genome-wide, ATRX moderately associates only with H3K9me3 and TRIM28 (Fig. 2E, left). In striking contrast, when focusing only on the ZNF genes, the correlation coefficients between ZNF274/TRIM28/SETDB1/ATRX denote a strong association, along with presence of H3K9me3 signal (Fig. 2E, right).

To investigate whether the ZNF274/TRIM28/SETDB1 complex interacts with ATRX, we performed co-immunoprecipitation experiments of chromatin-bound proteins in K562 cells. As show in Fig. 2F, ATRX is able to co-immunoprecipitate with all 3 factors. As a control for ATRX pull-down, we also confirmed ATRX's interaction with DAXX in K562. This data strongly suggests that ATRX physically interacts with ZNF274/TRIM28/SETDB1 at the chromatin level.

Finally, we performed a comprehensive motif analysis of ZNF274 binding sites (see Methods). We found 3 significant DNA motifs that were highly enriched at the 3' end of ATRX enriched Class I and Class II ZNF genes, but not at ATRX depleted Class III ZNFs (Fig. S4B). These data favors the idea that the ZNF274/TRIM28/SETDB1/ATRX complex could bind to ZNF genes at least in part through the recognition of ZNF274-binding motifs. Of note, the motifs that we found are in accordance with a previous ZNF274 motif analysis.⁵³

Collectively, we have identified a ZNF274/TRIM28/SETDB1/ATRX complex that localizes to the 3' region of ZNF genes and correlates with H3K9me3 enrichment in K562 cells.

ATRX enrichment at ZNFs is conserved across distinct human cell types

Next, we investigated whether ATRX is bound at ZNF 3' exons in other cell types. Because ATRX, SETDB1 and TRIM28 ChIP-seq datasets are not readily available, we first utilized available ZNF274 and H3K9me3 ENCODE ChIP-seq data for cell lines GM12878, H1-hESC, HeLa-S3, HepG2, and NT2-D1 (see Table S6 for details of data sources and analysis). ZNF274 binds to the 3' of ATRX-bound Class I and Class II ZNF genes (as defined in K562) in the majority of cell lines analyzed (Fig. S4C). Similar to other previous studies,⁴⁴ H3K9me3 at the 3' region of ZNFs was also conserved (Fig. S4C).

Using a panel of cell lines of diverse origin from ENCODE (K562, H1-hESC, HeLa-S3, and HepG2) and 2 neuroblastoma cell lines (LAN6 and SKNFI), we performed ATRX ChIP-qPCR for 11 randomly chosen Class I ZNFs distributed across different chromosomes, as well as 2 Class III ZNF genes as negative controls. We found significant enrichment of ATRX in Class I ZNFs across all cell lines analyzed, and none for Class III ZNFs or IgG control (Fig. 3A). Using available ChIP-seq data, we plotted the enrichment of H3K9me3 and ZNF274 at these specific ZNFs in the ENCODE cell line panel. As expected, all ATRX-enriched ZNFs also show relatively high levels of H3K9me3 and ZNF274 (Fig. 3B). Furthermore, these ZNFs show diverse expression levels independent of their ATRX binding status according to the normalized RNA-seq signal from ENCODE (Fig. 3C). This result substantiates the

notion that ATRX binding is independent of the transcriptional status of the ZNFs.

In order to corroborate our findings genome-wide, we performed ATRX and H3K9me3 ChIP-seq in LAN6, which showed robust ATRX enrichment by ChIP-qPCR (Fig. 3A). Strikingly, we found that the binding of both ATRX and H3K9me3 at the 3' of ZNF genes is conserved in LAN6 (Fig. 3D, compare with Fig. 2D). This binding was specific and highly significant for ZNF gene bodies (Fig. 3E), and as observed in K562 cells, ATRX and H3K9me3 demarcate the ZNF clusters on chromosome 19 in LAN6 (Fig. 3F,G).

Next, we re-analyzed the only other published human ATRX ChIP-seq dataset from primary human erythroblasts¹⁵ and 2 recent datasets from mouse embryonic stem cells (mESC) and mouse embryonic fibroblasts (MEF),⁴⁶ and assessed ATRX binding at ZNFs. In contrast to K562, LAN6, and all other cell lines analyzed, human erythroblasts, mESCs, and MEFs lack ATRX binding within ZNF gene bodies (Fig. S4D,E). We then re-analyzed available ChIP-seq data for H3K9me3 in mESCs and MEFs (Table S6) and found that H3K9me3 enrichment at ZNFs does not co-localize with ATRX, and only mESC showed H3K9me3 enrichment at the 3' exons of ZNFs (Fig. S4E). It is unclear if these differences are biologically significant and it will be interesting to define if the pattern observed in humans is also conserved in other vertebrates.

Overall, our analysis demonstrates that the enrichment pattern of ATRX, H3K9me3, and ZNF274 at ZNF 3' exons is conserved in several human cell lines of diverse lineages, including non-tumorigenic cells, suggesting a general mechanism for ZNF gene regulation via ATRX that is independent of transcription.

ATRX deficient cells have reduced H3K9me3 enrichment at 3' exons of ZNFs

To functionally investigate the role of ATRX, we generated 2 clonal ATRX knock out (KO) cell lines using CRISPR/Cas9 genome editing in K562 cells. Our KO lines (ATRX KO1 and KO2) were characterized in detail (Fig. S5A, Methods). As a control, we used a clonal cell line overexpressing Cas9 alone (referred to as V2). Immunoblot and sequencing analyses showed that ATRX KO1 cells retain residual ATRX from an allele with an in-frame deletion, while KO2 cells are completely devoid of ATRX protein (Fig. 4A, S5A). Moreover, the global levels of SETDB1, TRIM28, and ZNF274 in chromatin were largely unaffected by ATRX KO (Fig. S5B). We next performed ATRX ChIP-qPCR in these KO lines for the panel of Class I and Class III ZNFs. In accordance with our ChIP-seq data, Class III ZNFs lacked ATRX enrichment, while Class I ZNFs were enriched for ATRX in the control, reduced in KO1 and ablated in KO2 cells (Fig. 4B).

Loss of ATRX has been shown to promote genomic instability and defects in cell cycle.³⁷⁻⁴⁰ We tested these functional readouts in our KO cell lines and found that ATRX KO K562 cells have increased DNA damage as compared to control cells as assessed by Comet assays and immunoblots for γ H2A.X (Fig. S6A-C). Furthermore, ATRX KO cells displayed a slight,

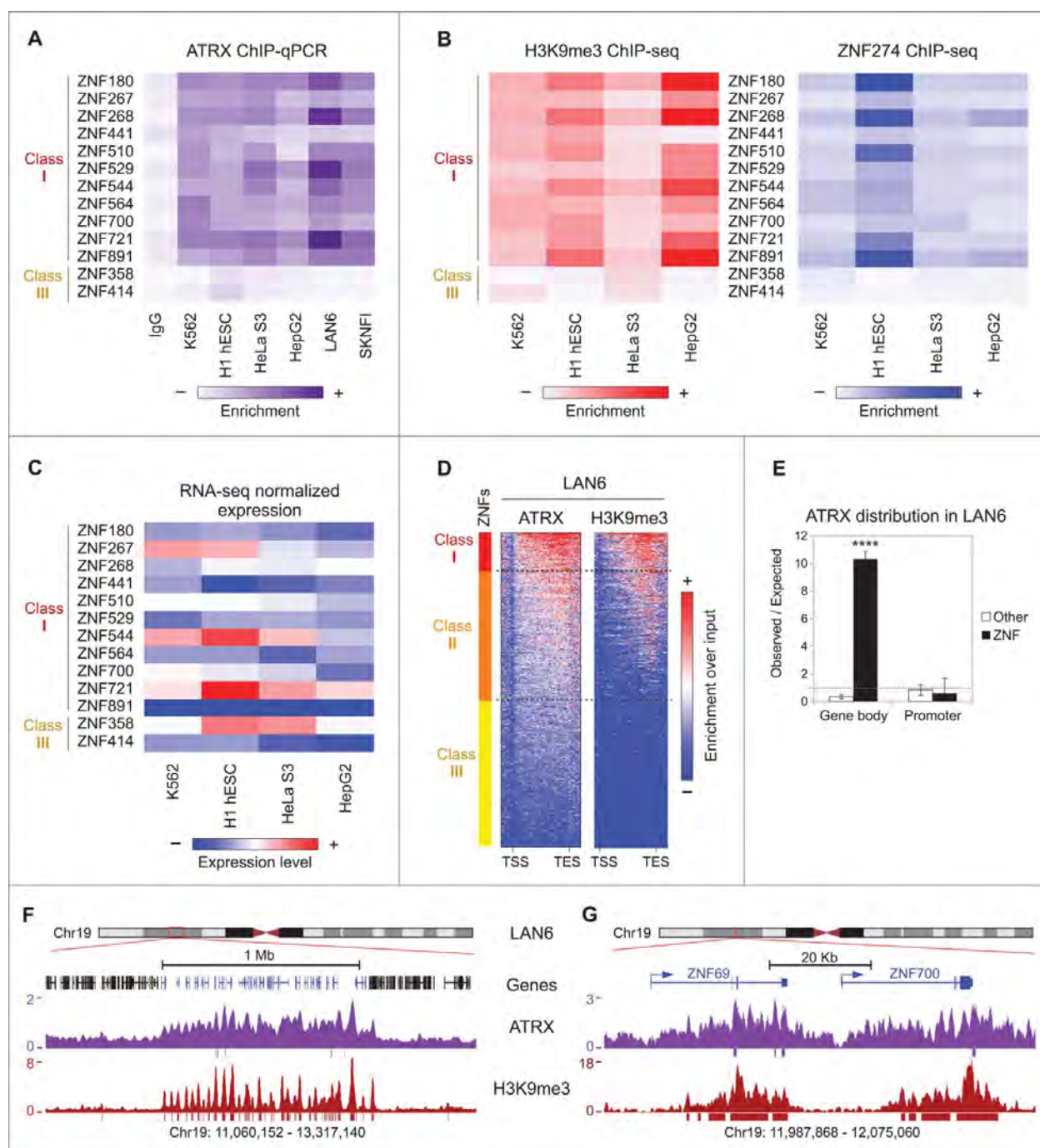


Figure 3. ATRX binding to ZNF 3' exons is conserved across human cell lines. (A) Heatmap of ATRX ChIP-qPCR enrichment over ZNF genes in several human cell lines. The color represents the average enrichment of at least 2 independent biological replicates per cell line. ZNF class I genes show significant enrichment as compared to IgG for all assessed cell lines (for details see Table S3). (B) H3K9me3 (left) and ZNF274 (right) ChIP-seq enrichment over the same panel of ZNFs as in (A) for cell lines with available ENCODE data. (C) RNA-seq normalized expression values ($\log_2(\text{RPKM}+1)$) from ENCODE for the panel of ZNF genes shown in (A) and (B). (D) Metagene profiles of ATRX and H3K9me3 ChIP-seq data over ZNFs gene bodies ± 1 kb in LAN6 neuroblastoma cell line. (E) Observed over expected random distribution of significant ATRX peaks in ZNFs and non-ZNF (other) promoters and gene bodies for LAN6. (F) UCSC Genome browser screenshot of a typical ZNF cluster (genes in blue) on chromosome 19. The enrichment over input signal for ATRX and H3K9me3 ChIP-seq in LAN6 is shown. Significant peaks represented as bars below enrichment tracks. (G) Zoomed in snapshot of 2 ZNFs contained within the ZNF cluster shown in (F).

but reproducible, cell cycle defect in the G1/S transition (Fig. S6D, E).

