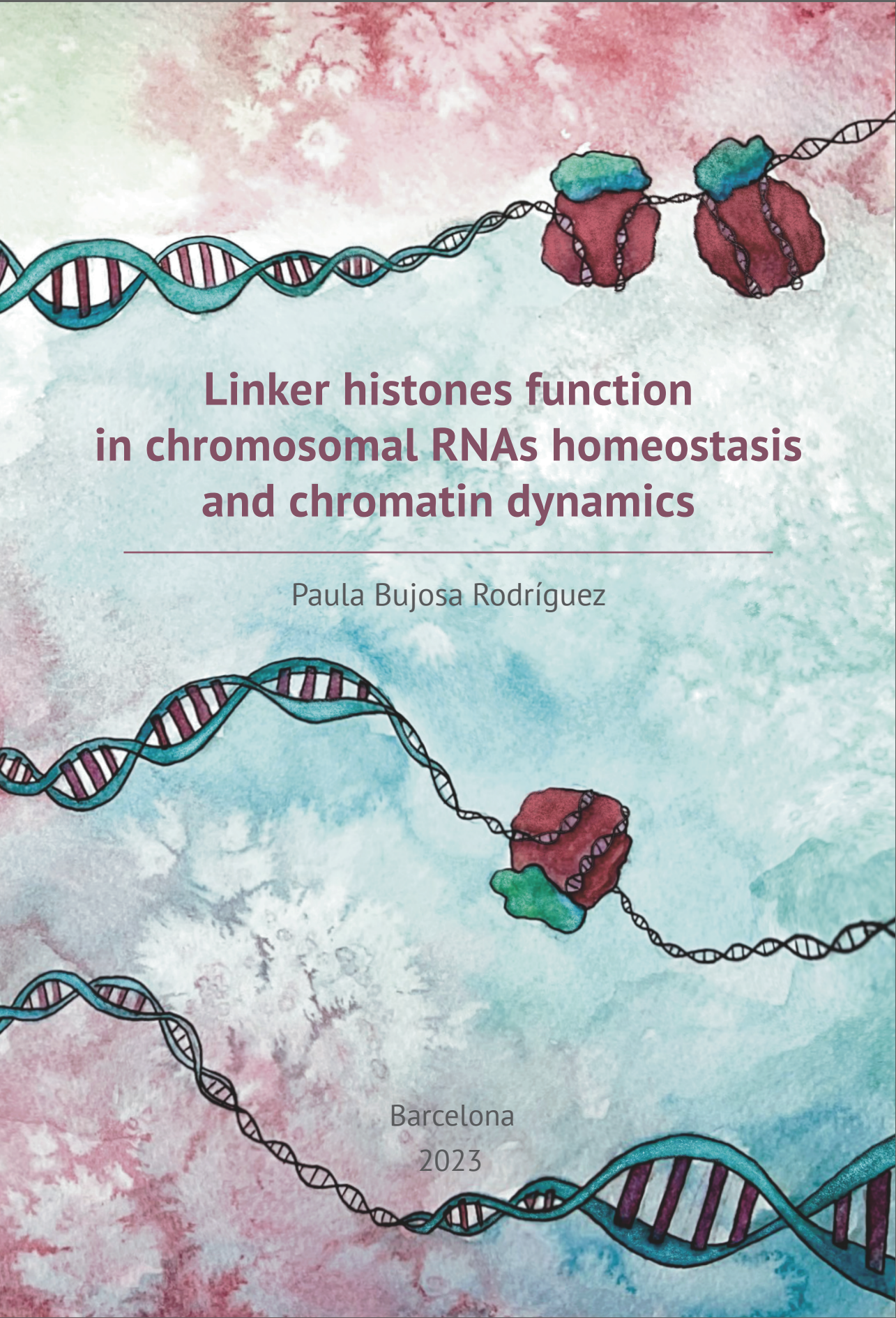
UNIVERSITAT DE BARCELONA

Linker histones function in chromosomal RNAs homeostasis and chromatin dynamics

Paula Bujosa Rodriguez



Linker histones function in chromosomal RNAs homeostasis and chromatin dynamics

Paula Bujosa Rodríguez

Barcelona
2023

Tesi Doctoral
Universitat de Barcelona

Linker histones function in chromosomal RNAs homeostasis and chromatin dynamics

Memòria presentada per
Paula Bujosa Rodríguez

Per optar al grau de Doctora per la Universitat de Barcelona
Programa de Genètica
Institut de Recerca Biomèdica de Barcelona, IRB Barcelona
Institut de Biologia Molecular de Barcelona, CSIC

Directors de la tesis,

Dr. Jordi Bernués Martínez

Dr. Ferran Azorín Marin

Tutor de la tesis,

Dr. Cristian Cañestro García

Doctoranda,

Paula Bujosa Rodríguez

Barcelona 2023

A mon pare, per ser es primer en donar resposta i alimentar sa meva curiositat científica, perquè – tot i no ser-hi – sempre hi ets.

ACKNOWLEDGEMENTS

M’agrada pensar que els agraïments són alhora el principi i el final de la tesi. I és que han estat molts anys que no hauria pogut viure i gaudir com ho he fet si no hagués tingut la sort d’estar acompanyada de tanta gent fantàstica a qui tinc tantes coses a agrair.

Vull començar agraint les tres altres persones que signen aquesta tesi. Ferran, sense el teu primer vot de confiança no podria estar escrivint aquests agraïments, per això i per haver dirigit la meva tesi sempre tenint en compte els meus interessos i aprenentatge, moltes gràcies. Jordi, no tinc paraules per agrair la direcció i acompanyament que m’has ofert durant la tesi i tot el que m’has ensenyat, tant a nivell experimental com a nivell de pensament científic i personal. Moltes gràcies per ser un mentor excepcional sempre amb temps i bona disposició per ajudar-me amb qualsevol problema o, senzillament, per parlar de ciència amb mi una estona. Cristian, moltes gràcies per després de ser el meu tutor de TFG i TFM ser també el meu tutor de tesi. Des de l’Oikolab he canviat molt de temàtica, però encara a dia d’avui recordo i aplico consells que em vas donar en aquell temps.

Mònica, el laboratori no s’entén sense tu, gràcies per sempre estar al cas de tot, i, encara que a vegades ens hagis de fer de mare de més, moltes gràcies també per sempre tenir paciència i ajudar-me amb tot!

Moltes gràcies també als meus companys inicials de U. Sry, thank you for your kindness and support. It was a pleasure to share time with you! Albert, no sé si sabré treballar sense tu assegut al meu costat. A qui li explicaré totes les meves preocupacions i alegries cada matí? Em vas dir un dia que la meva arribada al lab t’ha portat mala sort, així que ara que marxaré et desitjo que a la fi et vingui tota la sort que et mereixes! (I que ens puguem seguir veient de tan en tant).

Gràcies també a les Pauletes del laboratori: Paula Climent i Paula Escudero per tots els ànims i consells al llarg d’aquests anys! Escudero, va ser un privilegi i un plaer poder compartir amb tu la meva primera i la teva última

meitat de la tesi. Moltes gràcies per passar de ser la meva companya i suport diari al laboratori a la meva gran amiga fora d'ell.

Moltes gràcies també a les persones que heu entrat al lab després que jo. Maria, encara que a tots ens hauria agradat que pogués ser més, moltes gràcies pels dos anys compartits. Estic segura que allà on ets ara també ets una font d'alegria. Juan, no podría haber encontrado mejor compañero para mi etapa final de la tesis. Muchas gracias por siempre ayudarme tanto en las tareas del laboratorio y experimentos que hemos compartido, pero sobretodo, muchas gracias por escuchar siempre todas mis historias y comentar conmigo todas las series. Eres el siguiente... y no tengo ninguna duda de que irá todo genial!

Agraeixo molt també la arribada del Vicent Lab al nostre costat, en poc temps han passat moltes personetes aportant un extra de vida al laboratori. Especialment, Guillermo, Chari, Marta i Judit, encara que sigueu "bonitos però mentirosos" m'heu ajudat molt a estar bé i gaudir l'etapa final del doctorat, per això, moltes gràcies!

Fora del laboratori també vull agrair als meus amics predocs (tot i que ja casi tots sou doctors). Que consti que esteu per ordre alfabètic perquè no us baralleu, que això dels agraïments de tesi ha acabat semblant l'autoria d'un article! I jo us estimo molt a tots: Adrià, Clara, Hania, Marina i Nico. Aquest 2022-2023 ens hem anat veient acabar els uns als altres, moltes gràcies per sempre ser un grup de suport de tot el relacionat amb el doctorat, i més enllà!

Mireia, moltíssimes gràcies per tot el que m'has ensenyat i ajudat, la segona part de la tesi hauria estat complicadíssima sense tu!

No puc oblidar a les quatre membres del IRB que han fet possible part dels experiments i el seu anàlisi. Nacho, David, Oscar i Adrià moltíssimes gràcies per tot!

També vull mencionar i agrair els amics que m'han acompanyat fora de l'entorn científic: Marta, Dàhlia, Pablo, Miguel, Marina, Olaia, Alex, Dani,

Veronica, Marc. Hem anat canviant de bar però no ha canviat que passar temps amb vosaltres és una part fonamental del que em permet dedicar-me amb ganes i energia a la feina. Per tant, per la part que us toca respecte a la tesi, moltes gràcies! Fernanda, muchas gracias por, una vez más, acompanyarme en otra etapa muy importante de mi vida. Març, moltes gràcies per les rialles i reflexions que compartim!

També vull agrair a sa meva família Rodríguez-Frau, per acompanyar-me durant aquesta etapa de sa meva vida, com en totes, i tot i no entendre ben bé què faig sempre estar al meu equip animant-me i alegrant-vos per totes ses coses bones que me passen. Dins de l'equip d'animadors no puc deixar de fer-te una menció especial Kika, que tot i no ser sa meva predina jove sempre m'has cuidat com una, cobrint-me d'amor i de regals i mai faltant a cap data senyalada de sa meva vida, moltes gràcies per sempre tenir-me present.

Parlant de família, es nucli Bujosa-Rodríguez també mereix agraïments especials: Papa, has estat sa persona que més m'ha animat a ser investigadora des del canvi de carrera fins al moment de decidir iniciar una tesis doctoral, i, en realitat, des de molt abans, moltes gràcies. Mamà, mai has acabat d'entendre perquè serveix això que faig ni com viuré d'això, però moltes gràcies per s'esforç que has fet (i encara a vegades has de fer) per entendre com pens i ses coses que faig. Sobretot gràcies perquè no te faci falta estar d'acord amb mi per ser un gran suport que ha fet possible tot el que he fet. A sa meva germana, Aida, has estat més clau del que te penses per aquesta tesi, ja que si tu no haguessis estudiat medicina jo no hauria estudiat biotecnologia, per tant no puc deixar d'agrair-te moltíssim que m'obrissis sa primera porta cap a sa biologia i que fossis la primera en animar-me a canviar de camp. En general, moltes gràcies per ser sa meva estimada germana gran que sempre m'ha obert camí.

Per acabar, moltes gràcies també en es meu nucli familiar més petit, des qual només tu i jo en som part (esper que Bujosa-Rojo). Pep, te vaig conèixer a l'inici d'aquest camí, en el qual m'has acompanyat, cada vegada

més i de més a prop. Sent al final puc dir que no hauria volgut fer-lo amb ningú altre que no fossis tu. Gràcies per ser sa persona amb sa que més ric i jugo, sense mai menystenir ses meves preocupacions, fins i tot quan són una tonteria, i acompanyar-me en els moments realment dolents. Tens el poder de fer que qualsevol angoixa que porti a sobre desaparegui quan arribo/es a casa i estem junts. Moltes gràcies per compartir sa ciència i s'etapa predoc amb mi. Ja que jo te vaig cedir sa meva cita per la tesi, me prenc sa llibertat de robar-te sa teva darrera frase dels agraïments, i és que jo també t'estim com una fosfina a un centre metàl·lic.

ABSTRACT

Linker histones are essential proteins involved in higher-order chromatin structures and architectural functions. However, their roles extend beyond structure, encompassing various chromatin metabolic processes such as DNA replication, repair, and epigenetic modulation. To unravel the functional complexity of linker histones, we chose the model organism *Drosophila melanogaster* due to its peculiarity of only having a single somatic variant (dH1) and one germinal/embryonic variant (dBigH1).

First, we explored the consequences of dH1 depletion on chromatin structure and stability. Remarkably, depletion of dH1 resulted in genomic instability and the accumulation of RNA:DNA hybrids (R-loops) in heterochromatin. To explore the molecular mechanism behind these observed effects, we decided to study the chromosomal RNAs (cRNAs). Our findings revealed that cRNAs, a diverse group of RNA species co-purifying with chromatin, extensively covered approximately 28% of the genome. Notably, cRNAs exhibited a significant enrichment in heterochromatic transcripts and displayed strong associations with hnRNP A/B proteins hrp36 and hrp48, forming ribonucleoprotein (RNP) complexes. Intriguingly, the depletion of dH1 led to impaired assembly of hrp36 and hrp48 onto heterochromatic cRNAs, contributing to increased R-loop formation and abnormal cRNA retention. Furthermore, the altered chromatin organization resulting from dH1 depletion, characterized by reduced nucleosome occupancy and enhanced heterochromatin accessibility, facilitated the annealing of cRNAs to DNA templates. These findings provide unexpected insights into the involvement of linker histones in RNP assembly and cRNA homeostasis.

The second part of this study focused on investigating the biochemical properties and the phase separation capabilities of somatic dH1 and germinal/embryonic dBigH1 linker histones. Phase separation of chromatin is a fundamental process associated with diverse chromatin functions. Our analyses demonstrated the distinct characteristics of dH1 and dBigH1 in

promoting phase separation, suggesting their potential roles in different functional states of chromatin.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

3D	Three-dimensional
5hmC	5-hydroxymethylcy-tosine
6mA	6-methyladenosine
aa	Amino acid
ChIP	Chromatin immunoprecipitation
cRNA	Chromatin-associated RNA/ chromosomal RNA
CTD	C-Terminal Domain
DIC	Differential interference contrast
DOC	Sodium deoxycholate
DSB	Double-strand breaks
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
FC	Fold Change
FRAP	Fluorescence recovery after photobleaching
Fw	Forward
GD	Globular Domain
H1	Linker Histone
hnRNP	heterogeneous nuclear RNPs
HP1a/α	Heterochromatin protein 1a/α
IDR	Intrinsically disordered regions
IR	Intron retention
KD	Knock-Down
KO	Knock-Out
LLPS	Liquid-liquid phase separation
LSPS	Liquid–solid phase separation
m6A	N6-methyladenosine
Mnase	Micrococcal Nuclease
mtDNA	mitochondrial DNA
NCP	Nucleosome core particle
NRL	Nucleosome repeat length
NTD	N-Terminal Domain
O/N	Over night
PEV	Position-effect variegation
PTM	Post-translational modification
qPCR	Quantitative PCR

RBP	RNA binding protein
RE	Repetitive elements
RF	Replication fork
RNApol II	RNA polymerase II
RNP	Ribonucleoproteins
RT	Room temperature
Rv	Reverse
SN	Supernatant
ssDNA	single-stranded DNA
Suv3-9	Unified reference to both *D. melanogaster* Su(var)3-9 and mammals SUV39H1 and SUV39H2
TF	Transcription factor
THSS	Transposase hypersensitive site
TSS	Transcription start site
WB	Western blot
WT	Wild type

1. INTRODUCTION

1.1. CHROMATIN ORGANIZATION

Living organisms rely on nucleic acids to store genetic information, which determines their growth and biological structures. Eukaryotic organisms are characterized for having their DNA located in the nucleus, separated from the rest of the cell. Within the nucleus, DNA is packaged into chromatin, a complex of DNA and proteins that help to condensate, stabilize, and regulate the genetic information.

1.1.1. Chromatin compaction

Chromatin must be extremely condensed to fit into the nucleus, while maintaining a certain level of accessibility to the DNA, in order to facilitate and regulate essential processes such as transcription, replication, and repair. To achieve this, chromatin has different levels of organization associated with levels of packaging (**Figure 1.1**).

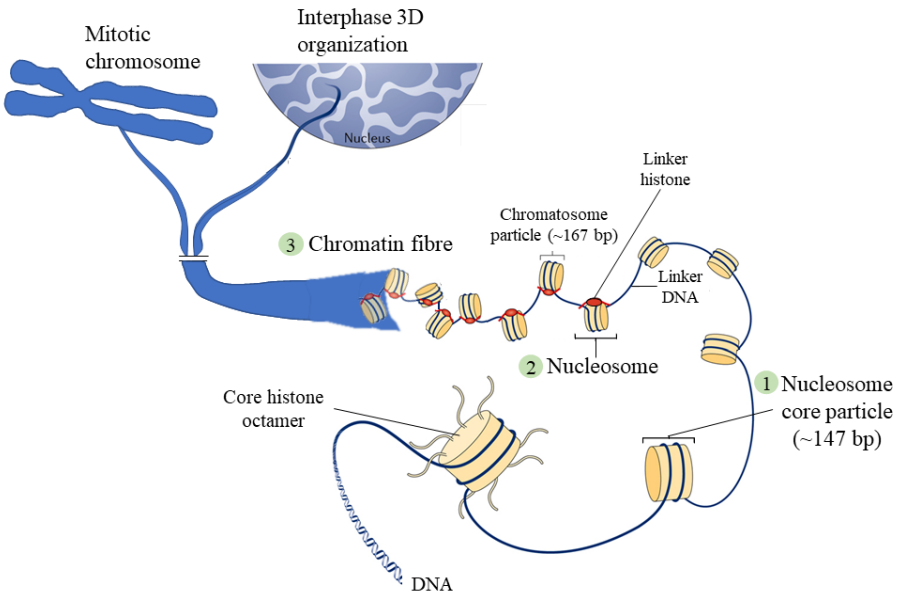


Figure 1.1. Levels of chromatin compaction. DNA compaction within the nucleus presents different structures: the nucleosome core particle (1), the nucleosome containing the linker DNA and the linker histone (2), and the chromatin fiber (3), formed by consecutive nucleosomes separated by the linker DNA. The folding and association of individual fibers in mitosis leads to formation of the condensed mitotic chromosome, and in interphase, to tertiary structures spatially organized. Adapted from (Fyodorov et al. 2018).

1.1.1.1. Nucleosome core particle: The core histones

The most basic packaging level of the chromatin is the nucleosome core particle (NCP). It is among the best conserved structures in metazoa (Clapier et al. 2008), consisting of ~147 bp of DNA that is tightly wrapped around the core histone octamer in about 1¾ turns of left-handed superhelix (Davey et al. 2002; Luger et al. 1997).

The core histone octamer is composed of two copies of each of the four core histones: H2A, H2B, H3 and H4 (**Figure 1.2A**). They are small proteins (~11-15 kDa), heavily basic, extremely conserved in eukaryotes and that have an α-helical structured C-terminal domain (CTD) and a small non-structured N-terminal domain (NTD). Although the primary sequences of the four core histones do not have a high degree of homology between each other, they have a central nearly identical structure that forms the "Histone fold" motif (Arents et al. 1991). This motif consists of three α-helices; one central and bigger α-helix (α2) flanked by two shorter helices (α1 and α3) (**Figure 1.2B**).

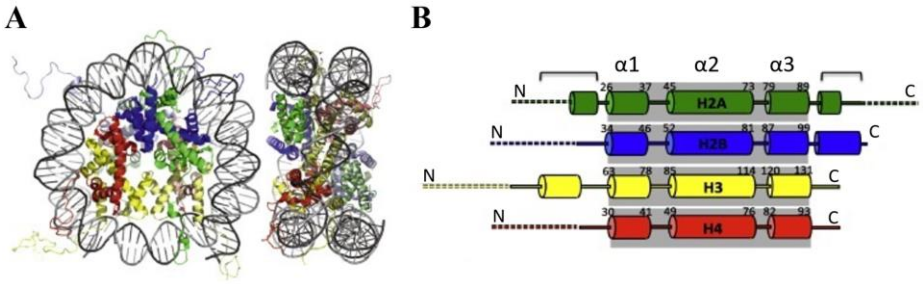


Figure 1.2. Structure of the nucleosome core particle. (**A**) Model of a nucleosome core particle from the frontal (left) and lateral (right) view. Core histones are painted as follows: H2A, green; H2B, blue; H3, yellow; H4, red. Proteins in lower half of nucleosome are painted with a lighter shade of the color. (**B**) Diagram of the secondary structure of the core histone proteins. Cylinders represent α-helices and the "Histone fold" is highlighted in grey. Additional helices outside the histone fold domain are indicated by brackets. The color code of the histones is maintained. Adapted from (Cutter & Hayes 2015).

The "Histone fold" motif is involved in the formation of specific heterodimers: H3/H4 and H2A/H2B. In order to form the octamer two

H3/H4 heterodimers self-associate forming a tetramer, then two H2A/H2B heterodimers bind to the tetramer concomitantly with the wrapping of the DNA, giving rise to the NCP. On the other hand, the unstructured NTDs of the core histones are also implicated in the stability of the NCP. The NTDs protrude from the octamer and electrostatically interact with the negatively charged DNA, thereby stabilizing the interaction (Cutter & Hayes 2015; Kornberg & Lorch 1999; McGinty & Tan 2015). Furthermore, the chemical modification of the NTDs of the core histones is key for the regulation of chromatin.

1.1.1.2. Nucleosome: The linker DNA and the linker histone

The NCP is the simplest structure of the chromatin. However, the nucleosome is what constitutes the fundamental structural repeating unit. It consists in the NCP, the linker DNA (DNA found in between two consecutive nucleosomes; **Figure 1.1**) and a protein called linker histone that binds to the nucleosome at the DNA entry/exit site (Kornberg 1977).

This repeating unit was first identified by digestion with Micrococcal Nuclease (MNase), an endonuclease that is able to preferentially cut the linker DNA. From this digestion two types of fragments are obtained: the NCP consisting only of 147bp DNA wrapped around the octamer (the linker DNA is fully degraded), and the chromatosome particle of about ~160-200 bp including the NCP, the linker histone and an extra ~10-50 bp of linker DNA protected from degradation by the linker histone (Simpson 1978).

While the evolutionary origins of linker histones can be traced back to bacteria, their abundance and structural conservation vary significantly across different taxa (Kasinsky et al. 2001). In metazoa, however, linker histones are typically part of the nucleosome. In those, the stoichiometry of the linker histone to the core histones ranges from 0.5 up to 1.3 depending on cell types, although in general there is approximately one linker histone per nucleosome (Woodcock et al. 2006). However, despite having a relative equivalent presence in chromatin as the core histones, linker histones are much more mobile: the residence time of the linker histone is of several

minutes (Lever et al. 2000; Misteli et al. 2000), in comparison to hours, days and longer for the core histones (Kimura & Cook 2001). Moreover, linker histones do not have such a high degree of conservation as core histones (Kasinsky et al. 2001). Nevertheless, they are also heavily basic and relatively small proteins, that share a tripartite structure consisting of a short N-terminal tail, a central globular domain, and a long C-terminal tail in metazoa (Harshman et al. 2013). Linker histones are fundamental to give rise and organize higher-order chromatin structures, yet they are not only architectural proteins, since their functions also include regulation of a wide variety of chromatin processes (Fyodorov et al. 2018).

The elucidation of the molecular structure of the chromatosome particle was a big challenge being only solved at a near-atomic resolution for the first time less than ten years ago (Zhou et al. 2015). Zhou et al. obtained the structure of the chromatosome with the globular domain of *G. gallus* linker histone H5. They found that the globular domain of H5 was binding on the dyad of the nucleosome, at the center of the entry/exit site of nucleosomal DNA interacting with both linker DNA chains (**Figure 1.3** Left). Currently, the obtention of chromatosome structures containing the globular domain of different linker histones (Öztürk et al. 2018), has unveiled that liker histones bind to the nucleosome in two major modes: establishing symmetric (as previously described) or asymmetric contacts with the linker DNA, named On-dyad or Off-dyad respectively (**Figure 1.3**) (Fyodorov et al. 2018; Zhou & Bai 2019).

These two modes depend on the linker histone sequence and offer different protection to the linker DNA, which combined to the variable linker DNA length, accounts for the range of different fragments obtained when digesting chromatin of different organisms and cell types. The measure of the average distance between the dyads of neighboring nucleosomes can be obtained from digestions with MNase, and it is referred to as the nucleosome repeat length (NRL), a biologically relevant property of the chromatin fiber (Perišić et al. 2010).

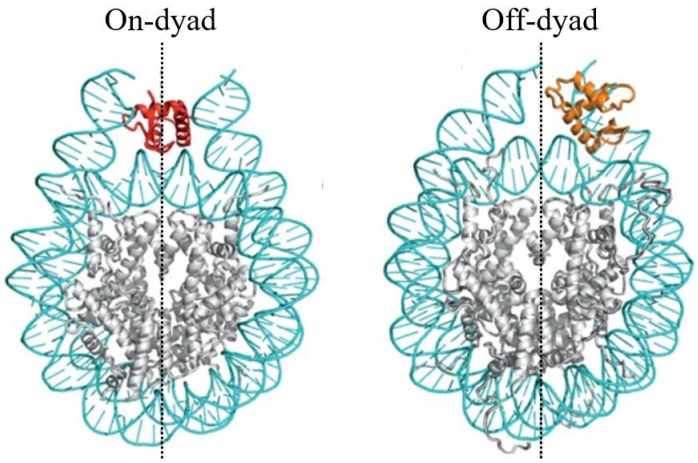


Figure 1.3. Structure of the chromatosome: On-dyad and Off-dyad mode of binding of the linker histones. Left, On-dyad binding of the globular domain of *G. gallus* H5 to the mono-nucleosome; Right, Off-dyad binding of the globular domain of *D. melanogaster* dH1 to the mono-nucleosome. The nucleosome dyad axes are represented with dashed vertical lines. Adapted from (Fyodorov et al. 2018).

1.1.1.3. Chromatin fiber

The simplest chromatin fiber consists of rows of nucleosomes connected by linker DNA (**Figure 1.1**), also known as the 10 nm fiber. This chromatin fiber was first observed in 1973 and was coined with the iconic name of 'beads on a string'(Olins & Olins 2003), where the nucleosomes would be the beads and the DNA the string connecting them (**Figure 1.4** Left).

How this chromatin fiber is further compacted and organized to fit 2 m of genomic DNA into a ~10 μm diameter nucleus has been controversial for many years. Initially, *in vitro* experiments showed 30 nm long and symmetric chromatin fibers, for which the two most famous structural models where the "one-start helix", also called solenoid, (Finch & Klug 1976) and the "two-start helix", also called zig-zag (Woodcock et al. 1984) (**Figure 1.4** Center). Many more models were proposed along the years since extensive studies were done on the 30 nm fibers (van Holde & Zlatanova 2007). Thus, the 30 nm fiber was widely accepted to be the higher-level folding of the 10 nm fiber, and there was the conception of an "hierarchical helical folding model" suggesting that a 30 nm fiber was

progressively folded into larger fibers, including ~100 nm and then ~200 nm (Horn & Peterson 2002). However, nowadays this has been discarded, since many experiments have revealed that the chromatin fiber of 30 nm does not exist as a long symmetric helical structure *in vivo*, neither in interphase nor in the mitotic chromosome (Dubochet et al. 1988; Fussner et al. 2012; Joti et al. 2012; Maeshima et al. 2014; McDowall et al. 1986; Ou et al. 2017). Instead, in interphase, the 10 nm fiber appears to be *trans*-interdigitated in irregular "clutches" of nucleosomes forming heterogenous and diameter-variable fibers (Maeshima et al. 2019; Ricci et al. 2015).

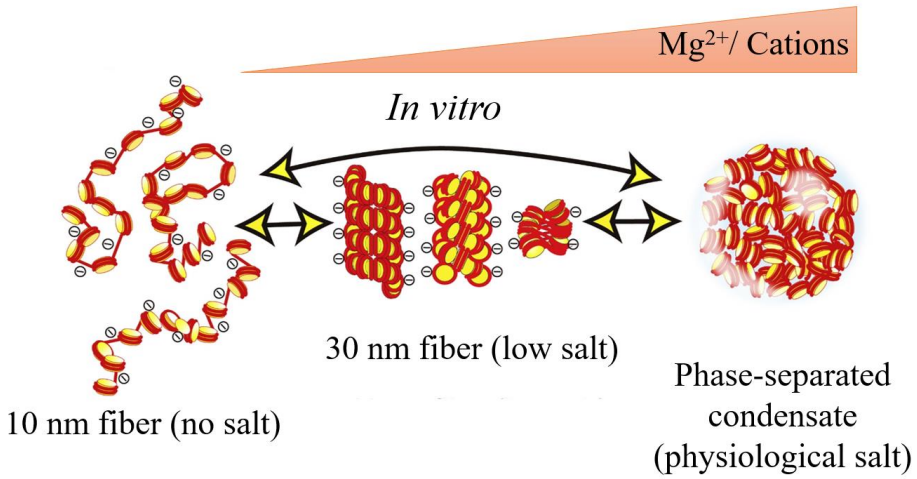


Figure 1.4. Structural variation of chromatin *in vitro* depending on salt condition. Left, 10 nm fiber formed in no salt conditions. Center, Different examples of 30 nm regular fibers formed in the presence of low salts. Right, large globular condensates with interdigitated 10-nm fibers (without the 30 nm structure) in more physiological salt condition. Adapted from (Maeshima et al. 2020).

This initial *in vitro* artifact was consequence of performing the experiments on ionic conditions far from physiological. Since DNA carries a high negative charge from its phosphate backbone, the core histones partially neutralize these charges in the assembly of the nucleosome core. Nevertheless, other elements such as the linker histone are fundamental so that the DNA can be tightly folded and does not produce electrostatic repulsion between adjacent DNA regions (Woodcock et al. 2006). However, even with linker histones, there are negative charges that need to

be compensated for higher levels of compaction, consequently the cationic content in the cell is key for the condensation of the chromatin (Maeshima et al. 2018; Strick et al. 2001). Therefore, when we observe the chromatin structure *in vitro* under no salt condition, the repulsion between nucleosomes is at is maximum, and we have the most stretched form of the 10 nm fiber that would correspond to the "beads on a string" observation (**Figure 1.4** Left); under low salt, the nucleosomes gently repel each other, favoring their binding to neighboring nucleosomes rather than distant ones, and giving rise to the regular helical 30 nm structures (**Figure 1.4** Center) (e.g., lower than ~1mM Mg^{2+}, depending on size and concentration of the fiber). On the other hand, if we place the chromatin in physiological salt condition, it produces large globular condensates where the electrostatic repulsion is reduced, allowing nucleosomes to interact with distal partners, and forming the irregular interdigitated rows of the 10 nm fiber that are observed *in vivo* (Maeshima et al. 2020). Moreover, those *in vitro* condensates are formed by phase separation (**Figure 1.4** Right). Oversimplifying, phase separation is a phenomenon similar to the demixing of oil and water that leads to the formation of oil droplets in water (Gibson et al. 2019; Muzzopappa et al. 2021).

All in all, the current conception of chromatin compaction in the nucleus is of a very dynamic structure that can be either folded and compacted or accessible depending on the needs of the cell.

1.1.2. Epigenetic modifications

Chromatin is also regulated by elements that go beyond the sequence of the DNA and locally alter their properties. These elements are what we call epigenetic modifications. Epigenetic means literally "added to the genetic information" ('epi': Greek for "on top of" or "in addition to"). It is clear that DNA is not sufficient to explain phenotypes. Paying attention at pluricellular organisms, one can become aware that despite all the cells of that organism have the same genetic information, they perform very different functions and have different morphological structures, giving rise

to largely different tissues such as the human skin, liver, or brain. Conrad Waddington defined the molecular mechanisms that transform the genetic information into observable phenotypes as "epigenetic landscape" (Waddington 1957). Nowadays, epigenetic modifications include stable yet reversible changes that lead to a change in the phenotype without changing the genotype (Portela & Esteller 2010). The four main elements implicated in the epigenetic regulation are: DNA methylation, chromatin remodeling, histone variants and histone modifications.

1.1.2.1. DNA methylation

DNA methylation in general refers to the chemical modification of adding a methyl group to the nucleobase cytosine to form 5-methylcytosine (5mC). It can happen in any cytosine, but the modification is most prevalent at CpG dinucleotides (Schmitz et al. 2019), that in mammalians tend to cluster in regions called CpG islands and function as promoters for approximately 60% of human genes (Portela & Esteller 2010). DNA methylation is a regulator of transcription repression. It can lead to repression by different mechanisms: the presence of 5mC can directly mask the motif of a transcription factor binding, it can attract methylcytosine-binding proteins that in turn will block the motif, or it can act as a scaffold for histone modification enzymes and chromatin remodelers that will make the genomic region inaccessible (Attwood et al. 2002).

Moreover, in addition to the extensively studied 5mC, there are also other DNA modifications that have received less attention and are sometimes overlooked. For instance, 6-methyladenosine (6mA) is a modification that has been characterized in prokaryotes, but it is also present in eukaryotes (Iyer et al. 2011). Furthermore, cytosines can also be modified into 5-hydroxymethylcytosine (5hmC), which can lead to their demethylation and potentially influence the regulation of DNA methylation (Dahl et al. 2011).

1.1.2.2. Chromatin remodeling and Histone variants

Chromatin remodeling is mediated by large protein complexes that are known to slide, destabilize, evict, or restructure nucleosomes in an ATP-

dependent manner (Eberharter & Becker 2004). Chromatin remodelers are crucial for the basic biology of the cell, as they fulfill several important functions. For example, they provide long stretches of DNA to the DNA repair machinery (Osley et al. 2007). They also ensure proper nucleosome spacing after replication (Yadav & Whitehouse 2016). Furthermore, they play a role in local changes in accessibility that can alter transcription, often in conjunction with other epigenetic modifications (Lorch & Kornberg 2017). In addition, chromatin remodelers are involved in restructuring nucleosomes with specific histone variants, necessary to mark certain events or give rise to specialized structures (Martire & Banaszynski 2020).

Histone variants are non-allelic variants of what we call canonical histones. We can distinguish canonical histones from variants because, in most metazoa, they are clustered in repeated arrays that are transcribed at very high levels during the S-phase of the cell cycle, while histone variants are single copy genes usually transcribed at lower levels and not coupled to the cell cycle (Bayona-Feliu et al. 2016; Marzluff et al. 2002). Histone variants present compositional and structural variations relevant for the regulation of processes such as DNA repair, meiotic recombination, chromosome segregation, transcription initiation/termination, and sex chromosome condensation (Henikoff & Smith 2015).

There is an abundant variety of histone variants, but there are four that can be considered “universal” since they are found in nearly all eukaryotes: the centromeric histone variant H3 (CenH3, also known as CENP-A in *H. sapiens* and centromere identifier (CID) in *D. melanogaster*), variant necessary for the centromere formation and essential for assembly of the kinetochore; H3.3, variant nearly identical to the canonical H3 that may confer some increased accessibility; H2A.Z, variant usually located at the TSS involved in the initiation of transcription; and H2A.X, variant involved in signaling in response to double-strand breaks (DSBs) and scaffold DNA repair proteins, histone modifying enzymes and chromatin remodeling complexes (reviewed in (Talbert & Henikoff 2010)). Remarkably, in *D.*

melanogaster the H2Av variant is the functional homolog of both H2A.Z and H2A.X (Madigan et al. 2002).

1.1.2.3. Histone modifications

Core histones can also acquire specific properties without being exchanged for variants, as they can get histone modifications. These are covalent post-translational modifications (PTM), the vast majority on the unstructured NTDs of the core histones, that are deployed, recognized, and removed by a large number of writers, readers, and erasers, respectively (Zhang et al. 2021). PTMs can take place in many different residues, meaning that in each nucleosome different histone tails can present different modifications, and the same histone tail can present multiple modifications. Moreover, they include acetylation, methylation, phosphorylation, ubiquitination, SUMOylation, ADP-ribosylation, among others (**Figure 1.5**) (Jenuwein & Allis 2001; Kouzarides 2007).

Histone modifications regulate chromatin by either changing the properties of the NTDs to disrupt the contacts between the DNA and the nucleosome core or by recruitment of non-histone proteins, working as scaffold for remodeling complexes and other proteins (Kouzarides 2007).

Figure 1.5. Histone modifications. Main PTMs in the NTDs of H3 and H4: acetylation (blue), methylation (red), phosphorylation (yellow) and ubiquitination (green). The number in gray under each amino acid represents its position in the sequence. Modifications associated with euchromatin, and heterochromatin are highlighted with green and red rectangles respectively. Adapted from (Portela & Esteller 2010).

Thus, combined with the other epigenetic molecular mechanisms, histone modifications are important for transcription regulation, DNA repair, DNA replication and chromosome condensation (Portela & Esteller 2010).

1.1.3. Chromatin states

The integration of the different levels of chromatin compaction and all the epigenetic modifications in the cell gives rise to functional chromatin states. The most canonical definition of chromatin states precedes the study of epigenetic modifications. Observations of the interphase nucleus using optical microscopy led to the identification of two distinct chromatin states, based on the staining pattern of chromatin during interphase: a condensed state of chromatin, characterized by strong staining, and a decondensed state, characterized by weak staining, corresponding to heterochromatin and euchromatin, respectively (Heitz 1928). Therefore, the first description of those two states was based on their different degree of compaction. Later, it has been demonstrated that those states are also associated with many other molecular features, such as specific histone marks: euchromatin presents high levels of acetylation, such as H3K27ac (Histone H3, lysine 27 acetylated), and methylation of H3K4me3 (Histone H3, lysine 4 tri-methylated), H3K36me3 and H3K79me3, while heterochromatin presents low levels of acetylation and high levels of H3K9me2, H3K9me3 and H3K27me3 (**Figure 1.5** and **1.6**) (Portela & Esteller 2010).

Systematic studies of histone modifications and the analysis of their combinatorial patterns have allowed to identify fine-tuned chromatin states that correlate with chromatin features (Ernst & Kellis 2010; Kharchenko et al. 2011). For instance, Kharchenko *et al*. (2011), studied the localization of eighteen different histone modifications in *D. melanogaster* cells and found combinatorial patters – applying a multivariate hidden Markov model – that defined nine chromatin states (**Figure 1.6A**):

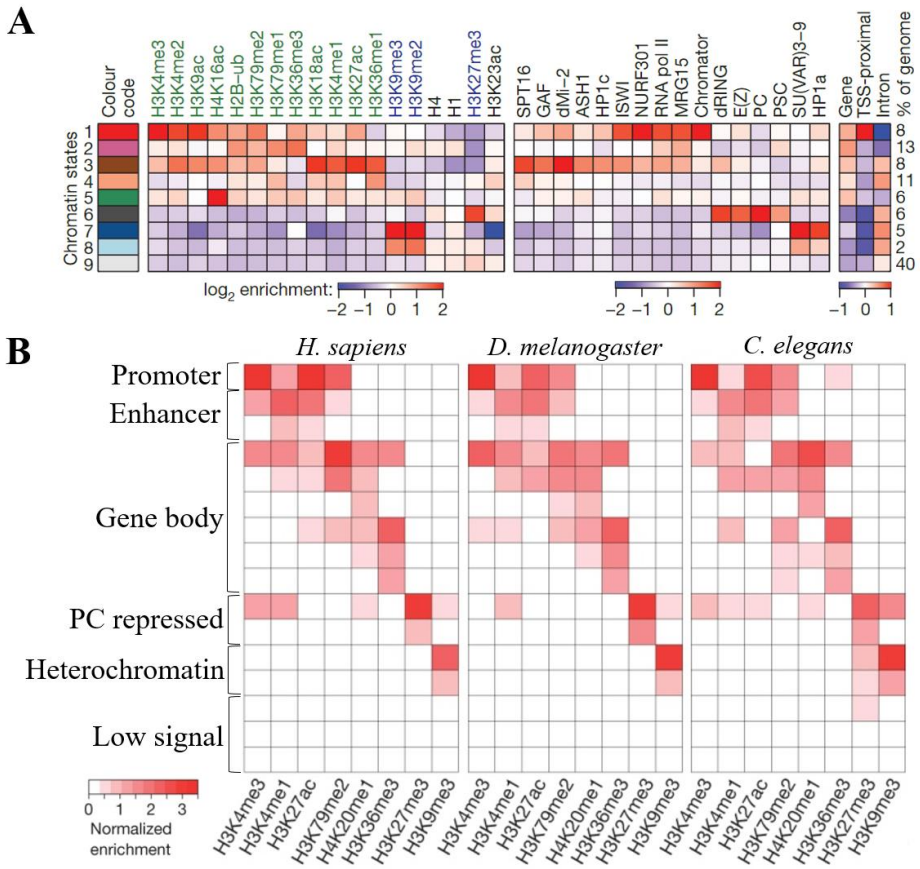


Figure 1.6. Histone modifications encode for chromatin states. (**A**) The first panel shows the enrichment patterns of eighteen histone marks that define nine chromatin states in the *D. melanogaster* genome. The second panel shows average enrichment of chromosomal proteins. The third panel over/under representation of the states in different genomic features. Adapted from (Kharchenko et al. 2011) (**B**) Enrichment patterns of eight histone marks that determine six chromatin states in *H. sapiens*, *D. melanogaster* and *C. elegans*. Adapted from (Ho et al. 2014).

State 1, associated to Promoters and TSS, marked by a strong enrichment in H3K4me3/me2 and H3K9ac; State 2, signature of transcription elongating genes, enriched in H3K36me3; State 3 and 4 are associated to active introns, enriched in H3K36me1, although state 3 is also enriched in H3K27ac, H3K4me1 and H3K18ac, and associates specifically to regulatory regions in introns; State 5, associated to active genes on the male X chromosome, enriched in H4K16ac and other "elongation marks" such as H3K36me3; State 6, corresponds to Polycomb-mediated repressed or

facultative heterochromatin, enriched in H3K27me3; State 7 defines the constitutive heterochromatin, enriched in H3K9me2/me3; State 8 consisting on heterochromatin-like regions with moderate levels of H3K9me2/me3; State 9 consists on transcriptionally silent intergenic domains, which is depleted for all chromatin marks.

Overall, different transcription-related features have associated different chromatin states and heterochromatin is mainly divided in two types. Those features and modifications distributions are very conserved between *H. sapiens* and *D. melanogaster*, and generally conserved with *C. elegans*, thus proving that the chromatin states are functional and conserved among metazoa (**Figure 1.6B**) (Ho et al. 2014).

In addition, chromatin states are not only described by histone modification, since they are also characterized by the distribution proteins bound to DNA (Filion et al. 2010; Kharchenko et al. 2011). Moreover, it has been shown that different chromatin states also correlate with different spacing and distribution of the nucleosomes. The differences are especially clear comparing actively transcribed regions with heterochromatin. In the first, linker length is shorter and nucleosome regularity is reduced thus translating into a reduced NRL, while in the latter, linker length is longer and nucleosome regularity is high resulting in longer NRL (Baldi et al. 2018).

All in all, the structure and the epigenetic modifications are not independent chromatin features, instead they are interrelated and have cross-talks that synergistically regulate the functionality of the genome.

1.1.3.1. Constitutive heterochromatin

As indicated in the chromatin states, there are two types of heterochromatins: facultative and constitutive. Although they both accomplish the general role of silencing transcription, they have fundamental and molecular differences.

Facultative heterochromatin is called facultative because it is formed in certain differentiated cell types, while it is not in other undifferentiated or

distinct differentiated cell types. Thus, facultative heterochromatin is strongly associated with differentiation, and its main molecular signature is H3K27me3, which is read and written by the Polycomb Group proteins that play a key role in establishing and maintaining facultative heterochromatin (Trojer & Reinberg 2007).

On the other hand, constitutive heterochromatin, aside from a chromatin state, is an essential nuclear compartment in eukaryotes that represents around 30 to 45% of *D. melanogaster* and *H. sapiens* genome, and it is heavily enriched in repetitive and transposable elements (Marsano & Dmitri 2022). Contrary to facultative heterochromatin, which is found spaced along the genome, constitutive heterochromatin is generally located in specific regions of the genome and conserved in different cell types. The largest portion of constitutive heterochromatin is surrounding the centromere (pericentromeric heterochromatin) and a smaller portion is at the end of the chromosomal arms (telomeres) (**Figure 1.7**) (Ho et al. 2014). On the other hand, it is found either forming perinucleolar domains and pericentromeric bodies, or large chromocenters located in the nuclear periphery at interphase (Janssen et al. 2018), displaying its characteristic high degree of compaction.

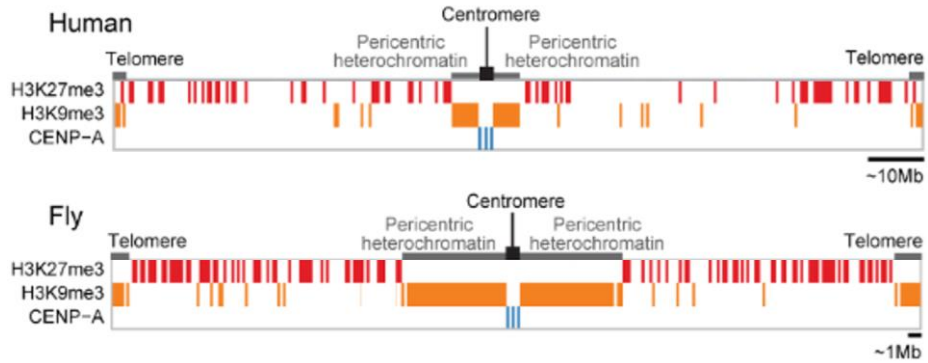


Figure 1.7. Distribution of facultative and constitutive heterochromatin canonical marks. Schematic diagrams of the distributions of enriched domains in H3K27me3 (red), H3K9me3 (yellow), and CENP-A (blue), in *H. sapiens* (top) and *D. melanogaster* (bottom). Adapted from (Ho et al. 2014).

The primary molecular signatures of constitutive heterochromatin are H3K9me2 and H3K9me3, which are written and read by the histone methyltransferase Suv3-9 (Su(var)3-9 in *D. melanogaster* and SUV39H1 and SUV39H2 in mammals) and it is strongly characterized by the occupancy of the Heterochromatin protein 1a/α (HP1a/α) (Allshire & Madhani 2018).

1.1.3.2. Pericentromeric heterochromatin formation and maintenance

Initiation of heterochromatin formation and maintenance have different molecular mechanisms, meaning that what is necessary for initiation does not influence the maintenance once heterochromatin is established.

Focusing on the pericentromeric heterochromatin (hereafter 'heterochromatin'), the initiation of heterochromatin is triggered by the transcription of the repetitive elements (Probst et al. 2010), which are retained as chromatin-associated RNAs (cRNAs). Those recruit Suv3-9 to the pericentromeric region either via the RNAi machinery (Allshire & Madhani 2018; Gu & Elgin 2013) or by direct association with Suv3-9 (Johnson et al. 2017). Moreover, in *D. melanogaster*, the presence of linker histone at the nucleation site has been identified as a relevant factor for the recruitment of Su(var)3-9 and the facilitation of H3 methylation within the chromatin substrate (Lu et al. 2013). After Suv3-9 is located, it initiates the deposition of methyl groups, thereby accomplishing the first step of heterochromatin nucleation. Once heterochromatin is nucleated, it has mechanisms to spread and maintain independently of DNA sequence (Allshire & Madhani 2018).

The canonical example of heterochromatin expansion is the "position-effect variegation" (PEV), that has been extensively studied in *D. melanogaster*. PEV occurs when an expressed gene normally located in euchromatin is rearranged, by transposition, nearby or inside heterochromatin, and the spread of the latter causes the transcriptional silencing of the gene (Elgin & Reuter 2013).

Molecularly, the spreading and maintenance of heterochromatin requires crosstalk among readers, writers, and erasers, which form positive feedback loops. The main actors are, once again, Suv3-9, that not only writes H3K9me3 but it is also a reader; HP1a/α, that is recruited by H3K9me3 and stabilized by Suv3-9; and histone deacetylases, erasers that are recruited by HP1a/α and condense the heterochromatin by removing the acetyl groups (Allshire & Madhani 2018; Grewal & Jia 2007; Janssen et al. 2018). This interplay of signals can expand the heterochromatin domain and make it inheritable, since histones containing the histone marks are randomly distributed in the two new chromatids after DNA replication, that would serve as nucleation spots/foci independent of *de novo* heterochromatin initiation. This will be spread by the reader–writer coupling (enabling its disassembly and assembly in each cell cycle) (Allshire & Madhani 2018). Nevertheless, this spreading needs to be controlled to maintain heterochromatin only in the corresponding regions. To achieve this, there are different mechanisms to create barriers to the expansion such as: controlling specific nucleosome-depleted regions or recruiting anti-silencing factors (Janssen et al. 2018).

Moreover, biophysical studies of HP1a/α have shown that heterochromatin compaction and compartmentation could potentially be accompanied by its phase separation modulated by HP1a/α (Keenen et al. 2021; Larson et al. 2017; Strom et al. 2017).

1.1.3.3. Constitutive heterochromatin and genomic stability

Genomic stability consists of ensuring the faithful transmission of genetic material from one generation, or somatic cell, to another. In contrast, genome instability refers to changes in the genome because of DNA damage, mutations, and other alterations of the chromosomal integrity caused by a wide range of factors. In this regard, the proper formation and maintenance of constitutive heterochromatin is key to maintain genomic stability.

First, the pericentromeric heterochromatin is surrounding the centromere, which is a region of the genome constituted by a very specialized chromatin – containing CENP-A and coding for arrays of tandem repeat sequences called satellites – different from euchromatin and heterochromatin (**Figure 1.7**) (Sullivan & Karpen 2004). However, it is not merely surrounding the centromere, since when pericentromeric heterochromatin is disrupted by either altering the H3K9me2/3 pattern or the functionality of HP1a/α, chromosome segregation presents errors (Ekwall et al. 1996; Peng & Karpen 2009; Peters et al. 2001). This indicates that pericentric heterochromatin integrity is necessary for centromeric function and mitotic fidelity (Janssen et al. 2018).

Moreover, constitutive heterochromatin maintains chromosomal integrity by having a specific molecular mechanism that spatially handles DSB, in order to repair them with reduced risk of aberrant repeat recombination from the abundant repeated sequences that it contains (Caridi et al. 2017; Janssen et al. 2018). In fact, when heterochromatin is not fully functional in *D. melanogaster*, it results in significantly elevated levels of extrachromosomal repeated sequences and aberrant intrachromosomal recombination among repeats (Peng & Karpen 2007).

Finally, the formation of heterochromatin is necessary to maintain the repetitive elements silenced. If this process fails, in *C. elegans*, it can give rise to an RNA containing potentially deleterious structure, called R-loops, which lead to DNA damage (Zeller et al. 2016). Moreover, it can destabilize the telomeres, or allow the potentially dangerous movement of mobile elements to other regions (Janssen et al. 2018).

All in all, proper functioning of constitutive heterochromatin is critical for the cell biology and absolutely required to avoid genomic instability, that in turn is related with aging and cancer progression (Janssen et al. 2018). The linker histone, as master regulator of chromatin compaction, has also been appointed as a key factor in establishing and maintaining the genomic stability of heterochromatin.

1.2. LINKER HISTONES

Linker histones are relatively small proteins, with a high positive charge, that bind to the DNA entry/exit site of the nucleosome.

1.2.1. Linker histones evolution and structure

The linker histone (H1) family, although it coincides in name with the core histones, has a completely independent evolutionary origin (Kasinsky et al. 2001). Furthermore, their sequence conservation in eukaryotes is drastically reduced compared to that of the core histones. For instance, the H1 homologs of *S. cerevisiae* and *H. sapiens* are 31% identical and 44% similar, while the core histone H4 homologs for those same organisms are 92% identical and 96% similar (Harshman et al. 2013). Metazoa also have poor sequence conservation: *D. melanogaster* and *H. sapiens* H1 homologs present 43% identity and 54% similarity. However, metazoa H1s present a conserved tripartite structure (**Figure 1.8** Bottom), in which a globular domain (GD) is flanked by two unstructured tails, the NTD and CTD (Harshman et al. 2013). The GD has a preference for nucleosome binding and is the most conserved region, particularly the positive residues that play an important role in nucleosome binding (**Figure 1.8** Top). Moreover, it gives rise to a conserved winged-helix motif with three alpha helices and a C-terminal beta hairpin (**Figure 1.8**) (Ramakrishnan et al. 1993). The flanking domains, NTD and CTD, are unstructured and poorly conserved. Indeed, the CTD – the larger domain – accounts for the major degree of variability, even among variants in the same species and it is involved in the modulation of many different functions of H1. Broadly, it contacts the linker DNA, and it is implicated in regulating the dynamics of the binding of H1 to the nucleosome and its residence time (Harshman et al. 2013). Moreover, it strongly influences the condensation of chromatin and mediates interactions with other proteins to regulate the chromatin state (Luque et al. 2014; McBryant et al. 2010; Perišić et al. 2019). On the other hand, the NTD is far less characterized, and it has been eventually stated as "non-essential", yet there is evidence of it affecting the binding affinity of

H1 to the nucleosome, although more research is needed to determine to what extent (McBryant et al. 2010; Vyas & Brown 2012).

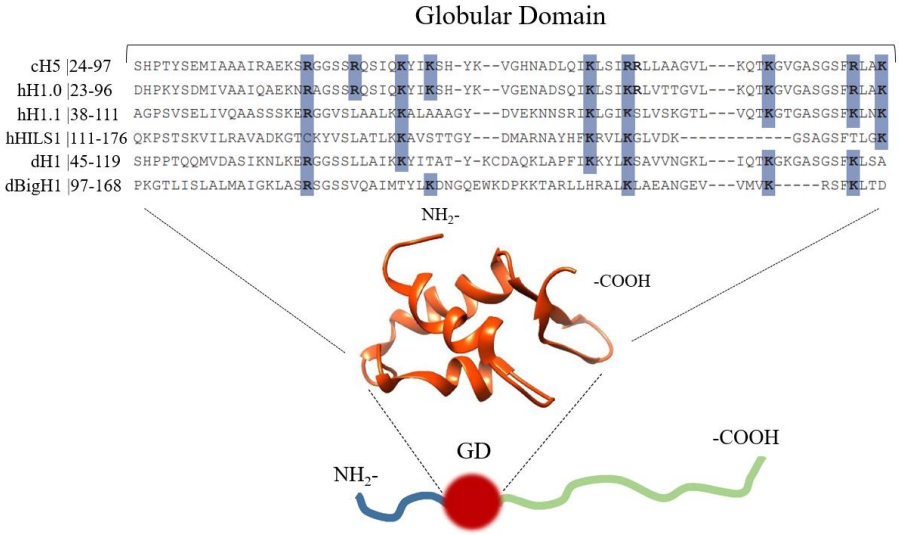
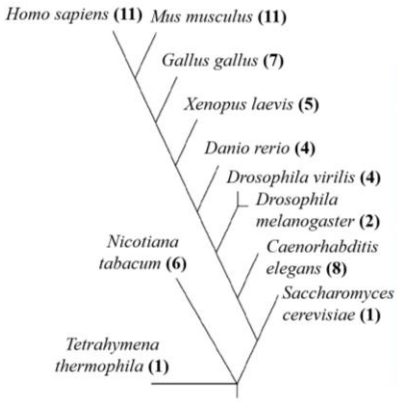


Figure 1.8. Linker histone structure and conservation of the globular domain. Top, alignment of the globular domain from *G. gallus* H5, *H. sapiens* H1.0, H1.1, HILS1, and *D. melanogaster* H1 and BigH1. Residues implicated in nucleosome biding are highlighted in blue. Bottom, tripartite H1 structure of a central structured globular domain (GD) flanked by short NTD and long CTD, both unstructured. Adapted from (Bednar et al. 2016; Fyodorov et al. 2018).

Similarly as for the core histones, all three domains are modified post-translationally. Despite there is a great variety of modifications, the most abundant PTMs in H1 are phosphorylation, methylation, and acetylation (Andrés et al. 2020), and they have a high impact in the multiplicity of H1 functions. However, the specific roles of the different PTMs are much less known for H1 than for core histones (Fyodorov et al. 2018).

1.2.2. Linker histone variants

Another common trait in metazoa, in addition to the reduced conservation of H1 sequence, is the existence of multiple variants in the same organism (**Figure 1.9**).



		Homo sapiens	*Mus musculus*	*Drosophila melanogaster*
Number of H1s		11	11	2
Cell-cycle dependent	**Somatic**	H1.1 H1.2 H1.3 H1.4 H1.5	H1a H1c H1d H1e H1b	dH1
Cell cycle independent	**Somatic**	H1x H1.0	H1.10 H1°	-
	Germline	H1t H1T2 HILS H1oo	H1t H1T2 HILS H1oo	dBigH1

Figure 1.9. H1 variants in different species and classification of linker histone variants. Left, species arranged in an evolutionary tree, the number of H1 variants is next to the species name in parenthesis. Right, classification and nomenclature of linker histone variants for *H. sapiens, M. musculus and D. melanogaster*. Adapted from (Izzo et al. 2008; Pérez-Montero et al. 2016; Prendergast & Reinberg 2021).

There are variants associated with the cell cycle, that are clustered and transcribed during the S-phase, when the majority of H1 is produced, and variants that are expressed independently of the S-phase, associated to specific differentiation states (Harshman et al. 2013; Izzo et al. 2008). Moreover, H1 variants that are cell cycle independent can also be differentiated into somatic or germline-specific variants (**Figure 1.9**). It is conserved among metazoa to have at least one germline variant, and some, like mammals, present specific sperm (H1t, H1T2, HILS) and oocyte (H1oo) variants. Germline variants are larger than somatic variants and are enriched in acidic residues (E, D). Furthermore, they exhibit highly specific expression patterns in the germline cells. Thus, they are involved in the spermatogenesis and oogenesis, and oocyte variants are retained during the early embryonic developmental stages. Generally, when zygotic transcription is activated, they are replaced by somatic variants (Reviewed in Pérez-Montero et al. 2016).

Studies of the mammalian somatic variants (using *H. sapiens* nomenclature) have shown that the relative expression and specific distribution of the different H1 variants are related to cell differentiation, given that different variants are enriched when comparing stem cells with differentiated cells

(Prendergast & Reinberg 2021). In fact, cell cycle-independent somatic variants – H1X and H1.0 – are characteristic of differentiated cells. However, they must be involved in different chromatin processes, since H1.0 is preferentially associated with transcriptionally silent genomic regions in terminally differentiated cells, while H1X is associated with actively transcribed regions (Mayor et al. 2015). Moreover, specific knockdowns (KDs) of H1 somatic variants lead to changes in expression of particular genes (Sancho et al. 2008). Therefore, different H1 variants are probably involved in specific functions.

However, in cell viability and development H1 variants present high redundancy, since mice knock-outs (KOs) of one or two different variants were able to develop and reproduce. Moreover, those KO showed total levels of H1s similar to the WT, due to the compensation of the remaining H1 variants (Fan et al. 2001). All things considered, the multiplicity of variants and redundancies have posed, and continue to pose, a challenge for functional studies of H1 in vertebrates.

1.2.2.1. *Drosophila melanogaster* H1 variants: a model of study

A peculiar yet convenient model organism regarding linker histone variants is *D. melanogaster*, since it has the particularity of having only two H1 variants (**Figure 1.9**): one somatic variant (dH1), encoded within the core histone cluster in ~100 copies per haploid genome (Nagel & Grossbach 2000), and one germline variant (dBigH1), present as a single copy gene (Pérez-Montero et al. 2013).

Remarkably, a single somatic variant is enough for complex developmental processes, and its study has contributed greatly to unveil functional properties of somatic linker histones (Bayona-Feliu et al. 2016) (discussed below).

On the other hand, the implication of dBigH1 in spermatogenesis and oogenesis has also been addressed (Carbonell et al. 2017; Climent-Cantó et al. 2021). Moreover, to evaluate its broader impact on chromatin, dBigH1 has been expressed ectopically in cultured somatic cells of *D. melanogaster*.

This results in the downregulation of gene expression, mediated by reduced binding of RNA polymerase II (RNApol II) and reduced levels of histone acetylation. These findings illustrate how dBigH1 alters the functional state of chromatin (Climent-Cantó et al. 2020).

Overall, *D. melanogaster* serves as an excellent model to investigate the fundamental functions of both somatic and germline H1s.

1.2.3. Functions of somatic H1s

Some H1 functions have been discussed above and are associated with specific contexts, for example, germline variants are involved in gametogenesis. However, H1 somatic variants are fundamental for the molecular biology of the cells. Despite compensation effects, the reduction of total levels to less than 50% is lethal in model animals (Lu et al. 2009; Yuhong et al. 2003). Nevertheless, the difficulty of its study caused H1 to be initially considered solely as an architectural protein. Indeed, H1 is structurally necessary for normal functioning of chromatin, as well as nucleosomes are structurally essential for chromatin formation. However, besides its fundamental role in maintaining the structural integrity of chromatin, H1 also plays a critical role in modulating several key cellular events. The mechanisms of action are either, on regulating chromatin compaction in answer to cell signals, or on blocking/recruiting specific epigenetic machineries. In general, a combination of both mechanisms is how complex processes are regulated (**Figure 1.10** Top) (Hergeth & Schneider 2015).

A good example of H1 changing the chromatin compaction and accessibility in response to cellular signals is the differential role of H1 in DNA replication and mitosis. H1 can modulate replication efficiency *in vitro*. It has the capacity of facilitating replication changes depending on its phosphorylation level, that responds to the cell cycle phase (Halmer & Gruss 1996; Talasz et al. 2009). Moreover, the implication of H1 in regulating the replication program *in vivo* has also been assessed, and H1 is linked to regulate replication timing (Andreyeva et al. 2017; Thiriet &

Hayes 2009). On the other hand, hyperphosphorylation of H1 is important for a high-degree condensation of chromatin during mitosis (Hergeth & Schneider 2015).

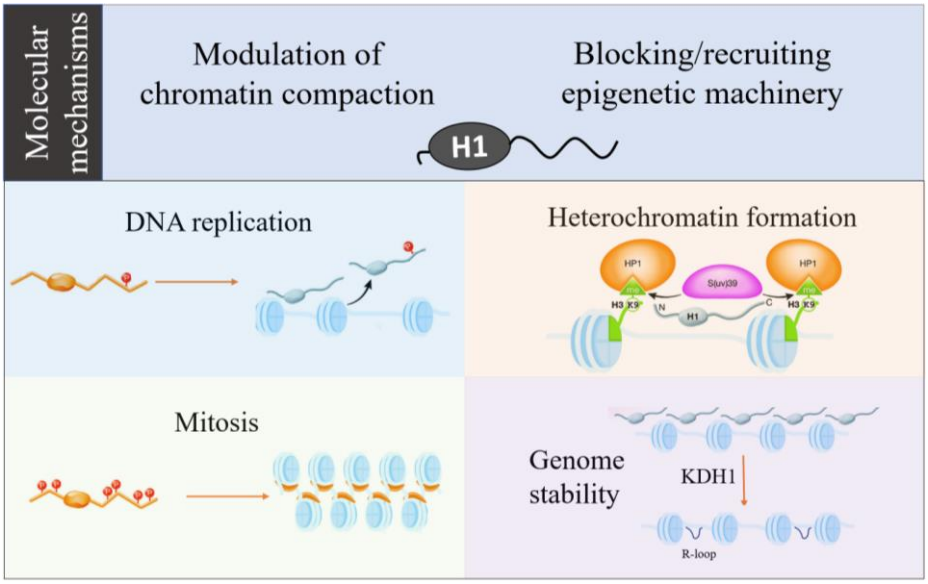


Figure 1.10. Model of H1 modes of action and overview of some functions of H1. Top, two molecular mechanisms that are involved in H1 functions. Bottom, representation of functions of H1 in DNA replication, mitosis, heterochromatin formation, and genome stability. Phosphorylation is represented as red circles. Detailed explanation in the main text. Adapted from (Hergeth & Schneider 2015).

Therefore, H1 phosphorylation can either be implicated in chromatin accessibility and chromatin compaction. This dual capacity is potentially mediated by a structural change of the protein with partial- or high-level of phosphorylation of H1 (Roque et al. 2008). Thus, in the context of replication, partial phosphorylation of H1 can provide local accessibility to chromatin regulating the access of the replication machinery to chromatin, while, in the mitotic context the hyperphosphorylated state of H1 leads to chromatin compaction (**Figure 1.10** Bottom). All in all, H1 is implicated in the modulation of those two cell-cycle dependent events by being modified by cell-cycle dependent kinases, facilitating the proper chromatin structure required at every phase (Hergeth & Schneider 2015).

In the context of the DNA damage response, H1 acts at both levels, altering accessibility to damaged regions and acting as a scaffold to the DNA repair machinery (Fyodorov et al. 2018; Thorslund et al. 2015).

1.2.3.1. Heterochromatin formation and silencing of repetitive elements

The role of H1 in heterochromatin was first addressed in *D. melanogaster* by Lu *et al.* (2009). They found that in dH1KD flies, genes translocated in the vicinity, or inside, pericentromeric heterochromatin are not silenced, thus dH1KD is by definition a suppressor of PEV. Moreover, genes that are characteristically expressed in pericentromeric heterochromatin, and that require a repressed environment to be expressed, are silenced in dH1KD flies, meaning that H1 stimulates silencing of pericentromeric heterochromatin and expression of heterochromatic genes. On the other hand, dH1KD does not appear to have a major effect in expression levels of euchromatic genes (Lu et al. 2009; Vujatovic et al. 2012). The effect seems to be somehow heterochromatin specific. However, macroscopically, chromatin structure is perturbed globally. This is observable in polytene chromosomes. Polytene chromosomes are a particular structure of flies, obtained from salivary glands, that results from repeated rounds of DNA replication without cell division that give rise to giant chromosomes. In normal polythene chromosomes, pericentromeric heterochromatin is under-replicated and fused together in the chromocenter. Remarkably, the chromocenter is the most affected structure in dH1KD flies, since it is lost as a well-defined structure. Regarding the molecular markers of heterochromatin, HP1a is not lost from chromatin. However, in normal conditions HP1a is located specifically in the chromocenter, and its disassembly in dH1KD polytene chromosomes creates two or more regions with HP1a signal. On the other hand, H3K9me2, a canonical histone mark of constitutive heterochromatin, is barely detectable in dH1KD flies. All in all, H1 is shown to be relevant for the establishment and/or maintenance of the molecular composition and structure of pericentromeric heterochromatin (Lu et al. 2009).

The implication of H1 in heterochromatin formation has been further explained in *D. melanogaster*. First, the presence of linker histone at the nucleation site has been identified as a significant factor in the recruitment of Su(var)3-9 and the facilitation of H3 methylation within the chromatin substrate (Lu et al. 2013). Moreover, aside from its implication in heterochromatin formation via Su(var)3-9, H1 is also involved in an alternative pathway for heterochromatin formation through STAT (Signal transducer and activator of transcription) (Xu et al. 2014).

Those mechanisms have not yet been directly addressed in mammals. However, H1 has been linked to several heterochromatin components in mammalian cells pointing to a conserved implication of H1 in heterochromatin (**Figure 1.10** Bottom). The best characterized interaction is between H1 and HP1α (Nielsen et al. 2001). The interaction is mediated by methylation of lysine K26 in NTD of H1.4, that provides a scaffold for HP1α. Moreover, the binding is also regulated by phosphorylation of serine S27, which disrupts the interaction and is cell-cycle modulated (Daujat et al. 2005; Hale et al. 2006).

A similar PTM exists for dH1 – K27me2 – necessary for heterochromatin genomic stability (Bernués et al. 2022), posing further evidence to the conserved role for H1 in heterochromatin in metazoa.

1.2.3.2. H1 and genomic stability

Further analysis of the consequences of dH1KD in *D. melanogaster*, showed that concomitantly with the upregulation of repetitive elements (RE), there is accumulation of DNA damage and genomic instability (Vujatovic et al. 2012).

How H1 is involved in genomic stability has been directly addressed by Bayona-Feliu *et al.* (2017). They found that upon dH1KD DNA damage appears preferentially at heterochromatic RE. Remarkably, the DNA damage is a consequence of the accumulation of R-loops in those same heterochromatic elements. This mechanism is specific of H1 and not an indirect consequence of the upregulation of RE, given that KD of HP1a –

which also leads to upregulation of RE – does not lead to accumulation of R-loops. Therefore, H1 specifically prevents accumulation of R-loops and maintains genome stability in heterochromatin (**Figure 1.10** Bottom) (Bayona-Feliu et al. 2017). However, how H1 and R-loops are molecularly related has not been addressed and poses the question of which is the relationship of H1 with the RNA associated to chromatin.

1.3. CHROMOSOMAL RNAS (CRNAS)

RNA is usually depicted soluble in the nucleoplasm or the cytoplasm, where it is most abundant. However, there is also a fraction of RNA that is associated to chromatin (Holmes et al. 1972; Huang & Bonner 1965).

1.3.1. Architectural roles of cRNAs

RNAs associated with chromatin are referred to as chromatin-associated RNAs or chromosomal RNAs (cRNAs). They are not only a byproduct of transcription, but cRNAs are also structural components of chromatin. RNA that is stably bound to chromatin account for 2%-5% of chromatin nucleic acids, and their removal by RNAseA treatments alter chromatin structure and increases its sensitivity to MNase (Rodriguez-Campos & Azorin 2007).

More molecular evidence for the architectural role of cRNAs involves the canonical examples of specific lncRNA that associate to chromatin with an architectural role, like mammalian *Xsist* and HOTAIR: *Xsist* is a lncRNA transcribed from the X chromosome that binds to it in *cis* and recruits epigenetic silencing machinery for female dosage compensation (Loda & Heard 2019). HOTAIR is an antisense lncRNA transcribed from a specific locus of the Hox genes – key genes in embryonic development. It binds in *trans* to another Hox genes locus and it is essential for its Polycomb-mediated repression (Rinn et al. 2007). Other examples are reviewed in (Thakur & Henikoff 2020).