We then queried whether ATRX deficiency alters the chromatin state of the atypical chromatin domains at ZNF genes. To address this question, we performed native ChIP for H3K9me3 and H3K36me3 in control and ATRX KO cells, followed by qPCR analysis for ZNFs. Strikingly, H3K9me3 levels decreased at Class I ZNFs in the ATRX KO cells compared to

control cells, and displayed a similar pattern to that of ATRX occupancy (Fig. 4B, C). In contrast, H3K36me3 levels remain stable upon ATRX loss (Fig. S7A). We also assessed whether ATRX depletion affected the transcription of ZNF genes. Consistent with our findings that ATRX enrichment at ZNFs does not correlate with their transcriptional activity (Fig. 2C, 3C), we did not observe changes in expression levels of ZNFs in the ATRX KO cell lines (Fig. S7B). Together, these data suggest

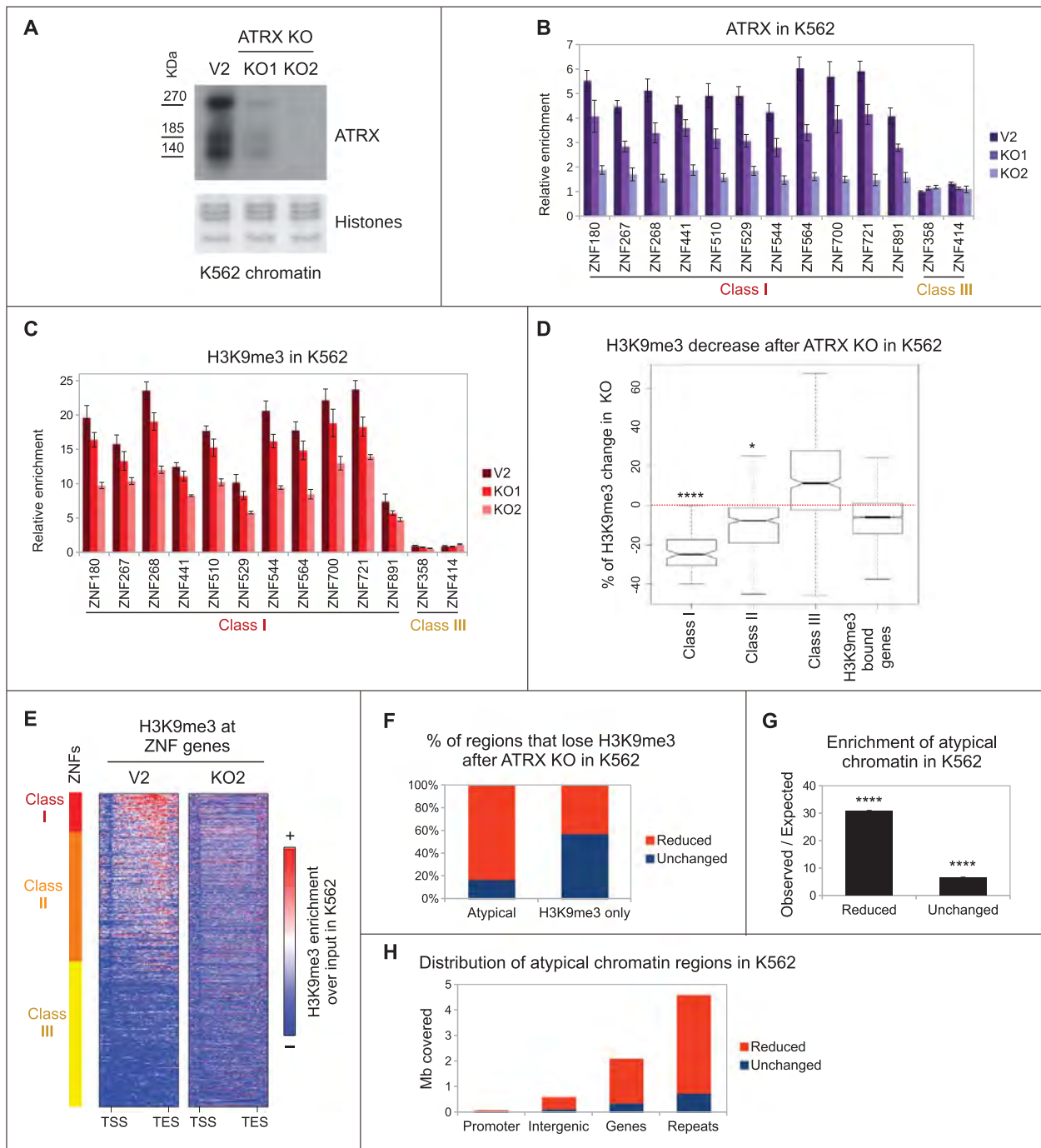


Figure 4. ATRX deficient cells have decreased levels of H3K9me3 at 3' exons of ZNFs and other atypical chromatin regions. (A) Western blot of ATRX in chromatin preparations from control (V2) and 2 CRISPR ATRX KO K562 cell lines. Amido Black staining of histones is shown as a loading control. (B) ATRX ChIP-qPCR over ZNF genes in control and ATRX KO K562 cell lines. (C) Same as in (B) with H3K9me3 native ChIP-qPCR. In both (B) and (C), bars represent average of at least 3 biological replicates and the error bars represent the SEM. (D) Box plot of the H3K9me3 changes (represented in % with respect to the control) observed in the ZNF classes and non-ZNF H3K9me3-bound genes upon ATRX KO. Asterisks show significant changes with respect to the non-ZNF genes (* P -value < 0.05; **** P -value < 1×10^{-4}). (E) ChIP-seq metagene profile of H3K9me3 in control (V2) and ATRX KO (KO2) K562 cell lines over ZNFs. (F) Distribution of reduced and unchanged H3K9me3 regions after ATRX KO in atypical chromatin and H3K9me3-only regions. (G) Observed over expected distribution of reduced and unchanged regions at atypical chromatin. Asterisks show significantly overrepresented regions (P -value < 1×10^{-4}) (H) Genomic distribution of atypical chromatin regions.

that ATRX regulates H3K9me3 levels at ATRX-bound ZNF genes, but that H3K36me3 and transcriptional regulation is ATRX-independent.

Because ZNF274 binds DNA and recruits TRIM28, which in turn recruits SETDB1, we speculated that the ZNF274/TRIM28/SETDB1 complex serves as a scaffold for ATRX at the 3' exons of ZNFs. Therefore, we posited that ATRX KO would

not affect the complex binding to the chromatin, but rather its capacity to deposit or maintain H3K9me3. As expected, ZNF274 and TRIM28 binding remains largely unchanged at ZNF genes after ATRX KO (Fig. S7C, D). While ATRX KO1 cell line did not affect SETDB1 binding, SETDB1 occupancy was increased in ATRX KO2 (Fig. S7E). This may reflect a compensatory effect for the complete loss of ATRX, although we

cannot rule out a clone-specific effect. These data broadly favors a model in which ZNF274/TRIM28/SETDB1 is able to bind to the 3' exons of ZNFs independently of ATRX, but ATRX is required to establish or maintain H3K9me3 at atypical chromatin of ZNF genes.

To investigate the genome-wide alterations of H3K9me3 after ATRX depletion, we performed native ChIP-seq for H3K9me3 in control and ATRX KO cells. We first analyzed the H3K9me3 enrichment at ZNF genes. In agreement with our ChIP-qPCR results, we found H3K9me3 levels significantly reduced in KO cells, specifically in ATRX-enriched Class I and II ZNFs (Fig. 4D). These results were confirmed by meta-gene analysis of ZNFs (Fig. 4E) and visualization of H3K9me3 peaks at the ZNF clusters on chromosome 19 (Fig. S8A, B). This demonstrates that ATRX is required for establishing and/or maintaining H3K9me3 levels at the 3' end of a subset of ZNFs.

H3K9me3 within atypical chromatin domains is sensitive to ATRX, but not DAXX, depletion

Besides ZNF genes, we identified additional genomic regions that show reduction of H3K9me3 upon ATRX depletion (Fig. S8C). To further understand the nature of the H3K9me3 regions that are sensitive to ATRX loss, we divided all H3K9me3 regions to “reduced” (n = 6,055) or “unchanged” (n = 8,331) based on whether the levels are decreased or remain stable in the ATRX KO cell line, respectively. We then analyzed the proportion of reduced regions in atypical chromatin and compared it to H3K9me3-only chromatin. Strikingly, a large proportion of reduced regions (~80%) were atypical chromatin in contrast to H3K9me3-only chromatin (~40%) (Fig. 4F). This trend is highly significant (Fig. 4G). Thus, our data suggest that the levels of H3K9me3 at atypical chromatin are particularly sensitive to ATRX depletion.

We next analyzed the genomic distribution of atypical chromatin. Interestingly, as shown in Fig. 4H, most atypical chromatin regions that lose H3K9me3 are located in repetitive sequences, followed by those located in genes (including ZNF genes). This is consistent with recent studies that show a role for ATRX in regulating H3K9me3 at repetitive elements; however, the levels of H3K36me3 at such regions have not been described.^{7,20-22,25}

ATRX has been reported to regulate H3K9me3 levels in concert with DAXX at a subset of repetitive and imprinted regions.^{7,20-22,25} In order to assess whether DAXX regulates H3K9me3 at ZNFs, we used CRISPR/Cas9 to generate DAXX KD and double DAXX KD + ATRX KO lines (see Methods for details). After confirming the knock down of DAXX (Fig. S9A), we performed H3K9me3 ChIP-qPCR for Class I and Class III ZNFs. We found little to no change in H3K9me3 levels in the single DAXX KD (Fig. S9B). The DAXX KD/ATRX KO shows a consistent decrease of H3K9me3 as observed with the single ATRX KO2 cell line (Fig. S9B). While we can not completely rule out a role for DAXX in the maintenance of H3K9me3 levels at ZNFs, our results suggests that its contribution is minor in comparison to other factors, such as ATRX and ZNF274 (see below). Overall, these results suggest that ATRX protects H3K9me3 levels at atypical chromatin regions genome-wide, and that ZNF genes comprise a subset of these affected regions.

ZNF274 recruits ATRX to a subset of ZNF genes to regulate H3K9me3 levels and genomic stability

To further understand the interplay between ATRX and ZNF274, we depleted ZNF274 using CRISPR/Cas9 in K562 cells. We generated 3 clonal cell lines (see Methods for details): a single ZNF274 KO cell line (referred to as Z274 KO), a double ZNF274/ATRX KO cell line (Z274/ATRX KO) and a control cell line expressing a non-specific sgRNA (Random). We characterized the mutation status of the ZNF274 gene (Fig. S10) and confirmed that the Z274 KO abrogated ZNF274 binding to ZNF genes by ChIP-qPCR (Fig. 5A).

We next analyzed ATRX binding in the ZNF274 KOs by ChIP-qPCR. Strikingly, we observed a decreased binding of ATRX in the Z274 KO cells for most Class I ZNFs. As expected, ATRX is depleted from all ZNFs in the Z274/ATRX KO (Fig. 5B). These results strongly suggest that ZNF274 tethers ATRX to ZNFs. However, because ATRX is not lost at all ZNFs examined, we hypothesize that there may be other (potentially compensating) factors that can recruit ATRX to these regions.

To determine the impact of ZNF274 depletion on H3K9me3, we performed ChIP-qPCR in the ZNF274 KO lines. In order to compare the effect of the double Z274/ATRX mutant, we included the ATRX KO2 cell line in our analysis. As shown in Fig. 5C, the ZNFs that have lost ATRX binding in the ZNF274 KO cells also lose H3K9me3 enrichment. In contrast, the ZNFs that still retain some ATRX binding show a more modest reduction in H3K9me3 levels. Strikingly, in many cases, the Z274/ATRX KO cells showed a stronger decrease of H3K9me3 than either individual KO, suggesting that knocking out a single component of the complex may not entirely abrogate its function.

Finally, we determined the functional consequence of H3K9me3 reduction in the ZNF274 KOs. We noted an increase in γ H2A.X in the Z274 KO cell line as compared to the control, which was even stronger in the Z274/ATRX KO cells (Fig. 5D). Because the ATRX KO cells show a defect in cell cycle, we queried whether the ZNF274 KO would have a similar phenotype. Indeed, ZNF274 KO cells show a delay in G1/S (Fig. 5E,F), comparable to that observed in the ATRX KO2 cells.

Collectively, we propose a model in which ATRX is tethered to the 3' exons of ZNF genes by the ZNF274/TRIM28/SETDB1 complex to establish or maintain/protect H3K9me3 at these transcriptionally active regions (Fig. 6). ATRX and ZNF274 depletion leads to reduction of H3K9me3 levels at ZNF 3' exons and is associated with higher DNA damage and cell cycle defects. Furthermore, ATRX acts as a genome-wide regulator of H3K9me3 levels at atypical chromatin that is not restricted to ZNFs. Finally, we propose that impairing ATRX function has important consequences for the genomic stability and evolution of ZNFs clusters, as the loss of H3K9me3 at ZNF 3' exons may increase the probability of HR between them (Fig. 6; see discussion below).

Discussion

ATRX is an important chromatin regulator involved in diverse processes, such as transcriptional regulation, maintenance of imprinted loci, replication, genome stability, and chromatin looping.^{7-9,30,54} Here we report that ATRX binds