Furthermore, there is also evidence of the architectural implication of cRNA in other specialized structures in metazoa. For instance, centromeric satellites are transcribed and retained in the centromere as cRNAs, and their removal leads to disassembly of centromere proteins, impeding proper chromosome segregation (McNulty et al. 2017; Shatskikh et al. 2020). Moreover, as mentioned above, cRNAs from pericentromeric heterochromatin are necessary for SUV39H1 binding, suggesting that cRNAs are necessary for the binding of histone modifying enzymes and

H3K9me3 deposition on heterochromatin (Johnson et al. 2017; Velazquez Camacho et al. 2017).

In addition, soluble RNA is known to be a major component and modulator of phase separation events in membrane-less RNA-filled nuclear bodies (Chujo & Hirose 2017). Overall, propensity of RNA to form many versatile secondary structures and establish multivalent weak interactions, on top of its negative charge, make RNA a key candidate to modulate phase separation events. In this regard, cRNAs have also been appointed to potentially have a role in LLPS heterochromatin condensation (Thakur & Henikoff 2020).

1.3.2. Modes of RNA-chromatin interaction and other associated roles

cRNAs can interact with chromatin mainly in two ways: in *cis* or in *trans*. In *cis*-interactions, nascent RNAs remain at their synthesis site, while, in *trans*-interactions, RNAs are released from where they were transcribed to interact with other regions, away from where they were synthesized. Those modes are not mutually exclusive, indeed in many situations they act synergistically to perform specific regulatory functions.

Two *trans*-interaction modes have been described up to date. First, through the formation of RNA:DNA triplex, formed via Hoogsteen base-pair interactions of a single RNA strand with the major groove of dsDNA (**Figure 1.11a**). An example for this binding mechanism is the promoter-associated RNA – complementary to rDNA promoter – that hybridizes to the promoter region forming a triplex, and acts as scaffold for a DNA methyltransferase, which methylates CpG dinucleotide sequences of rRNA genes (Schmitz et al. 2010). In general terms, the formation of RNA:DNA triplex has been appointed to represent a broad mechanism for lncRNA to target DNA sequences (Li et al. 2016). The second mode of *trans*-interaction is through protein-mediated interaction with the DNA, in which the RNA acts as scaffold to bind different elements of which at least one is an RNA binding protein (RBP), and another a DNA-binding factor (**Figure 1.11b**) (Li & Fu 2019).

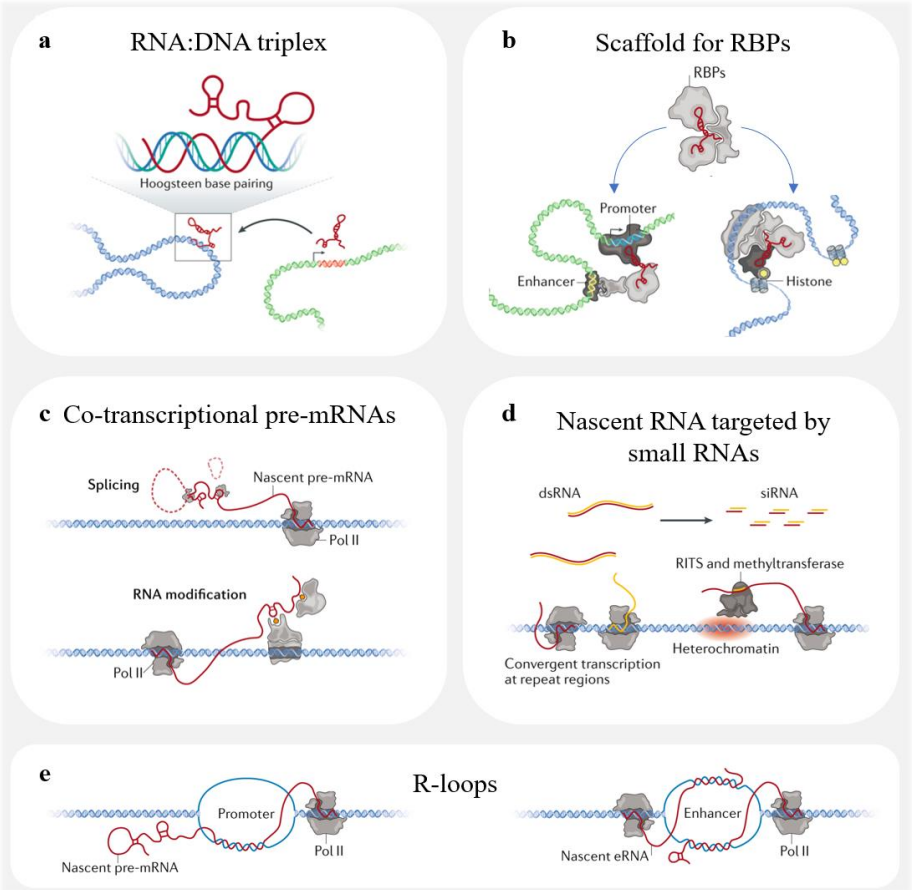


Figure 1.11. Modes of RNA-chromatin interaction. Distinct forms of cRNA found in metazoa retained in chromatin: (**a**) RNA:DNA triplex, RNA is transcribed and hybridize in *trans* a different genomic region through Hoogsteen base-pairing with the major grove of dsDNA; (**b**) Scaffold for RBPs, bringing close different elements, one of which is a DNA-binding factor; (**c**) Co-transcriptional pre-mRNAs are retained by RNApol II while being processed; (**d**) Nascent RNA targeted by small RNAs through the RNAi machinery; (**e**) R-loops, co-transcriptional formation of DNA:RNA hybrids involved in different cellular processes. Detailed explanation in the main text. Adapted from (Li & Fu 2019).

On the other hand, the largest *cis*-interacting RNAs are protein-coding pre-mRNAs, which are being processed co-transcriptionally (**Figure 1.11c**) (Bentley 2014). Tethering through RNApol II is a key mechanism of RNA retention also mediating CheRNAs – a special class of lncRNA – enrichment in chromatin (Werner & Ruthenburg 2015). Moreover, nascent RNA can also be retained by *trans*-acting amplified RNA, via the RNAi

machinery (**Figure 1.11d**). This mode is the one involved in the silencing of RE and initiation of heterochromatin establishment mentioned above (Allshire & Madhani 2018; Gu & Elgin 2013). It is common for RE, since distant RE have perfect complementarity. This mechanism is an example of combined *cis*- and *trans*-interacting cRNAs (Li & Fu 2019). Finally, one of the most characterized *cis*-interacting mechanisms for cRNAs is the formation of R-loops, which involves the co-transcriptional reannealing of nascent RNA to the template DNA, forming a double-stranded RNA:DNA hybrid and displacing the non-template single-stranded DNA (ssDNA) (**Figure 1.11e**). The most accepted mechanism for R-loop formation is the "thread-back" model, where the nascent RNA and the template DNA exit through independent channels from the RNApol II, to be reannealed as soon as they exit (Roy et al. 2008). R-loops have a dual nature, since they are needed for many cellular processes but are also dangerous for the cell.

It must be taken into consideration that despite this description of modes of cRNA association to chromatin has been categorized as *cis*- or *trans*-interacting, some of the elements here classified as *cis*-interacting can occasionally act as *trans*-interacting and the other way around.

1.3.3. R-loops

R-loops are involved in many different physiological processes, and they are clearly biologically relevant. However, when they are generated unscheduled or the control mechanisms to prevent and resolve R-loops fail, they accumulate aberrantly and interfere with DNA replication, transcription, and DNA repair, thus leading to DNA damage and compromising genome integrity (García-Muse & Aguilera 2019).

1.3.3.1. Physiological R-loops

R-loops have been characterized as required intermediates for different processes, such as, in the lagging-strand during DNA replication or within the transcription bubble (Aguilera & García-Muse 2012). They are also a prerequisite structure for mitochondrial DNA (mtDNA) replication, in a

mechanism conserved in the mitochondria from yeast to humans (Pohjoismäki et al. 2010; Xu & Clayton 1996).
Furthermore, it has been extensively characterized the need for R-loops in the Ig class-switch recombination. In this process, the R-loop formation displaces the ssDNA that is then targeted by the B cell-specific cytidine deaminase AID, which is the first fundamental step for the Ig class-switch recombination (Aguilera & García-Muse 2012; Roy et al. 2008). R-loops are also intermediates for the CRISP-Cas9 recognition system, which acts as an immunity mechanism for prokaryotes and has become extensively used as a genome-editing tool. There, the guide RNA forms an R-loop to identify the target for cleavage (Jinek et al. 2012).

On the other hand, there are also evidences for the implication of R-loops in gene transcription regulation. First, in *Arabidopsis thaliana*, a set of antisense transcripts called COOLAIR, which modulate the FLC gene – controller of plant flowering – are regulated by the formation of an R-loop in the promoter. That R-loop is further stabilized by a homeodomain protein binding to the ssDNA, thus, controlling the transcriptional silencing of COOLAIR (Sun et al. 2013). In addition, R-loops have been suggested to play a role in regulating the process of transcriptional activation by forming at CpG islands. By physically blocking the action of methyltransferases, R-loops prevent methylation from occurring at CpG islands, thereby helping to maintain the active transcription of genes (Ginno et al. 2012). Furthermore, R-loops are appointed to be critical for transcriptional termination, specially over G-rich pause sites downstream of poly(A)+ genes, where the resolution of the R-loop is coupled with termination and facilitating the RNApol II release (Skourti-Stathaki et al. 2011). Similarly, R-loops have also been correlated with transcriptional pausing at TSS (Chen et al. 2017).

All in all, R-loops are abundant physiological forms that occupy up to 5% of the genome (Sanz et al. 2016). They answer to and regulate different needs of the cell.

1.3.3.2. R-loops resolution and prevention factors

It must be taken into consideration that RNA:DNA hybrids are more stable than dsDNA (Thomas et al. 1976). Therefore, when formed uncontrolled and unscheduled they require specific mechanisms to be resolved. The most studied and characterized factor with the capability to resolve R-loops is RNaseH, a ribonuclease conserved from bacteria to humans that specifically degrades the RNA strand of RNA:DNA hybrids (**Figure 1.12A** left). RNaseH main function is to remove RNA primers of Okazaki fragments during replication (Cerritelli & Crouch 2009). Moreover, it has been extensively used in experiments as a rescue to the phenotype of R-loop accumulation (Aguilera & García-Muse 2012; Drolet et al. 1995). Still, degrading the RNA is unlikely to be the only mechanism to resolve R-loops. In this regard, it has been hypothesized that RNA-dependent ATPases (e.g., SETX, FANCM, DDX5…) could also be involved in R-loop resolution (**Figure 1.12A** left). However, it has not yet been addressed *in vivo* (García-Muse & Aguilera 2019).

The difficulty of R-loop resolution shows the relevance of what is the most extended mechanism to avoid their pathological formation: their prevention. Several factors have been identified as preventing R-loop formation, most of them have been identified because of the R-loops accumulation phenotype found when depleted (Aguilera & García-Muse 2012). On one hand, factors influencing DNA topology are relevant for the prevention of R-loops. When the transcription machinery operates, it creates negative supercoiling behind, which promotes the formation of R-loops by causing the dsDNA opening. Thus, the topoisomerase – helicase capable of resolving that torsion – is a relevant factor in R-loop prevention (**Figure 1.12A** right) (Drolet et al. 1995; El Hage et al. 2010; Liu & Wang 1987). Similarly, general chromatin accessibility is also associated to R-loop formation (Sanz et al. 2016), accordingly elements associated with chromatin compaction, such as histone deacetylases, have also been appointed to prevent R-loop accumulation (García-Muse & Aguilera 2019).

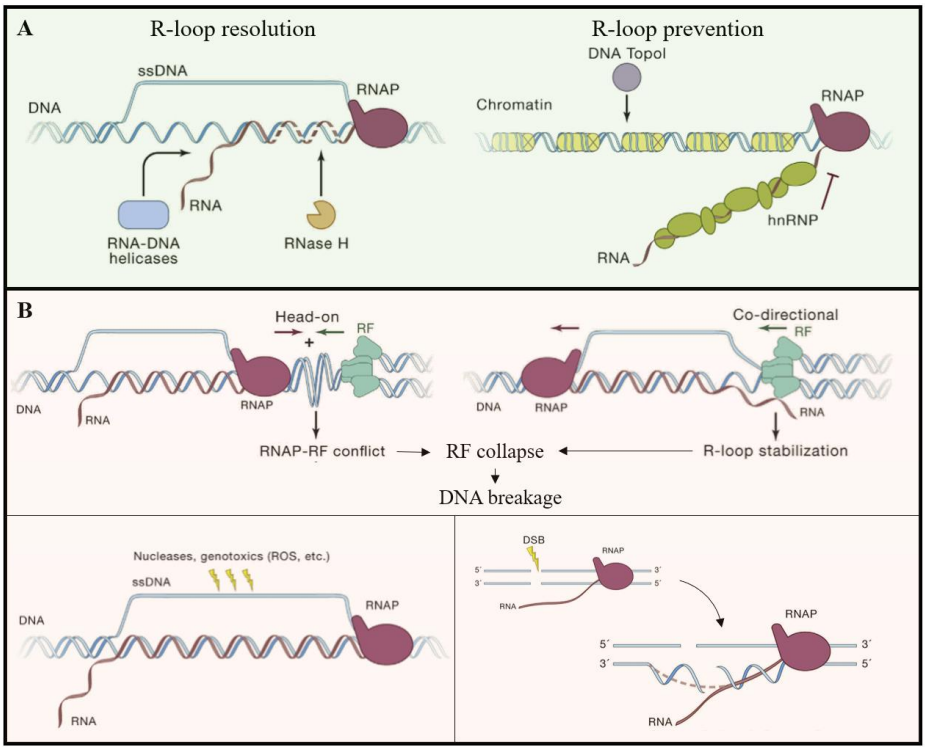


Figure 1.12. Mechanisms to control pathological R-loop formation and DNA damage associated with their accumulation. A) R-loop accumulation is avoided either by resolving the already formed R-loops (**left**), by DNA-RNA helicases or RNase H; or by preventing R-loop formation (**right**), by specific hnRNPs binding to RNA or by DNA Topoisomerase maintenance of adequate DNA topology. **B)** R-loops are a source of DNA damage. (**top**) By transcription-replication conflicts, (**bottom, left**) ssDNA is target of enzymes and mutagens, (**bottom, right**) DNA break can also favor R-loops formation. Adapted from (García-Muse & Aguilera 2019).

On the other hand, another key factor in prevention is the packaging of the nascent RNA molecule with proteins involved in processing and exporting mRNA. This role is not a general role of all RNA binding proteins (RBP), but a function of a subset of factors involved in the co-transcriptional assembly of pre-mRNAs, such as hnRNPs (**Figure 1.12A** right) (García-Muse & Aguilera 2019).

hnRNPs (heterogeneous nuclear RNPs (ribonucleoproteins)) are modular proteins, always containing at least one RNA recognition motif, that bind to the pre-mRNA produced by RNApol II, and that do not stably associate

with other RNPs (Han et al. 2010). Among them, the hnRNPA/B family is the most abundant group, which is also well evolutionarily conserved. They play important regulatory functions in RNA splicing, export, and localization (Dreyfuss et al. 2002). *D. melanogaster* encodes for five hnRNPA/B proteins, yet there are three mainly expressed (hrp36, hrp40, and hrp48) (Matunis et al. 1992). From those, there is evidence that hrp36 is implicated in telomere maintenance (Singh & Lakhotia 2016). Moreover, both hrp36 and hrp48 localize in chromatin when analyzed in polythene chromosomes (Matunis et al. 1993). However, their relationship with R-loops has not been addressed.

On a final note, it has been recently proposed that N6-methyladenosine (m6A) – the most abundant reversible RNA modification that plays critical roles in various post-transcriptional processes – is involved in regulating co-transcriptional R-loop formation and/or resolution (Abakir et al. 2020; Yang et al. 2019).

1.3.3.3. Pathological R-loops and genomic instability

Despite the control mechanisms for R-loop prevention and resolution, certain conditions can inevitably lead to pathological R-loop accumulation. Their accumulation leads to DNA damage and genomic instability. The most relevant mechanism by which R-loops cause genome instability is by stalling the replication fork (RF) progression and consecutively causing DNA breakage (**Figure 1.12B** top). This is demonstrated, among other things, by the fact that genomic instability is strongly suppressed if active replication is impaired in an R-loop prone genetic background (Gan et al. 2011). The conflicts with the RF can occur in two modes: due to head-on collision of the RF with the RNA polymerase, or due to stabilization of an R-loop that was transcribed in co-directional orientation respect to the RF (**Figure 1.12B** top) (García-Muse & Aguilera 2019).

An additional source of DNA damage from pathological R-loops, is the vulnerability of the ssDNA displaced strand (**Figure 1.12** bottom, left). That can be targeted by enzymes like AID (mentioned above), which can lead to

either mutagenesis or even DSB (Sollier & Cimprich 2015). Moreover, ssDNA is in general more susceptible to chemical reactions, thus, genotoxics could also easily lead to DNA damage (Aguilera 2002).

Overall, there is a link between DNA breaks and R-loops. As stated above, R-loops can cause DNA breaks in several ways. However, DNA breaks can also favor R-loops formation. In a DNA break, a free DNA end relieves the torsional stress of the dsDNA molecule, thus facilitating the entanglement of the invading RNA with its template (**Figure 1.12B** bottom, right) (Aguilera & Gómez-González 2017), in a similar way, although more extreme, than how negative supercoiling favors R-loop formation. Nevertheless, the impact that an R-loop in a DNA break has on its repair is unclear. It had been proposed that R-loops were necessary intermediates in DSB repair, yet recent evidence has questioned that possibility. Whether R-loops are consequences of DNA breaks, functional intermediates in repair, or both is still unclear (García-Muse & Aguilera 2019).

1.4. PHASE SEPARATION IN BIOLOGY

In eukaryotes, the compartmentalization by lipid-membrane organelles of specific functions, such as the production of ATP in the mitochondria or protein degradation in lysosome, poses an obvious advantage. However, such compartmentalization is not only mediated by membranes, since there are canonical membrane-less organelles, such as the nucleolus, that are compartmentalized by phase separation – also referred to as biomolecular condensates (Hyman et al. 2014).

Moreover, as stated above, an increasing amount of research is starting to suggest that phase separation could also be involved as an organizational force in terms of general chromatin organization, such as in the establishment and maintenance of heterochromatin.

1.4.1. Principles of phase separation

The physical explanation of phase separation in biology lies in the interplay between macromolecules and the solvent. When macromolecules are in solution, they interact with the solvent, and the system tends to achieve the maximum entropy by forming a uniform mixture of the components. However, if the macromolecules can interact with each other in a more energetically favorable way than with the solvent, their demixing from the solvent can reduce the free energy of the system and become the lowest-energy state, thereby becoming the most stable (**Figure 1.13**). This is the event of phase separation, and it is in direct proportion to the macromolecule concentration in the solution. Thus, as the concentration increases, the entropic penalty decreases. Therefore, there is a critical concentration below which the mixed state is more stable and above which phase separation occurs, resulting in the formation of a dilute solution phase and a condensed macromolecule phase. If the interactions are of high affinity, the condensates will have the material properties of a solid, which is referred to as liquid–solid phase separation (LSPS). Conversely, if the interactions are weak and of low affinity, the condensates will behave as a liquid, leading to

liquid-liquid phase separation (LLPS) (Hansen et al. 2021; Hyman et al. 2014; Zhang et al. 2020).

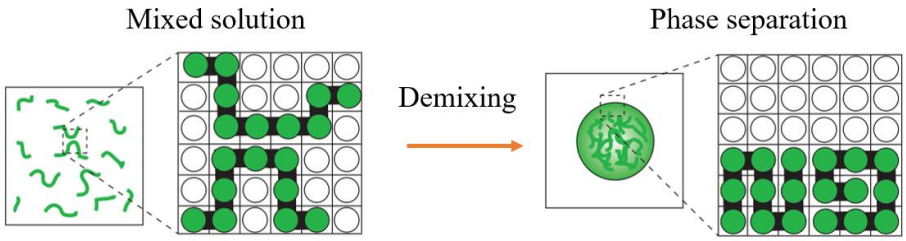


Figure 1.13. Polymer phase separation from a mixed solution state. In the left the polymer is in solution, thus there are many interactions with the polymer and the solvent molecules and the distribution in solution is homogeneous. Past the critical point of demixing, the polymer maximizes the interactions with the polymer and minimizes interactions with the solvent, giving rise to a solution with two phases in the form of a spherical polymer concentrated phase and a diluted phase. Adapted from (Brangwynne et al. 2015).

LLPS events are characterized by their liquid properties: LLPS objects, referred to as droplets, have a spherical shape driven by surface tension; They exchange material with the diluted phase and there is internal mobility of the molecules, while maintaining the spherical shape. Two droplets can fuse after making contact and the fused droplet will maintain a spherical shape. On the other hand, LSPS share some properties with the LLPS, although their key difference is the lack of internal mobility in the LSPS condensates (Hyman et al. 2014).

Nowadays there are several examples of LLPS compartments in the cell, for instance in the cytosol we can find stress granules, that assemble in response to certain stress conditions (Protter & Parker 2016), and in the nucleus the canonical membrane-less compartment is the nucleolus, which is the site of ribosome production and contains hundreds of RNAs and proteins (Lafontaine et al. 2021). Moreover, it is emerging the idea that in the nucleus, aside from the soluble fraction, there are also phase transition events involving the chromatin itself.

1.4.2. Phase separation in chromatin

The first report of phase transition involving chromatin was described as aggregation. It was observed that increasing the salt concentration in solution led to the formation of insoluble precipitates of chromatin in a reversible fashion. Thus, it was hypothesized that chromatin solubility changed as a function of salt concentration (Jensen & Chalkley 1968). Since then, this property has been extensively used for chromatin fractionation, and it has been described that transcriptionally active chromatin has different physical properties, because after DNA digestion and promoting aggregation, transcriptionally active chromatin still remains soluble (Gottesfeld et al. 1974). Later, it has been further characterized that the soluble fraction of chromatin, and other salt dependent fractions, are associated with epigenetic marks, making of salt fractionation a good method to separate certain chromatin states (Henikoff et al. 2009). However, for a long time the physical nature of chromatin aggregates was not really addressed, instead, as discussed above, much attention was given to the theoretical structure of the 30-nm fiber in low salt conditions. The condensates formed when salt was increased to physiological concentrations were ignored and described as insoluble precipitants instead of a phase separation event of LSPS (Hansen et al. 2021). Nevertheless, a turning point in the study of chromatin physicochemical properties was the description of *in vitro* reconstituted chromatin. It was obtained by the construction of a DNA molecule consisting of an array of tandem repeats of nucleosome positioning sequences that could be assembled with purified core histones into reconstituted chromatin (Simpson et al. 1985). With this relevant tool, there were many studies addressing the aggregated form of chromatin, which was then referred to as self-association or oligomerization to better reflect that it is a reversible event driven by chromatin fiber self-interaction (Hansen et al. 2021). Remarkably, reconstituted chromatin forming condensates has transcriptional activity *in vitro* proving that promoters in chromatin condensates are accessible to RNA polymerases (Zhou et al. 2007). Further studies have shown that reconstituted chromatin

in vitro can undergo LSPS and LLPS, depending on the solvent conditions (Gibson et al. 2019; Strickfaden et al. 2020). A systematic analysis of the implication of the length of the DNA molecule in the *in vitro* phase separation properties has unveiled that the material properties of the condensed phase driven by chromatin are more liquid-like when DNA molecules are shorter than 1 kb, while above this length the properties progressively change towards a solid-like condensate (Muzzopappa et al. 2021).

Overall, there is extensive evidence demonstrating that chromatin aggregation *in vitro* is a reversible phase separation process, and that those condensates formed *in vitro* with Mg^{2+} reflect properties of condensed chromatin *in vivo* (Hansen et al. 2021). However, how accurate is this representing the properties of chromatin *in vivo* is still unclear. It is possible that chromatin physical properties are scale dependent, since in the mesoscale we find chromatin displaying a solid-like behavior (Strickfaden et al. 2020), which at the chromosome level agrees with the fact that chromosomes occupy a specific region and are not mixed in the nucleus, a feature known as chromosome territories (Cremer & Cremer 2010). Nevertheless, when analyzing chromatin dynamics at the nanoscale, the mobility of nucleosomes is clear and points towards a more liquid-like behavior (Ashwin et al. 2020; Hihara et al. 2012; Nozaki et al. 2017). All in all, the current model – still under construction – is of a viscoelastic chromatin that has solid properties at the mesoscale and liquid properties at the nanoscale. In this model both natures of the chromatin coexist, but where do the liquid-like and the solid-like start is yet to be characterized (Hansen et al. 2021).

1.4.2.2. Linker histone phase separation

The relevance of core histones in chromatin phase separation was clear even when it was described as aggregation. It was described that reconstituted chromatin with core histones lacking the unstructured N-terminal tail did not undergo aggregation, showing the necessity of the core histone NTD for

chromatin phase separation (Schwarz et al. 1996). Later it has been confirmed that without the histone tails chromatin does not phase separate (Gibson et al. 2019).

However, another clear nucleosomal element modulating chromatin phase separation is the linker histone. Thus, when H1 is added to reconstituted chromatin, phase separation occurs at lower salt conditions (**Figure 1.14A**).

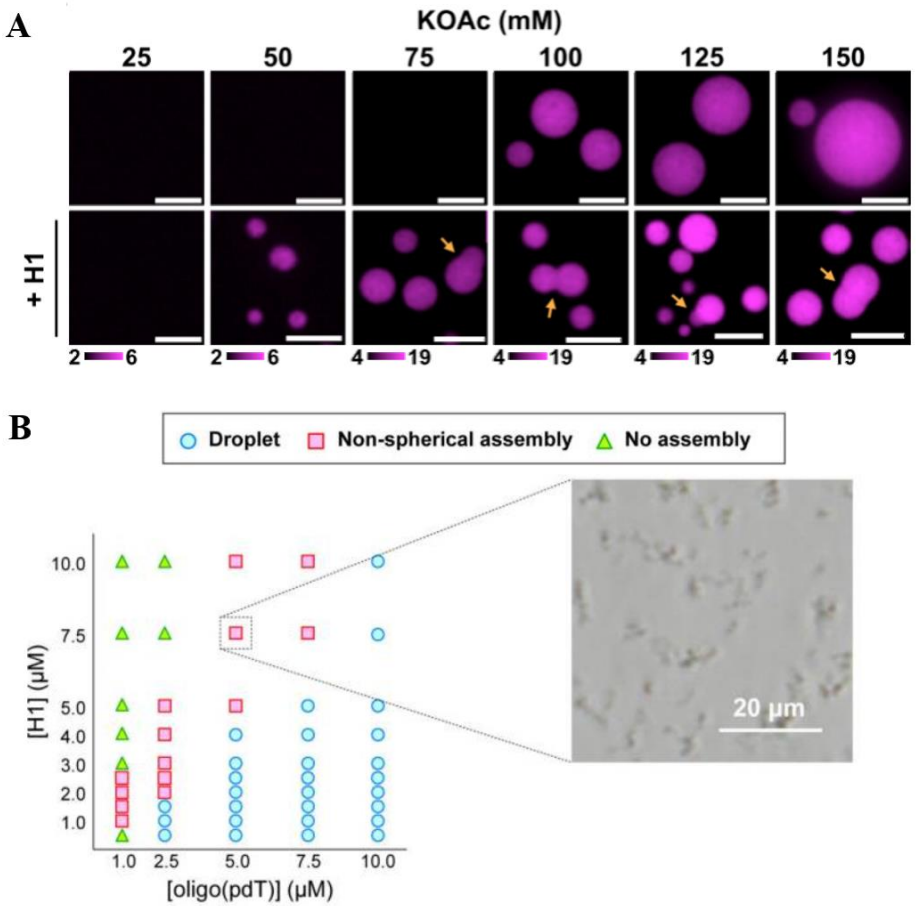


Figure 1.14. Linker histone phase separation. A) Fluorescence microscopy of labeled reconstituted chromatin following titration of potassium acetate with (bottom) or without (top) bovine linker histones. Adapted from (Gibson et al. 2019) **B)** Phase diagrams of aqueous mixtures of oligomers and H1 (right) square highlighting phase-contrast microscopy image of typical non-spherical assemblies. Adapted from (Mimura et al. 2021).

Moreover, the droplets dynamics decrease in presence of H1, becoming more solid-like, and the concentration of nucleosomes per droplet increases (Gibson et al. 2019; Hansen et al. 2021). This property has been argued to have a role in heterochromatin formation, and, in general, to be associated to linker histone mediated compaction (Shakya et al. 2020).

Remarkably, the linker histone can form LLPS *in vitro* in presence of dsDNA, ssDNA and RNA (Leicher et al. 2022; Mimura et al. 2021; Shakya et al. 2020). As in the modulation of reconstituted chromatin phase separation, H1 phase separation is driven by its long and unstructured CTD, and PTMs, such as phosphorylation, can interfere with the LLPS (Gibson et al. 2019; Turner et al. 2018). Moreover, the effect of the ratio between nucleic acids and H1 is critical. Mimura *et al*. (2021), addressed this by generating phase diagrams of aqueous mixtures of oligomers and H1 (**Figure 1.14B**). In these, three outputs where clearly differentiated: First, when [oligomer]/[H1] < 1, there is no phase separation. It is in agreement with the fact that H1 cannot phase separate on its own (Shakya et al. 2020); Second, [oligomer]/[H1] ~1 gives rise to solid-like non-spherical assemblies; Finally, [oligomer]/[H1] > 1 promotes LLPS.

It appears that when the ratio of relative charges is compensated H1 can give rise to LLPS droplets with different DNA lengths (Shakya et al. 2020). However, when added to nucleosomal arrays the mixtures shift towards the formation of solid-like irregularly shaped condensates (Hansen et al. 2021; Leicher et al. 2022; Shakya et al. 2020).

All in all, those studies will facilitate our understanding of the biophysical properties behind chromatin organization. However, systematic comparisons of the implication of H1 PTMs or the contribution of H1 variants to the phase separation driven by H1 have not been addressed yet.

2. OBJECTIVES

In this project, we aimed to analyze the structural and functional properties of histones H1 in *D. melanogaster*, using two different approaches:

First, to study the role of somatic histone dH1 in the control of genomic stability and chromatin structure. Making special emphasis in studying the RNA associated to chromatin in order to understand the molecular mechanism that links the depletion of dH1 with the accumulation of R-loops.

Second, to characterize and compare the biochemical properties of embryonic dBigH1 and somatic dH1 linker histones to contribute to the understanding of their functional differences.

Ultimately our main objectives were the following:

1. The study of the molecular mechanism of the linker histone dH1 in the maintenance of genomic stability
    a. Characterization of chromosomal RNAs (cRNAs) and their packaging in chromatin
    b. Involvement of dH1 in the regulation of cRNAs and the relationship between R-loops accumulation and cRNAs
    c. Contribution of dH1 to chromatin structure

2. The characterization of the physico-chemical properties of the somatic dH1 and the embryonic dBigH1 linker histones.
    a. Study of the impact of dBigH1 to the Nucleosome Repeat Length in cells
    b. Description and characterization of phase separation properties for dH1 and dBigH1

3. RESULTS

3.1. STUDY OF THE MOLECULAR MECHANISM OF THE LINKER HISTONE DH1 IN THE MAINTENANCE OF GENOMIC STABILITY

3.1.1. CHARACTERIZATION OF CRNAS AND THEIR PACKAGING IN CHROMATIN

3.1.1.1. cRNAs are enriched in heterochromatin

To characterize the catalog of cRNAs in *D. melanogaster* S2 cells we adapted the RNA extraction protocol from fractionated chromatin (Werner & Ruthenburg 2015). By d*e novo* transcriptome assembly, we detected ~14,000 cRNA species that covered ~28% of the Drosophila genome, distributing across all chromosomes and in both strands (**Figure 3.1.1A**).

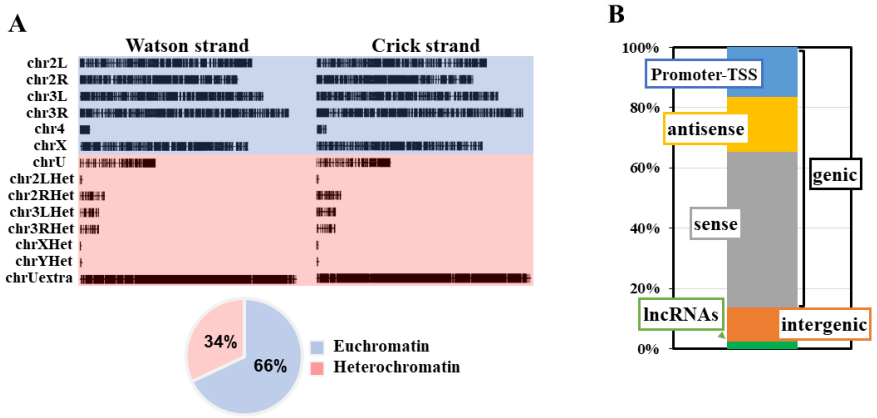
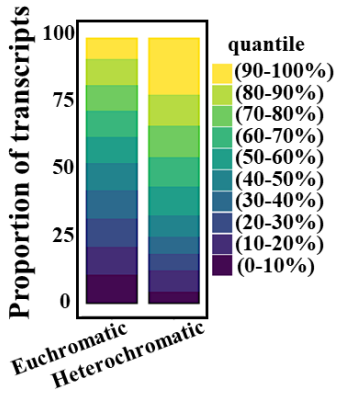


Figure 3.1.1. cRNAs constitute an heterogeneous group of RNA species covering the entire genome. (**A**) Chromosomal distribution of cRNAs along *D. melanogaster* chromosomes separated by the Watson and Crick strands on the dm3 genome assembly: chr2L and chr2R, and chr3L and chr3R correspond to chromosome 2 and 3 left and right arms respect to the position of the centromere, respectively. chr4 and chrX are oriented with the centromere to the right. chr2LHet, chr2RHet, chr3LHet, chr3RHet, chrXHet and chrYHet correspond to partially assembled pericentromeric heterochromatin regions of the indicated chromosomes. chrU and chrUextra correspond to unassembled highly repetitive heterochromatic scaffolds. Euchromatic and heterochromatic regions are highlighted in blue and red, respectively. On the bottom, pie chart of the proportion of cRNAs mapping to euchromatic and heterochromatic regions. (**B**) Proportion of cRNAs matching lncRNAs, intergenic, genic sense and antisense, and regulatory (UTRs, promoter, TSS) regions.

Moreover, we classified the detected cRNAs and found that the most represented category corresponded to genic sense cRNAs, category associated to potential transcribing pre-mRNAs. However, other categories such as genic antisense, and non-genic cRNAs were also detected (**Figure 3.1.1B**). Remarkably, ~34% of cRNA transcripts mapped at pericentromeric heterochromatin (assembled in 2LHet, 2RHet, 3LHet, 3RHet, YHet and XHet) and to highly repetitive elements (assembled in the artificial scaffolds U and Uextra).

To address the extent of the accumulation of cRNAs to heterochromatin, we analyzed the distribution by expression quantiles of euchromatic and heterochromatic cRNAs separately. While euchromatic cRNAs presented a rather uniform distribution, heterochromatic cRNAs were enriched in the highest decile of expression (**Figure 3.1.2A**).

A



B

REPEAT MASKER	% of transcripts
Total repeated elements	72.57%
Retroelements	42.75%
LINES	17.66%
LTR-elements	25.09%
DNA transposons	2.35%
Satellites	19.74%
Simple repeats	1.34%
Other	6.39%

Figure 3.1.2. cRNAs are enriched in heterochromatin and RE. (**A**) Distribution by quantile of expression for euchromatic and heterochromatic cRNAs. (**B**) Proportion of heterochromatic cRNAs mapping different classes of repetitive elements analyzed with the RepeatMasker.

Deeper analysis of the heterochromatic cRNAs showed a very high content in RE (as anticipated by their coverage of the U and Uextra scaffolds (**Figure 3.1.1A**)). Heterochromatic cRNAs where composed, from highest to lowest, of LTR-elements, Satellites, LINES, DNA transposons and Simple repeats (**Figure 3.1.2B**).

All in all, the analysis of the cRNA composition of S2 cells provided a portrayal of a cRNA-enriched heterochromatin in WT conditions.

3.1.1.2. hrp36 and hrp48 are on chromatin through tethered RNA

It is well known that RNAs need to be packaged with RBPs in the cell. Therefore, we hypothesized that those heterochromatic cRNAs needed to be assembled into RNPs as well. To address this possibility, we studied the detection on chromatin of two of the most abundant hnRNPs of *D. melanogaster*, hrp36 and hrp48, which had been described to localize in chromatin when analyzed in polytene chromosomes (Matunis et al. 1993).

First, we confirmed, as it had been described on flies, that we could detect them on S2 cells chromatin (**Figure 3.1.3**). Moreover, since hrp36 and hrp48 are RBPs, we prepared RNaseA treated chromatin samples to see whether hrp36 and hrp48 binding to chromatin was mediated by RNA bound to chromatin. Indeed, the removal of RNA drastically reduced the signal of hrp36 and hrp48 on chromatin by ~4-fold (**Figure 3.1.3**). Thus, we found that the wide majority of hrp36 and hrp48 detected on chromatin are bound through tethered RNA.

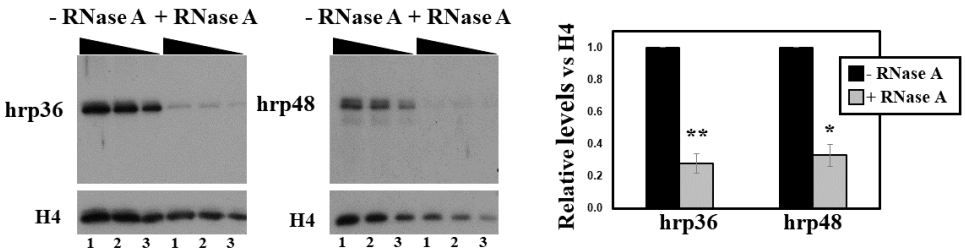


Figure 3.1.3. hrp36 and hrp48 are on chromatin through tethered RNA. Western blot (WB) analysis with α-hrp36 (left) and α-hrp48 (right) antibodies of increasing amounts (lanes 1-3) of crosslinked chromatin prepared from S2 cells treated or not with RNase A after crosslinking. α-H4 antibodies were used for loading control. Quantification of the results is shown on the right. Results are the average of 3 independent experiments. Error bars are standard deviation. (p-values: **< 0.01, *< 0.05; two-tailed paired Student's test).

3.1.1.3. hrp36 and hrp48 are packaging the heterochromatic cRNAs

To further study the location of hrp36 and hrp48 on chromatin, we performed ChIP-seq experiments in S2 cells. We detected ~1000 peaks for

both proteins that widely overlap (**Figure 3.1.4B**). Moreover, their distribution was clearly shifted towards heterochromatic regions, containing a 66% and a 77% of the peaks for hrp36 and hrp48, respectively (**Figure 3.1.4A**).

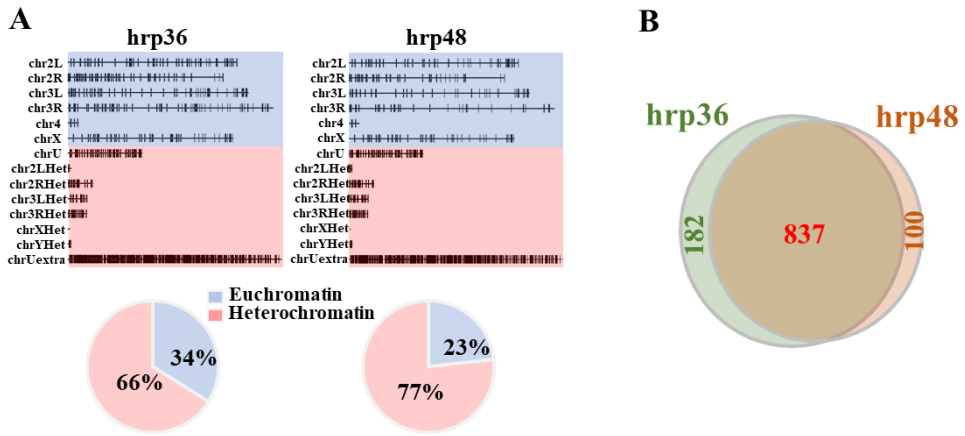


Figure 3.1.4. Peak distribution of hrp36 and hrp48 in *D. melanogaster* genome. (**A**) Top, chromosomal distribution of hrp36 (left) and hrp48 (right) ChIP-seq peaks. Chromosome definition and orientation are as in Figure 3.1.1A. Euchromatic and heterochromatic regions are highlighted in blue and red, respectively. Bottom, pie charts showing the proportion of hrp36 (left) and hrp48 (right) enriched regions mapping to euchromatic and heterochromatic regions. (**B**) Venn diagram showing the overlap between hrp36 and hrp48 peaks.

In addition, classification of the common peaks for hrp36 and hrp48 according to the nine chromatin states model (Kharchenko et al. 2011), showed that the most represented state was constitutive heterochromatin (state 7), with ~37% of the enriched regions. This enrichment was statistically significant (permutation test, p-value: 0.022) (**Figure 3.1.5A**). The enrichment was also tested individually for each protein (**Figure 3.1.5B**). Moreover, the following highest represented state was the one corresponding to silent/intergenic chromatin (state 9), with a ~27% of the enriched regions (**Figure 3.1.5A**).

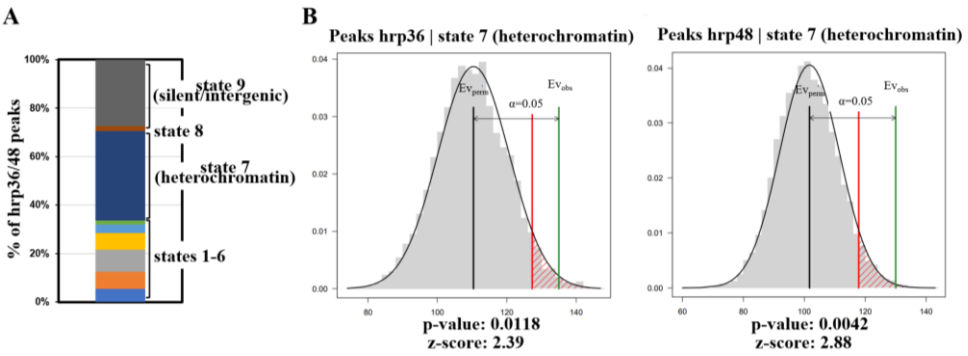


Figure 3.1.5. hrp36 and hrp48 are enriched in heterochromatin. (**A**) Proportion of common hrp36 and hrp48 peaks assigned to each of the nine chromatin epigenetic states (Kharchenko et al. 2011). (**B**) Permutation tests showing statistical significance of the localization of hrp36 (left) and hrp48 (right) on constitutive heterochromatin. The expected value is marked with a black bar. The experimentally observed value is marked with a green bar. The value corresponding to a statistical significance of α=0.05 is marked with a red bar. p-value and z-score is indicated, 5000 permutations.

Thus, our results showed that hrp36 and hrp48 are specifically enriched in heterochromatin.

To determine whether hrp36 and hrp48 enriched regions corresponded to the regions where we detected cRNAs, we analyzed the relative localization on chromatin of the hrp36 and hrp48 peaks compared to the cRNA signal. We found that 88% and 90% of the peaks for hrp36 and hrp48, respectively, overlapped with cRNA signal, and the overlap was statistically significant (permutation test, p-value< 0.001) (**Figure 3.1.6** Left). On the other hand, since there are ~10 times more cRNAs than hrp36 and hrp48 peaks, many cRNAs did not overlap with hrp36 and hrp48 peaks. Nevertheless, when we compared the frequency of cRNAs non-overlapping with the frequency of cRNAs overlapping with hrp36 and hrp48 peaks, we found that the abundance of the overlapping cRNAs was significantly higher (p-value< 0.001, permutation test) (**Figure 3.1.6** Right).

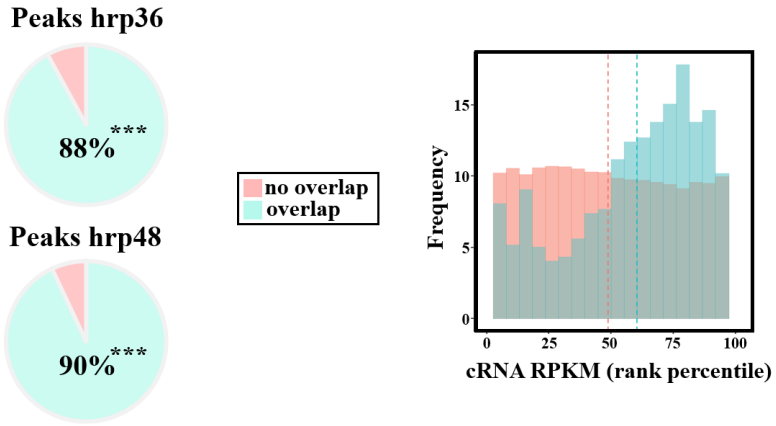


Figure 3.1.6. hrp36 and hrp48 colocalize with cRNAs on chromatin. Left, pie charts showing the proportion of hrp36 (top) and hrp48 (bottom) ChIP-seq peaks overlapping (light blue) or not (pink) with cRNAs (p-values: ***< 0.001, permutation test). Right, frequency of RPKM counts ranked by percentile of cRNAs at regions overlapping (light blue) or not (pink) with common hrp36 and hrp48 ChIP-seq peaks.

3.1.2. THE INVOLVEMENT OF DH1 IN THE REGULATION OF CRNAS AND THE RELATIONSHIP BETWEEN R-LOOPS AND CRNAS

Based on the results we obtained, we concluded that heterochromatin is enriched in cRNAs packaged by hrp36 and hrp48. Next, we wanted to address how depletion of dH1 affected to these elements.

3.1.2.1. hrp36 and hrp48 binding to chromatin is dependent on dH1

We performed WB analysis to study the effect of the depletion of dH1 on hrp36 and hrp48 on chromatin samples. Our findings indicated that under ~50% depletion of H1, both hrp36 and hrp48 signals displayed a similar and significant reduction in chromatin (**Figure 3.1.7A**). This result was further confirmed by ChIP-seq of dH1KD ($dsRNA^{dH1}$), where we observed a drastic and significant decrease in the intensity of hrp36 and hrp48 binding compared to the control sample ($dsRNA^{LacZ}$) (**Figure 3.1.7B**).

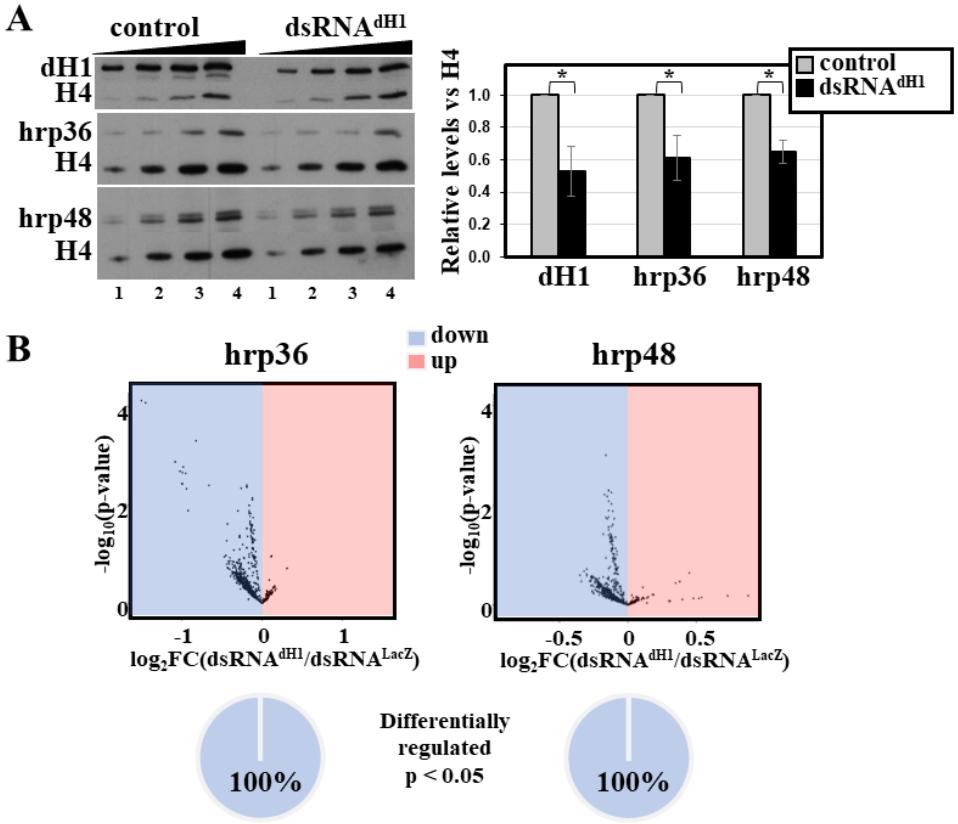


Figure 3.1.7. hrp36 and hrp48 binding to chromatin is dependent on dH1. (**A**) WB analysis with α-dH1 (top), α-hrp36 (center) and α-hrp48 (bottom) antibodies of increasing amounts (lanes 1-4) of crosslinked chromatin prepared from control (dsRNA^{LacZ}) (left) and dH1KD (dsRNA^{dH1}) (right) cells. α-H4 antibodies were used for loading control. Quantification of the results is shown on the right. Results are the average of 5 independent experiments. Error bars are standard deviation. (p-values: *< 0.05; two-tailed paired Student's test). (**B**) Top, Volcano plots showing the change in intensity of hrp36 (left) and hrp48 (right) ChIP-seq peaks in dH1KD (dsRNA^{dH1}) in comparison to control (dsRNA^{LacZ}) cells. Down-regulated and up-regulated peaks are shown in blue and red, respectively. Bottom, pie chart showing the proportion of significantly down-regulated and up-regulated peaks (p-value< 0.05).

Finally, to confirm that the binding of hrp36 and hrp48 was directly affected we analyzed the expression levels of hrp36 and hrp48 in the studied conditions. We found that in dH1KD the mRNA levels of hrp36 and hrp48 were unaltered respect to the control by RT-qPCR (**Figure 3.1.8**) and suggesting that depletion of dH1 is indeed altering the binding of hrp36 and hrp48 to cRNA, and consequently their packaging.

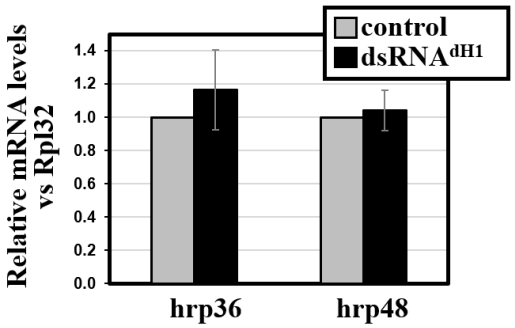


Figure 3.1.8. mRNA levels of hrp36 and hrp48 are unaltered in KD of dH1 cells. RT-qPCR experiments showing the mRNA levels of hrp36 (left) and hrp48 (right) relative to Rpl32 in control ($dsRNA^{LacZ}$) (grey) and KD of dH1 ($dsRNA^{dH1}$) (black) cells. Results are the average of 3 independent experiments. Error bars are standard deviation. (p-values> 0.05; two-tailed unpaired Student's test).

3.1.2.2. Depletion of dH1 leads to accumulation of heterochromatic cRNAs

Next, we studied the effects that dH1KD had on cRNAs. To do so, we performed cRNA-seq on dH1KD ($dsRNA^{dH1}$) and control ($dsRNA^{LacZ}$) cells. First, we found that regions where we could detect cRNA signal in both conditions were widely overlapping. About ~95% of control cRNAs overlapped with ~88% of dH1KD (**Figure 3.1.9A**). Despite dH1KD had slightly more regions containing cRNAs, when we analyzed the differential expression globally (p-value < 0.05), we found a similar proportion of upregulated and downregulated cRNAs, 49% and 51%, respectively (**Figure 3.1.9B**). Thus, there was no clear general tendency towards accumulation or reduction of cRNAs. Nevertheless, when we studied specifically the changes of heterochromatic cRNAs, we found that there was a clear upregulation of cRNAs, with a proportion of ~78% in dH1KD (**Figure 3.1.9C**), and this tendency of accumulation was statistically significant (p-value< 0.001, permutation test). Moreover, dH1KD also exhibited a high proportion of Intron Retention (IR) (**Figure 3.1.9D**), suggesting an additional malfunctioning of co-transcriptional splicing of cRNAs.

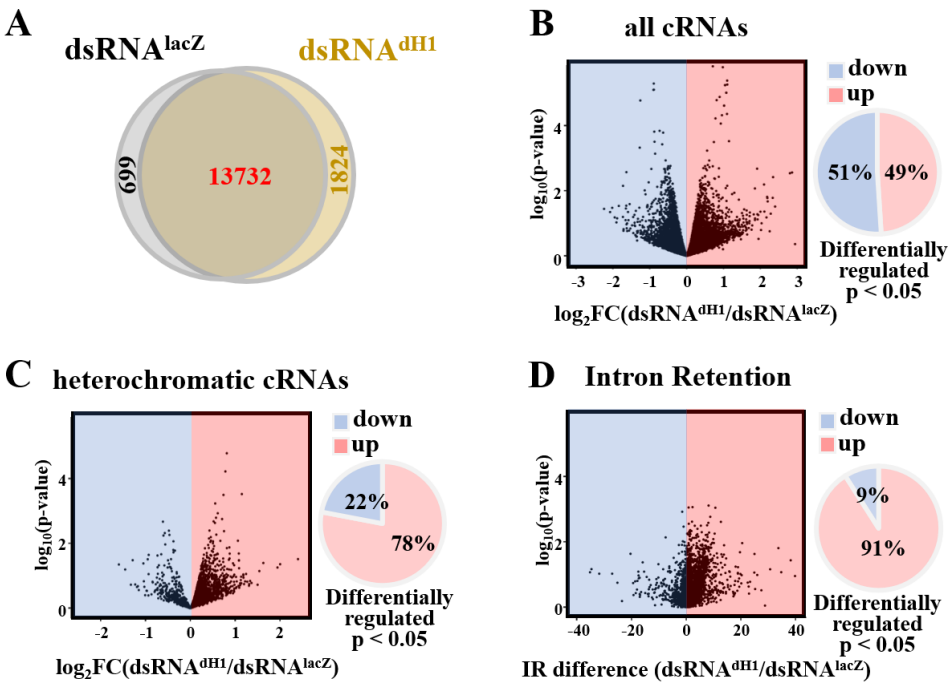


Figure 3.1.9. dH1 depletion leads to the accumulation of heterochromatic cRNAs and intron retention. (**A**) Venn diagram showing the overlap between cRNAs detected in control (dsRNA^{LacZ}) and dH1KD (dsRNA^{dH1}) cells. (**B**) Left, Volcano plot showing the change in intensity of all cRNAs detected in dH1KD in comparison to control cells. Down-regulated and up-regulated transcripts are shown in blue and red, respectively. Right, pie chart showing the proportion of significantly down-regulated and up-regulated cRNAs (p-value< 0.05). (**C**) As in **B**, but for heterochromatic cRNAs. (**D**) Left, Volcano plot showing the frequency of intron retention (IR) of cRNAs detected in dH1KD respect to control. Decreased and increased IR are shown in blue and red, respectively. Right, pie chart of the proportion of cRNAs showing significantly increased and decreased IR (p-value< 0.05).

In addition, specific categories of cRNAs, such as intergenic or lncRNAs also had a higher proportion of upregulated cRNAs (**Figure 3.1.10A** and **B**), although only the accumulation of intergenic cRNAs was statistically significant (Intergenic p-value < 0.001; lncRNAs test p-value = 0.110, permutation test).

Regarding the remaining categories, genic antisense had a similar proportion of upregulated and downregulated cRNAs (**Figure 3.1.10C)**, while genic sense cRNA showed a statistically significant increased

proportion of downregulated cRNAs in dH1KD (p-value < 0.001, permutation test) (**Figure 3.1.10D)**. Intrigued by the result of the genic sense cRNAs, we also analyzed the changes in cRNAs mapping to active chromatin states. Accordingly, promoter TSS, regulatory regions and transcription elongation cRNAs had a tendency to reduce in dH1KD that in all cases was statistically significant (p-value < 0.001, permutation test) (**Figure 3.1.10E-G)**

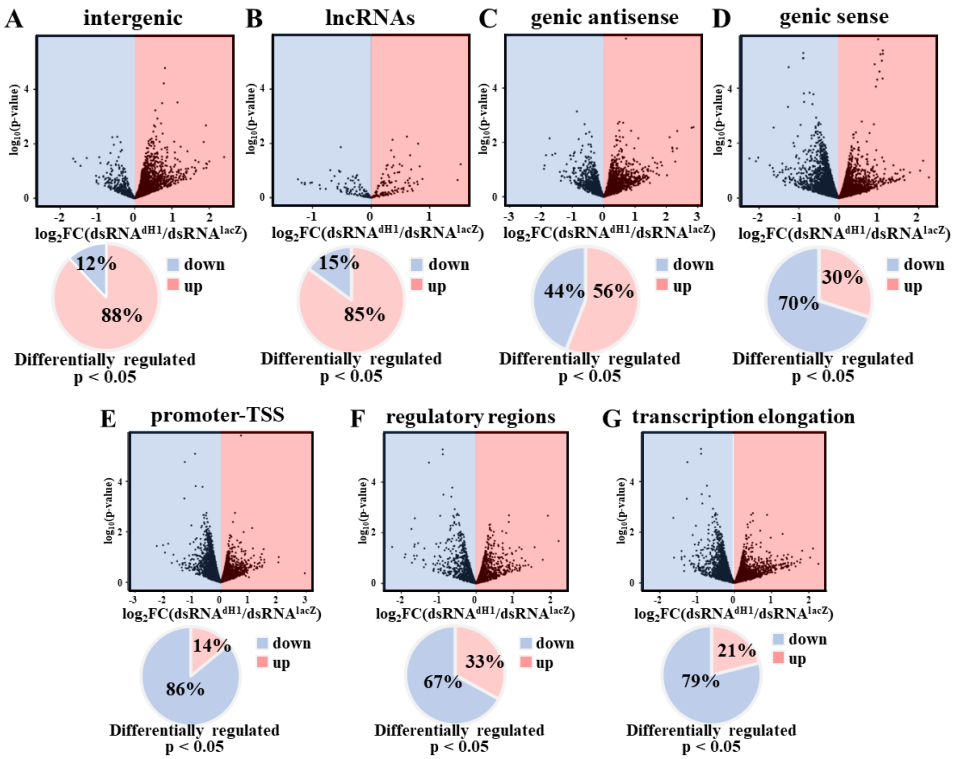


Figure 3.1.10. Changes in cRNA categories in dH1KD cells. (**A-D**) Top, Volcano plots displaying the changes in intensity in dH1KD ($dsRNA^{dH1}$) respect to control ($dsRNA^{LacZ}$) cells of intergenic (**A**), lncRNAs (**B**), genic antisense (**C**) and genic sense (**D**) cRNAs. Down-regulated and up-regulated transcripts are shown in blue and red, respectively. Bottom, pie charts showing the proportion of significantly down-regulated and up-regulated cRNAs (p-value< 0.05). (**E-G**) As in A-D, but for cRNAs mapping to active chromatin states (Kharchenko et al. 2011): promoter-TSS (**E**), regulatory regions (**F**) and transcription elongation (**G**).

Remarkably, we noticed that the accumulation of heterochromatic cRNAs in dH1KD cells was in agreement with the previous results of accumulation of R-loops in heterochromatin in dH1KD cells (Bayona-Feliu et al. 2017).

3.1.2.3. The impaired packaging of cRNAs gives rise to R-loops

Based on the fact that depletion of dH1 leads to the disruption of the packaging of heterochromatic cRNAs and to an accumulation of R-loops, we hypothesized that the impaired assembly of heterochromatic cRNAs by hrp36 and hrp48 could be a relevant factor to facilitate the formation of R-loops in heterochromatin. To test this, we studied the effects of directly reducing the packaging of the cRNAs by depleting either hrp36 or hrp48.

First, we studied the changes in cRNAs by cRNA-seq in hrp36KD ($dsRNA^{hrp36}$) and in hrp48KD ($dsRNA^{hrp48}$) compared to control ($dsRNA^{LacZ}$) cells. Again, we detected a similar total number of cRNAs, which strongly overlapped with the regions previously identified for dH1KD (**Figure 3.1.11A**).

For hrp48KD, like in dH1KD, there was a significant accumulation of heterochromatic cRNAs (p-value < 0.001) (**Figure 3.1.11B**). On the contrary, hrp36KD had a similar proportion of down-regulated and up-regulated cRNAs, without a clear tendency (**Figure 3.1.11C**). Remarkably, neither hrp48KD nor hrp36KD significantly affected intron retention (**Figure 3.1.11D** and **3.1.11E**).

These results evidenced that despite hrp36 and hrp48 have a strong colocalization on chromatin, they must have at least partially independent roles, since the phenotype of their KDs is clearly differentiated.

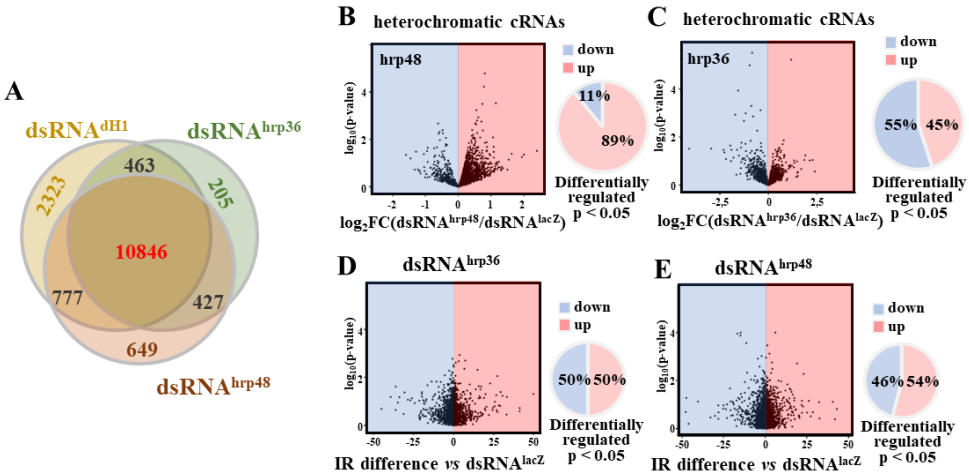


Figure 3.1.11. cRNA-seq in KD of hrp36 and KD of hrp48 cells. (**A**) Venn diagram showing the overlap between cRNAs detected in dH1KD (dsRNA^{dH1}), hrp36KD (dsRNA^{hrp36}) and hrp48KD (dsRNA^{hrp48}) cells. (**B**) Left, Volcano plot showing the change in intensity of heterochromatic cRNAs detected in hrp48KD in comparison to control (dsRNA^{LacZ}). Down-regulated and up-regulated transcripts are shown in blue and red, respectively. Right, pie chart showing the proportion of differentially down-regulated and up-regulated cRNAs (p-value < 0.05). (**C**) As in **B** but for hrp36KD. (**D**) Left, Volcano plot showing the frequency of IR of genic cRNAs detected in hrp48KD respect to control cells. Decreased and increased IR are shown in blue and red, respectively. Right, pie charts of the proportion of cRNAs showing differentially increased and decreased IR (p-value < 0.05). (E) As in D but for hrp36KD.

We further confirmed that, experimentally, their KDs were mutually independent. By WB analysis of chromatin samples, we found that hrp36KD did not affect protein levels of hrp48 on chromatin and, the other way around, hrp48KD did not affect to protein levels of hrp36 on chromatin (**Figure 3.1.12**).

Next, we directly addressed R-loops formation in hrp36KD and hrp48KD cells. We performed DRIP-seq experiments, and from the results we called R-loops to regions significantly enriched after immunoprecipitation with S9.6 antibodies that, upon treatment with RNase H before immunoprecipitation, decreased coverage by FC < -1.25.

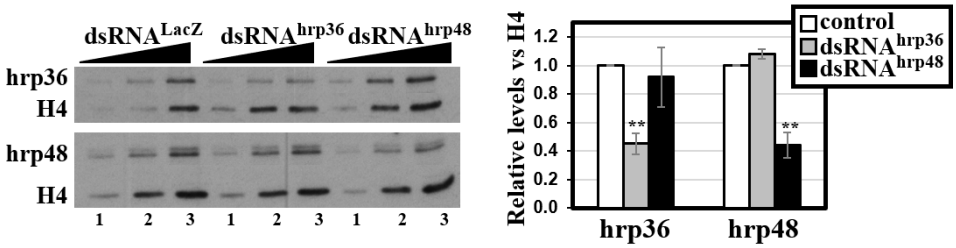


Figure 3.1.12 Depletions of hrp36 and hrp48 are independent of each other. WB analysis with α-hrp36 (top) and α-hrp48 (bottom) antibodies of increasing amounts (lanes 1-3) of crosslinked chromatin prepared from control (dsRNA^{LacZ}) (left), hrp36KD (dsRNA^{hrp36}) (center) and hrp48KD (dsRNA^{hrp48}) (right) cells. α-H4 antibodies were used for loading control. Quantification of the results is shown on the right. Results are the average of 3 independent experiments. Error bars are standard deviation (p-values: **<0.01; two-tailed paired Student's test).

We called ~2000 R-loops for control (dsRNA^{LacZ}) and hrp36KD (dsRNA^{hrp36}), and ~5000 R-loops for hrp48KD (dsRNA^{hrp48}), which were distributed along the whole genome (**Figure 3.1.13A-C**). In agreement with previous results, the total number of R-loops for hrp36 did not increase compared to the control, while the KD of hrp48 presented more than twice R-loops than control cells (**Figure 3.1.13D**). However, despite the difference in magnitude comparing the absolute numbers of R-loops between the KD of hrp36 and the KD of hrp48, their distribution on the genome presented a similar shift towards heterochromatin, for both the vast majority of R-loops detected only in the KDs mapped to heterochromatin, similarly to common peaks between the KDs and the control, while the control only R-loops were largely mapping in euchromatin (**Figure 3.1.13E** and **Figure 3.1.13F**).

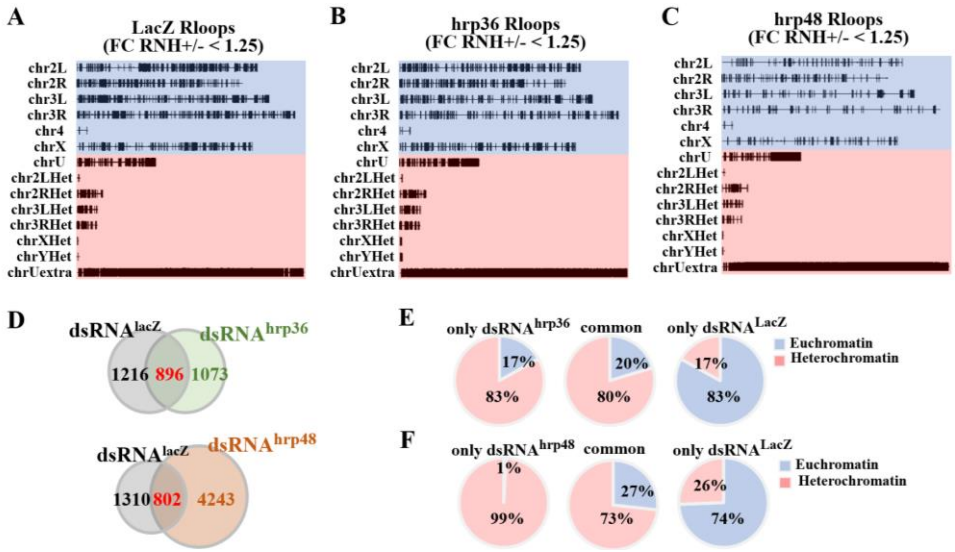


Figure 3.1.13. Depletion of hrp36 and hrp48 are enriched in heterochromatic R-loops. (**A-C**) Chromosomal distribution of DRIP-seq peaks RNaseH sensitive for control (dsRNA^{LacZ}) (**A**), hrp36KD (dsRNA^{hrp36}) (**B**) and hrp48KD (dsRNA^{hrp48}) (**C**). Chromosome definition and orientation are as in **Figure 3.1.1A**. Euchromatic and heterochromatic regions are highlighted in blue and red, respectively. (**D**) Venn diagrams showing the overlap between R-loops detected in control and hrp36KD cells (top) and between control and hrp48KD cells (bottom). (**E**) Pie charts showing the proportions of R-loops mapping at euchromatin (blue) and heterochromatin (red) detected only in hrp48KD (left), only in control cells (right), and both in hrp48KD and in control cells (center). (**F**) As in **E**, but for R-loops detected in hrp36KD cells.

Remarkably, Volcano plot analysis for hrp36KD and hrp48KD showed that both KDs strongly increased global R-loops intensity in comparison to control cells (**Figure 3.1.14A** and **B**), thus confirming that impairing the packaging of the cRNAs through depletion of hrp36 of hrp48 enhanced the formation of R-loops. In the same direction of previous results, although both KDs presented the same R-loop-accumulation phenotype, hrp48KD showed a stronger phenotype, and the accumulated R-loops in hrp48KD were more enriched in heterochromatin than those accumulated in hrp36KD (**Figure 3.1.14C** and **D**).