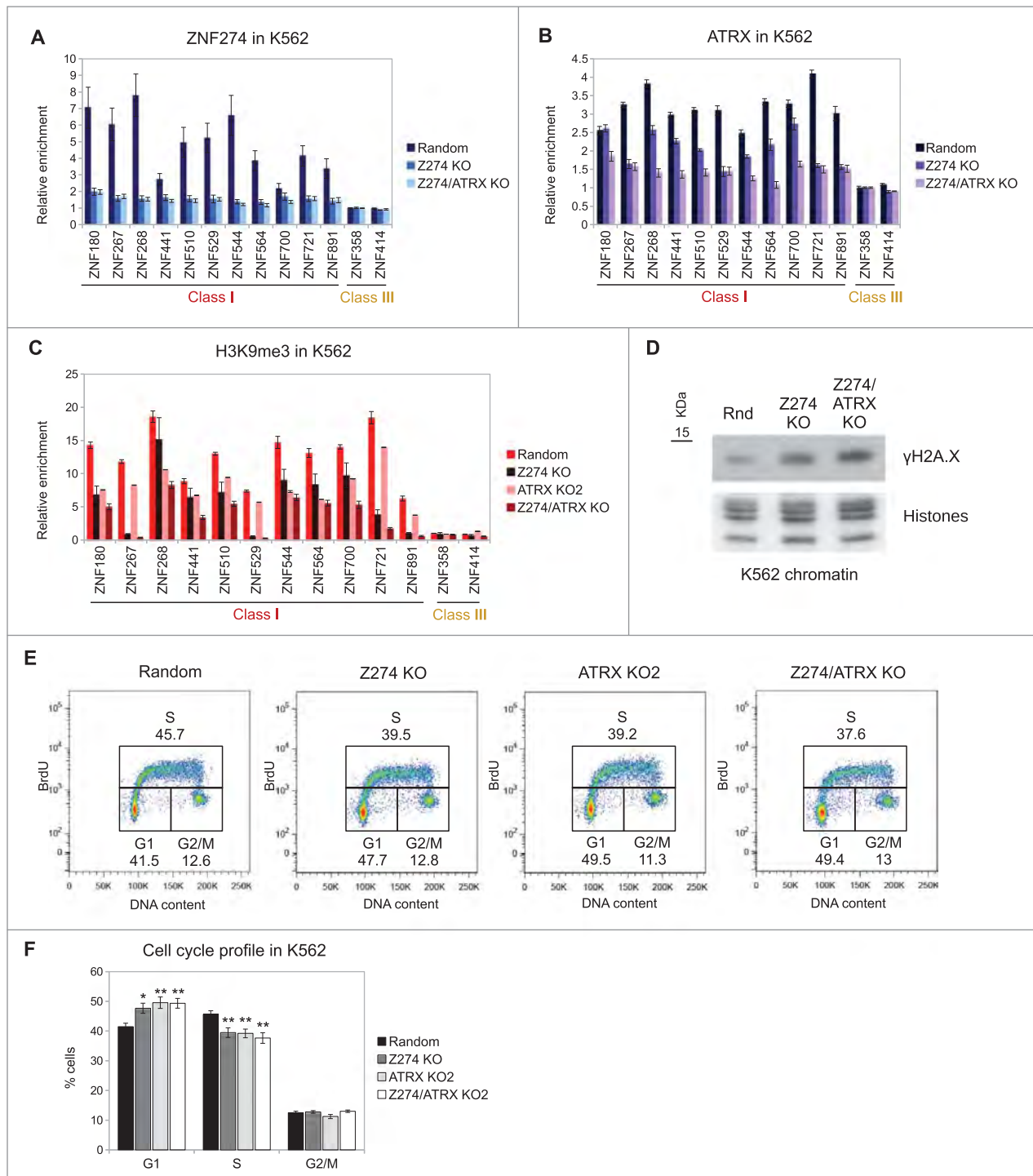


Figure 5. ZNF274 KO reduces ATRX and H3K9me3 levels at ZNFs. (A) ZNF274, (B) ATRX and (C) H3K9me3 ChIP-qPCR in K562 ZNF274 KO and K562 double ZNF274/ATRX KO at ZNF genes. Single ATRX KO2 cells are used for the H3K9me3 ChIP for comparison. In all graphs, the bars represent the average of at least 2 independent biological replicates. Error bars depict SEM. Results of statistical comparisons in Table S3. A non-specific sgRNA (random) used as control (see Table S7). (D) Chromatin immunoblot of γ H2A.X in control (Rnd) and ZNF KO K562 cells. Histones used as loading control. (E) Representative K562 cell cycle profiles of control (random), ATRX and ZNF274 single and double KO assessed by BrdU/PI staining. $n \geq 6$ biological replicates. (F) Graph depicting quantifications of (E). The bars show the average % of cells in each phase, error bars depict SEM. Asterisks show significant changes compared to the control (* P -value < 0.05; ** P -value < 0.01).

the 3' exons of ZNFs with an atypical chromatin signature to establish or maintain H3K9me3 (Fig. 6). ZNFs represent the fastest expanding gene family encoding transcription factors in humans, with the greatest diversity of target sequences.^{41-43,55} This property of ZNFs to recognize a multitude of targets is generated through diverse combinations of their zinc

finger motifs.⁵⁵ The average ZNF gene contains 9 independent zinc finger motifs and the number of motifs varies from gene to gene. Interestingly, these motifs are almost always located at the 3' exons of ZNF gene loci and harbor an atypical chromatin signature consisting of both H3K9me3 and H3K36me3 (Fig. 6).^{42,44}

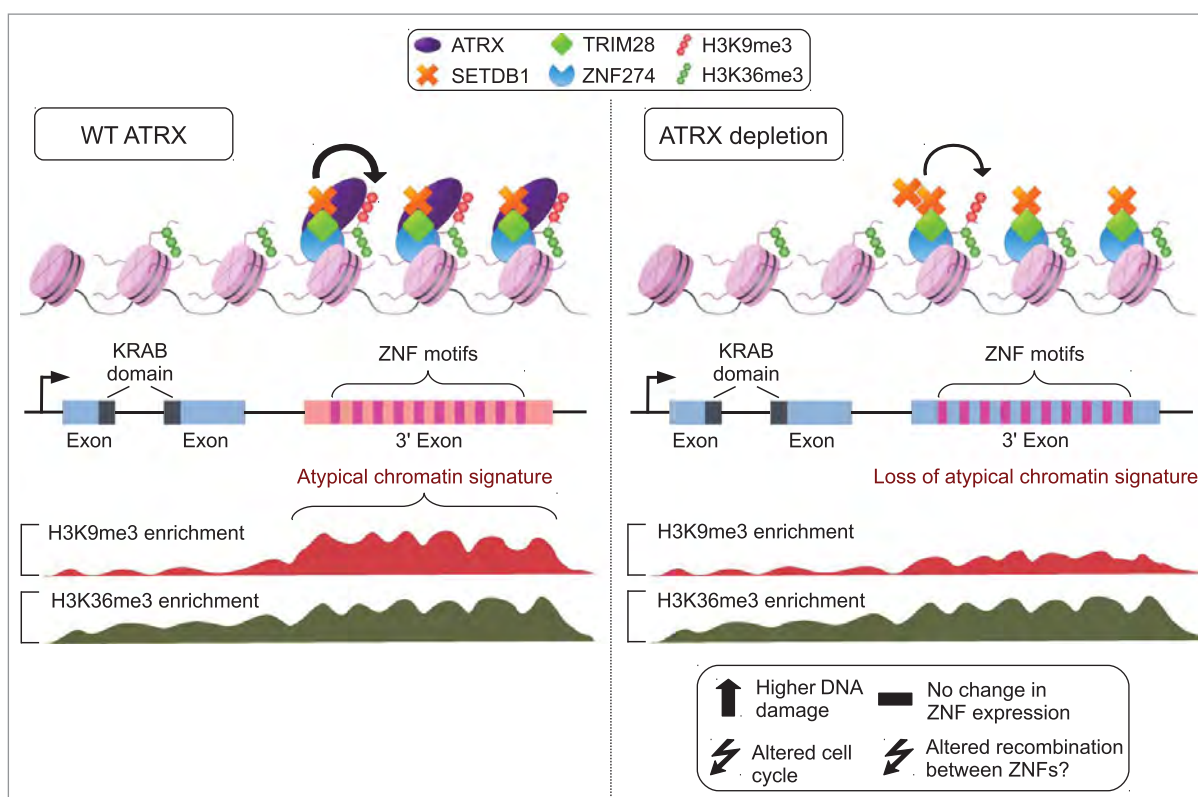


Figure 6. Model of ATRX regulation at ZNF 3' exons. Left: ATRX forms a complex with ZNF274, TRIM28 and SETDB1 to facilitate the deposition and maintenance of H3K9me3 at ZNF 3' exons. The presence of the mark establishes an atypical H3K9me3/H3K36me3 domain. Right: Upon ATRX depletion, H3K9me3 and the atypical chromatin domains at ZNF 3' exons are lost. Loss of ATRX induces altered cell cycle, increased DNA damage and possibly recombination between ZNFs.

Through an unbiased ChIP-seq approach and analysis of ENCODE data, we found that ATRX preferentially binds to the 3' exons of a subset of ZNF genes containing this atypical chromatin signature. These ZNF genes are distinguished by the presence of a KRAB repressor domain, a higher than average number of zinc finger motifs, low levels of G-content and low probability of forming G-quadruplexes. Previous studies have shown that ATRX recognizes and resolves G-quadruplexes, relevant in the context of gene regulation.¹⁵ However, our data suggests that ATRX binds and regulates these atypical chromatin regions by an alternative mechanism(s).

We found that ATRX co-localizes and interacts with the previously reported ZNF274/TRIM28/SETDB1 complex⁵³ at ZNF 3' exons with an atypical chromatin signature. Our data suggests a model whereby ZNF274 recognizes a DNA motif that is common in certain ZNF 3' exons (i.e. those bound by ATRX), and recruits the TRIM28/SETDB1/ATRX complex to establish, maintain or protect H3K9me3. Of course, there are questions about this model that require further clarification. For example, we wonder why in the absence of ZNF274, ATRX still retains its ability to bind a subset of ZNFs. These results, coupled to the fact that we find atypical chromatin regions that are enriched in TRIM28/SETDB1/ATRX but not ZNF274 (data not shown), suggests that other ZNF transcription factors may be involved in targeting ATRX to specific loci. Thus, the TRIM28/SETDB1/ATRX complex may be guided to unique loci or chromatin environments via ZNF proteins that may have cell- or stage-specific binding patterns. Moreover, while it has been established that KRAB-ZNF proteins recruit TRIM28

through their KRAB domain, and that TRIM28 in turn recruits SETDB1,⁵⁶ the dynamics of ATRX recruitment to the complex are unclear. One possibility is that SETDB1 establishes basal H3K9me3 levels that are recognized and bound by the ATRX ADD domain, which in turn promotes further H3K9 methylation forming a feedback loop to maintain H3K9me3 levels.

Our loss-of-function studies revealed that H3K9me3 levels at atypical chromatin regions, in particular those at ZNF 3' exons, are decreased after ATRX loss. Recent studies have reported similar effects upon ATRX loss at known ATRX target regions such as ERVs and imprinted loci.^{20-22,25} However, the mechanisms underlying H3K9me3 loss in ATRX-deficient cells remain unclear. Based on our observations, we hypothesize 3 non-exclusive scenarios: 1) ATRX may facilitate SETDB1-mediated H3K9me3 deposition by promoting an optimal nucleosome structure for SETDB1 function, 2) ATRX binds H3K9me3 and blocks demethylase activity, and 3) ATRX may help to reestablish H3K9me3 after transcription or replication. These scenarios would also explain why H3K9me3 levels decrease after ATRX KO, despite the fact that SETDB1 binding remains unchanged.

Functionally, loss of H3K9me3 in ATRX and ZNF274 KO cells may impact on the genomic stability of ZNFs. The ZNF family has expanded in the primate lineage through gene duplication.⁴¹⁻⁴³ ZNFs contain long stretches of highly similar sequences in their 3' exons and high levels of H3K36me3, a mark that enhances the resolution of double strand breaks (DSB) through the HR pathway.^{57,58} Therefore, ZNF regions may be prone to recombine. As H3K9me3-rich

heterochromatin regions are refractory to repair of DSB and HR,^{59,60} we hypothesize that ATRX-mediated H3K9me3 enrichment at ZNF 3' exons protects them from HR. In this regard, recent studies demonstrated that ATRX can act as a suppressor of recombination at telomeres.^{22,36,47,61} Furthermore, the notion that H3K9me3 at ZNF 3' regions protects from HR has been previously suggested in at least 2 independent studies.^{44,49} If this hypothesis is correct, ATRX may function, at least partially, as an epigenetic modulator of ZNF gene expansion, balancing the positive selective pressure for ZNF duplication and the deleterious effects that an excessive recombination rate has for genomic stability. Moreover, altered recombination levels at ZNFs, telomeres and other ATRX-regulated regions may be responsible, at least in part, for the defects in cell cycle and increased DNA damage we observed in ATRX KO cells.

Despite the evidence that ATRX suppresses recombination at certain genomic regions, there are alternative explanations for its role at ZNFs. For instance, we find that Class I ZNFs tend to be late replicating. Thus, it is possible that ATRX facilitates DNA replication at atypical chromatin regions. This idea is in line with the finding that ATRX KO cells subjected to replicative stress have a higher accumulation of stalled replication forks.^{39,40} Moreover, a replication defect can also explain the higher levels of DNA damage and the cell cycle delay observed in ATRX and ZNF274 KO cells. Furthermore, it has been suggested that the chromatin environment can affect RNA polymerase processivity, in turn affecting the RNA splicing machinery.⁶² Thus, the levels of H3K9me3 at ZNFs could affect alternative splicing. Because the RNA-seq analyses performed in this study were focused only on global expression levels, potential changes in isoform expression, polymerase processivity, and other related mechanisms, remain to be explored.

Several recent studies have found that the ATRX/DAXX complex deposits the H3.3 histone variant at several H3K9me3-containing regions, such as telomeres, imprinted regions, and retrotransposons.^{7,20-22,25} Interestingly, H3K9me3 levels in those regions are frequently decreased upon ATRX inactivation. How similar these H3.3/H3K9me3 regions are to the H3K9me3/H3K36me3 atypical chromatin signature remains an open question. While our data suggests that DAXX has a minimal role in maintaining H3K9me3 at ZNF 3' regions, we cannot exclude the possible involvement of H3.3.

In summary, we demonstrate here that ATRX regulates H3K9me3 levels at the 3' exons of ZNFs and other loci containing an atypical chromatin signature. This unexpected function sheds light onto the complex genomic regulatory pathways that ATRX participates in, and may be important for the future understanding of diseases in which ATRX is mutated or altered.

Materials and methods

XL-MNase ChIP

Cross-linked-MNase ChIP was performed with Cell Signaling SimpleChIP Enzymatic Chromatin IP Kit (cat. #9003) following manufacturer's instructions with modifications. Briefly, K562 and LAN6 ($2-4 \times 10^6$) cells were cross-linked with 1% formaldehyde in PBS, 10 mM NaCl, 0.1 mM EDTA pH 8.0,

5 mM HEPES pH 8.0, for 10 min at room temperature. Reaction was quenched with 125 mM glycine. Cells were lysed to obtain nuclei and chromatin was digested with Micrococcal Nuclease (MNase) (NEB, cat. #M0247S) at 37°C for 20 min. Nuclei were disrupted by brief sonication (4 cycles, 20 sec ON/OFF, high power) in a Bioruptor Twin. Chromatin was quantified and 40–80 μ g was incubated with antibody at 4°C overnight with 1% taken as input sample. After incubation, Magna protein A/G magnetic beads (Millipore, cat. #16-663) were added for 3 h at 4°C. Beads were washed following the manufacturer's protocol, followed by an extra LiCl buffer wash (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 1% sodium deoxycholate, 1% Igepal, 250 mM LiCl). DNA bound to complexes was eluted at 65°C for 30 min, treated with RNase A (10 mg/ml) for 1 h at 37°C, with Proteinase K (20 mg/ml) for 3 h at 55°C and then cross-linking was reversed for 4–6 h at 65°C. DNA was purified using the Qiagen MinElute PCR purification kit and subsequently analyzed and quantified using an Agilent 2100 Bioanalyzer High Sensitivity Kit.