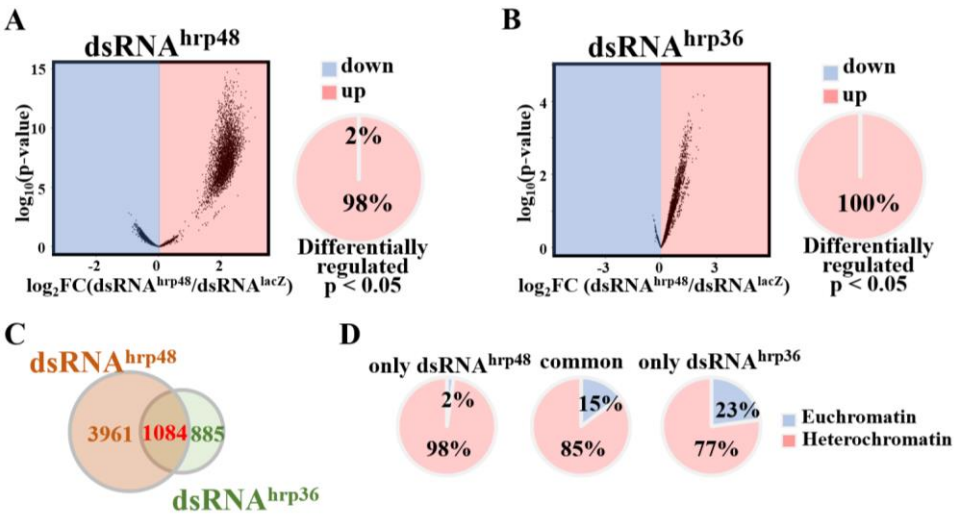


Figure 3.1.14. Depletion of hrp36 and hrp48 induce R-loops accumulation in heterochromatin. (**A**) Left, Volcano plot showing the change in intensity of R-loops detected in hrp48KD (dsRNA^{hrp48}) in comparison to control (dsRNA^{LacZ}) cells. Down-regulated and up-regulated R-loops are shown in blue and red, respectively. Right, pie chart showing the proportion of differentially down-regulated and up-regulated R-loops (p-value< 0.05). (**B**) As in **A**, but for R-loops detected in hrp36KD (dsRNA^{hrp36}) cells. (**C**) Venn diagram showing the overlap between R-loops detected by DRIP-seq in hrp48KD and hrp36KD cells. (**D**) Pie charts showing the proportion of R-loops mapping at euchromatin (blue) and heterochromatin (red) detected only in hrp48KD (left), only in hrp36KD (right), and in both (center).

Ultimately, we also wanted to address whether those R-loops where *de novo* cRNAs or cRNAs stabilized in the form of R-loops. Therefore, we compared the localization of cRNAs of WT samples with the accumulated R-loops in hrp36KD and hrp48KD, and found that they had a remarkable overlap, which was statistically significant (p-values < 0.001, permutation test) (**Figure 3.1.15A**). This further confirmed that the altered packaging of the cRNAs promotes R-loop formation. Moreover, R-loops were detected in ~50% of the hrp36 and hrp48 enriched regions in WT conditions and this overlap was also statistically significant (p-values < 0.001, permutation test) (**Figure 3.1.15B**).

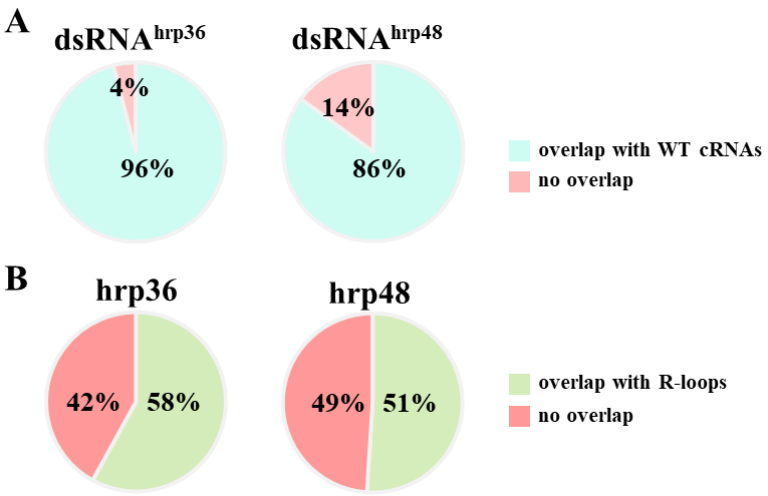


Figure 3.1.15. R-loops induced by depletion of hrp36 and hrp48 colocalize with WT cRNA and hrp36 and hrp48 ChIP-seq peaks. (**A**) Pie charts showing the proportions of R-loops detected by DRIP-seq in hrp36KD ($dsRNA^{hrp36}$) (left) and hrp48KD ($dsRNA^{hrp48}$) (right) cells overlapping (light blue) or not (pink) with WT cRNAs. (**B**) Pie charts showing the proportions of hrp36 and hrp48 ChIP-seq peaks overlapping (green) or not (red) with R-loops detected by DRIP-seq in hrp36KD (left) and hrp48KD (right).

All in all, our results showed that the accumulation of R-loops in heterochromatin in dH1KD is somehow mediated by the disruption of the packaging of the heterochromatic cRNAs.

3.1.3. THE CONTRIBUTION OF DH1 TO CHROMATIN STRUCTURE

Finally, with the well-known structural role of linker histone H1 in mind, we aimed to assess the impact of dH1KD on chromatin structure. Specifically, we investigated changes in chromatin accessibility concomitant to the cRNA changes to gain insight into the effects of dH1KD.

3.1.3.1. Depletion of dH1 leads to an opening of heterochromatin

We performed ATAC-seq in dH1KD ($dsRNA^{dH1}$) and control ($dsRNA^{LacZ}$) cells and analyzed the proportion of the different fragment sizes obtained. From this analysis we could distinguish transposase hypersensitive sites (THSS) (fragments <120 bp) and nucleosomal insert sizes (fragments of 120-250 bp). When comparing the relative proportion of each insert size between the dH1KD and control cells, we found an enrichment of THSS in dH1KD cells, indicating a general increased accessibility (**Figure 3.1.16A**). On the other hand, the nucleosomal fragment sizes were in general depleted, with an especially strong depletion of the two main nucleosomal fragments detected (138 bp and 151 bp) in dH1KD cells (**Figure 3.1.16A**).

As a direct measure of the changes in accessibility, we first addressed the changes in the THSS fragments, and performed peak calling with them. We found that, despite the enrichment of THSS fragments in dH1KD, there was a reduction of the total number of peaks. Nevertheless, from those detected in dH1KD ~80% overlapped with those of the control cells (**Figure 3.1.16B)**. Furthermore, the common peaks in both conditions displayed a distribution of a similar proportion of peaks between euchromatin and heterochromatin (**Figure 3.1.16C** center). In contrast, peaks specific to the control condition or to dH1KD had a biased distribution. On one hand, the control was biased towards euchromatin (91%) (**Figure 3.1.16C** left), while dH1KD was biased towards heterochromatin (87%) (**Figure 3.1.16C** right).

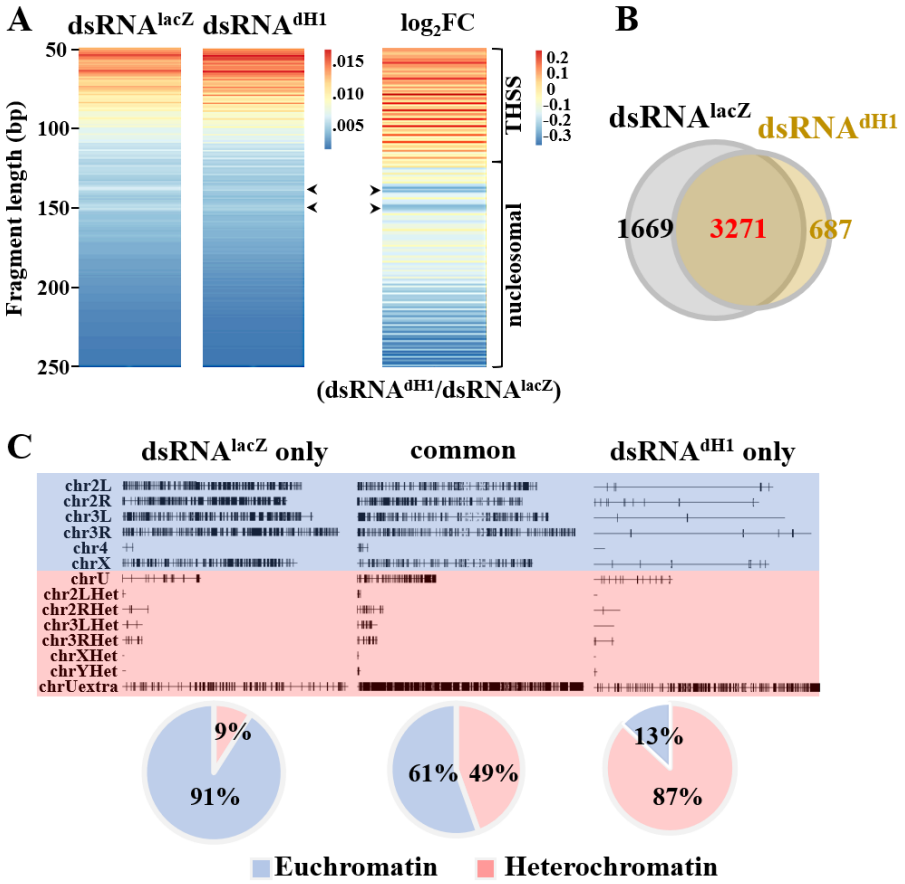


Figure. 3.1.16. Changes in fragment length enrichment and distribution of THSS in KD of dH1. (**A**) Left, heat maps showing the size (bp) distributions of ATAC-seq fragments in control (dsRNA^{LacZ}) and dH1KD (dsRNA^{dH1}) cells. Right, differential heat map of the Fold Change (FC) of the intensities of the ATAC-seq fragments between dH1KD and control cells. Fragment sizes corresponding to transposase hypersensitive (THSS) and nucleosomal fragments are indicated. Arrowheads indicate the two major nucleosomal fragments detected. (**B**) Venn diagram displaying the overlap between peaks called from the THSS fragments in control and dH1KD cells. (**C**) Top, chromosomal distribution of peaks called from the THSS fragments detected only in control cells (left), only in dH1KD cells (right), or common in both conditions (center). Chromosome definition and orientation are as in **Figure 3.1.1A**. Euchromatic and heterochromatic regions are highlighted in blue and red, respectively. Bottom, pie charts showing the proportion of peaks mapping to euchromatic and heterochromatic regions.

Next, we compared the average intensity of the peaks between dH1KD and control cells grouped by the chromosomal regions. This analysis showed that in heterochromatic regions the overall intensity of the THSS fragments

was higher in dH1KD, while in euchromatic regions the intensity was reduced (**Figure 3.1.17**), confirming that there is an opening of heterochromatin when dH1 is depleted.

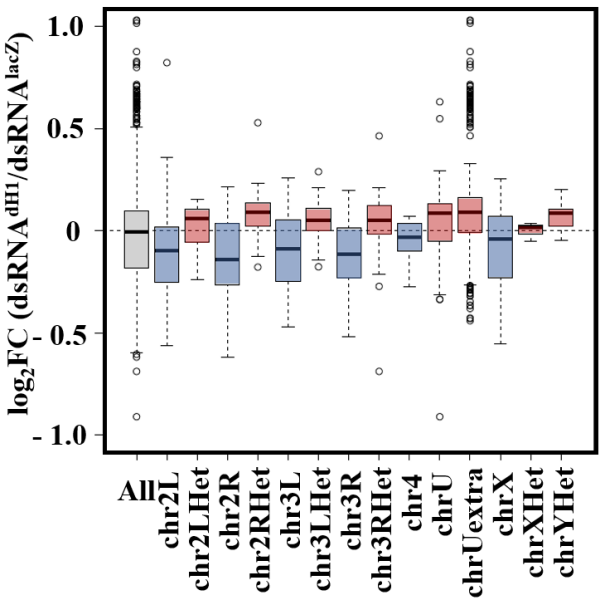


Figure 3.1.17. There is a specific opening of heterochromatin in KD of dH1 cells. Box plot showing the average FC in intensity of THSS fragments in dH1KD ($dsRNA^{dH1}$) and control ($dsRNA^{LacZ}$) cells for all the scaffolds together (All, in grey) and for each chromosome and scaffold separately. Euchromatic and heterochromatic regions are highlighted in blue and red, respectively.

3.1.3.2. A general increased accessibility in dH1KD cells

The analysis by peak calling of the THSS showed, aside from the opening of heterochromatin, a reduction of peaks and intensity in euchromatic regions in dH1KD cells. Thus, to further study the changes in euchromatic regions we focused on the distribution of those fragments around the TSS, since TSS was the most represented category in the control peaks. We generated two-dimensional occupancy plots centered on the TSS for the THSS fragments. These plots showed that, while control cells had a narrow peak of accessibility ~100bp upstream the TSS, in dH1KD the accessible region was not constricted upstream the TSS but spread along the gene body

(**Figure 3.1.18**). Showing that rather than being less accessible the signal became wider in dH1KD.

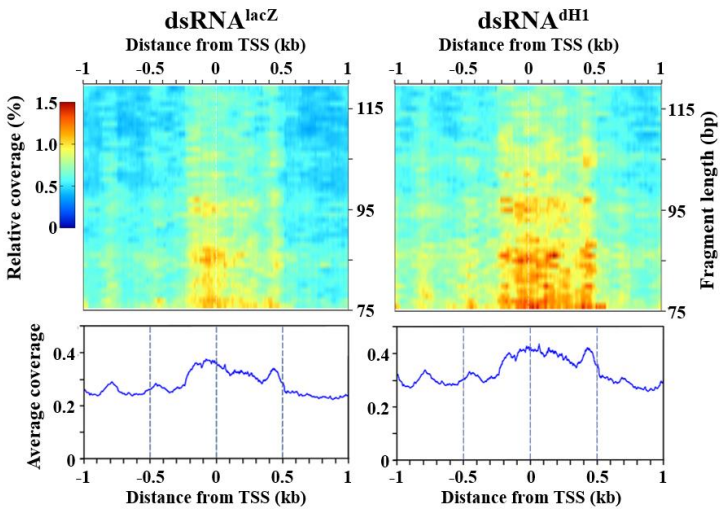


Figure 3.1.18. There is increased accessibility at TSS in dH1KD cells. Top, 2D occupancy plots of THSS coverage around the TSSs in control ($dsRNA^{LacZ}$) (left) and dH1KD ($dsRNA^{dH1}$) cells (right). Bottom, plots of average coverage around the TSS. Distance from TSS (kb) is indicated.

3.1.3.3. dH1KD reduces nucleosome occupancy

Taking into consideration the previous results, where we observed that nucleosomal fragment sizes were in general depleted (**Figure 3.1.16A**) and accessibility downstream of the TSS was increased in dH1KD (**Figure 3.1.18**), potentially indicating an alteration of nucleosome distribution, we decided to study the nucleosomal fragments. Thus, we generated two-dimensional occupancy plots centered on the TSS for the nucleosomal fragments. We detected, again, two main nucleosomal fragments, of 138bp and 151bp, and a less abundant subnucleosomal fragment of 125 bp (**Figure 3.1.19A** and **B**). When comparing the control and dH1KD cells, we observed that nucleosomal fragments presented a reduced intensity in dH1KD (**Figure 3.1.19A** top) and that the percentage of nucleosomal fragments was also reduced and subnucleosomal fragments were enriched in dH1KD compared to control cells (**Figure 3.1.19B**). Altogether it indicated a destabilization and reduction in nucleosome occupancy around

the TSS in dH1KD. Furthermore, plotting the average dyad density, we found that the positioning of the +1 nucleosome was also affected in dH1KD, as the peak was less well-positioned and slightly broader, in combination with a broader peak of nucleosome +2 (**Figure 3.1.19A** bottom).

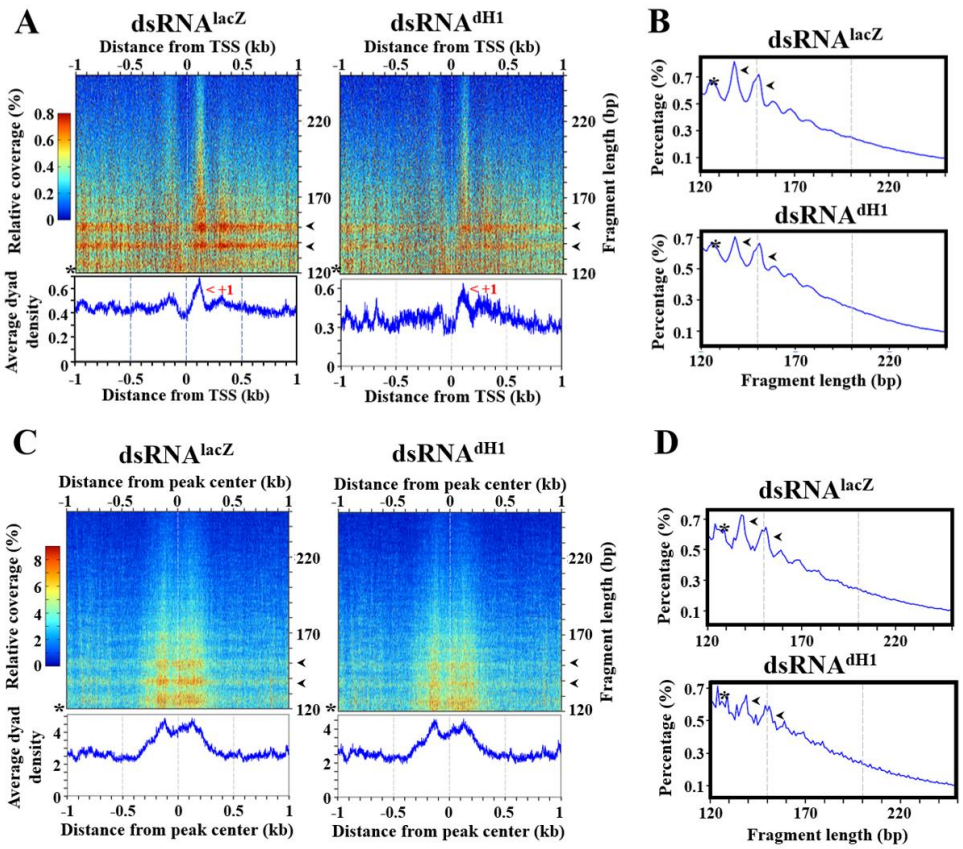


Figure 3.1.19. Changes in nucleosomal fragments and nucleosome occupancy in dH1KD cells. (**A**) Top, occupancy plots of coverage for nucleosomal fragments around the TSSs in control (dsRNA^{LacZ}) (left) and dH1KD (dsRNA^{dH1}) cells (right). The asterisk indicates a subnucleosomal fragment of 125 bp. Arrowheads indicate the two major nucleosomal bands detected, of 138 bp and 151 bp. Bottom, plots of average dyad density around the TSSs. Distance from the TSS and position of the +1 nucleosome are indicated. (**B**) Proportion of nucleosomal fragments in **A** is shown as a function of fragment size in control (top) and dH1KD cells (bottom). The asterisk indicates the subnucleosomal fragment. Arrowheads indicate the two major nucleosomal bands detected. (**C**) As in **A**, but around the center of the common peaks of the ChIP-seqs of hrp36 and hrp48. (**D**) As in **B**, but around the center of the common peaks of the ChIP-seqs of hrp36 and hrp48.

Next, we wanted to assess whether this change was global or specific for the TSS. To do so, we decided to study the enriched regions in hrp36 and hrp48, which were mostly heterochromatic. We generated two-dimensional occupancy plots focused on the center of the common ChIP-seq peaks of hrp36 and hrp48 for the nucleosomal fragments. We observed that, in the same direction that around the TSS, comparing the control and dH1KD cells, the nucleosomal fragments presented a reduced intensity in dH1KD (**Figure 3.1.19C** top). In this scenario the three fragment sizes detected were the same as before, although the subnucleosomal fragment was present in a higher proportion, and, again, the percentage of nucleosomal fragments was reduced and shifted towards the subnucleosomal fragment in dH1KD compared to the control cells (**Figure 3.1.19D**). In this case, average dyad density plots were less informative in these regions because they are a mixture of different nucleosomal positions. Consequently, we could not find many differences between conditions with this approach (**Figure 3.1.19C** bottom).

In summary, all these results suggested that the dH1KD leads to a general destabilization of the nucleosomes, resulting in a global increase in chromatin accessibility, with a specific significant change in heterochromatin.

3.2. COMPARISON OF THE BIOCHEMICAL PROPERTIES OF SOMATIC DH1 AND EMBRYONIC DBIGH1

3.2.1. STRUCTURAL PROPERTIES OF DH1 AND DBIGH1

The two linker histones variants of *D. melanogaster*, dH1 and dBigH1, present the metazoa conserved tripartite structure consisting of a short N-terminal tail, a central globular domain (GD), and a long C-terminal tail (Ramakrishnan et al. 1993). Nevertheless, they have fundamental structural differences that are somewhat representative of the differences that somatic and germline H1 variants present in metazoa. On one hand, dBigH1 is globally ~100 aa bigger than dH1 (**Figure 3.2.1** Top). Even though the globular domains are very similar in length, both dBigH1 tails are ~50 aa longer. This is especially critical for the N-terminal domain, since the N-terminal tail of dBigH1 is twice the size of dH1 N-terminal tail. Moreover, the same larger N-terminal tail of dBigH1 has a large patch enriched in acidic residues (**Figure 3.2.1** Center), which makes the electrostatic nature of both linker histones very different.

However, the predicted distribution of unstructured and globular domains in both isoforms is conserved (**Figure 3.2.1** Bottom). As it has been described and conserved among metazoa, the N-terminal domain (NTD) and C-terminal domain (CTD) are disorganized while the only ordered region corresponds to the winged-helix motif of the GD.

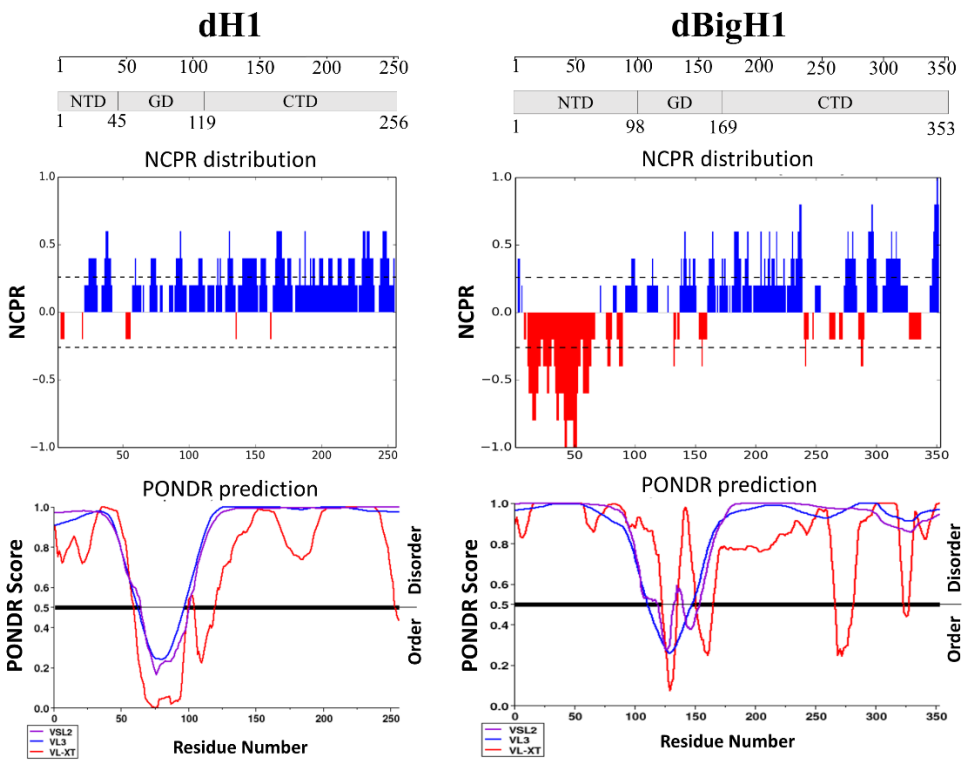


Figure 3.2.1. dH1 and dBigH1 domains, charge distribution and potential disordered regions. Comparison of different features for the proteins H1 (left) and BigH1 (right). **Top**, scheme showing the three domains and their relative length in aa. **Center**, plot showing the linear net charge per residue generated with CIDER (Holehouse et al. 2017). **Bottom**, graphical output from a PONDR computational prediction of ordered and disordered regions, using the VSL2, VL3, and VL-XT algorithms (Xue et al. 2010).

The complete structure for dH1 and dBigH1 has not yet been experimentally solved. However, using AlphaFold (Jumper et al. 2021) to predict the 3D structure we found that, as expected, dH1 presented the conserved three alpha helices with a C-terminal beta hairpin of the GD and the unstructured NTD and CTD (**Figure 3.2.2A**). dBigH1 showed the conserved GD likewise, yet it also showed a strikingly long α-helix (**Figure 3.2.2B**). This long α-helix marks another relevant difference between these proteins, potentially having implications in binding affinity and/or dimerization events.

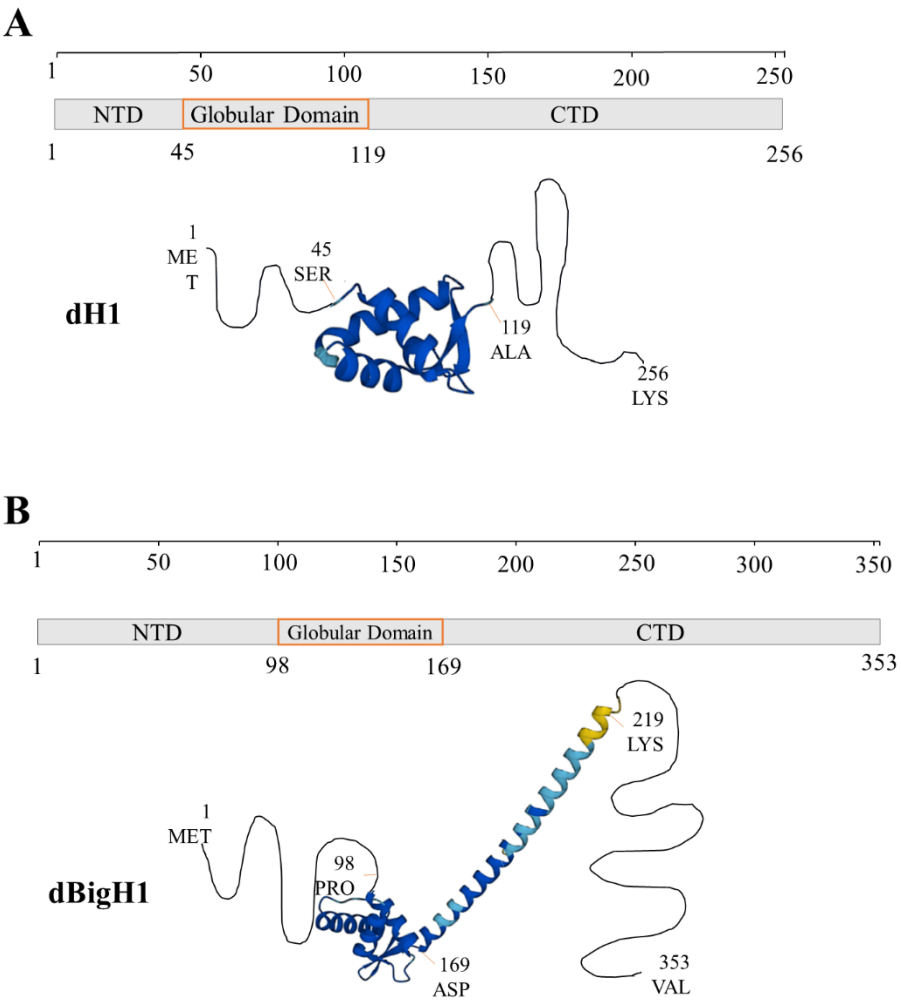


Figure 3.2.2. Prediction of the secondary structure of dH1 and dBigH1. Structure generated by AlphaFold for (A) dH1 of (B) dBigH1. The AlphaFold output has been modified to restrict it to the regions that in the model of confidence present only a Very High or a Confident value. The regions with Very Low or Low confidence are schematically drown. The aa that limits each structure are highlighted.

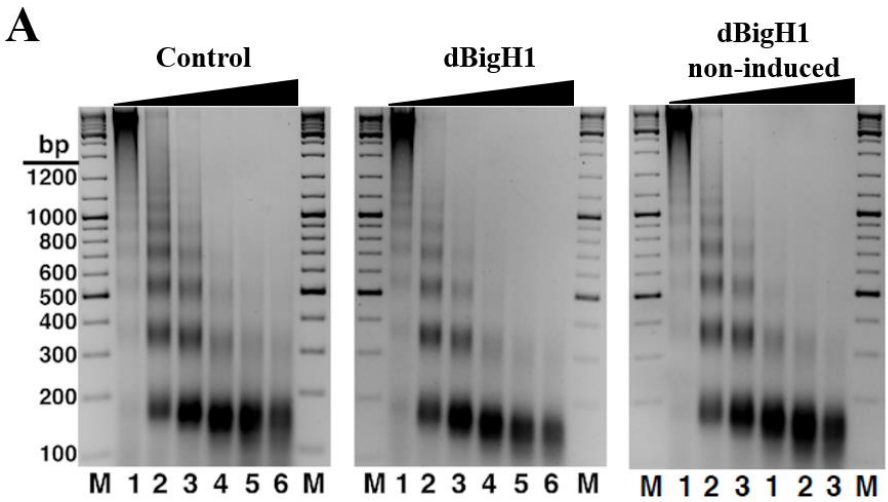
Therefore, dH1 and dBigH1 present key structural differences that could alter how they interact with chromatin. Because work from another member from our laboratory showed that dH1 and dBigH1 had a differential impact on the functional state of chromatin (Climent-Cantó et al. 2020), we reasoned that this could be based on their biochemical differences.

3.2.2. EFFECT OF THE TWO LINKER HISTONES ON THE NUCLEOSOME REPEAT LENGTH (NRL)

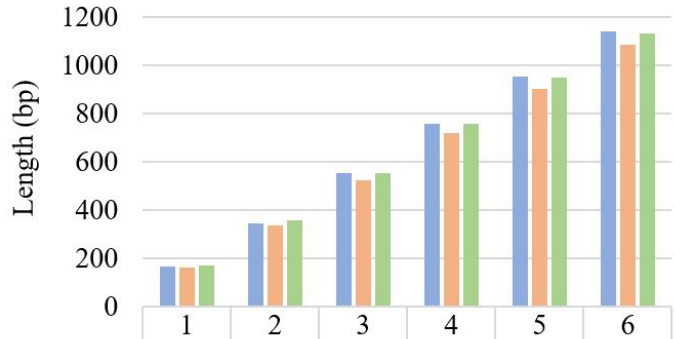
3.2.2.1. The NRL is reduced in presence of dBigH1 in somatic cells

Prior to a detailed *in vitro* characterization of these proteins, we first assayed the properties of the two linker histones in cells to ensure how their structural differences may be responsible for not only a differential regulation but also for their different physical interaction with chromatin. Thus, we started by analyzing the changes in the nucleosome repeat length (NRL).

Comparison of the NRLs in presence or absence of dBigH1 was performed using the stable cell line: dBigH1::FLAG, which has a tagged form of dBigH1 regulated by a metallothionein promoter (pMT) that is inducible by its exposure to some metal ions. We used $CuSO_4$ to induce its expression. For these experiments we used three different samples, S2 WT cells treated with copper (referred to as Control), dBigH1::FLAG cells treated with copper (referred to as dBigH1), and dBigH1::FLAG cells not treated with copper (referred to as dBigH1 non-induced) and obtained their corresponding MNase treated ladders (**Figure 3.2.3A**). While for Control and dBigH1 non-induced conditions band sizes were similar, measurements of the nucleosomal bands were clearly shorter for the dBigH1 induced condition. These differences became clearer as arrays of nucleosomes were longer (from trinucleosomes to hexanucleosomes) (**Figure 3.2.3B**).While Control and non-induced dBigH1 showed NRLs of 196.6bp and 193.5bp, respectively, induced dBigH1 presented a NRL of 185.8bp (**Figure 3.2.4**). Therefore, the presence of BigH1 on cells reduced the overall NRL by ~10 bp (i.e. about one helix turn).



B



	1	2	3	4	5	6
Control	168.9	347.1	553.0	757.3	954.7	1140.0
dBigH1	165.3	336.8	525.9	718.3	904.9	1086.6
dBigH1 non-induced	172.7	356.8	553.9	756.6	948.5	1131.7

Nucleosomes (n)

Figure 3.2.3. MNase digestion ladder and measure of the nucleosomal sizes in presence or absence of dBigH1. (A) MNase digestion for increasing time (lanes 1–6) of nuclei obtained from control cells (left), dBigH1-expressing cells (center) and dBigH1 non-induced cells (right). Lanes M correspond to MW markers (the sizes in bp are indicated). **(B)** Quantitative analysis of the nucleosomal sizes in bp of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes. The measure is from samples showing equivalent extent of digestion (lanes 2 in A).

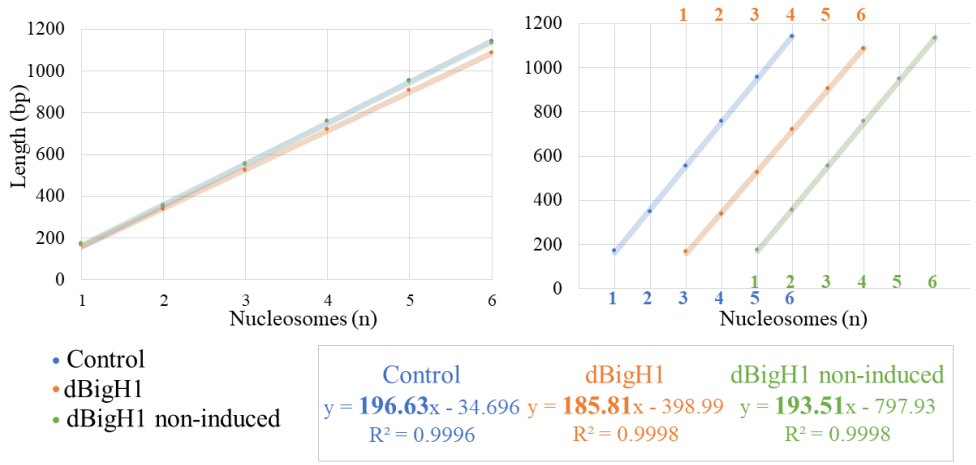


Figure 3.2.4. NRL is reduced in presence of dBigH1 in cells. NRL analysis of the samples from **Figure 3.2.3**. The size in bp of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes, are plotted against the number of nucleosomes. Represented with the data from the control (blue), dBigH1 (orange) and dBigH1 non-induced (green) samples overlapping (left) and separated with a different X axis reference per sample. following the color code (right). The correlation coefficients (R2) and slopes, which correspond to the apparent NRL, are indicated.