Native ChIP

Native ChIP was performed as previously described⁶³ with minor changes. Briefly, nuclei isolated from K562 and LAN6 ($4-6 \times 10^6$) cells were treated with MNase (Affymetrix Cat #70196Y) and ~40–80 μ g of digested chromatin was immunoprecipitated with specific antibodies. The immunoprecipitated material was treated with Proteinase K for 3 h at 56°C and purified using the Qiagen MinElute PCR purification kit.

Antibodies

See Table S4 for a full list of antibodies and concentrations used for each assay.

Library preparation and ChIP-seq

ChIP-seq libraries were prepared as previously described⁶³ and libraries were sequenced single-end on Illumina HiSeq 2500 and NextSeq 500. See Table S5 for detailed descriptions (number and size of reads, etc.) of the sequenced samples.

ChIP-seq analysis

Sequenced reads were aligned to the GRCh37 (hg19) assembly using Bowtie 1.0.0.⁶⁴ Redundant reads were eliminated using MACS2 (2.1.0)⁶⁵ filterdup option with default parameters. The estimated background reads and optimal normalization factor between ChIP and Input samples was calculated with the R NCIS package 1.0.1⁶⁶ using a shift size of 75bp. Peak calling was performed with MACS2 callpeak using the $-\text{ratio}$ option with the estimated value from NCIS. ChIP/Input fold enrichment pileups were created with the MACS2 bdgcmp tool using the $-\text{m FE}$ option and converted to bigWig files using the bedGraphToBigWig program (v4) from the UCSC binaries. For some samples, a second peak calling was performed using SICER⁶⁷ 1.1 using a window of 200 bp, allowing gaps of 400 bp and filtering for q-values $< 1 \times 10^{-8}$ (see Table S6 for details). All peaks overlapping with ENCODE blacklisted regions⁴⁵

were eliminated. When available, a final list of peaks overlapping the MACS2 and SICER peaks was obtained from the intersectBed program from bedtools 2.17.0.⁶⁸ See Table S6 for GEO accessions, details of analysis parameters (q-values, etc.), and results for all the datasets analyzed. For all the ATRX-related analyses, the K562 ATRX-H300 ChIP-seq sample was used due to its high quality, low percentage of background, and high number of reads (Table S5).

Gene and coordinate analyses

All analyses were performed using Ensembl genes (putative genes, pseudogenes and genes unmapped to chromosomes were excluded). For human the Ensembl genes 75 version (GRCh37.p13) was used. For mouse the Ensembl genes 77 version (GRCm38.p3) was used. The lists of ZNF genes were downloaded from Biomart using the InterPro id for the Zinc Finger C2H2 domain (IPR007087) as a filter, manually analyzed and curated. Repeats coordinates were obtained from the repmask table from the UCSC Genome Browser for hg19. Bed files for subsequent analysis were generated using the following coordinates: promoters (−3 kb to TSS), intergenic (regions falling outside genes or promoters).

Correlation analysis

Correlation heatmaps were generated with the bigwigCorrelate program from the DeepTools suite (v 1.5.9.1)⁶⁹ using the spearman correlation method. Fold enrichment over input bigwig files were used as inputs. For genome-wide analysis, all chromosomes were divided in 10kb non-overlapping bins. Bins falling in the blacklisted regions from ENCODE were excluded. For gene-specific analysis, a bed file with the coordinates from TSS to TES was used.

Analysis of genomic distribution and correlation with chromatin states

Bed files with the coordinates of the chromatin states for K562 calculated in⁴⁸ were generated. Similar chromatin states were merged into single categories (states 1–3 were fused as promoters, states 4–7 as enhancers, 10–11 as transcription, 14–15 as repetitive). The probability of overlap between ATRX peaks and the HMM states, genes, promoters and intergenic regions was calculated with The Genomic HyperBrowser.⁷⁰ ATRX peaks were randomized 10,000 times preserving its segment length and inter-segment gaps. Observed/expected values were calculated by dividing the overlap of the ATRX peaks over the overlap of the randomized regions.

Gene ontology analysis

Genes overlapping ATRX significant peaks were obtained using intersectBed from the Bedtools suite. Gene Ontology analysis was performed with DAVID⁷¹ using default parameters. See Table S1 for the list of genes overlapping ATRX peaks.

Generation of ZNF classes

Average ATRX ChIP/Input enrichment per gene was obtained with the computeMatrix program from the DeepTools suite. The genes were then clustered by kmeans into 3 groups using R (v 3.0.1). Bed files from each ZNF Class are provided in Additional File 1.

Protein domain analysis

The protein sequence of the ZNFs was obtained with Biomart. The sequences were matched to all the Prosite motifs using the ScanProsite tool with default parameters. In-house scripts were used to parse the ScanProsite output and count the number of motifs per ZNF gene.

G-content and G-quadruplex analysis

G-content at the 3' region of ZNF genes (last 3 kb) was calculated with in-house scripts. The probability of G-quadruplex formation was calculated using the quadparser program⁷² with default parameters.

Genetic features plot

The color plots showing the distribution of genetic features among ZNF classes were drawn using the image function of R. Darker colors represent presence of KRAB domains, a higher number of zinc finger motifs, high G content at the C-terminal ZNF region (last 3 kb of the gene) and presence of sequences predicted to form G-quadruplexes. For RNA-seq data, the Z-score of the RNA-seq normalized signal ($\log_2(\text{RPKM}+1)$) from K562 ENCODE data was and plotted. The calculated values used to generate color plots of [Figure 2C](#) are provided in Additional File 2.

Metagene analysis

ZNF genes were ordered by ATRX enrichment from high to low. The metagene plots were generated with DeepTools using the ordered gene file and the ChIP/Input bigwig tracks. All genes from TSS to TES were scaled to a 5 kb region \pm 1 kb with sliding windows of 100 bp. Metagene values were calculated using computeMatrix. Values were plotted with heatmapr and the average enrichment profiles were plotted with profiler. The matrices for all heatmaps and profile plots are provided in Additional File 3. The data sources can be found in Table S6.

Motif analysis

The ZNF274-bound regions were obtained from the summit peaks generated by MACS2. A fasta file from these coordinates was created using in-house scripts and used as input for MEME-ChIP (v. 4.1.0.0).⁷³ The sequences were compared to the JASPAR vertebrate database; other options were set as defaults. The regions matching the motifs (fimo output) were parsed, sorted by class, and counted using in-house scripts.

Repli-seq analysis

Repli-seq bigwig files generated by the ENCODE project for K562 were analyzed using DeepTools. The average signal per ZNF was calculated with computeMatrix and then plotted in R using the heatmap.2 function. Calculated Repli-seq values are provided in Additional File 4.

Co-immunoprecipitation of chromatin bound proteins

K562 cells were harvested for each IP (50×10^6) using PBS supplemented with protease inhibitor without EDTA, 1 mM DTT and 0.05% NP-40 and incubated on ice for 3 min and centrifuged at $500 \times g$ for 5 min at 4°C. The nuclei were washed with PBS and re-suspended in EX 100 buffer (10 mM Hepes pH 7.6, 100 mM NaCl, 1.5 mM MgCl₂, 0.5 mM EGTA, 10% glycerol) and digested with MNase (NEB, cat. #M0247S) at 37°C for 10 min. The reaction was stopped with EGTA (10 mM) followed by spinning ($10,000 \times g$ for 10 min at 4°C). The collected supernatant containing mononucleosome particles was used for IP. ATRX antibody (HPA001906; 4 μg) or Rabbit IgG were added to each IP and incubated overnight (rotating at 4°C). Protein A+G magnetic bead slurry (Millipore, cat. #16-663; 25 μl) were washed twice with PBS and once with EX 100 buffer and incubated with the immunocomplexes for 2.5 h at 4°C followed by 2 × 5 min washes with buffer G250 (50 mM Tris pH 7.5, 250 mM NaCl and 0.5% NP-40) and one 5 min wash with G150 (50 mM Tris pH 7.5, 150 mM NaCl and 0.5% NP-40). Laemmli sample buffer was added to each sample, boiled at 95°C, and run on NuPAGE 4–12% Bis-Tris gels (Thermo Fisher).

Generation of ATRX and ZNF274 KO cell lines by CRISPR/Cas9

Single guide RNAs (sgRNAs) targeting the exons of the ATRX and ZNF274 genes were designed⁷⁴ and cloned into the lentiCRISPR V2 vector (addgene #52961, see sgRNAs sequences in Table S7). Lentiviral production using HEK293T cells was performed using standard laboratory protocols. To generate stable cell lines, K562 ($\sim 1 \times 10^6$) cells were infected with virus for each CRISPR guide and either an empty lentiCRISPR V2 or a lentiCRISPR V2 containing a non-specific sgRNA as control. Infected cells were selected (puromycin 2 μg/ml; 3 weeks) and subsequently single clones were sorted into 96-well plates using an IMI5L cell sorter (BD Biosciences). Clones were grown in selection for 3–4 weeks and tested for KO by DNA sequencing and immunoblot analysis. For the generation of ATRX KO cell lines 2 individual clones from 2 independent sgRNAs were selected and expanded (KO1 and KO2) for further analyses. For the generation of ZNF274 KO cells, 2 independent sgRNAs were used; clones could only be obtained from one. One clone from each CAS9-only infection (V2) and non-specific sgRNA (random) were randomly chosen for controls. The four most likely off-target loci of each sgRNA were also sequenced to assure specificity with no off-target mutations observed in any of the clones (data not shown). For the generation of double KO clones, the ATRX KO2 clone was infected with lentiCRISPR V2 targeting ZNF274, selected, sorted and analyzed as above. In order to determine the allele-specific mutations, gDNA was extracted from each clone, amplified

with specific primers to the target sgRNA sequence (Table S7), cloned into the pCR4-TOPO vector, transformed into bacteria, and single bacterial colonies sequenced.

Generation of DAXX KD cell lines by CRISPR/Cas9

K562 cells were infected and selected as described above, with the exception that single cells were not sorted. Pools of KD cells were maintained in constant selection (puromycin 1 μg/ml) and analyzed by immunoblotting to assess the efficiency of the KD. The double ATRX KO/DAXX KD was obtained by infecting and selecting ATRX KO2 cell line. Pools of cells infected with random sgRNA were used as control.

Comet assay

The alkaline comet assay was performed as described with modifications.⁷⁵ K562 (1×10^5) cells were washed with cold PBS, resuspended, and diluted in 100 μL of 0.5% low melting point agarose to be pipetted onto slides covered with 1.5% agarose. Cells were lysed (2.5 M NaCl, 100 mM EDTA, 10 mM Tris, pH 10, 1% Triton and 10% DMSO) for 24 h at 4°C, incubated in electrophoresis buffer (300 mM NaOH, pH 13.1 mM EDTA) for 30 min and subjected to electrophoresis in the dark for 25 min at 25 V. Slides were neutralized three times with Tris buffer (0.4 M Tris, pH 7.5) for 5 min, dried with 100% ethanol and stained with ethidium bromide (20 μg/mL). Cells were imaged on a Nikon Eclipse® microscope and ≥ 100 random cells were analyzed with CellProfiler software. Cell images were segmented using pixel intensity of 0.5 as threshold to generate masks matching the nucleoid. The comet tail was calculated by subtracting the nucleoid-integrated intensity from the comet-integrated intensity. For each sample, a positive control with cells treated with hydrogen peroxide (H₂O₂) (100 μM for 30 min at 25°C) was analyzed concurrently. Experimental analysis was performed in a blinded fashion.

Cell cycle analysis

K562 (3×10^6) cells were pulsed with 10 μM of BrdU for 45 min, washed with PBS, and fixed in 70% cold ethanol for 1 h on ice. The cells were washed twice with cold PBS and incubated in 2M hydrochloric acid for 30 min at room temperature. After incubation, cells were washed twice with cold PBS and once with cold PBS-T solution (PBS, 0.2% Tween, 1% BSA), then stained with anti-BrdU FITC PBS-T (eBioscience, cat #11-5071-41) for 30 min in the dark. Cells were washed once with cold PBS-T, once more with cold PBS and then stained for 30 min on ice with a PI/RNase solution (PBS, 20 μg/ml PI 10 μg/ml RNase). FACS analysis (10,000 cells per sample) was performed using BD FACS Canto II and FlowJo 10.

H3K9me3 enrichment analysis

The significantly reduced H3K9me3 domains in the K562 ATRX KO2 H3K9me3 ChIP-seq were calculated with the SICER-df program.⁶⁷ Only regions overlapping significant H3K9me3 peaks in the V2 control were taken into account. H3K9me3 peaks in the H3K9me3 that were not significantly

reduced compared to V2 were marked as unchanged. Bed files with the reduced and unchanged regions coordinates are provided in Additional File 5.

Statistical analyses

The hypergeometric test and permutation tests implemented in the R coin package (v 1.0–24) were the primary statistical tests utilized (See Table S2 for details). See Table S3 for the results of all the performed tests.

ChIP-qPCR and RT-qPCR

ChIP-qPCR and RT-qPCR were performed as described.⁷⁶ See Table S7 for the list of primers used in our ChIP-qPCR and RT-qPCR experiments.

Data Access

The datasets supporting the results of this article are available in the Gene Expression Omnibus (GEO) repository with the accession number GSE70920.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.



Acknowledgments

The authors thank the Genomics Core Facility, the Flow Cytometry Center, the Department of Scientific Computing and the Pluripotent Stem Cell Core Facility at Mount Sinai for technical assistance. We thank Matthew O'Connell and Kajan Ratnakumar for helpful discussions and guidance. Funding was provided by the DGAPA-PAPIIT, UNAM (IN209403, IN203811 and IN201114), CONACyT (42653-Q, 128464 and 220503) and Fronteras de la Ciencia-2015 CONACyT (Grant 290) to FR-T, Graduate fellowship from CONACyT (239663, CVU 257385) to DV-G, NCI T32-CA078207 to ZAQ, NIH EY014867, EY018599 and CA168875, Cancer Center Support from the NCI (CA21765), support from the American Lebanese Syrian Associated Charities (ALSAC) and a grant from Alex's Lemonade Stand Foundation for Childhood Cancer to MAD, and NCI/NIH R01CA154683 to EB. This article is part of the doctoral thesis of DV-G from the Doctorado en Ciencias Biomédicas, UNAM.

Author's contributions

DV-G and EB conceived this study with guidance and support from FR-T. DV-G and ZAQ performed ChIP-seq with assistance and support from DH and MAD. DV-G performed data analysis with help from DSM and ZAQ. DV-G derived the all KO cell lines and performed molecular characterization and cell cycle analysis with help from DSM. FGG performed comet assays and AHC performed co-IPs. DV-G and EB wrote the manuscript with assistance from ZAQ. All authors read and approved the final manuscript.

ORCID

David Valle-García  <http://orcid.org/0000-0002-3397-3103>
 Zulekha A. Qadeer  <http://orcid.org/0000-0002-1335-660X>
 Domhnall S. McHugh  <http://orcid.org/0000-0002-0255-5873>
 Flávia G. Ghiraldini  <http://orcid.org/0000-0001-8658-5270>
 Félix Recillas-Targa  <http://orcid.org/0000-0002-4336-0407>
 Emily Bernstein  <http://orcid.org/0000-0001-6533-8326>

References

- Sinha M, Peterson CL. Chromatin dynamics during repair of chromosomal DNA double-strand breaks. *Epigenomics* 2009; 1:371–85; PMID:20495614; <http://dx.doi.org/10.2217/epi.09.22>
- Hargreaves DC, Crabtree GR. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res* 2011; 21:396–420; PMID:21358755; <http://dx.doi.org/10.1038/cr.2011.32>
- Euskirchen G, Auerbach RK, Snyder M. SWI/SNF chromatin-remodeling factors: Multiscale analyses and diverse functions. *J Biol Chem* 2012; 287:30897–905; PMID:22952240; <http://dx.doi.org/10.1074/jbc.R111.309302>
- Gospodinov A, Herceg Z. Chromatin structure in double strand break repair. *DNA Repair (Amst)* 2013; 12:800–10; PMID:23919923; <http://dx.doi.org/10.1016/j.dnarep.2013.07.006>
- Längst G, Manlyte L. Chromatin Remodelers: From Function to Dysfunction. *Genes (Basel)* 2015; 6:299–324; PMID:26075616; <http://dx.doi.org/10.3390/genes6020299>
- Clynes D, Gibbons RJ. ATRX and the replication of structured DNA. *Curr Opin Genet Dev* 2013; 23:289–94; PMID:23453691; <http://dx.doi.org/10.1016/j.gde.2013.01.005>
- Voon HPJ, Wong LH. New players in heterochromatin silencing: histone variant H3.3 and the ATRX/DAXX chaperone. *Nucleic Acids Res* 2016; 44:1496–501.gkw012; PMID:26773061; <http://dx.doi.org/10.1093/nar/gkw012>
- Ratnakumar K, Bernstein E. ATRX: the case of a peculiar chromatin remodeler. *Epigenetics* 2013; 8:3–9; PMID:23249563; <http://dx.doi.org/10.4161/epi.23271>
- Watson LA, Goldberg H, Bérubé NG. Emerging roles of ATRX in cancer. *Epigenomics* 2015; 7:1365–78; PMID:26646632; <http://dx.doi.org/10.2217/epi.15.82>
- Park DJ, Pask AJ, Huynh K, Renfree MB, Harley VR, Marshall Graves J A. Comparative analysis of ATRX, a chromatin remodeling protein. *Gene* 2004; 339:39–48; PMID:15363844; <http://dx.doi.org/10.1016/j.gene.2004.06.025>
- Eustermann S, Yang J-C, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D, et al. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 2011; 18:777–82; PMID:21666677; <http://dx.doi.org/10.1038/nsmb.2070>
- Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, et al. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* 2011; 18:769–76; PMID:21666679; <http://dx.doi.org/10.1038/nsmb.2062>
- Xue Y, Gibbons R, Yan Z, Yang D, McDowell TL, Sechi S, Qin J, Zhou S, Higgs D, Wang W. The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc Natl Acad Sci USA* 2003; 100:10635–40; PMID:12953102; <http://dx.doi.org/10.1073/pnas.1937626100>
- McDowell TL, Gibbons RJ, Sutherland H, O'Rourke DM, Bickmore W A, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ, et al. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc Natl Acad Sci USA* 1999; 96:13983–8; PMID:10570185; <http://dx.doi.org/10.1073/pnas.96.24.13983>
- Law MJ, Lower KM, Voon HPJJ, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A, et al. ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. *Cell* 2010; 143:367–78; PMID:21029860; <http://dx.doi.org/10.1016/j.cell.2010.09.023>
- Drané P, Ouararhni K, Depaux A, Shuaib M, Hamiche A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* 2010; 24:1253–65; PMID:Can't; <http://dx.doi.org/10.1101/gad.566910>
- Goldberg AD, Banaszynski L A, Noh K-M, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 2010; 140:678–91; PMID:20211137; <http://dx.doi.org/10.1016/j.cell.2010.01.003>

18. Wong LH, McGhie JD, Sim M, Anderson M A, Ahn S, Hannan RD, George AJ, Morgan K A, Mann JR, Choo KHA. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Res* 2010; 20:351–60; PMID:20110566; <http://dx.doi.org/10.1101/gr.101477.109>
19. Lewis PW, Elsaesser SJ, Noh K-M, Stadler SC, Allis CD. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci USA* 2010; 107:14075–80; PMID:20651253; <http://dx.doi.org/10.1073/pnas.1008850107>
20. Sadic D, Schmidt K, Groh S, Kondofersky I, Ellwart J, Fuchs C, Theis FJ, Schotta G. Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep* 2015; 16:836–50; PMID:26012739; <http://dx.doi.org/10.15252/embr.201439937>
21. Elsässer SJ, Noh K-M, Diaz N, Allis CD, Banaszynski LA. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* 2015; 522:240–4; PMID:25938714; <http://dx.doi.org/10.1038/nature14345>
22. He Q, Kim H, Huang R, Lu W, Tang M, Shi F, Yang D, Zhang X, Huang J, Liu D, et al. The Daxx/Atrx Complex Protects Tandem Repetitive Elements during DNA Hypomethylation by Promoting H3K9 Trimethylation. *Cell Stem Cell* 2015; 17:273–86; PMID:26340527; <http://dx.doi.org/10.1016/j.stem.2015.07.022>
23. Bérubé NG, Smeenk CA, Picketts DJ. Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum Mol Genet* 2000; 9:539–47; PMID:10699177; <http://dx.doi.org/10.1093/hmg/9.4.539>
24. Lechner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem Biophys Res Commun* 2005; 331:929–37; PMID:15882967; <http://dx.doi.org/10.1016/j.bbrc.2005.04.016>
25. Voon HPJ, Hughes JR, Rode C, De La Rosa-Velázquez IA, Jenuwein T, Feil R, Higgs DR, Gibbons RJ. ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. *Cell Rep* 2015; 11:405–18; PMID:25865896; <http://dx.doi.org/10.1016/j.celrep.2015.03.036>
26. Ray-Gallet D, Woolfe A, Vassias I, Pellentz C, Lacoste N, Puri A, Schultz DC, Pchelintsev NA, Adams PD, Jansen LET, et al. Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3.3 to Maintain Chromatin Integrity. *Mol Cell* 2011; 44:928–41; PMID:22195966; <http://dx.doi.org/10.1016/j.molcel.2011.12.006>
27. Schneiderman JI, Orsi GA, Hughes KT, Loppin B, Ahmad K. Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. *Proc Natl Acad Sci USA* 2012; 109:19721–6; PMID:23150573; <http://dx.doi.org/10.1073/pnas.1206629109>
28. Filipescu D, Szenker E, Almouzni G. Developmental roles of histone H3 variants and their chaperones. *Trends Genet* 2013; 29:630–40; PMID:23830582; <http://dx.doi.org/10.1016/j.tig.2013.06.002>
29. Ratnakumar K, Duarte LF, LeRoy G, Hasson D, Smeets D, Vardabasso C, Bönisch C, Zeng T, Xiang B, Zhang DY, et al. ATRX-mediated chromatin association of histone variant macroH2A1 regulates α -globin expression. *Genes Dev* 2012; 26:433–8; PMID:22391447; <http://dx.doi.org/10.1101/gad.179416.111>
30. Clynes D, Higgs DR, Gibbons RJ. The chromatin remodeller ATRX: a repeat offender in human disease. *Trends Biochem Sci* 2013; 38:461–6; PMID:23916100; <http://dx.doi.org/10.1016/j.tibs.2013.06.011>
31. Levy M a., Kernohan KD, Jiang Y, Bérubé NG. ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions. *Hum Mol Genet* 2015; 24:1824–35; PMID:25452430; <http://dx.doi.org/10.1093/hmg/ddu596>
32. Heaphy CM, de Wilde RF, Jiao Y, Klein AP, Edil BH, Shi C, Bettegowda C, Rodriguez FJ, Eberhart CG, Hebbar S, et al. Altered telomeres in tumors with ATRX and DAXX mutations. *Science* 2011; 333:425; PMID:21719641; <http://dx.doi.org/10.1126/science.1207313>
33. Schwartzentruber J, Korshunov A, Liu X-Y, Jones DTW, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang D-AK, Tönjes M, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 2012; 484:130–130; PMID:22286061; <http://dx.doi.org/10.1038/nature11026>
34. Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, de Lange T, De S, Petrini JHJ, Sung P A, Jasin M, et al. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* 2012; 8:e1002772; PMID:22829774; <http://dx.doi.org/10.1371/journal.pgen.1002772>
35. Bower K, Napier CE, Cole SL, Dagg RA, Lau LMS, Duncan EL, Moy EL, Reddel RR. Loss of wild-type ATRX expression in somatic cell hybrids segregates with activation of Alternative Lengthening of Telomeres. *PLoS One* 2012; 7:e50062; PMID:23185534; <http://dx.doi.org/10.1371/journal.pone.0050062>
36. Clynes D, Jelinska C, Xella B, Ayyub H, Scott C, Mitson M, Taylor S, Higgs DR, Gibbons RJ. Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nat Commun* 2015; 6:7538; PMID:26143912; <http://dx.doi.org/10.1038/ncomms8538>
37. Huh MS, Price O'Dea T, Ouazia D, McKay BC, Parise G, Parks RJ, Rudnicki MA, Picketts DJ. Compromised genomic integrity impedes muscle growth after Atrx inactivation. *J Clin Invest* 2012; 122:4412–23; PMID:23114596; <http://dx.doi.org/10.1172/JCI63765>
38. Watson LA, Solomon LA, Li JR, Jiang Y, Edwards M, Shin-Ya K, Beier F, Bérubé NG. Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *J Clin Invest* 2013; 123:2049–63; PMID:23563309; <http://dx.doi.org/10.1172/JCI65634>
39. Leung JW-CC, Ghosal G, Wang W, Shen X, Wang J, Li L, Chen J. Alpha thalassemia/mental retardation syndrome X-linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress. *J Biol Chem* 2013; 288:6342–50; PMID:23329831; <http://dx.doi.org/10.1074/jbc.M112.411603>
40. Clynes D, Jelinska C, Xella B, Ayyub H, Taylor S, Mitson M, Bachrati CZ, Higgs DR, Gibbons RJ. ATRX dysfunction induces replication defects in primary mouse cells. *PLoS One* 2014; 9:e92915; PMID:24651726; <http://dx.doi.org/10.1371/journal.pone.0092915>
41. Emerson RO, Thomas JH. Adaptive evolution in zinc finger transcription factors. *PLoS Genet* 2009; 5; PMID:19119423; <http://dx.doi.org/10.1371/journal.pgen.1000325>
42. Dhvani Tadeppally H, Aubry M. Evolution of C2H2 Zinc-finger Gene Families in Mammals. In: *Encyclopedia of Life Sciences*. Chichester, UK: John Wiley & Sons, Ltd; 2010; 1–9; <http://dx.doi.org/10.1002/9780470015902.a0021738>
43. Nowick K, Fields C, Gernat T, Caetano-Anolles D, Kholina N, Stubbs L. Gain, loss and divergence in primate zinc-finger genes: a rich resource for evolution of gene regulatory differences between species. *PLoS One* 2011; 6:e21553; PMID:21738707; <http://dx.doi.org/10.1371/journal.pone.0021553>
44. Blahnik KR, Dou L, Echipare L, Iyengar S, O'Geen H, Sanchez E, Zhao Y, Marra MA, Hirst M, Costello JF, et al. Characterization of the contradictory chromatin signatures at the 3' exons of zinc finger genes. *PLoS One* 2011; 6:e17121; PMID:21347206; <http://dx.doi.org/10.1371/journal.pone.0017121>
45. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis C A, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489:57–74; PMID:22955616; <http://dx.doi.org/10.1038/nature11247>
46. Sarma K, Cifuentes-Rojas C, Ergun A, Del Rosario A, Jeon Y, White F, Sadreyev R, Lee JT. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell* 2014; 159:869–83; PMID:25417162; <http://dx.doi.org/10.1016/j.cell.2014.10.019>
47. Ramamoorthy M, Smith S. Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. *Cancer Cell* 2015; 28:357–69; PMID:26373281; <http://dx.doi.org/10.1016/j.ccell.2015.08.003>
48. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011; 473:43–9; PMID:21441907; <http://dx.doi.org/10.1038/nature09906>
49. Vogel MJ, Guelen L, De Wit E, Peric-Hupkes D, Lodén M, Talhout W, Feenstra M, Abbas B, Classen AK, Van Steensel B. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res* 2006; 16:1493–504; PMID:17038565; <http://dx.doi.org/10.1101/gr.5391806>