3.2.2.2. The NRL changes are caused by the acidic rich domain of dBigH1

Having confirmed that the presence of dBigH1 altered the spacing of the nucleosomes we wanted to assess whether this feature could be caused by the acidic residues from dBigH1 NTD, which is one of the main differences that dH1 and dBigH1 present. Thus, we analyzed the NRL using the stable cell line dBigH1ΔED::FLAG (referred to as dBigH1ΔED), that upon induction by metal ions expressed a mutated form of dBigH1 missing the region of the NTD where the acidic residues accumulate (**Figure 3.2.5**). When induced, dBigH1ΔED::FLAG was expressed to a similar level than full-length dBigH1ΔED::FLAG (Climent-Cantó et al. 2020).

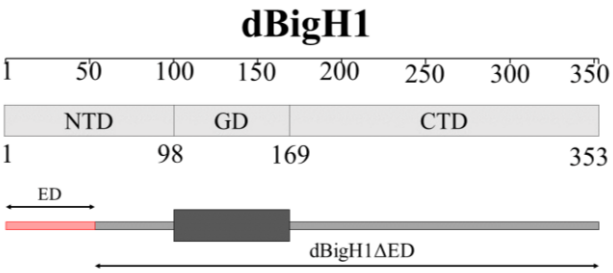
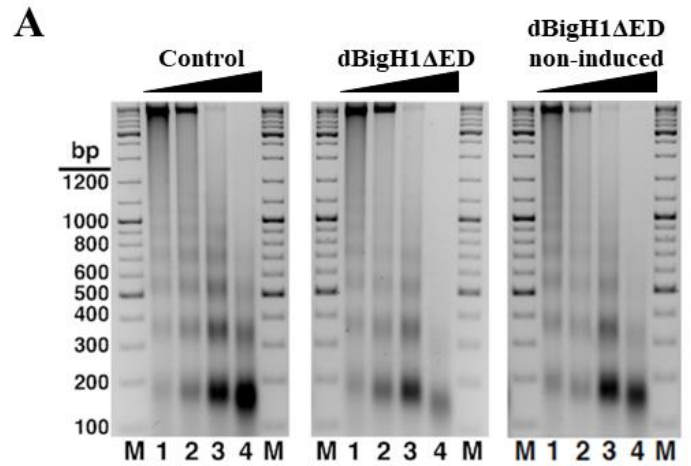
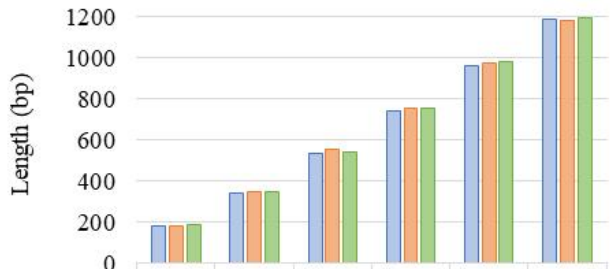


Figure 3.2.5. Construct of dBigH1ΔED. Schematic representation of the domain organization of dBigH1 and the excised region in dBigH1ΔED. ED, removed the acidic domain. Numbers indicate aa position in the primary sequence.

Again, we used as control samples S2 WT cells treated with copper, dBigH1ΔED cells treated with copper and dBigH1ΔED cells non-treated with copper (referred to as dBigH1ΔED non-induced) and obtained their corresponding MNase ladders (**Figure 3.2.6A**). In this scenario, measurements of the bands showed that the nucleosomal sizes of the dBigH1ΔED condition were neither larger or smaller than the two controls, but very similar to the other two conditions (**Figure 3.2.6B**).



B



	1	2	3	4	5	6
Control	176.3	337.7	534.5	736.9	962.5	1183.9
dBigH1ΔED	181.5	346.6	550.5	755.4	972.7	1179.4
dBigH1ΔED non-induced	185.6	346.1	541.8	753.4	977.2	1192.6

Nucleosomes (n)

Figure 3.2.6. MNase digestion ladder and measure of the nucleosomal sizes in presence or absence of dBigH1ΔED. (A) MNase digestion for increasing time (lanes 1–4) of nuclei obtained from control cells (left), dBigH1ΔED -expressing cells (center) and dBigH1ΔED non-induced cells (right). Lanes M correspond to MW markers (the sizes in bp are indicated). (B) Quantitative analysis of the nucleosomal sizes in bp of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes. The measure is from samples showing equivalent extent of digestion (lanes 1 in B).

As anticipated by the nucleosomal sizes, the NRL of the three conditions was fairly similar, with values of 203.29, 202.08 and 203.99 for the Control, the dBigH1ΔED and the dBigH1ΔED non-induced, respectively (**Figure 3.2.7**). We conclude that in these conditions the acidic region of the N-terminal domain accounted for the differences in NRL observed in cells.

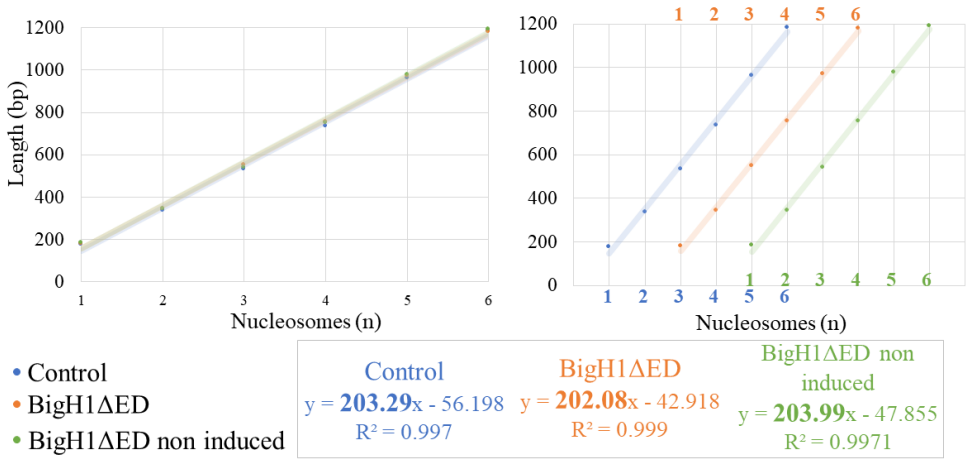


Figure 3.2.7. In absence of the acidic domain the presence of dBigH1 does not change the NRL on cells. Quantitative analysis of the results. The size in bp of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes, are plotted against the number of nucleosomes. Represented with the data from the control (blue), dBigH1ΔED (orange) and dBigH1ΔED non-induced (green) samples overlapping (left) and separated with a different X axis reference per sample following the color code (right). The correlation coefficients (R2) and slopes, which correspond to the apparent NRL, are indicated.

Having confirmed how a structural difference of the two linker histones was responsible for a differential nucleosome spacing, we argued that a deeper understanding of their physicochemical properties would provide useful insight to understand their biological differences.

3.2.3. STUDY OF THE PHASE SEPARATION PROPERTIES OF BOTH LINKER HISTONES

It has been described that H1 has an impact on chromatin phase separation, altering its dynamics (Gibson et al. 2019; Hansen et al. 2021). Moreover, the intrinsic phase-separation properties of H1 have also been characterized (Leicher et al. 2022; Mimura et al. 2021; Shakya et al. 2020). However, only one mammalian somatic variant has been included in all the studies, while H1s from other organisms and linker histone variants have not been

studied. We argued that, based on their structural differences and the fact that they give rise to different states of chromatin, BigH1 and H1 phase separation properties could be different and contribute to explain their biological differences. Thus, to study their phase separation properties we expressed and purified the two recombinant proteins as described in the section "**5.2.7. Protein production and purification**".

3.2.3.1. Phase separation conditions for dH1: the presence of DNA

First, we focused on the phase separation properties of dH1. It has already been described that H1 cannot phase separate without nucleic acids (Shakya et al. 2020). Moreover, that it can give rise to different non-spherical or spherical phase separated elements depending on the relative proportion of protein and DNA concentration (Mimura et al. 2021). Thus, to confirm the reproducibility with the somatic linker histone of *D. melanogaster*, dH1, we made a phase diagram maintaining the protein concentration of dH1 at 10 µM and changing the DNA concentration in a range from 0 to 2 µM, at increasing NaCl concentrations (**Figure 3.2.8A/B**). As DNA we used a dsDNA fragment amplified by PCR (primers and sequence described in "**5.1.2.4. Primers for dsDNA production**") of 281bp. Samples were analyzed immediately after preparation to avoid their evolution over time. Samples were studied using the projection of Differential interference contrast (DIC) images from 40µm segments, acquired at 4 µm intervals, that we then analyzed as described in the section "**5.2.9.4. LLPS analyses**" to obtain the quantitative information of the count and size of the phase-separated objects (**Figure 3.2.9**). The results we obtained recapitulated what was described for other somatic linker histones. On one hand, we did not observe phase separation without DNA (**Figure 3.2.8A/B j-l**). Next, at the lower concentration of DNA (0.25 µM) we observed non-spheric aggregates (**Figure 3.2.8A/B g-i**). At the concentration of DNA corresponding to 0.5 µM we observed a mixture of non-spheric and spheric phase-separated objects (**Figure 3.2.8A/B d** and **e**), while at the highest DNA concentration (2 µM) there was phase-separation with exclusively spheric droplets (**Figure 3.2.8A/B a-c**).

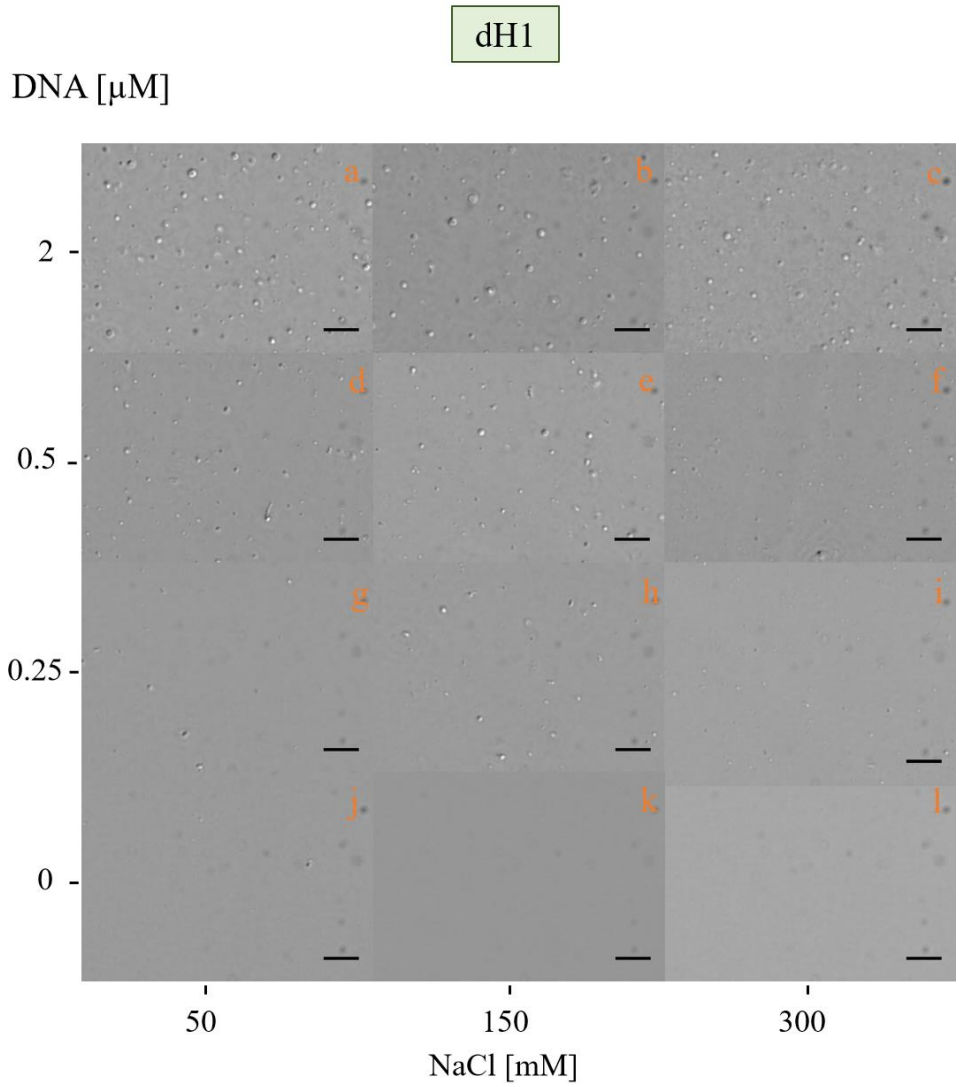


Figure 3.2.8A. Phase separation of dH1 is dependent on DNA. A phase diagram for dH1-DNA is shown, composed of projections of DIC images from 40µm segments, acquired at 4 µm intervals. dH1 concentration is of 10 µM with varying DNA and NaCl concentrations, as indicated. Images were acquired immediately after mixing. Scale bars correspond to 10 µm.

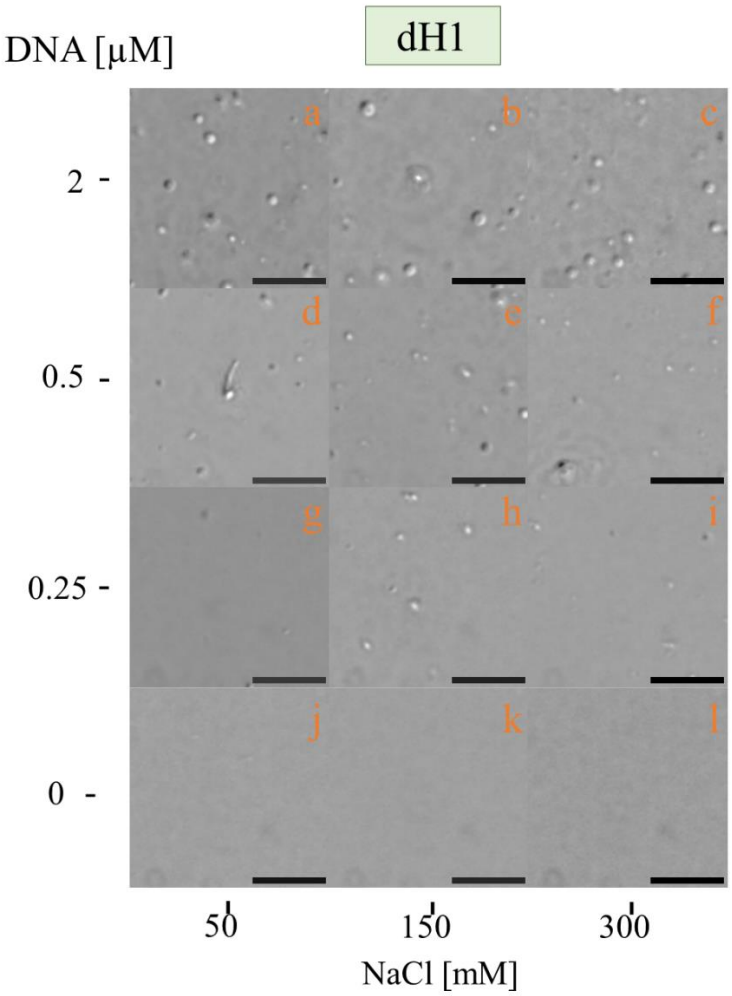


Figure 3.2.8B. Phase separation of dH1 is dependent on DNA. Zoom-in of a central region of phase diagram from **Figure 3.2.8A.** Scale bars correspond to 10 µm.

Thus, we found that there is a direct relation between the DNA content and the shape of the dH1 phase-separated objects and also with the amount of these objects and the DNA concentration (**Figure 3.2.9** left). On the other hand, the size of the objects presented a threshold between 0.25 to 0.5 µM DNA concentrations, potentially related to the spheric phase-separated droplets in contrast to the non-spherical aggregates (**Figure 3.2.9** right). In addition, at the highest NaCl concentration tested (300 mM) an increase in the number of phase-separated elements was observed (**Figure 3.2.9** left)

concomitant with an abrupt reduction in average size (**Figure 3.2.9** right). Remarkably, the 300 mM NaCl and 0.5 µM DNA condition drives a qualitative change in the shape of the phase separated objects, from a mixture of spherical and non-spherical to homogeneous small spherical droplets (**Figure 3.2.8A/B f**).

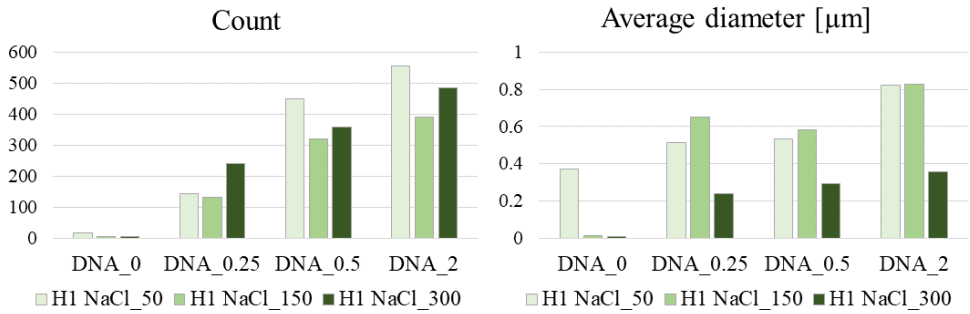


Figure 3.2.9. Count and size analysis of phase separated aggregates and droplets from dH1-DNA phase diagram. Graph bars showing the count (**left**) and the average diameter, in µm (**right**) of phase-separated elements in the different conditions displayed in the phase diagram from **Figure 3.2.8**.

Finally, we wanted to further analyze the change of the non-spherical aggregates when increasing the protein concentration. Therefore, at a concentration of 0.7 µM DNA and 50 mM NaCl we tested increasing protein concentrations and observed that aggregates increased size proportionally to protein concentration, displaying clearly irregular forms (**Figure 3.2.10**).

All in all, we could recapitulate the properties of phase separation described for other somatic linker histones, confirming the need for DNA of dH1 to phase-separate and the critical proportions of dH1-DNA to form aggregates or to phase-separate spherical droplets.

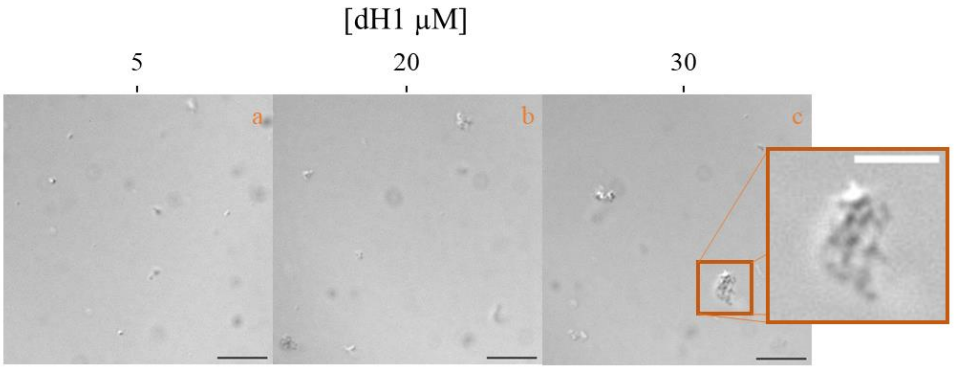


Figure 3.2.10. dH1 aggregates change with high protein concentration. Three DIC images are shown. DNA and NaCl concentrations are constant and 0.7 μM and 50 mM, respectively. Protein concentration is indicated. Images were acquired immediately after mixing. Black scale bars correspond to 10 μm; White scale bar corresponds to 5 μm.

3.2.3.2. Phase separation conditions for BigH1: a linker histone that can phase separate in absence of DNA

Next, we performed the analogous experiments with dBigH1, and generated a phase diagram maintaining the protein concentration of dBigH1 at 10 μM and changing the DNA concentration in a range from 0 to 2 μM, at increasing NaCl concentration (**Figure 3.2.11A/B**). Strikingly, we found that dBigH1 is able to phase separate without DNA (**Figure 3.2.11A/B j**). Moreover, phase-separation driven by dBigH1 – contrary to that driven by dH1 – never presented irregularly shaped aggregates, but it was composed of spherical droplets. Furthermore, compared to dH1, phase separation of dBigH1 was greatly hindered at increasing NaCl concentration (**Figure 3.2.11A/B**). The quantitative information of the count and size of the phase-separated droplets for dBigH1, clearly showed how globally, for all the different DNA concentrations, both parameters decreased at increasing NaCl concentrations (**Figure 3.2.12**). Remarkably, the 300 mM NaCl condition showed almost a complete abolition of BigH1 phase separation (**Figure 3.2.11A/B e**, **f**, **i** and **l** and **Figure 3.2.12**). On the other hand, the 150 mM condition effect on the droplet count was dependent on the DNA content, being more resistant to the increased salt concentration at higher DNA containing conditions, ranging from an almost total abolition in the 0

µM DNA condition to a small reduction at 2 µM DNA (**Figure 3.2.12** left). On the other hand, a strong reduction on the droplet size was observed at increasing NaCl concentration regardless of DNA concentration (**Figure 3.2.12** right).

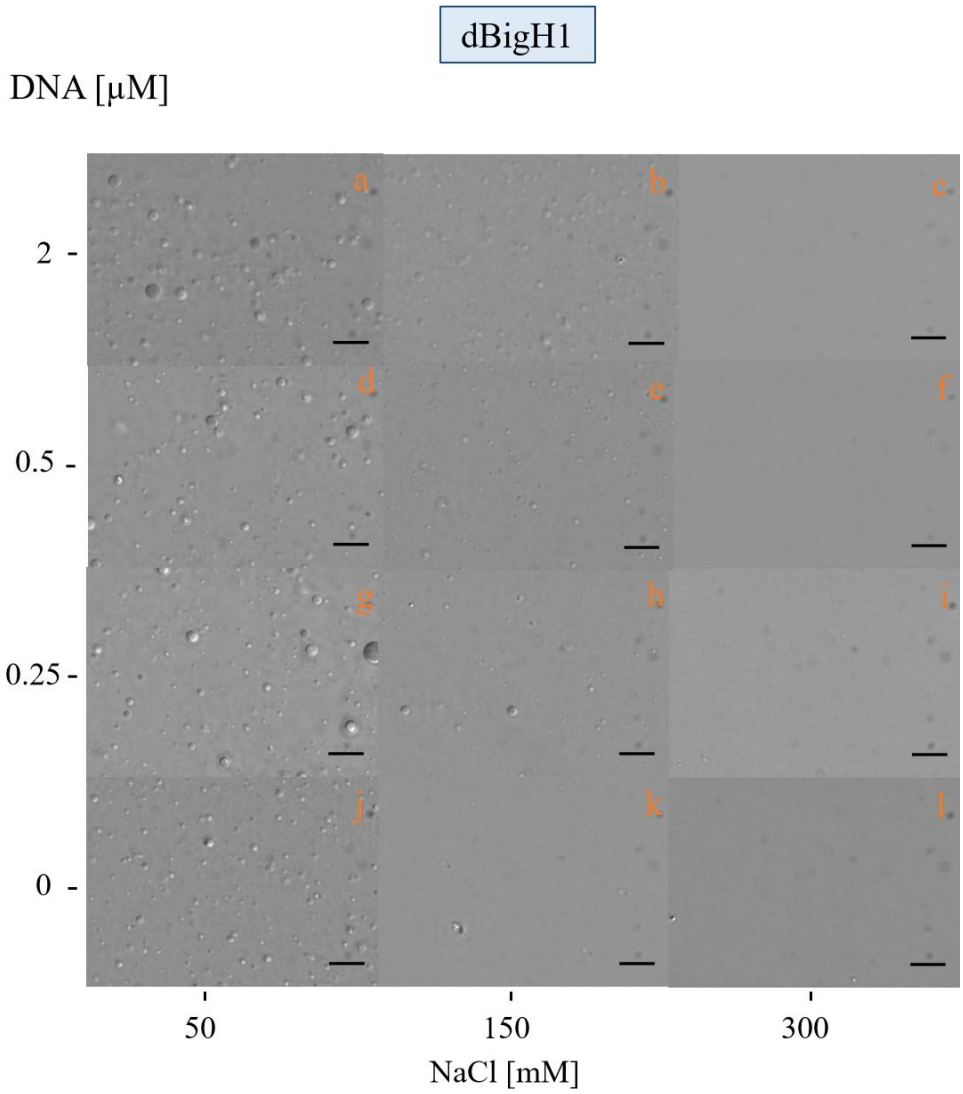


Figure 3.2.11A. dBigH1 phase separation is independent on DNA. A phase diagram for dBigH1-DNA is shown, composed of projections of DIC images from 40µm segments, acquired at 4 µm intervals. dBigH1 concentration is of 10 µM with varying DNA and NaCl concentrations, as indicated. Images were acquired immediately after mixing. Scale bars correspond to 10 µm.

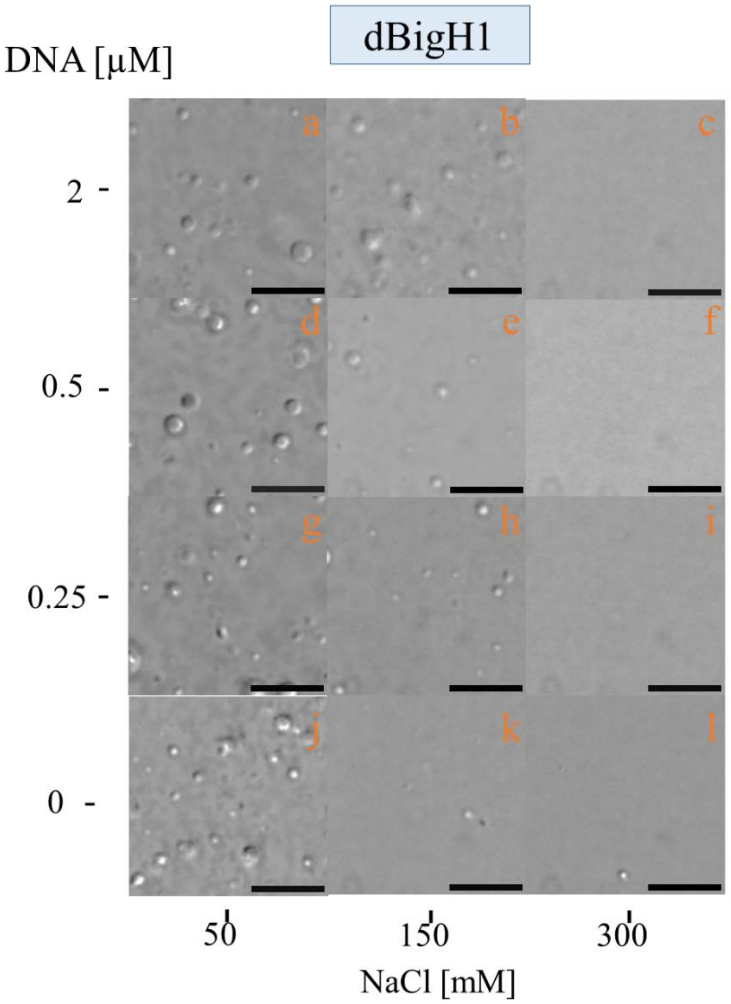


Figure 3.2.11A. dBigH1 phase separation is independent on DNA. Zoom-in of a central region of phase diagram from **Figure 3.2.11A**. Scale bars correspond to 10 μm.

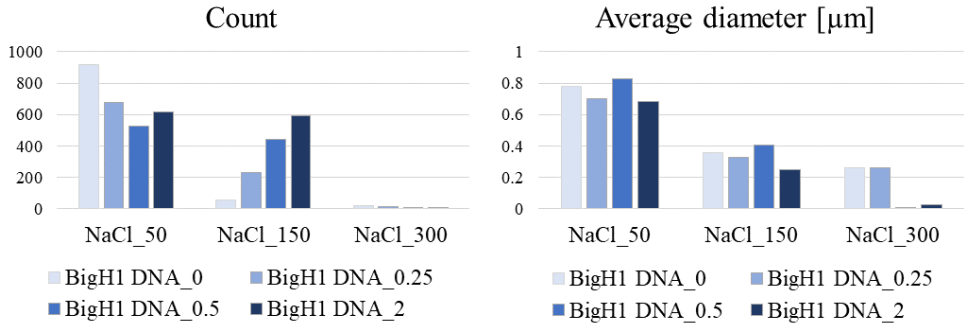


Figure 3.2.12. Count and size analysis of phase separated droplets from dBigH1-DNA phase diagram. Graph bars showing the count (**left**) and the average diameter, in µm (**right**) of phase-separated elements in the different conditions displayed in the phase diagram from **Figure 3.2.11**.

3.2.3.3. Comparison of dBigH1 and dH1 phase diagrams: effect of protein concentration and monovalent salt

To further characterize the differences between dH1 and dBigH1 we performed an additional phase diagram. We decided to use a condition in which dH1 did not present the irregular-shaped aggregates. Thus, in these phase diagrams we maintained the DNA concentration at 2 µM, while changing protein concentration in a range from 5 to 20 µM and increasing NaCl concentrations (**Figure 3.2.13**). We found that both dH1 and dBigH1 phase separation is impaired when increasing ionic strength. At the highest NaCl concentration tested (300mM), we could observe a threshold protein concentration below which there was no phase-separation and above which there was. For dH1 it was between 5 and 10 µM (**Figure 3.2.13a-b**), and for dBigH1 it was between 10 and 20 µM (**Figure 3.2.13k-l**).Therefore, as stated before, the impact that increasing NaCl concentrations had on dBigH1 count and average size was stronger than on dH1 (**Figure 3.2.14**), showing that dH1 phase-separation, although affected, was more resistant to increasing salt concentration.

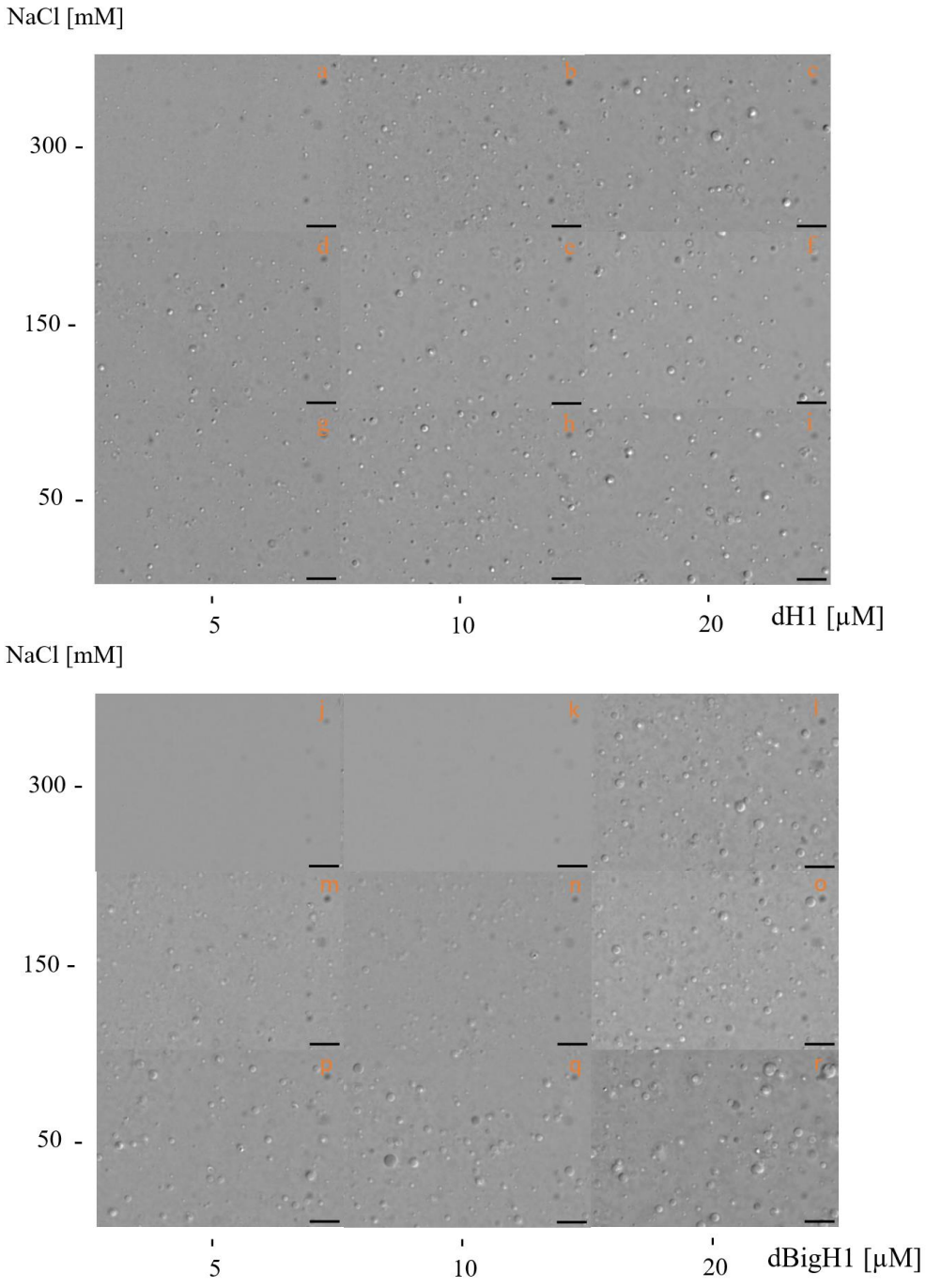


Figure 3.2.13. Phase diagram for dH1-NaCl and dBigH1-NaCl. Phase diagrams for dH1-NaCl (top) dBigH1-NaCl (bottom) are shown, composed of projections of DIC images from 40 µm segments, acquired at 4 µm intervals. DNA concentration is constant of 2 µM. Images were acquired immediately after mixing. Scale bars correspond to 10 µm.

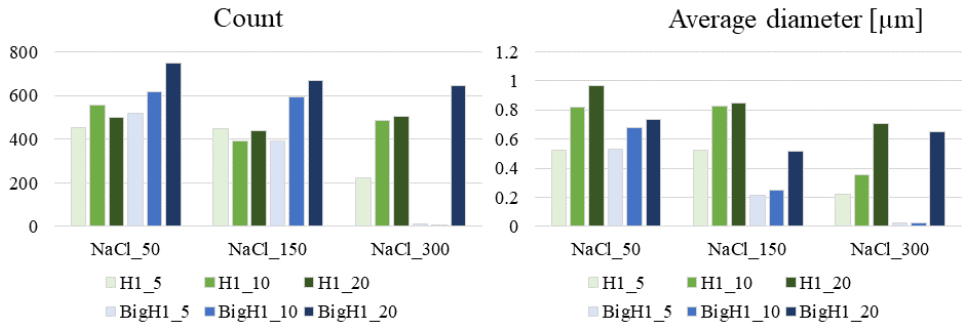


Figure 3.2.14. Count and size analysis of phase separated droplets from dH1-NaCl and dBigH1-NaCl phase diagrams. Graph bars showing the count (**left**) and the average diameter, in µm (**right**) of phase-separated elements in the different conditions displayed in the phase diagram from **Figure 3.2.13**.