50. Julien H, Zoufir A, Audit B, Arneodo A. Human genome replication proceeds through four chromatin states. *PLoS Comput Biol* 2013; 9: e1003233; PMID:24130466; <http://dx.doi.org/10.1371/journal.pcbi.1003233>
51. Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, Libbrecht M, Giardine B, Ellenbogen PM, Bilmes J A, Birney E, et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res* 2013; 41:827–41; PMID:23221638; <http://dx.doi.org/10.1093/nar/gks1284>
52. Iyengar S, Farnham PJ. KAP1 protein: An enigmatic master regulator of the genome. *J Biol Chem* 2011; 286:26267–76; PMID:21652716; <http://dx.doi.org/10.1074/jbc.R111.252569>
53. Frieze S, O'Geen H, Blahnik KR, Jin VX, Farnham PJ. ZNF274 recruits the histone methyltransferase SETDB1 to the 3' ends of ZNF genes. *PLoS One* 2010; 5:e15082; PMID:21170338; <http://dx.doi.org/10.1371/journal.pone.0015082>
54. Bérubé NG. ATRX in chromatin assembly and genome architecture during development and disease. *Biochem Cell Biol* 2011; 89:435–44; PMID:Can't; <http://dx.doi.org/10.1139/o11-038>
55. Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, Albu M, Weirauch MT, Radovani E, Kim PM, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat Biotechnol* 2015; 33:555–62; PMID:25690854; <http://dx.doi.org/10.1038/nbt.3128>
56. Iyengar S, Ivanov A V, Jin VX, Rauscher FJ, Farnham PJ. Functional analysis of KAP1 genomic recruitment. *Mol Cell Biol* 2011; 31:1833–47; PMID:21343339; <http://dx.doi.org/10.1128/MCB.01331-10>
57. Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP, et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* 2014; 21:366–74; PMID:24658350; <http://dx.doi.org/10.1038/nsmb.2796>
58. Pfister SX, Ahrabi S, Zalmas L-P, Sarkar S, Aymard F, Bachrati CZ, Helleday T, Legube G, La Thangue NB, Porter ACG, et al. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep* 2014; 7:2006–18; PMID:24931610; <http://dx.doi.org/10.1016/j.celrep.2014.05.026>
59. Murray JM, Stiff T, Jeggo P A. DNA double-strand break repair within heterochromatic regions. *Biochem Soc Trans* 2012; 40:173–8; PMID:22260685; <http://dx.doi.org/10.1042/BST20110631>
60. Kalousi A, Hoffbeck A-S, Selemenakis PN, Pinder J, Savage KI, Khanna KK, Brino L, Delleire G, Gorgoulis VG, Soutoglou E. The Nuclear Oncogene SET Controls DNA Repair by KAP1 and HP1 Retention to Chromatin. *Cell Rep* 2015; 11:149–63; PMID:25818296; <http://dx.doi.org/10.1016/j.celrep.2015.03.005>
61. Napier CE, Huschtscha LI, Harvey A, Bower K, Noble JR, Hendrickson EA, Reddel RR. ATRX represses alternative lengthening of telomeres. *Oncotarget* 2015; 6:16543–58; PMID:26001292; <http://dx.doi.org/10.18632/oncotarget.3846>
62. Naftelberg S, Schor IE, Ast G, Kornbliht AR. Regulation of Alternative Splicing Through Coupling with Transcription and Chromatin Structure. *Annu Rev Biochem* 2015; 84:165–98; PMID:26034889; <http://dx.doi.org/10.1146/annurev-biochem-060614-034242>
63. Hasson D, Panchenko T, Salimian KJ, Salman MU, Sekulic N, Alonso A, Warburton PE, Black BE. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat Struct Mol Biol* 2013; 20:687–95; PMID:23644596; <http://dx.doi.org/10.1038/nsmb.2562>
64. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; 10:R25; PMID:19261174; <http://dx.doi.org/10.1186/gb-2009-10-3-r25>
65. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008; 9:R137; PMID:18798982; <http://dx.doi.org/10.1186/gb-2008-9-9-r137>
66. Liang K, Keleş S. Normalization of ChIP-seq data with control. *BMC Bioinformatics* 2012; 13:199; PMID:22883957; <http://dx.doi.org/10.1186/1471-2105-13-199>
67. Zang C, Schonnes DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* 2009; 25:1952–8; PMID:19505939; <http://dx.doi.org/10.1093/bioinformatics/btp340>
68. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26:841–2; PMID:20110278; <http://dx.doi.org/10.1093/bioinformatics/btq033>
69. Ramírez F, Dündar F, Diehl S, Grüning B a., Manke T. DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 2014; 42:187–91; PMID:24799436; <http://dx.doi.org/10.1093/nar/gku365>
70. Sandve GK, Gundersen S, Johansen M, Glad IK, Gunathasan K, Holden L, Holden M, Liestøl K, Nygård S, Nygaard V, et al. The Genomic HyperBrowser: an analysis web server for genome-scale data. *Nucleic Acids Res* 2013; 41:W133–41; PMID:23632163; <http://dx.doi.org/10.1093/nar/gkt342>
71. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; 4:44–57; PMID:19131956; <http://dx.doi.org/10.1038/nprot.2008.211>
72. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 2005; 33:2908–16; PMID:15914667; <http://dx.doi.org/10.1093/nar/gki609>
73. Machanick P, Bailey TL. MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* 2011; 27:1696–7; PMID:21486936; <http://dx.doi.org/10.1093/bioinformatics/btr189>
74. Ran F, Hsu P, Wright J, Agarwala V. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013; 8:2281–308; PMID:24157548; <http://dx.doi.org/10.1038/nprot.2013.143>
75. Singh NP, McCoy MT, Tice RR, Schneider EL. A simple technique for quantitation of low levels of DNA damage in individual cells. *Exp Cell Res* 1988; 175:184–91; PMID:3345800; [http://dx.doi.org/10.1016/0014-4827\(88\)90265-0](http://dx.doi.org/10.1016/0014-4827(88)90265-0)
76. Vardabasso C, Gaspar-Maia A, Hasson D, Pünzeler S, Valle-Garcia D, Straub T, Keilhauer EC, Strub T, Dong J, Panda T, et al. Histone Variant H2A.Z.2 Mediates Proliferation and Drug Sensitivity of Malignant Melanoma. *Mol Cell* 2015; 75–88; PMID:26051178; <http://dx.doi.org/10.1016/j.molcel.2015.05.009>

RESEARCH

Open Access

The ATRX cDNA is prone to bacterial IS10 element insertions that alter its structure

David Valle-García^{1,2†}, Lyra M Griffiths^{3†}, Michael A Dyer^{3,4*}, Emily Bernstein^{2*} and Félix Recillas-Targa^{1*}

Abstract

The SWI/SNF-like chromatin-remodeling protein ATRX has emerged as a key factor in the regulation of α -globin gene expression, incorporation of histone variants into the chromatin template and, more recently, as a frequently mutated gene across a wide spectrum of cancers. Therefore, the availability of a functional ATRX cDNA for expression studies is a valuable tool for the scientific community. We have identified two independent transposon insertions of a bacterial IS10 element into exon 8 of *ATRX* isoform 2 coding sequence in two different plasmids derived from a single source. We demonstrate that these insertion events are common and there is an insertion hotspot within the ATRX cDNA. Such IS10 insertions produce a truncated form of ATRX, which significantly compromises its nuclear localization. In turn, we describe ways to prevent IS10 insertion during propagation and cloning of ATRX-containing vectors, including optimal growth conditions, bacterial strains, and suggested sequencing strategies. Finally, we have generated an insertion-free plasmid that is available to the community for expression studies of ATRX.

Keywords: ATRX; Insertion element; Chromatin remodeling; Cloning vector; ATRX over-expression; IS10 element

Background

ATRX is an ATP-dependent chromatin-remodeling helicase of the SWI/SNF family and was originally identified because mutations in its coding region give rise to a complex disorder known as α -Thalassemia, Mental Retardation X-linked (ATRX) syndrome. Its role in the α -Thalassemia feature of this syndrome is through its regulation of the chromatin state of the α -globin genes (Higgs et al. 2005; Gibbons 2006; Law et al. 2010; Ratnakumar et al. 2012). However, its role in other ATRX syndrome traits remains largely unknown. Mechanistically, ATRX has been implicated in heterochromatin maintenance and formation, DNA methylation, regulation of gene expression, alternative lengthening of telomere suppression and histone variant deposition (for review see: Bérubé 2011; Ratnakumar and Bernstein 2013; Clynes et al. 2013).

Recently, large-scale genome sequencing efforts have identified *ATRX* mutations and deletions in a range of tumor types, including glioblastoma multiforme, neuroblastoma and pancreatic neuroendocrine tumors (Jiao et al. 2011; Schwartzentruber et al. 2012; Cheung et al. 2012; Molenaar et al. 2012; Clynes et al. 2013). These alterations appear to result in loss of functional ATRX, suggesting that ATRX acts as tumor suppressor. However, the mechanisms by which such mutations result in ATRX loss-of-function remain unclear.

For the reasons noted above, ATRX has emerged as an important chromatin remodeler to study during normal development and disease. However, the cloning and overexpression of *ATRX* cDNA has been challenging due to its large size: 2,492 amino acids encoded in an 11,167 bp cDNA (isoform 1) and 2,454 amino acids encoded in an 11,031 bp cDNA (isoform 2). Here, we report two independent transposon insertions of an IS10 element into exon 8 of the *ATRX* isoform 2 coding sequence in two different plasmids derived from a single original source. Our data suggest that *ATRX* contains at least one insertion hot spot that is highly active when the plasmid is grown under standard bacterial conditions. Therefore, we recommend strategies to prevent IS10 insertions during propagation and cloning of *ATRX*-containing

* Correspondence: michael.dyer@stjude.org; emily.bernstein@mssm.edu; frecilla@correo.ifc.unam.mx

†Equal contributors

³Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN, USA

²Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA

¹Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Ciudad Universitaria, México DF, México

Full list of author information is available at the end of the article

vectors. Furthermore, we analyzed the functional effect of this insertion and conclude that IS10 insertions produce a truncated form of ATRX that compromises its cellular distribution.

Results and discussion

We sequenced two different ATRX plasmids that have been widely used for overexpression experiments (Additional file 1) (Bérubé et al. 2002, 2008; McFarlane and Preston 2011; Ratnakumar et al. 2012; Newhart et al. 2012). One of the plasmids contains the full length ATRX cDNA sequence in a pEGFP (GFP-ATRX) (Bérubé et al. 2008) backbone and the other is cloned in a lentiviral backbone (ATRX-YFP) and was derived from the sequence contained in GFP-ATRX (Newhart et al. 2012). We found that the ATRX cDNA sequence that has been cloned corresponds to isoform 2, which lacks exon 6 in comparison to isoform 1. Exon 6 codes for 38 highly conserved amino acids. The functional relevance, if any, between isoform 1 and isoform 2 has yet to be reported. In addition, we found that E124 is missing in ATRX isoform 2. This amino acid is the last residue encoded by exon 5.

To our surprise, we uncovered transposon insertions of IS10 elements within the ATRX coding sequence in both plasmids analyzed (Figure 1A-B). The transposable prokaryotic IS family has been reported to be a common source of mutations in the *Escherichia coli* genome (Rodriguez et al. 1992). Likewise, it has also been shown that IS10-related sequences may lead to cloning aberrations and sequence contamination of plasmids hosted in *E. coli* (Rood et al. 1980; Müller et al. 1989; Kovarik et al. 2001; Kobori et al. 2009; Umenhoffer et al. 2010). Interestingly, the IS10 insertion in both plasmids - which we name IS10-GFP-ATRX and IS10-ATRX-YFP - was found in exon 8. However, sequencing and PCR analyses showed that both the insertion site and the direction of the IS10 element are different in each plasmid, suggesting independent insertion events (Figure 1A,B, Additional file 1). The insertion site within IS10-GFP-ATRX closely resembles the 9-bp canonical IS10 insertion sequence (Additional file 1, see Kovarik et al. 2001), and may represent a hotspot for IS10 transposition. The insertion site within IS10-ATRX-YFP has a more divergent sequence. It is interesting to note that IS10-ATRX-YFP was derived from IS10-GFP-ATRX, however there is no sequence evidence of an IS10 excision in IS10-ATRX-YFP. This suggests that the ATRX cDNA of IS10-GFP-ATRX used to derive IS10-ATRX-YFP was lacking the IS10 insertion, and that the insertion observed in the latter was acquired during the cloning or plasmid expansion process. This highlights the dynamic feature of this transposable element.

In order to assess if the expression of the truncated form of ATRX derived from the IS10-GFP-ATRX plasmid may

retain some of the characteristics of full length ATRX, we sub-cloned the ATRX cDNA sequence from the IS10-GFP-ATRX plasmid to generate an 'Insertion-Free' vector that we call IF-GFP-ATRX (see details in Methods). IF-GFP-ATRX was further analyzed by PCR and restriction analysis (data not shown) and the ATRX cDNA was fully sequenced to ensure that the insertion was completely removed (Additional file 1).

Next we investigated if the insertions are an isolated event or a common phenomenon. IF-GFP-ATRX was transformed into both DH5 α and Stbl4 *E. coli* strains. The plasmids were analyzed by colony PCR with primers spanning the IS10 insertion sites found in IS10-ATRX-YFP and IS10-GFP-ATRX (Additional file 1, see Methods). We found that 18% of the DH5 α colonies (26 out of 139 analyzed colonies) present evidence of insertions within exon 8. Sequencing and restriction analysis of the plasmid containing insertions revealed that the observed change is caused by *de novo* insertions of the IS10 transposable element (Figure 1C, Additional file 1). Conversely, the plasmid derived from the Stbl4 bacteria - a strain specifically designed for cloning unstable inserts - showed a significant lower insertion rate (3 out of 80 analyzed colonies - p -value < 0.01, binomial test). Furthermore, we found evidence of other aberrations in the DH5 α colonies by restriction analyses (6 out of 30 analyzed colonies that did not contain the exon 8 IS10 insertion; data not shown). This suggests that the ATRX plasmid may be subjected to IS10 insertions in other regions of the ATRX cDNA or other types of alterations such as recombination events. This data strongly suggests that insertions within the ATRX coding sequence in *E. coli* are a common event that is influenced by the bacterial strain used to amplify ATRX-containing plasmids. Moreover, our results suggest that ATRX presents an IS10 insertion hotspot within exon 8. The causes that trigger the insertion of the IS10 element within the ATRX cDNA remain unclear. However, the homogeneity of the restriction patterns in the IF-GFP-ATRX with *de novo* IS10 insertions (Figure 1C) suggests that the IS10 transposition is an early clonal event. In fact, it has been proposed that the stress induced by the transformation process may be a trigger for IS10 transposition (Kovarik et al. 2001).

We next transfected IS10-GFP-ATRX and IF-GFP-ATRX into HEK 293T cells and assessed ATRX expression and chromatin association. We found that the transfection efficiency was approximately double for IS10-GFP-ATRX (60%) compared to IF-GFP-ATRX (30%), and that the sub-cellular localization was dramatically different (data not shown, Figure 2A). As expected, IF-GFP-ATRX transfected cells show a focal nuclear localization of GFP-ATRX (Figure 2A) in agreement with previous reports (Xue et al. 2003; Tang et al. 2004; Ishov et al. 2004; Bérubé et al. 2008). However, we observed a clear cytoplasmic GFP signal in

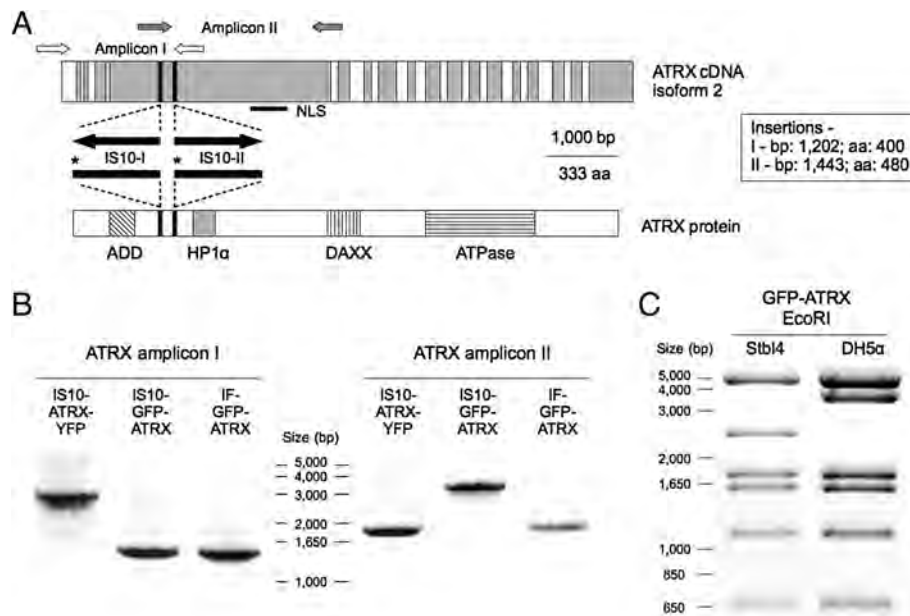


Figure 1 ATRX overexpression plasmids contain IS10 insertions. **(A)** Schematic map of the insertions found in IS10-GFP-ATRX and IS10-ATRX-YFP. The exon structure (alternate white and gray bars) of ATRX cDNA isoform 2 (top) and some relevant domains from the protein sequence (bottom) are shown. The insertion sites are denoted by black bars and their positions in the cDNA and the protein are described in the right box. Black arrows show the direction of the IS10 elements found. The stop codons introduced by IS10 in the protein sequence are indicated with asterisks. The position of two of the primer pairs used for analyzing ATRX sequence is shown (white and gray arrows). The NLS within ATRX cDNA is indicated with a black bar. **(B)** PCR analysis of IS10-GFP-ATRX, IS10-ATRX-YFP and IF-GFP-ATRX using the primers shown in (A). Both amplicons of IF-GFP-ATRX have the expected size (1,449 bp for amplicon I and 1,935 bp for amplicon II) while amplicon I shows an insertion in IS10-ATRX-YFP and amplicon II shows an insertion in IS10-GFP-ATRX (both have additional 1,336 bp). **(C)** Representative EcoRI digestion patterns of the IF-GFP-ATRX plasmid grown in two different bacteria strains. The Stbl4-derived plasmid has the expected pattern while the plasmid derived from DH5 α shows an insertion in the 2,400bp fragment (shift from a 2,468 bp fragment to ~3,800 bp).

the IS10-GFP-ATRX transfected cells and a more diffuse nuclear localization (Figure 2A). According to our sequence analysis, a stop codon is introduced right after the IS10 insertion site in both plasmids containing the IS10 element (Figure 1A). As ATRX is N-terminally GFP-tagged, the fluorescence observed in the IS10-GFP-ATRX transfectants is likely due to the overexpression of a truncated GFP-ATRX fusion that contains only the ADD domain (Figure 1A).

Although we cannot exclude the possibility that the re-localization of the IS10-GFP-ATRX is due to an overexpression effect, ATRX nuclear localization is dependent on nuclear localization sequences (NLS) (Bérubé et al. 2008) and our sequence analysis predicts that the IS10-GFP-ATRX lacks the NLS (Figure 1A). However, we note that although ATRX cytoplasmic localization is clear in the IS10-GFP-ATRX transfected cells, cells still do show some nuclear localization (Figure 2A). Because the ADD domain alone is capable of recognizing and binding the histone H3 tail *in vitro* by reading the combinatorial methylation state of histone H3 lysines 4 and 9 (Eustermann et al. 2011; Iwase et al. 2011) a small fraction of the expressed protein product may still be tethered to chromatin.

To probe this further and assay the chromatin association of the ATRX protein products, we prepared both whole cell and chromatin extracts from the transfected HEK 293T cells and performed Western blot analyses. As predicted, we found a truncated ATRX product from the transfections performed with the IS10-GFP-ATRX plasmid (Figure 2B,C). This truncated product is ~70 kDa and corresponds to the ATRX ADD domain coupled to the N-terminal GFP. This is further supported by the fact that the IS10-GFP-ATRX product is detectable only with a GFP antibody and not with an ATRX antibody whose epitope recognizes the C-terminal portion of the protein (data not shown).

While we observed chromatin-bound full-length ATRX from the IF-GFP-ATRX transfections, we noted that the IS10-GFP-ATRX truncated product also binds to the chromatin fraction (Figure 2C). This supports the idea that the ADD domain allows binding to chromatin. Whether the IS10-GFP-ATRX product has any effect on the chromatin landscape remains to be explored, for example, it could act as a dominant negative. However, it is interesting to note that in comparison to full-length ATRX, the IS10-GFP-ATRX fragment has a greater accumulation in whole cell extracts (Figure 2B) than

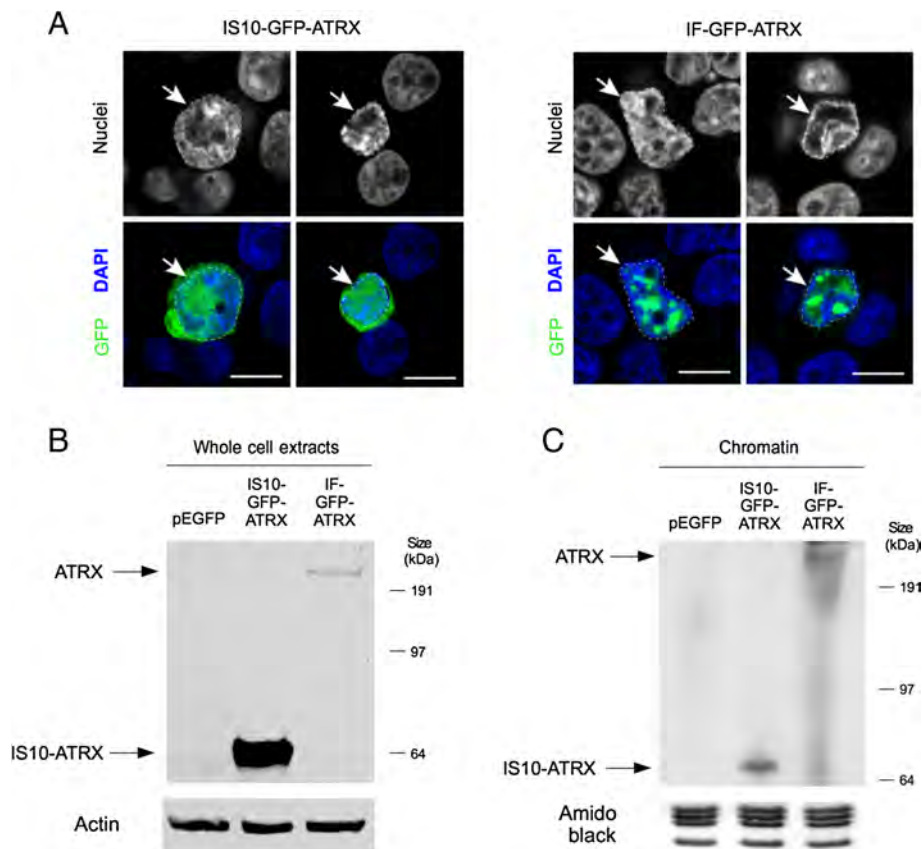


Figure 2 IS10-element alters GFP-ATRX localization and incorporation into chromatin. **(A)** HEK-293T cells were transiently transfected with IS10-GFP-ATRX (left) or IF-GFP-ATRX (right). The localization of GFP-ATRX was determined by confocal fluorescence microscopy 72 hours post-transfection. Representative images are shown. Dashed lines on the highlighted cells indicate nuclei area. IS10-GFP-ATRX displays a diffuse nuclear and cytoplasmic localization. Conversely, ATRX-GFP has a nuclear localization in the IF-GFP-ATRX transfected cells with a focal pattern. **(B)** Western analysis of whole cell protein extracts from the IS10-GFP-ATRX and IF-GFP-ATRX transfections against GFP (top) and actin (bottom). Bands corresponding to IF-GFP-ATRX and IS10-GFP-ATRX products are indicated with arrows. **(C)** Chromatin extracts from the IS10-GFP-ATRX and IF-GFP-ATRX transfections were analyzed with GFP antibodies (top). Bands corresponding to IF-GFP-ATRX and IS10-GFP-ATRX products are highlighted with arrows. Amido black staining of histones was used as a loading control (bottom). A truncated product is observed in the IS10-GFP-ATRX transfections.

in chromatin (Figure 2C). Taken together, these results strongly suggest that the IS10 element provokes the production of a truncated ATRX form that has a diffuse cellular localization, but still retains the capacity to bind to chromatin. Importantly, during experimental conditions, if one were to perform immunofluorescence of IS10-GFP-ATRX transduced cells without Western blot analysis, one might conclude that full length ATRX is expressed.

Conclusions

Here we show that two ATRX overexpression plasmids, utilized in the ATRX research community, contain IS10-element insertions around a probable hot spot in exon 8 of *ATRX*. Further, we demonstrate that IS10 transposition events are common if *ATRX*-containing vectors are grown in standard bacteria strains. We note here that the dynamic nature of the insertions makes it likely that any single lab may have batches of contaminated and

insertion-free plasmid. We further highlight that the published studies of ATRX function using these plasmids may have indeed been insertion-free, as we have confirmed for our own studies (Ratnakumar et al. 2012). However, because the IS10-element has a significant effect on ATRX sub-cellular localization and allows its incorporation into chromatin, experiments performed with IS10-ATRX-containing plasmids could generate inaccurate conclusions.

It has not escaped our attention that despite all the mutations identified in ATRX syndrome patients and, more recently, in distinct cancer types (Jiao et al. 2011; Schwartzenuber et al. 2012; Cheung et al. 2012; Clynes et al. 2013), few ATRX over-expression experiments mimicking such mutations have been published. It is our belief that to understand the contribution of ATRX mutations in disease etiology, we need to develop reliable over-expression systems. However, our results show that

these experiments are challenging and subjected to technical hurdles that the community should be aware of.