3.2.3.3. Fundamental difference: the phase separation dynamics of dH1 and dBigH1

Finally, we addressed the evolution of the samples over time, and the potential nature of the phase separation for dH1 and dBigH1. To do so, we selected a condition of high protein and DNA concentration, and low NaCl concentration, to promote droplet formation for each protein. Thus, we used 20µM dH1 or dBigH1, 5µM DNA and 50mM NaCl, and observed the samples immediately after preparation and every 30 minutes up to two hours (**Figure 3.2.15**). Remarkably, we found that droplets generated by dBigH1 tended to fuse giving rise to bigger spherical droplets, like liquid-like droplets would. On the other hand, dH1 droplets were not able to completely fuse, giving rise to droplets that partially fused together but that did not recover the spherical fused form (**Figure 3.2.15**). This observation was further confirmed by the quantification of the sizes of those droplets (**Figure 3.2.16**). We observed that at t = 0 the dispersion of sizes in BigH1 was wider despite having a very close mean to dH1. Moreover, over time, the size of dH1 phase-separated objects only slightly increased.

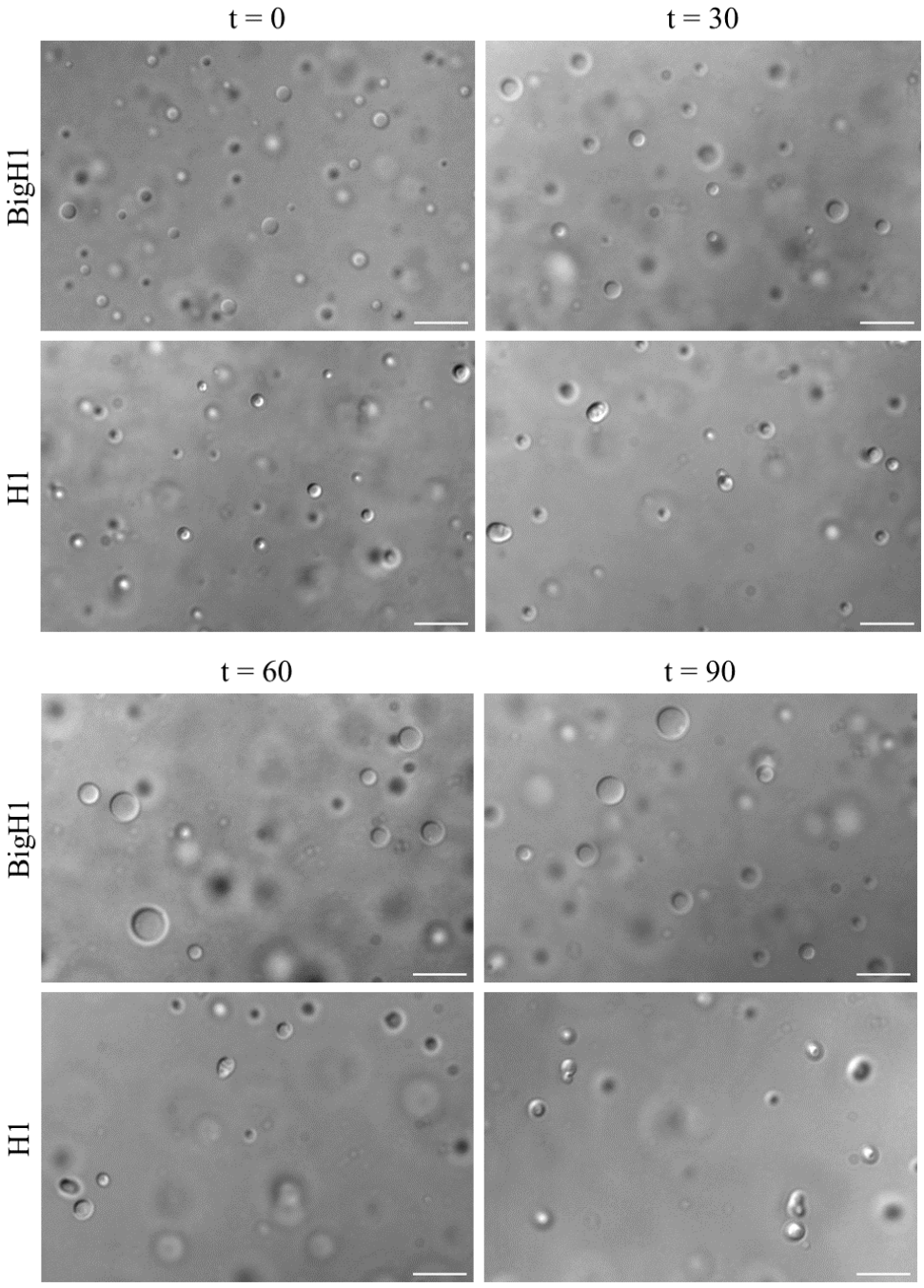


Figure 3.2.15. Sample evolution for dH1 and dBigH1. DIC images showing the dBigH1 (top) and dH1 (bottom) phase-separated objects in different indicated time points. Scale bars correspond to 10 μm.

On the other hand, dBigH1 size linearly increased over time. We also observed that in 120 minutes there was an apparent reduction in size (**Figure 3.2.16**). However, this was a because the biggest droplets reached such a big size that sedimented and did not reflect an actual reduction of the droplet size in the sample.

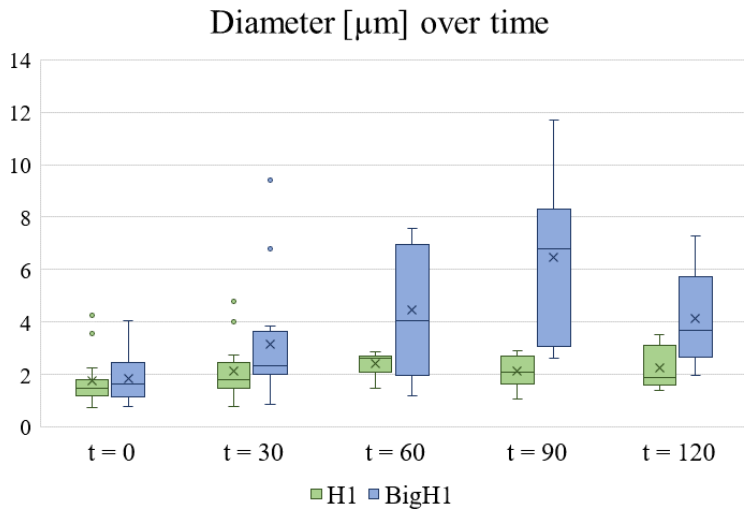


Figure 3.2.16. Size analysis of the sample evolution for dH1 and dBigH1. Box plots showing the diameter, in µm of phase-separated objects over time for dH1 (green) and dBigH1 (blue).

Overall, the difference in their sample evolution poses a critical difference to the nature of the phase separation condensates that dH1 and dBigH1 give rise to.

4. DISCUSSION

4.1.1. cRNAs ARE STRUCTURAL ELEMENTS OF CHROMATIN

The eukaryotic cell nucleus contains a large number of nucleic acids, including RNA molecules. It is well-established that a portion of this RNA is soluble in the nucleoplasm, including mRNA, rRNA, and smaller soluble RNAs. Additionally, another fraction of the RNA is somehow attached to chromatin, forming chromosomal RNA (cRNA). In the first part of this thesis, we have provided a complete catalogue of the cRNAs in *D. melanogaster* S2 cells, showing that they are present across the genome. The relevance of the cRNAs, at the structural level, has already been shown (Rodriguez-Campos & Azorin 2007; Thakur & Henikoff 2020). We found that ~28% of the *D. melanogaster* genome was covered with cRNA. It had been previously reported that cRNA account for 2%-5% of chromatin nucleic acids (Rodriguez-Campos & Azorin 2007). Nevertheless, this does not come into conflict with our result since in that study the sample preparation lead to include only the cRNAs very strongly associated to chromatin and were not nascent transcripts. On the other hand, in our study, we found that the most represented category is the one that corresponds to pre-mRNAs. Of note, the large abundance of pre-mRNAs in the cRNAs, due to their very active transcription, had already been reported (Li & Fu 2019). Moreover, a recent study performed in S2 cells, which fractionated the cRNA using a different methodology than ours, obtained similar results regarding the pre-mRNA fraction in the cRNAs (Planells et al. 2023). Overall, these differences point to the idea that different classes of cRNAs have different affinities and residence times in chromatin.

4.1.1.2. Heterochromatic cRNAs are enriched in chromatin and packaged by hrp36 and hrp48

We have found that, surprisingly, cRNAs are specially enriched at heterochromatic regions. In addition, they show a bias towards the highest quantile of expression when compared to euchromatic cRNAs, possibly pointing out to a higher stability and/or slower turn-over of heterochromatic cRNAs. In this regard, it has been described for several model organisms –

including *S. cerevisiae*, *D. melanogaster*, and *M. musculus* – that the establishment of heterochromatin is mediated by those repetitive heterochromatic transcripts (Alekseyenko et al. 2014; Allshire & Madhani 2018; Johnson et al. 2017; Probst et al. 2010). Moreover, the specific structural contribution of those heterochromatic cRNAs to the structure of *M. musculus* pericentromeric heterochromatin has been reported (Thakur et al. 2019), further validating the existence and relevance of heterochromatic cRNAs and suggesting that they are a general feature of metazoa.

Furthermore, we addressed the packaging of heterochromatic cRNAs in chromatin and showed that heterochromatic cRNAs are packaged by hrp36 and hrp48. The idea that cellular RNA must be necessarily packaged with hnRNPs to be properly processed has been strongly validated in the context of pre-mRNA (Dreyfuss et al. 2002). Thus, we hypothesized that this feature could be extensible to all cRNAs. We addressed this point by studying the *D. melanogaster* hnRNPs hrp36 and hrp48, since there was evidence of their localization on chromatin (Matunis et al. 1993), as well as for their mammalian orthologs (hnRNPA/B family) (Han et al. 2010). Moreover, hrp36 and hrp48 were reported to localize in heterochromatin as well (Matunis et al. 1993; Piacentini et al. 2009). We have confirmed that their association with chromatin is mediated by RNA. Furthermore, hrp36 and hrp48 are packaging heterochromatic cRNAs, since both extensively overlap with cRNAs that are likely retained on chromatin in *cis*. The relative enrichment of overlapping versus non overlapping cRNAs might indicate a higher stability and/or slower turnover of these cRNAs.

Previously described functions for hrp36 were telomere maintenance (Singh & Lakhotia 2016); heterochromatin formation (Piacentini et al. 2009); development and stress response as a core constituent of omega speckles (Singh & Lakhotia 2012); and modulating the alternative splicing of the pre-mRNA of Prospero, a transcription factor involved in neuronal differentiation and axonal outgrowth (Borah et al. 2009). On the other hand, hrp48 has been involved in X-chromosome dosage compensation by inhibiting msl-2 translation (Szostak et al. 2018); modulating notch

signaling (Suissa et al. 2010), repressing oskar mRNA (a posterior determinant during embryogenesis) during transport, restricting oskar activity to the posterior part of the oocyte (Yano et al. 2004); and in modulation of splicing (Blanchette et al. 2005). With our work, we have reported a novel and shared function for hrp36 and hrp48. Interestingly, their targets on chromatin are different of those associated to their soluble-RNA targets, since on chromatin they are specifically enriched in non-coding heterochromatic regions. However, despite their strong colocalization, the association of hrp36 with chromatin is independent of the association of hrp48 with chromatin. Taking into consideration their small differences in localization and the different phenotype regarding cRNAs, and specifically R-loops accumulation, it appears as if hrp48 has a stronger effect on heterochromatic cRNAs homeostasis. However, we cannot ensure that hrp36 milder effects are not a consequence of partial redundancy with other hnRNPs. In fact, it is likely that other hnRNPs are also involved in packaging the same and/or additional cRNAs, since hrp36 and hrp48 do not account for the packaging of all heterochromatic cRNAs. This is indicated by cRNA-containing regions that are not covered by these two proteins.

Overall, our results paint an intriguing picture of heterochromatin that is enriched with cRNAs packaged by hrp36 and hrp48

4.1.2. dH1 IS FUNDAMENTAL TO MAINTAIN HETEROCHROMATIN INTEGRITY

4.1.2.1. dH1 prevents R-loop accumulation by facilitating heterochromatic cRNA packaging by hrp36 and hrp48

Our results suggest that dH1 is a modulator of the packaging of heterochromatic cRNAs. A ~50% dH1 reduction in dH1KD is capable to reduce the binding of hrp36 and hrp48 to chromatin to a similar extent,

while not altering their mRNA levels, showing the need of dH1 for hrp36 and hrp48 to package heterochromatic cRNAs.

Moreover, dH1KD leads to the accumulation of heterochromatic cRNAs. In this regard, it has been confirmed that depletion of H1, both in *D. melanogaster* and in *H. sapiens* cells, leads to the deregulation and aberrant expression of repetitive heterochromatic sequences (Izquierdo-Bouldstridge et al. 2017; Vujatovic et al. 2012). Moreover, cRNA changes in *M. musculus* cells have been addressed, and triple-H1knockout embryo derived cells show a great accumulation of lncRNAs and coding cRNAs. Yet, when tested in KD cells they only recapitulate mild lncRNA accumulation (Fernández-Justel et al. 2022). However, while our results also demonstrate a mild accumulation of lncRNAs in dH1KD cells, the conservation of heterochromatic cRNAs change in *M. musculus* cells cannot be fully compared since studies in mice did not analyze heterochromatic regions.

Nevertheless, heterochromatic cRNA accumulation in dH1KD is in complete agreement with the previously reported R-loop accumulation in heterochromatin in dH1 depleted conditions (Bayona-Feliu et al. 2017). We have found that hrp36KD and hrp48KD present the same phenotype of R-loop accumulation in heterochromatin, indicating that R-loop formation in dH1KD is, at least partially, the consequence of the impaired packaging of cRNAs by hrp36 and hrp48. In accordance, in *H. sapiens* cells hnRNPs have been established as R-loop preventing factors, since a genome-wide siRNA screen identified several hnRNPs that lead to accumulation of γH2AX (H2AX-phosphorylation, mark for DSB), and that phenotype was rescued by overexpression of RNAse H (García-Muse & Aguilera 2019; Paulsen et al. 2009).

To further validate our results, other members from our laboratory performed immunostainings in polytene chromosomes using control flies (w^{1118}), dH1KD ($dH1^{RNAi}$), hrp36KD ($hrp36^{RNAi}$) and hrp48KD ($hrp46^{RNAi}$) flies. They confirmed the phenotype of loss of hrp36 and hrp48 from

polytene chromosomes in dH1-depleted conditions, and the accumulation of R-loops in the chromocenter (heterochromatin) in hrp36KD and hrp48KD (**Figure 4.1.1**), highlighting the relevance of this results also in the whole organism, not exclusively in cells.

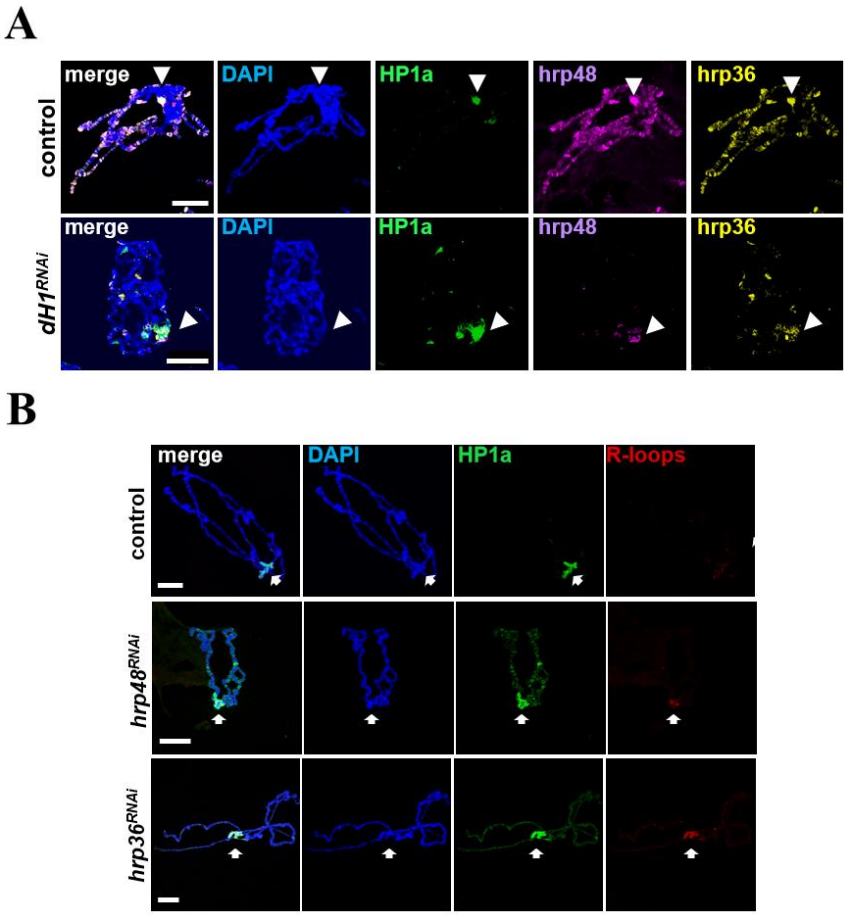


Figure 4.1.1. Phenotypes in cells are recapitulated in polytene chromosomes from flies. (**A**) Immunostainings of polytene chromosomes from control w^{1118} (top) and dH1KD ($dH1^{RNAi}$) (bottom) flies with α-hrp36 (in yellow) and α-hrp48 (in magenta) antibodies. (**B**) Immunostainings with S.9.6 antibodies to detect R-loops (in red) of polytene chromosomes from control w1118 (top), hrp48KD ($hrp48^{RNAi}$) (center) and hrp36KD ($hrp36^{RNAi}$) (bottom) flies. (**A** and **B**) Immunostaining with α-HP1a antibodies (in green) is shown to identify the heterochromatic chromocenter (arrowheads). DNA was stained with DAPI. Scale bars correspond to 20μm.

Interestingly, regarding the molecular link between dH1, hrp36 and hrp48, co-immunoprecipitation of H1 with hrp36 and hrp48 has been reported (Ni et al. 2006) and confirmed in our group (Casas-Lamesa 2018).

On the other hand, the study of cRNAs in *M. musculus* cells also reported that reduced levels of H1 lead to reduced levels of the RNA modification m6A (Fernández-Justel et al. 2022). Moreover, m6A has been reported to modulate R-loops formation and their stability (Abakir et al. 2020; Yang et al. 2019), at least in part by facilitating the binding of hnRNPs to their target RNAs (Alarcón et al. 2015b,a). Thus, it could potentially be a decreased level of m6A in H1-depleted conditions what prevented the binding of hrp36 and hrp48 to the heterochromatic cRNAs and promoted the formation of R-loops. However, further work is needed to elucidate whether a direct contact between dH1 and hrp36/hrp48 or the alteration of m6A is mediating the dependence of hrp36 and hrp48 for dH1 to package the cRNAs.

Additionally, we characterized that the R-loops that accumulate in hrp36KD and in hrp48KD colocalize with WT cRNAs. This fact implies that, as established, heterochromatic cRNAs are present in normal conditions and when they are not packaged by hrp36 and hrp48, they are more frequently retained on chromatin and give rise to the formation of R-loops. Since R-loops are mostly described to form in *cis* (Li & Fu 2019) the colocalization of those R-loops with the cRNAs, again indicates that the heterochromatic cRNAs that we have studied are present on chromatin mostly in *cis*.

4.1.2.2. dH1 maintains heterochromatin compaction

Histone H1 is a structural component of chromatin. Thus, in addition to elucidating the effects of H1 depletion on the packaging of heterochromatic cRNAs, we also analyzed the structural changes of chromatin in H1-depleted conditions.

Using ATAC-seq, we identified two nucleosomal sizes (of 151 and 138bp) and a subnucleosomal size (of 125bp). Subnucleosomal fragments are probably associated to further digested nucleosomal fragments. Previous studies have reported that a higher A/T content correlates with increased

digestion levels within the same timeframe, and similar fragments to those we have identified were observed in S2 cells using MNase digestion (Chereji et al. 2019). Our results show that depletion of dH1 leads to a global reduction of nucleosome occupancy, in agreement with the fact that dH1 stabilizes nucleosomes (Fyodorov et al. 2018). Moreover, concomitant with occupancy reduction, destabilization of the nucleosomes is probably the cause of the increased digestion of dH1KD samples.

Regarding euchromatin accessibility, our analysis revealed that the canonical image of an accessible region upstream the TSS (Bai & Morozov 2010) is perturbed when dH1 is removed. In dH1KD the existing accessible regions upstream of the TSS became broader compared to the control conditions. Consequently, we observed mild phenotypes in euchromatin. These included a reduction of euchromatic cRNAs and an increased intron retention in dH1KD, and a decrease in euchromatic R-loops in hrp36KD and hrp48KD. These events are potentially related to the destabilization of an already accessible region. Nonetheless, further work is necessary to examine the impact of H1 depletion in euchromatin.

The association of dH1 with chromatin compaction have been long known (Fyodorov et al. 2018; Perišić et al. 2019; Woodcock et al. 2006). Nevertheless, we found that, in dH1KD, heterochromatin is showing the greatest change in accessibility. Similar results were reported in *M. musculus* cells, that after H1 depletion showed the greatest change in accessibility in facultative and constitutive heterochromatin (Willcockson et al. 2021). On the other hand, a study performing KD of multiple H1 variants in tumoral human breast cancer cells displayed changes in accessibility not only restricted to heterochromatic regions (Serna-Pujol et al. 2022). The reason for these differences could be that in the latter study they were able to directly reduce H1 variants mainly found in euchromatin. In addition, tumor cells already have lower levels of H1 than WT cells that result in less well-formed heterochromatin and overall, less compacted chromatin (Scaffidi 2016).

4.1.3. THE LINKER HISTONE DH1 IS KEY FOR THE HOMEOSTASIS OF CRNAS AND HETEROCHROMATIN INTEGRITY

Taking all the results in consideration, we propose a model in which at WT conditions heterochromatin is well formed with the linker histone present and the nucleosomes well-spaced, containing heterochromatic cRNAs, that are packaged by hrp36 and hrp48 and properly processed (**Figure 4.1.2** Top).

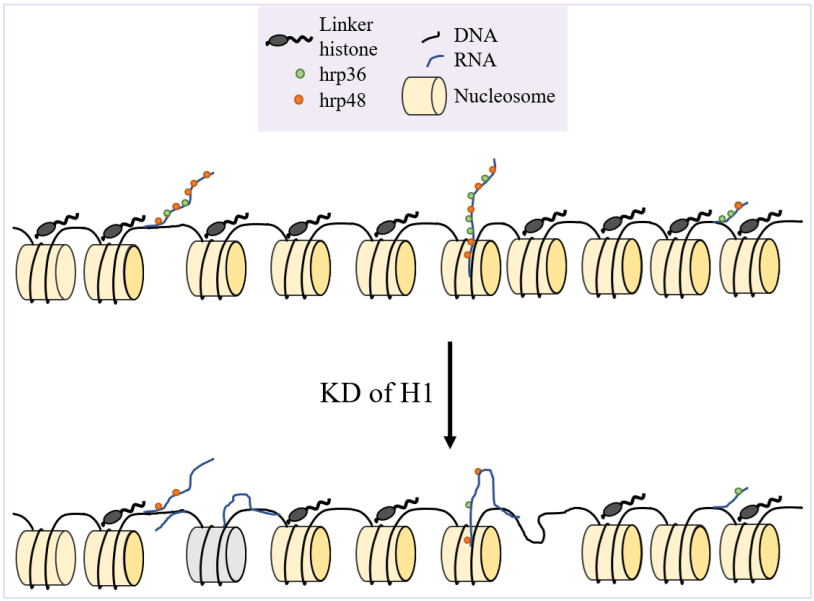


Figure 4.1.2. Model for the contribution of linker histone dH1 to the accumulation of cRNAs and the formation of R-loops in heterochromatin. Schematic representation of heterochromatin in WT conditions (top), with well-spaced nucleosomes, and heterochromatic cRNAs packaged by the hrp36 and hrp48. Upon KD of dH1 (bottom), nucleosomes are destabilized, heterochromatic cRNAs are accumulated and poorly packaged by hrp36 and hrp48. The increased accessibility and unpackaging of the cRNAs promotes the formation of R-loops.

Next, in parallel to the absence of H1 resulting in chromatin destabilization and increased accessibility, particularly in heterochromatic regions, we observe that hrp36 and hrp48 are unable to package cRNAs. This leads to the accumulation of heterochromatic cRNAs and facilitates the formation

of unscheduled R-loops. (**Figure 4.1.2** Bottom). Indeed, increased accessibility has been linked to promote R-loop formation (García-Muse & Aguilera 2019). Thus, the structural phenotype is probably also contributing to the cRNAs retention in heterochromatin.

The fact that DNA:RNA hybrids are more stable than dsDNA (Thomas et al. 1976) and that there is saturation of the cellular mechanisms to resolve the unscheduled accumulated R-loops lead to conflicts with the RF (Gan et al. 2011; García-Muse & Aguilera 2019), that end up giving rise to DSB, genomic instability and cell death. Interestingly, a complete disruption of heterochromatin in *C. elegans*, mediated by the KO of the H3K9 histone methyltransferases, also leads R-loop accumulation (Zeller et al. 2016). However, *C. elegans* heterochromatin exhibits remarkable differences, both structurally and molecularly, that make very complicated to establish a clear parallelism.

Overall, the role of dH1 is unveiled as a dual element, that, on one hand, maintains the structural integrity of heterochromatin and, on the other hand, promotes the proper packaging and processing of the heterochromatic cRNAs.

Regarding the homeostasis of cRNAs, a recent study has shown that a direct destabilization of the correct processing of cRNAs, by impairing the action of the RNA exosome, aside from having many consequences in euchromatin, also presents a similar phenotype of accumulation of cRNAs and increased accessibility of pericentromeric heterochromatin (Planells et al. 2023). Remarkably, the exosome is only marginally present on heterochromatin, while heavily present in euchromatin (Piacentini et al. 2009). Thus, these results confirm that cRNAs are especially relevant for the integrity of the pericentromeric heterochromatin and that, despite having a slower turnover than actively transcribed genes, they also need to be synthesized and degraded. Moreover, the fact that they also observe accessibility changes poses the perspective of a positive-feedback between the increased accessibility and cRNA accumulation (Planells et al. 2023),

as anticipated in our model. However, a key difference in the KD of the exosome and the KD of dH1 is the specificity of the accumulation of cRNA and changes in accessibility on heterochromatin in our context. This suggests that modulation of heterochromatic cRNAs by dH1 could be somehow specific.

The biased effects towards heterochromatin that we observe could be a consequence of the limited KD we have (~50%). We can not rule out the possibility that higher levels of reduction or the capacity to specifically affect the euchromatic H1 content would have greater consequences for euchromatin. Furthermore, it could also be a consequence of the fact that heterochromatin is enriched in H1 compared to actively transcribed regions (AeRi & Ann 2003; Climent-Cantó et al. 2020; Kharchenko et al. 2011; Weintraub 1984; Willcockson et al. 2021). In this regard, another example of the differential consequence of dH1KD in euchromatin and heterochromatin is the change in NRL. The NRL of heterochromatin suffers a reduction in dH1KD, while the NRL of actively transcribed genes (which is already smaller in normal conditions) remains unaltered (Baldi et al. 2018). On the other hand, it is also possible that the localized effects are associated to a specific dH1 PTM. Regarding this question, dH1K27me2 has been reported to be necessary for heterochromatin genomic stability and prevention of R-loop accumulation in *D. melanogaster* (Bernués et al. 2022). It is tempting to connect this phenotype to the molecular mechanism herein described. In any case, further work will be needed to specifically study this link.

In summary, in the first part of the thesis we have shown that the somatic variant of *D. melanogaster*, dH1, is key to maintain the homeostasis of cRNAs, promote their packaging with hrp36 and hrp48, and avoid their accumulation in heterochromatin, thus preventing the formation of deleterious R-loops.

4.2.1. dH1 AND dBigH1 STRUCTURAL DIFFERENCES HAVE BIOLOGICAL CONSEQUENCES

Our work has shown that, average NRL changed upon dBigH1 expression in S2 cells, but not upon dBigH1ΔED expression, suggesting that the acidic domain of dBigH1 leads to a structural change in chromatin. Upon induction, dBigH1 represented ~23% of total linker histones. Thus, the change we have observed is remarkable and reinforces the idea that even small quantities of dBigH1 impact global structural changes in chromatin. The functional consequences of the expression of dBigH1 had been further characterized by another member from our laboratory (Climent-Cantó et al. 2020). And, indeed, this change has biological consequences such as in the chromatin acetylation levels and in transcription.

It must be taken into consideration, that the bipolar structure of dBigH1 is specific of *Drosophila*. However, it is a conserved feature that embryonic H1s are generally more acidic than somatic ones (Pérez-Montero et al. 2016). Moreover, it has been reported that other mammalian embryonic linker histones also alter chromatin organization and dynamics (Becker et al. 2005; Saeki et al. 2005; Ura et al. 1996). Thus, despite its peculiar charge distribution, the study of the differences between the somatic and the embryonic linker histones of *D. melanogaster* could bring insight into conserved roles in metazoa.

4.2.2. THE PHASE-SEPARATION PROPERTIES OF dH1 AND dBigH1 PRESENT FUNDAMENTAL DIFFERENCES

Our results have shown that *D. melanogaster* dH1 behaves similarly to mammalian somatic linker histones regarding its capacity to form condensates in presence of DNA but not in its absence (Mimura et al. 2021; Shakya et al. 2020). Moreover, we also recapitulated the formation of non-spherical assemblies at lower DNA concentrations. However, conditions where dH1 gives rise to phase-separated condensates present differences

with those previously described, since the dH1-DNA condensates are not typically liquid-like. Nevertheless, as we have not performed FRAP (Fluorescence recovery after photobleaching) studies we cannot confirm whether there is diffusion of material inside the droplets, even though we observed that these droplets are not able to fuse. In this regard, previous studies showed that even the non-spherical assemblies, generated with a somatic linker histone H1 and ssDNA, presented recovery of fluorescence in FRAP experiments, showing that those phase-separated condensates are not completely solid (Mimura et al. 2021). Overall, this difference in the dynamism of the phase-separated condensates that we obtained compared to those observed in previous studies could be due to the fact that all those studies used a linker histone purified from bovine thymus while we worked with a recombinant dH1. In this regard, linker histones are described to present a great abundance and variety of PTMs (Andrés et al. 2020; Wiśniewski et al. 2007). In particular, dH1 has also been reported to present extensive PTMs including phosphorylations, methylations, acetylations and ubiquitinations (Bonet-Costa et al. 2012), while acetylation and phosphorylation have been reported to alter LLPS properties (Monahan et al. 2017; Saito et al. 2019). However, even in the LLPS contexts previously reported, linker histone H1 has been described as an element to reduce fluidity of reconstituted chromatin droplets in a concentration dependent manner (Gibson et al. 2019), as well as of ssDNA-containing droplets (Mimura et al. 2021).

Herein, we have addressed for the first time the phase-separation capacity of an embryonic H1 variant. Remarkably, we found that its characteristics are clearly different to those of the somatic variant. dBigH1 has the capacity to phase separate even in the absence of nucleic acids. Moreover, the phase-separated droplets formed by dBigH1 either with or without DNA are highly dynamic and present many features that characterize them as LLPS. This suggests that, while dH1 leads to less fluid more static compartments, dBigH1 could lead to more dynamic states of chromatin. About this point, it has been reported that aside from heterochromatin, transcription is also,

to some extent, associated with compartmentalization and compaction of the actively transcribed loci (Chen et al. 2018; Di Stefano et al. 2020; Germier et al. 2017). Moreover, dBigH1 has been described to lead to transcriptional repression *in vitro* and *in vivo* (Climent-Cantó et al. 2021). Thus, the phase-separation features observed could be involved in the modulation of the chromatin dynamics and the transcriptional output. On the other hand, in the case of heterochromatin, H1-mediated compaction is associated with silencing the region rather than promoting transcription. However, the embryonic contexts where dBigH1 is present do not exhibit a well-defined heterochromatin structure like somatic cells. In this regard, it has been observed that the three-dimensional (3D) structure of the genome is significantly different during the initial stages of embryonic development, when only dBigH1 is present, compared to the final stages, when dBigH1 has been replaced by dH1. The difference consists in the absence of well-defined topologically associated domains (TADs) and the undefined 3D structure of the dBigH1-containing embryos compared to the dH1-containing stages, that resemble a compartmentalized somatic 3D genomic structure. Notably, this change in structure is not influenced by transcription (Hug et al. 2017). Thus, the differential content of dH1 and dBigH1 and how they interact with chromatin may account for the observed changes. Experiments on the effect of dH1 and dBigH1 on chromatin-driven phase-separation are scheduled to be performed soon to directly address these questions.

Finally, as anticipated by the charged nature of dH1 and dBigH1 their phase separation properties are driven by electrostatic interactions. While in both cases phase-separation events are negatively affected by increasing ionic strength, our results have shown that BigH1 presents a much higher sensitivity than dH1, as it would be expected from its higher content in charged residues.

Overall, our results have shown that dH1 and dBigH1 have different phase-separation properties: dBigH1 is able to phase separate in absence of nucleic acids, while dH1 can not, and promotes phase-separated droplets much

more fluid-like and dynamic than dH1. Thus, these clearly differentiated properties give rise to the inference that their respective impacts on chromatin should differ. Furthermore, such distinctions hold the potential to enhance our comprehension of the disparities observed between the embryonic and post-embryonic states, as well as the transitional stage where both isoforms coexist.

5. CONCLUSIONS

1. cRNAs are enriched in heterochromatin in S2 cells.

2. hrp36 and hrp48 are present on chromatin where they largely overlap. Their binding to chromatin is mutually independent, tethered through RNA and enriched in heterochromatin.

3. hrp36 and hrp48 are packaging heterochromatic cRNAs.

4. Binding of hrp36 and hrp48 to cRNAs is dependent on H1.

5. Depletion of H1 in S2 cells leads to accumulation of heterochromatic cRNAs, accompanied by a slight reduction of cRNAs corresponding to actively transcribed regions and an increase in intron retention.

6. Depletion of hrp48 in S2 cells leads to an accumulation of heterochromatic cRNAs that is not observed with depletion of hrp36. In both cases depletion results in an increase in R-loops in heterochromatin.

7. Depletion of H1 in S2 cells leads to a reduced nucleosome occupancy concomitant to a global decompaction of chromatin, and to a relevant increased accessibility of heterochromatin.

8. The acidic NTD of BigH1 causes a reduction in the NRL in S2 cells.

9. dH1 behaves similarly to other somatic linker histones and form condensates *in vitro* only in presence of DNA. The relative proportion of DNA and dH1 decides the formation of irregular aggregates or spherical phase-separated condensates.

10. dBigH1 has the ability to form condensates *in vitro*, either with or without nucleic acids; these condensates undergo LLPS and exhibit higher sensitivity to increasing ionic strength than dH1 condensates.