Finally, we have made our IF-GFP-ATRX plasmid available through the AddGene database (plasmid 45444). Also, we provide technical advice for the propagation and cloning of ATRX-containing vectors (see Methods section). The sequences obtained from the IS10-GFP-ATRX and IS10-ATRX-YFP plasmids, as well as the primers used for the ATRX sequencing, cloning and analysis are included in the Additional file 1. Thus, we recommend that each laboratory should take the necessary steps to avoid IS10 contamination and test each plasmid preparation for IS10 insertion elements. We are hopeful that this technical resource will be useful for future studies and will help to avoid propagation and usage of IS10-ATRX-containing plasmids.

Methods

ATRX-containing plasmid propagation and handling

The ATRX plasmids were propagated using ElectroMAX™ Stbl4™ Competent Cells (Life Technologies, 11635-018). Top10 (Invitrogen, C404003) and DH5α bacteria were also used and are not recommended for growing ATRX-containing plasmids. We recommend growing bacteria at 30°C to avoid plasmid recombination. Using strains lacking IS-elements such as MDS42 is also recommended.

ATRX PCR analysis and sequencing

PCR analyses were performed using Platinum® Pfx DNA Polymerase (Invitrogen 11708-013), according to manufacturer's indications. The primers used span across the ATRX cDNA and their sequences can be found in Additional file 1. ATRX cDNA sequencing was performed using the same primers and one additional primer specific for exon 8 that can also be found in Additional file 1. We recommend analyzing the plasmid by PCR (or restriction digest) following culture in *E. coli* and on a frequent basis when working with ATRX-containing vectors.

ATRX transformation and growth in different bacterial strains

ElectroMAX™ Stbl4™ competent cells (Life Technologies, 11635-018) and DH5α competent cells (Life Technologies 18258-012) were transformed according to manufacturer's indications with 50 ng of IF-GFP-ATRX. Cells were plated and growth at 37°C. Colonies from three independent transfections were picked and diluted in 100 μl of PBS. 5 μl were used for colony PCR using the ATRX IS10 primers from Additional file 1. Colonies that presented evidence of insertion, as well as control colonies were miniprep using the QIAprep Spin Miniprep Kit (Qiagen, 27106) 2 μg of the plasmid were digested with 2 units of EcoRI (New England Biolabs, R0101S) for 1 hour at 37°C.

IF-GFP-ATRX construct generation

IS10-GFP-ATRX was digested with BamHI and AflII (NEB) to produce a 1200 bp band (Fragment A). A plasmid containing internal ATRX sequence was generated using GenScript services: (5'-ggtagcTACGTCAGCTTAAGGCTTTTAAGTCTGTGTTGGCTGATATTAAGAAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTAACAAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTAAAAAAAAGCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCTAAA CTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAACAAAGAACAAA TAAAAGTACCGGTGGTGAACATaagctt-3). A restriction digestion product was generated using the GenScript plasmid and enzymes, AflII and AgeI (Fragment B). A restriction digestion product of IS10-GFP-ATRX was generated using enzymes, AgeI (NEB) and XbaI (Fragment C). Finally, BamHI and XbaI (NEB) were used to digest pEGFP-C2 plasmid, generating a backbone for insertion of the ATRX fragments. The three ATRX-containing fragments (A, B, and C) were ligated first (1200 bp + 300 bp + 6000 bp). Then the backbone pEGFP vector was added for a second ligation reaction. Ligation product was transformed into dam⁻/dcm⁻ competent *E. coli* (NEB C2925H). Colonies were mini-prepped and sequenced.

Transfections

Transfections were performed in the HEK-293T cell line (ATCC: CRL-11268) using Lipofectamine® 2000 transfection reagent according to manufacturer's indications. 14 μg of each plasmid were transfected in 70% confluent 10 cm plates. Cells were harvested 48 hours post-transfection for chromatin preparation or fixed for microscopy (see below).

Microscopy

24 hours post-transfection, cells were seeded onto chamber well slides. 48 hours later, cells were fixed with 4% paraformaldehyde/PBS (Sigma Aldrich) at 4°C overnight. Slides were stained with 4',6-diamidino-2-phenylindole (DAPI) for 10 minutes at room temperature. Images were collected with a Zeiss LSM700 confocal microscope.

Chromatin isolation and western analysis

Chromatin isolation was performed as described previously (Méndez and Stillman 2000) and chromatin samples were boiled at 100°C for 10 minutes instead of being sonicated. Western analysis was performed with the following antibodies: GFP (Santa Cruz sc-8334, 1:200 or Roche 1181460001, 1:10,000), Actin (Sigma-Aldrich A1978, 0.75 μg/mL). Histones from chromatin extracts were stained with amido black staining solution (Sigma A8181-1EA).

Additional file

Additional file 1: Plasmid and primer sequences. This files contains the sequences of the IS10 insertions, all the primers used for the analyses described on the paper, and the full IF-GP-ATRX sequence.

Abbreviations

ATRX: α -Thalassemia, Mental Retardation X-linked gene.

Competing interest

The authors have no conflicts-of-interest or financial disclosures to report.

Authors' contributions

DV-G carried out the molecular biology analysis and ATRX transfections and WB. LMG sub cloned ATRX into IF-ATRX-GFP and performed microscopy and WB analysis. FR-T, EB and MAD conceived, designed and guided the study. EB and DV-G wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We acknowledge Dr. David Picketts and Dr. Susan Janicki for providing reagents and Georgina Guerrero for technical assistance. This work was supported by the Dirección General de Asuntos del Personal Académico-Universidad Nacional Autónoma de México (IN203811 and IN201114), Doctorado en Ciencias Biomédicas-UNAM and Consejo Nacional de Ciencia y Tecnología, México (CONACYT; 42653-Q and 128464) to F.R-T, a Graduate fellowship from CONACYT (239663) to D.V-G, NIH R01EY014867-06 and NIH R01EY018599-01A1 to L.M.G. and M.A.D, and NCI/NIH R01CA154683 to E.B. This article is part of the doctoral thesis of D.V-G from the Doctorado en Ciencias Biomédicas, UNAM.

Author details

¹Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Ciudad Universitaria, México DF, México. ²Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ³Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN, USA. ⁴Howard Hughes Medical Institute, Chevy Chase, MD, USA.

Received: 16 December 2013 Accepted: 22 April 2014

Published: 2 May 2014

References

- Bérubé NG (2011) ATRX in chromatin assembly and genome architecture during development and disease. *Biochem Cell Biol* 89:435–444, doi:10.1139/o11-038
- Bérubé NG, Jagla M, Smeenk C, De Repentigny Y, Kothary R, Picketts DJ (2002) Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Hum Mol Genet* 11:253–261
- Bérubé NG, Healy J, Medina CF, Wu S, Hodgson T, Jagla M, Picketts DJ (2008) Patient mutations alter ATRX targeting to PML nuclear bodies. *Eur J Hum Genet* 16:192–201, doi:10.1038/sj.ejhg.5201943
- Cheung N-KV, Zhang J, Lu C, Parker M, Bahrami A, Tickoo SK, Heguy A, Pappo AS, Federico S, Dalton J, Cheung IY, Ding L, Fulton R, Wang J, Chen X, Becksfors J, Wu J, Billups C, Ellison D, Mardis ER, Wilson RK, Downing JR, Dyer M (2012) Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA* 307:1062–1071, doi:10.1001/jama.2012.228
- Clynes D, Higgs DR, Gibbons RJ (2013) The chromatin remodeler ATRX: a repeat offender in human disease. *Trends Biochem Sci* 38:461–466, doi:10.1016/j.tibs.2013.06.011
- Eustermann S, Yang J-C, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D, Higgs DR, Neuhaus D (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 18:777–782, doi:10.1038/nsmb.2070
- Gibbons R (2006) Alpha thalassaemia-mental retardation, X linked. *Orphanet J Rare Dis* 1:15, doi:10.1186/1750-1172-1-15
- Higgs DR, Garrick D, Anguita E, De Gobbi M, Hughes J, Muers M, Vernimmen D, Lower K, Law M, Argentaro A, Deville M, Gibbons R (2005) Understanding alpha-globin gene regulation: aiming to improve the management of thalassemia. *Ann N Y Acad Sci* 1054:92–102, doi:10.1196/annals.1345.012
- Ishov AM, Vladimirova OV, Maul GG (2004) Heterochromatin and ND10 are cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX. *J Cell Sci* 117:3807–3820, doi:10.1242/jcs.01230
- Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, Shi Y (2011) ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* 18:769–776, doi:10.1038/nsmb.2062
- Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti M, Velculescu VE, Diaz L, Vogelstein B, Kinzler KW, Hruban RH, Papadopoulos N (2011) DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* 331:1199–1203, doi:10.1126/science.1200609
- Kobori S, Ko Y, Kato M (2009) A target site for spontaneous insertion of IS10 element in pUC19 DNA located within intrinsically bent DNA. *Open Microbiol J* 3:146–150, doi:10.2174/1874285800903010146
- Kovarík A, Matzke M, Matzke A, Koulaková B (2001) Transposition of IS 10 from the host *Escherichia coli* genome to a plasmid may lead to cloning artefacts. *Mol Genet Genomics* 266:216–222, doi:10.1007/s004380100542
- Law MJ, Lower KM, Voon HPJ, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A (2010) ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* 143:367–378, doi:10.1016/j.cell.2010.09.023
- McFarlane S, Preston CM (2011) Human cytomegalovirus immediate early gene expression in the osteosarcoma line U2OS is repressed by the cell protein ATRX. *Virus Res* 157:47–53, doi:10.1016/j.virusres.2011.02.002
- Méndez J, Stillman B (2000) Association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late. *Mol Cell Biol* 20:8602–8612, doi:10.1128/MCB.20.22.8602-8612.2000. Updated
- Molenaar JJ, Koster J, Zwijnenburg D, van Sluis P, Valentijn LJ, van der Ploeg I, Hamdi M, van Nes J, Westerman B, van Arkel J, Ebus ME, Haneveld F, Lakeman A, Schild L, Molenaar P, Stroeken P, van Noesel MM, Ora I, Santo EE, Caron HN, Westerhout EM, Versteeg R (2012) Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 483:589–593, doi:10.1038/nature10910
- Müller J, Reinert H, Malke H (1989) Streptokinase mutations relieving *Escherichia coli* K-12 (pRI4) of detriments caused by the wild-type skc gene. *J Bacteriol* 171:2202–2208
- Newhart A, Rafalska-Metcalf IU, Yang T, Negorev DG, Janicki SM (2012) Single-cell analysis of Daxx and ATRX-dependent transcriptional repression. *J Cell Sci* 125:5489–5501, doi:10.1242/jcs.110148
- Ratnakumar K, Bernstein E (2013) ATRX: the case of a peculiar chromatin remodeler. *Epigenetics* 8:3–9, doi:10.4161/epi.23271
- Ratnakumar K, Duarte LF, LeRoy G, Hasson D, Smeets D, Vardabasso C, Bönisch C, Zeng T, Xiang B, Zhang DY, Li H, Wang X, Hake SB, Schermelleh L, Garcia B, Bernstein E (2012) ATRX-mediated chromatin association of histone variant macroH2A1 regulates α -globin expression. *Genes Dev* 26:433–438, doi:10.1101/gad.179416.111
- Rodríguez H, Snow E, Bhat U, Loechler E (1992) An *Escherichia coli* plasmid-based, mutational system in which supF mutants are selectable: insertion elements dominate the spontaneous spectra. *Mutat Res* 270:219–231
- Rood JI, Sneddon MK, Morrison JF (1980) Instability in tyrR strains of plasmids carrying the tyrosine operon: isolation and characterization of plasmid derivatives with insertions or deletions. *J Bacteriol* 144:552–559
- Schwartzentruber J, Korshunov A, Liu X-Y, Jones DTW, Pfaff E, Jacob K, Sturm D, Fontebasso AM, Quang D-AK, Tönjes M, Hovestadt V, Albrecht S, Kool M, Nantel A, Konermann C, Lindroth A, Jäger N, Rausch T, Ryzhova M, Korbel JO, Hielscher T, Hauser P, Garami M, Klekner A, Bogner L, Ebinger M, Schuhmann MU, Scheurlen W, Pekrun A, Frühwald MC et al (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* 482:226–231, doi:10.1038/nature10833
- Tang J, Wu S, Liu H, Stratton R, Barak OG, Shiekhattar R, Picketts DJ, Yang X (2004) A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. *J Biol Chem* 279:20369–20377, doi:10.1074/jbc.M401321200
- Umenhoffer K, Fehér T, Balikó G, Ayaydin F, Pósfai J, Blattner FR, Pósfai G (2010) Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for

molecular and synthetic biology applications. *Microb Cell Fact* 9:38,
doi:10.1186/1475-2859-9-38

Xue Y, Gibbons R, Yan Z, Yang D, McDowell TL, Sechi S, Qin J, Zhou S, Higgs D,
Wang W (2003) The ATRX syndrome protein forms a chromatin-remodeling
complex with Daxx and localizes in promyelocytic leukemia nuclear bodies.
Proc Natl Acad Sci U S A 100:10635–10640, doi:10.1073/pnas.1937626100

doi:10.1186/2193-1801-3-222

Cite this article as: Valle-García *et al.*: The ATRX cDNA is prone to
bacterial IS10 element insertions that alter its structure. *SpringerPlus*
2014 **3**:222.

**Submit your manuscript to a SpringerOpen[®]
journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