6. MATERIALS AND METHODS

6.1. MATERIALS

6.1.1. DNA CONSTRUCTS

For dsLacZ production, we used pBSsk (pBluescriptSK, Promega) plasmid.

For dsH1 production, we used a pMultiBac-1xHis repeat containing one copy of the HIS locus repeat (kindly provided by Dr Duronio).

For protein expression, initial constructs of pET30a containing the coding sequence for BigH1 and pET29b containing the coding sequence for H1 previously obtained by Dr. Eva Satovic were modified by directed mutagenesis.

6.1.2. OLIGONUCLEOTIDES

The list of the oligonucleotides used during this thesis is presented in the following tables.

6.1.2.1. Primers to synthesize PCR templates for dsRNA production

The T7 RNA polymerase promoter is underlined.

Name	Sequence (5'-3')
T7LacZ-F	TAATACGACTCACTATAGGGATGACCATGATTACGCCAAGC
T7LacZ-R	TAATACGACTCACTATAGGGCAATTTCCATTCGCCATTCAG
f-H1	TAATACGACTCACTATAGGGGATGTCTGATTCTGCAGTTGC
r-H1	TAATACGACTCACTATAGGGGGGCTTCGACTTTATGATTCC
T7hnRNP36-F	TAATACGACTCACTATAGGGAGACCAACGGAAACTACGACGAT
T7hnRNP36-R	TAATACGACTCACTATAGGGAGATGCTTGGCAATAGCCTTCTT
T7hnRNP48-F	TAATACGACTCACTATAGGGAGAACGAGAGGGGCAAACTTTTT
T7hnRNP48-R	TAATACGACTCACTATAGGGAGAGCGGGACTTCTTCTTCTCCT

6.1.2.2. Primers for RT-qPCR

Name	Sequence (5'-3')
Rpl32_Fw	GCTAAGCTGTCGCACAAATG
Rpl32_Rv	ACCAAGGAACTTCTTGAATCCG
Hrp36_all_isoforms_Fw	GATGGCCTGAAGGCTCACTT
Hrp36_all_isoforms_Rv	GCATTCTGCGCATTGTCGAT
Hrp48_all_isoforms_Fw	ACATGCCACCTAACTCTGCC
Hrp48_all_isoforms_Rv	GCCGTAGTCGTACTCAGAGC

6.1.2.3. Primers for directed mutagenesis

Name	Sequence (5'-3')
H1_trom_Forward	AGAGGTTCCCAGCTTGCGGCCGCACTCG
H1_trom_Reverse	TGGGACAAGGGCCTTTTTGGCAGCCGTAG
BigH1_plasm_Forward	ATGAAACTAAAGCCGGTTGAAC
BigH1_plasm_Reverse	AGAACCGCGTGGCACCAG

6.1.2.4. Primers for dsDNA production

Name	Sequence (5'-3')
H281_F	AATGTCGCTGTGTTGTTGCT
H281_R	CCAGAACACCATCAACACCC

Fragment amplified (281bp):

AATGTCGCTGTGTTGTTGCTGTGCGAAAGAGAGAATCGTACTGATCACCGA
CGCGATGCAGGCAGCCGGGATGCCGGATGGTCGCTATACGTTATGTGGCGAA
GAAGTGCAGATGCACGGTGGCGTTGTCCGTACCGCGTCCGGTGGGCTGGCGG
GCAGTACGCTGTCTGTTGATGCGGCAGTGCGCAACATGGTCGAGTTGACGGG
CGTAACGCCTGCGGAAGCCATTCATGGTGTTCTGTATGGCGTCGCTGCATCCG
GCGCGAATGCT**GGGTGTTGATGGTGTTCTGG**

6.1.3. ANTIBODIES

The primary and secondary antibodies used for western blot, ChIP and DRIP experiments are presented in the following tables.

6.1.3.1. Primary antibodies

Name	Specie	Clonality	Experimental condition	Source
αdH1	Rabbit	Polyclonal	WB: 1/10,000	Provided by Dr Kadonaga (Vujatovic et al. 2012)
αhrp36	Mouse	Monoclonal	WB: 1/400 ChIP: 4 µg	Provided by Dr Saumweber (Hovemann et al. 2000)
αhrp48	Rabbit	Polyclonal	WB: 1/5,000 ChIP: 6 µL	Provided by Dr Gebauer (Szostak et al. 2018)
αH4	Rabbit	Polyclonal	WB: 1/1,000	Abcam (ab10158)
αS9.6	Mouse	Monoclonal	DRIP: 10 µg	(Boguslawski et al. 1986)
αRNAPolII	Mouse	Monoclonal	ChIP: 5 µL	BioLegend (664906)
αdBigH1	Rabbit	Polyclonal	WB: 1/10,000	Home made (Pérez-Montero et al. 2013)

6.1.3.2. Secondary antibodies

Name	Dilution	Source
Peroxidase AffiniPure Donkey αMouse IgG	WB: 1/10000	Jackson, 715-035-150
Peroxidase AffiniPure Goat αRabbit IgG	WB: 1/10000	Jackson, 111-035-144
Peroxidase AffiniPure Donkey αRat IgG	WB: 1/10000	Jackson, 712-035-150

6.1.4. CELL LINES

Drosophila S2 cells (SL2, Schneider 2; ATCC CRL-1963); The cell line was derived from a primary culture of late stage (20-24 hours old) Drosophila melanogaster embryos (Schneider 1972).

6.1.4.1. Stable cell lines

The stable cells lines used in this thesis are shown in the following table. Both had been previously established in the lab using original S2 cells. The name of each of them is presented in the first row. In the second row we see the protein that is stably expressed.

Name	Expressed protein	Comments
dBigH1	dBigH1::FLAG	Copper induction: 1mM $CuSO_4$, 24h Generated by Dr. Carbonell Described in (Carbonell et al. 2017)
dBigH1ΔED	dBigH1ΔED::FLAG	Copper induction: 1mM $CuSO_4$, 24h Generated by Dr. Climent-Cantó Described in (Climent-Cantó et al. 2020)

6.2. METHODS

6.2.1. MANIPULATION OF CELLS

6.2.1.1. Culture and maintenance

Drosophila S2 cells were grown at 25ºC in Complete-Medium (Schneider's Insect Medium (L0207-500, Biowest) supplemented with 10% heat-inactivated FBS (10270, Gibco) and 1% penicillin-streptomycin (15140-122, Gibco)). The stable cell lines were grown at 25ºC in Complete-Medium supplemented with 0.3 mg/mL Higromicin B (Sigma).

Cells were collected by pipetting, diluted 1/5 and seeded in T25 flasks (Corning) every 3-4 days for maintenance.

Cells were frozen to keep cell storage and used to a maximum of 20 passages. To freeze cells, two to three T-175 flasks (Corning), 80% confluent, were collected and pelleted. Then they were resuspended in FBS + 10% DMSO and aliquoted in 1 mL of the cell suspension per vial. Cells were frozen gradually using a polystyrene container at -80°C for 24 h. Finally, they were transferred to liquid nitrogen for long-term storage.

6.2.1.2. Double-strand RNA treatment

For hrp36 and hrp48 KD, 1.5×10^7 cells were resuspended in 4 mL of Serum-Free medium with 50 µg in vitro-transcribed double-stranded RNA (dsRNA) and seeded in a T-75 flask (Corning). After 1 hour at room temperature, 11 mL of 14%-Medium (Complete-Medium supplemented with 14% FBS) was added to the cells. After 3 days cells were collected and processed.

For dH1 after 3 days from the first treatment of dsRNA (described above) cells were collected, and the dsRNA treatment repeated on 1.5×10^7 collected cells. After another 3 days cells were collected and processed.

A control for each KD treatment was performed using the same experimental design with dsRNA for the bacterial LacZ gene.

6.2.1.3. $CuSO_4$ induction

The expression of the constructs used in this thesis is regulated by a metallothionein promoter (pMT) that is inducible by its exposure to some metal ions. Therefore, $CuSO_4$ was added to the medium of cells to express the construct.

A stock solution of 1M $CuSO_4$ (C8027, Sigma) was prepared and kept at 4ºC. A 1:10 dilution of the stock medium was freshly prepared right before using it by diluting it with complete growth medium and the appropriate amount was added to the culture to obtain a final concentration of 1mM $CuSO_4$. After 24 h cells were collected and processed.

6.2.2. MANIPULATION OF DNA

6.2.2.1. Directed mutagenesis

The appropriate primers were previously phosphorylated with polynucleotide kinase (20350427, Roche). The Reaction Mix (1 µL of the 100 µM oligonucleotide, 1 µL 10x of the polynucleotide buffer, 2 µL 10mM of ATP, 10 u (1 µL) of polynucleotide kinase, 5 µL of nuclease-free water) was prepared and the reaction was allowed to take place for 2 h at 37ºC. Then the enzyme was inactivated by incubating the mixture 10 min at 70ºC.

The full plasmid was amplified with primers containing the desired modifications (described in section "5.1.2.3. Directed mutagenesis") using the Phusion High Fidelity DNA polymerase (M0530, New England Biolabs). The PCR reaction (10 ng of the template DNA, 2 µL of the 10 µM phosphorylated forward oligonucleotide, 2 µL of the 10 µM phosphorylated reverse oligonucleotide, 2 µL of 10 mM dNTPs, 10 µL of the 5X fusion HF buffer, 0.5 µL of the DNA polymerase) was set up to a final volume of 50 µL with nuclease-free water. The amplification program used: Pre-denaturation: 98°C – 30 sec; 30 cycles of (Denaturation: 98°C – 15 sec; Annealing: 58°C – 30 sec; Extension: 72°C – 2 min); Final elongation: 72°C – 10 min. After amplification, 12 µL of PCR product were run in an agarose gel. The corresponding band was cut from the gel and eluted in 40 µL by using the PCR DNA and gel band purification kit (28903470, Cytiva). 40

µL of the eluted PCR product was ligated with the T4 DNA ligase (EL0016, Fermentas), O/N at 18ºC. The ligated product was transformed in bacteria and the subsequent colonies were validated by sequencing.

6.2.2.2. Transformation of competent cells

The competent cells were incubated with 10 µL of ligation product for 30 min on ice. A thermic shock was performed in thermostatic bath at 42ºC for 2 min. Quickly followed by the addition of 1 mL of sterile LB. Cells were then grown with agitation at 37ºC for 30 min. The preculture was then seeded at different concentrations in LB plates supplemented with appropriated antibiotic for selection. Plates were grown at 37ºC O/N.

6.2.2.3. Minipreparation of plasmid DNA

Isolation of plasmid DNA was performed using the "alkaline minipreparation" method.

A single colony of the transformed bacteria was seeded in 4 mL of LB media supplemented with the corresponding antibiotic. The liquid culture was grown with agitation at 37ºC O/N. Next day, 1.5 mL of culture were pipetted in an eppendorf that was centrifuged at 3000g for 1 min. The supernatant (SN) was discarded, and the sample was resuspended in 100 µL of GTE solution (50 mM Glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA) by vortexing. Then 200 µL of Lysis solution (0.2N NaOH, 1% SDS) were added. Samples were mixed by inversion and incubated on ice 5 min. Then 150 µL of 3M Sodium Acetate were added. Samples were mixed by inversion and incubated on ice 5 min. Samples were centrifuged at max speed for 5 min. Supernatant was recovered in a new eppendorf. DNA was extracted by phenol-chloroform and shortly precipitated with ethanol. The pelleted DNA after the precipitation was dried at 37ºC and resuspended with 40 µL of TE (10 mM Tris-HCl 8.0, 1 mM EDTA) supplemented with 1 µg/mL of RNaseA. Digestion was carried at 37ºC for 20 min.

Plasmids were validated by sequencing.

6.2.3. ANALYSIS OF DNA

6.2.3.1. Chromatin preparation

10^8 cells were collected and fixed with 1.8% formaldehyde for 10 min at room temperature by gently mixing. Cross-linking was stopped by adding 1.25 M glycine dissolved in PBS to a final concentration of 125 mM. After 5 min, cells were spun down for 2 min at 1500g and washed with 5 mL of PBS. Then, cells were resuspended in 10 mL of ChIP wash A (10 mM HEPES pH 7.9, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100) and incubated for 10 min in a rotating wheel at 4°C. Cells were spun down again and resuspended in 10 mL of ChIP wash B (10 mM HEPES pH 7.9, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100), incubated for 10 min in a rotating wheel at 4°C and spun again for 2 min at 1500g. Later, cells were lysed in 5 mL of TE (10 mM Tris–HCl pH 8.0, 1 mM EDTA) with 1% SDS. Chromatin was washed three times with 5 mL of TE, resuspended in TE, 1 mM PMSF, 0.1% SDS and sonicated in a Bioruptor Twin Sonication System (Diagenode) to obtain fragments of 200–500 bp. After sonication, lysates were adjusted to 1% Triton X-100, 0.1% sodium deoxycholate (DOC), 140 mM NaCl, and incubated for 10 min in a rotating wheel at 4°C. Only soluble chromatin sample was pooled and then aliquoted in 0.5 mL volumes in Eppendorf tubes. 100 µL were set aside to decrosslink, purify, run in an 1% agarose gel, and check for the DNA size.

6.2.3.2. Chromatin RNaseA treatment

Cells were fixed and permeabilized with ChIP wash A and ChIP wash B as described above. Then cells were resuspended in TE, the treated sample was supplemented with 100µg/mL RNaseA, and both control and treated sample were incubated at 37ºC for 30 min. Later on, cells were lysed with 1% SDS and chromatin was washed, sonicated, and adjusted to the proper salt and detergent concentration as described above.

6.2.3.3. Chromatin immunoprecipitation - seq (ChIP-seq)

For each experiment, 400 µL of chromatin was used for the immunoprecipitation (IP) and 40 µL for the input sample. IPs were carried

out in RIPA buffer (140 mM NaCl, 10 mM Tris–HCl pH 8.0, 1 mM EDTA, 1% Triton X-100, 0.01% SDS, 0.1% DOC). A pre-clearing step was carried out in a rotating wheel for 1 h at 4°C with 30 µL of 50% (v/v) protein A-Sepharose CL4B beads (GE Healthcare, 17-0780-01) previously blocked with RIPA–1% BSA. Then, the corresponding antibodies (described in section "5.1.3.1. Primary antibodies") were added and incubated O/N at 4°C in a rotating wheel. IPs were performed by adding 40 µL of 50% (v/v) protein A-Sepharose CL4B beads previously blocked with RIPA–1% BSA and incubated in a rotating wheel for 3 h at 4°C. Beads were washed five times for 5 min in 1 mL of RIPA buffer, once for 5 min in 250 mM LiCl, 10 mM Tris–HCl pH 8.0, 1 mM EDTA, 0.5% NP-40, 0.5% DOC and twice for 5 min in TE. Beads were then resuspended in 40 µL of TE, DNase-free RNaseA was added at 0.25 µg/mL to the IPs and input samples, and incubated for 30 min at 37°C. To purify the DNA, samples were adjusted to 1% SDS, 0.1 M $NaHCO_3$, 0.2 mg/mL Proteinase K and incubated O/N at 65°C. DNA was purified by phenol-chloroform extraction and ethanol precipitation.

Libraries were generated using NEBNext Ultra II DNA Library Prep Kit (Illumina) and sequenced by 1 x 50 single reads at the CNAG Barcelona.

6.2.3.4. Genomic preparation for DRIP

S2 cells were washed with PBS (5 min at 200g at 4˚C) twice, resuspended in 400 µL of Lysis Buffer DRIP (10 mM Tris-HCl pH 7.4, 10 mM EDTA, 10 mM NaCl, 0.5% SDS, 0.5 µg/µL proteinase K) and incubated for 1 h at 50˚C. Then the DNA was purified by phenol-chloroform extractions followed by ethanol precipitation. DNA was resuspended in 50 µL of ultrapure water and the concentration was measured by Qubit. 50 µg of sample were adjusted to 1xTE and sonicated in a Bioruptor Twin Sonication System (Diagenode) approximately 4 to 6 cycles, on "High" Power setting, 30s ON/ 30s OFF, to obtain fragments of 400-500 bp. The DNA size was checked in a 1% agarose gel.

6.2.3.5. DRIP

Each sample was separated in two Eppendorf tubes containing 5 µg of genomic DNA in the RNaseH Reaction Buffer (50 mM Tris-HCl pH 8.0, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT). One of the replicates was untreated (the one used to detect the signal for the R-loops, referred to as RNH-) and the other one was treated adding 6 u RNaseH (18021-071, Invitrogen) as a control for the specificity of the signal (referred to as RNH+). Both were incubated at 37ºC O/N. Then 500 µL of Binding Buffer DRIP (10 mM Phosphate buffer pH 7.3, 140 mM NaCl, 0.05% TritóX-100) were added to each sample. 10 µg of mouse monoclonal S9.6 antibody was added and incubated in a rotatory wheel at 4ºC O/N. The following day, 35 µL of preblocked Protein G Dynabeads (10003D, Thermo Fisher Scientific) (with Binding Buffer DRIP - 0.5% BSA) were added to each sample and incubated in a rotatory wheel at 4ºC for 2 h. Beads were washed three times with Binding Buffer DRIP. To purify the DNA, samples were eluted with three sequential washes of DRIP Elution Buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5% SDS). An input sample was prepared with 10% of initial sample in DRIP Elution Buffer. Proteinase K was added at 0.2 mg/mL and incubated at 55ºC for 45 min. DNA was purified by phenol-chloroform extraction and ethanol purification.

Libraries were generated using NEBNext Ultra II DNA Library Prep Kit (Illumina) and sequenced by 2 x 150 paired-end reads on a NovaSeq 6000 (Illumina) at the CNAG Barcelona.

6.2.3.6. ATAC-seq

ATAC-seq experiments were performed as described (Buenrostro et al. 2013). Briefly, 100,000 nuclei were resuspended in 50 µL Tn5 transposase mixture (2.5 µL Tn5, 25 µL 2X TD buffer, 22.5 µL RNase Free Water), and incubated at 37°C for 30 min. After the reaction the DNA was purified using the MinElute PCR Purification Kit (QIAGEN). The PCR cycles for the library preparation were determined by qPCR, and cycles corresponding to ¼ of maximum fluorescent intensity were used to amplify the library

(ranging from 8 to 9 cycles). Following the amplification, the size selection of the library was performed using Ampure beads (Agencourt RNAClean XP Beads, S01307).

Libraries were sequenced by 2 x 50 paired-end reads at CRG Barcelona.

6.2.3.7. NRL sample preparation

For NRL determination, 2×10^7 cells of each condition were grown and treated with 1 mM $CuSO_4$ for 24 h. The cells were collected, washed twice with PBS and fixed with 1.1% of formaldehyde for 10 min at room temperature by gently mixing. Cross-linking was stopped adding glycine to a final concentration of 125 mM. After 5 min, cells were spun down for 8 min at 1000g at 4°C and washed twice with 10 mL of cold PBS before pelleting and flash freezing in liquid nitrogen. For MNase digestion, the pellet was resuspended in PBS, 0.1% Triton X-100 (PBS-TX). The digestion was performed at 37°C in a volume of 400 µL PBS–TX with 4×10^6 cells per digestion time. $CaCl_2$ was adjusted to 1 mM, cells were pre-warmed and, after adding 0.8 u of MNase (Sigma-Aldrich), incubated for increasing time from 30 s to 5 min. Digestion was stopped in ice by adding EDTA and EGTA to a final concentration of 20 mM each. To purify the DNA, samples were adjusted to 10 mM Tris-HCl pH 8, 0.4% SDS, 0.4 mg/mL Proteinase K and incubated O/N at 65°C. DNA was extracted by phenol-chloroform and precipitated with ethanol. Then samples were treated with RNase A for 30 min at 37°C, ran on 2% agarose gels at 90 V for 5h and stained with ethidium bromide.

6.2.4. MANIPULATION OF RNA

6.2.4.1. dsRNA synthesis

6.2.4.1.1. dsRNA production

dsRNA was produced using a MEGAscript T7 kit (AM1334, Thermo Fisher Scientific). The templates were produced as described in (Rogers & Rogers 2008) (primers described in section “5.1.2.1. dsRNA”). To set up the MEGAscript reaction (total volume 20 µL) the following components were

mixed: 8 µL NTPs (2 µL of each), 2 µL 10X Buffer, 1.5 µL Enzyme mix, and 8.5 µL PCR product flanked with T7 promoter regions. The reaction was incubated at 37°C O/N.

6.2.4.1.2. dsRNA purification

The double strand RNA generated was purified with RNeasy Mini Kit (74104, Qiagen). First, 2 µL TURBO DNase were added and incubated for 15 minutes at 37°C to get rid of the DNA template. From here on, RNeasy Mini Kit was used according to the manufacturer's instructions. At the last step samples were eluted using 30 µL of RNase-free water. The RNA concentration was measured using nanodrop and an aliquot was saved to be checked on an agarose gel before keeping the dsRNA at -20°C.

6.2.5. ANALYSIS OF RNA

6.2.5.1. cRNA-seq

cRNA samples were obtained adapting the protocol from (Werner & Ruthenburg 2015) for S2 cells.

10^8 cells were collected, washed twice with PBS and resuspended in 400 µL of Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol). Triton X-100 was added to a final concentration of 0.1% and cells were incubated for 8 min on ice. Nuclei were collected by centrifugation through sucrose cushion (24% sucrose in lysis Buffer A), rinsed with ice-cold PBS + 1mM EDTA, and resuspended in 400 µL of ice-cold Glycerol Buffer (20 mM Tris-HCl pH 7.9, 75 mM NaCl, 0.5 mM EDTA, 50% glycerol). Nuclei were lysed by adding 400 µL of ice-cold Lysis Buffer B (10 mM HEPES pH 7.9, 7.5 mM $MgCl_2$, 3 mM EDTA, 0.2 mM EGTA, 0.3M NaCl, 1M urea, 1% NP40), shaken vigorously and kept on ice for 10 min, with periodic shaking. Chromatin was sedimented by centrifugation at 14000g for 2 min at 4ºC, rinsed twice with cold PBS + 1mM EDTA, and resuspended in 150 µL PBS + 1mM EDTA.

RNAs were purified using RNAzol RT adding 400 µL per sample, then 100 µL of chloroform, shaking the samples vigorously, incubating them on ice

3 min and centrifuging for 5 min at maximum speed at 4ºC. The aqueous phase was recovered and purified with the RNeasy Mini Kit (74104, Quiagen) following manufacturer's instructions. Before library preparation, ribosomal RNA was depleted from the samples with Ribo-Zero Plus rRNA Depletion Kit (Illumina).

Libraries were generated using NEBNext Ultra II Directional RNA Library Prep Kit (Illumina) and sequenced by 2 x 150 paired-end reads at the CNAG Barcelona.

6.2.5.2. Total RNA extraction

10^6 cells were collected, washed twice with PBS. Cells were lysed and RNA extracted by adding 1mL RNAzol RT, then 200 µL of chloroform, shaking the samples vigorously, incubating them on ice 3 min and centrifuging for 5 min at maximum speed at 4ºC. The aqueous phase was recovered and purified with the RNeasy Mini Kit (74104, Quiagen) following manufacturer's instructions, and eluted with 14 µL of RNase-free water.

6.2.5.3. RT-qPCR

Retrotranscription of mRNA to cDNA was accomplished by using Transcription First Strand cDNA Synthesis Kit (Roche) following the manufacturer's instructions. Briefly, starting reaction mix was set up (1 µg RNA, 1 µL oligo-dT, RNAse-free H_2O up to 13 µL) and incubated at 65°C for 10 min. Then, the following components were added to the mix: 4 µL Transcription Reverse Transcriptase Reaction Buffer 5X, 0.5 µL Protector RNase Inhibitor, 2 µL Deoxynucleotide Mix [10 mM each], 0.5 µL Transcriptor Reverse Transcriptase. The reaction was mixed gently and incubated with the following program: 50°C – 1 h; 85°C – 5 min; pause at 4°C. Then the retrotranscribed products were diluted 1/10 and stored at -20°C or continued directly with qPCR.

Quantitative PCR (qPCR) was conducted in a final reaction volume of 10 µL per well. The following primer/SYBR Green master mix was prepared for each specific pair of primers: 0.5 µL Fwd primer [10 µM], 0.5 µL Rev

primer [10 µM], 5 µL 2X SYBR Green I Master (Roche), and 2 µL water. Then, 2 µL of cDNA and 8 µL of the mix were added per well in a 96-well plate. A genomic control, from the initial RNA sample not retrotranscribed, with a proportional dilution was also added. The following standard PCR program was used in a QuantStudio 5 - 96-Well 0.2-mL Block: 95°C – 5 min; 45 cycles of (95°C – 10 sec; 60°C – 10 sec; 72°C – 10 sec); Final elongation: 72°C – 10 min. Quantitative determination of RNA levels was performed in triplicate. In all cases, values were normalized to the expression of the housekeeping gene Rpl32 and relative expression levels were calculated using the standard curve method.

6.2.6. ANALYSIS OF PROTEINS

6.2.6.1. Western Blot (WB) (total sample or chromatin)

WB analyses were performed by standard procedures using total cell extracts obtained by resuspending cells in PLB (25 mM Tris-HCl pH 6.8, 4.35% Glycerol, 1% SDS) or crosslinked chromatin obtained as described above supplemented with PLB. After proteins were separated by size, they were transferred to a nitrocellulose membrane (GE Healthcare Life Sciences) in a wet transfer system with Transfer Buffer (25 mM Tris-HCl, 40 mM Glycine, 0.05% SDS, 20% methanol) at 80 V for 1 h + 30 min. Then, membranes were blocked with 5% powdered skimmed-milk diluted in PBS-0.1% Tween (PBS-T) at room temperature for 1 h or O/N at 4˚C. Incubation with primary antibodies was performed for either 1 h at room temperature or O/N at 4˚C. The membrane was washed 3 times for 5 min with PBS-T and incubated with peroxidase-conjugated secondary antibody for 1 h at room temperature. The dilution used of the different primary antibodies varied between them (described in section “5.1.3.1. Primary antibodies”) and for secondary antibodies was 1:10,000. Finally, the membrane was washed 5 times with PBS-T for 5 min and bound antibodies were detected by the ECL chemiluminescence assay (Amersham) and exposed to autoradiographic films.

6.2.6.2. Coomassie blue staining

After electrophoresis, the SDS-PAGE gel was stained with Coomassie Brilliant Blue R250 (CBB) staining solution (0.1% CBB, 10% acetic acid, 30% methanol). Then, it was washed out with 10% acetic acid as required.

6.2.7. PROTEIN PRODUCTION AND PURIFICATION

6.2.7.1. dBigH1

The DNA sequence coding for Drosophila melanogaster Linker histone variant BigH1 (GeneBank: AE014297.3:14663591-14664746 Drosophila melanogaster chromosome 3R) from the base 1 to 697, 791 to 1156 was cloned in pET30a with the S-Tag removed and in phase with the N-terminal His-tag and thrombin cut site. This was modified by directed mutagenesis (primers specified in section "5.1.2.2. Directed mutagenesis primers").

BigH1 transformed *E. coli Rosetta* cells were grown at 37°C in LB medium with the appropriated antibiotics. Protein expression was induced at OD (600 nm) = 0.7 with 1 mM IPTG and cells were cultured at 37°C for 3h. Cells were harvested by centrifugation at 5000g for 20 min and resuspended in a phosphate buffered saline (PBS), containing 0.05% NaN_3. Cells were then treated with DNaseI and disrupted twice by sonication for 5 min at 36% intensity with a pulse 5 sec on/10 sec off. Supernatants were recovered by centrifugation at 20,000 r.p.m. (rotor JA-30.50) for 45 min and the pellets were discarded.

The supernatant was cleared with syringe filters of 0.22µm and applied to a 5 mL HisTrap HP column (GE healthcare) at 4°C. Samples were eluted from the column using a gradient of 500mM imidazole in binding buffer. An additional wash with 1M KCl was added prior to elution to remove any DNA contaminants in the sample.

Samples were then dialyzed twice in a dialysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT). His6-tag was cleaved by thrombin protease digestion added to the second dialysis step. The protein obtained

by proteolytic cleavage were separated from the tags by reverse Ni^{2+} affinity.

To further purify the protein without any degraded form, the sample was applied to a 5 mL HiTrap Q HP anion exchange column (GE healthcare) with the buffer at pH7 at 4ºC. Samples were eluted from the column using a gradient of 0.25M to 1M NaCl in the binding buffer.

6.2.7.2. dH1

The DNA sequence coding for Drosophila melanogaster Linker histone variant H1 (GeneBank: AE014134.6: 21427075-21427845 Drosophila melanogaster chromosome 2L) from the base 1 to 768 was cloned in pET29b with the S-Tag removed and adding a C-terminal thrombin cut site and a His-tag. This was modified by directed mutagenesis (primers specified in section "5.1.2.2. Directed mutagenesis primers").

Expression and purification of dH1 was performed by Dr Roque lab (UAB). Shortly, the expression was performed following the BigH1 protocol, the purification was performed from inclusion bodies using guanidinium chloride at room temperature and Ni_{2+} affinity purification in denaturing conditions, and the protein was dialyzed and lyophilized.

Samples were then dialyzed twice in a dialysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT). His6-tag was cleaved by thrombin protease added in the second step of the dialysis. The protein obtained by proteolytic cleavage were separated from the tags by reverse Ni^{2+} affinity.

6.2.7.3. Final preparation of the samples and storage

Both H1 and BigH1 proteins were dialyzed at the same time to Sample storage buffer (50 mM Tris-HCl, 200 mM NaCl, 1mM Tcep). The proteins were concentrated using 3000 MWCO 15 mL Amicon 3K centrifugal filters, filtered using sterile 0.22 µm filters, adjusted to a protein concentration of 80 µM and stored at −80°C.

The protein concentrations were determined by measuring UV absorbance at 280 nm using Nanodrop 8000 (Thermo Scientific). Extinctions

coefficients used were of 4470 M^{-1} cm^{-1} and 8480 M^{-1} cm^{-1} for dH1 and dBigH1 respectively.

6.2.8. LLPS ASSAYS

Purified proteins were diluted to desired conditions from storage buffer. When protein and salt concentration was adjusted, the appropriate volume of ds DNA was added and finally 10% Ficoll 400 (Sigma) as a crowding agent. The order of addition was maintained through all sample preparation

For each condition, 1.5 µL of sample was deposited in a sealed chamber comprising a slide and a coverslip sandwiching double sided tape (3 M 300 LSE high-temperature double-sided tape of 0.17 mm thickness). The images were taken by DIC microscopy in an automated inverted Olympus IX81 microscope with a 60x using the Xcellence rt 1.2 software.

6.2.9. DATA ANALYSIS AND VISUALIZATION

6.2.9.1. NRL analysis

The images obtained from DNA ladders for the NRL sample preparation were analyzed using Fiji software and the size of each fragment was determined from the maximum of the corresponding size distribution using MW markers. To determine the apparent NRL, the size of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes, was plotted against the number of nucleosomes and the apparent NRL calculated from the slope of the corresponding regression curve. To compare NRLs of the different conditions, samples showing similar extent of digestion were analyzed.

6.2.9.2. WB quantification and analysis

For quantification, autoradiographic films were digitalized using a GS-800 Calibrated Laser Densitometer (Bio-Rad) and quantified using ImageJ Gel Analyzer plugin. Signal from each lane was normalized to the corresponding loading control. Two to three lanes were quantified per biological sample, and at least three biological replicates were quantified

per experiment. Two-tailed paired Student's t-tests were performed using GraphPad software.

6.2.9.3. ATAC-seq fragment size analyses

Fragment size and distributions of ATAC-seq samples were analyzed using plot2DO (https://github.com/rchereji/plot2DO) including the assembled heterochromatic regions of chromosomes 2LHet, 2RHet, 3LHet, 3RHet, YHet and XHet, and the unassembled repetitive heterochromatic elements of the artificial chromosomes U and Uextra.

6.2.9.4. LLPS analyses

z-stacks of 40 µm at a step size of 4 µm were collected using the project manager from the software Xcellence rt 1.2. z-stacks bright field projections were made using an ImageJ macro. Projections were then analyzed with a macro including the Trainable Weka Segmentation plugin (https://imagej.net/plugins/tws/) to quantify and measure phase-separated objects in each condition. Macros were kindly provided by Anna Lladó (Advanced Digital Microscopy Facility, IRB Barcelona).

6.2.9.5. Bioinformatic analysis (*Performed by the Biostatistics/ Bioinformatics IRB Facility)

6.2.9.5.1. cRNA-seq

For analysis of the retrieved reads, Illumina adapters were removed using cutadapt (v1.18) (Martin, 2011). Trimmed reads were aligned to *Drosophila* rRNA sequences (Quast et al., 2013) using Bowtie2 (Langmead and Salzberg, 2012) with parameters q --local -N 1 -k 1. rRNA unaligned reads were then aligned to dm3 genome using tophat2 (v2.1.1) (Kim et al., 2013) with options --library-type fr-firststrand -g 1. Aligned reads were separated by strand (Crick vs Watson strands) using samtools (v1.3) (Li et al., 2009). Independently for the two strands, reads were merged by condition and assembled into *de novo* transcripts using Cufflinks (Mv2.2.1) (Roberts et al., 2011). Counts per transcript were computed with the R package

Rsubread (v2.8.2) (Liao et al., 2013), function featureCounts with default parameters. For every condition, low abundance (<20 reads) transcripts were removed. Overlap plots (Venn diagrams) between conditions and chromosome location plots were represented for the transcriptomes obtained at this point. For differential expression analysis, the consensus transcriptomes between control and experimental conditions were considered. These transcriptomes were further curated: transcripts that overlapped partially with an annotated gene (±100bp from start/end coordinates) in the RefGene database were split, taking both the split transcripts located outside genic regions and all genic regions defined in RefGene. Reads per transcript were recalculated with featureCounts, keeping as final transcripts those with 10 or more reads. Transcriptomes were annotated with Homer (Heinz et al., 2010) leading to 5 categories of transcripts: genic-sense were defined genes in RefGene database; genic-antisense were transcripts that overlap with genes of the complementary strand; intergenic and lncRNAs were defined by Homer; and NAs were defined by those transcripts that were not labeled as intergenic by Homer but where not annotated in the RefGene database. These included mostly less curated genic regions. Differential expression analysis was performed using DESeq2 (v1.34) (Love et al., 2014). When intron retention was determined, for genic regions annotated in the refGene database, gene-exon coverage was calculated with featureCounts (Liao et al., 2013) with options GTF.featureType="exon", GTF.attrType = "gene_id", similarly gene-transcript coverage was obtained with featureCounts with options GTF.featureType="transcript", GTF.attrType = "gene_id", allowMultiOverlap = TRUE. Intron retention differences were assessed with DESeq2 (v1.34) (Love et al., 2014) by comparing exon coverage *vs* transcript coverage.

6.2.9.5.2. ChIP-seq

For analysis of the retrieved reads, NGMerge was used to automatically detect and remove sequencing adapters using options -a -n 18 -v (Gaspar, 2018). FastQ files were aligned to the dm3 genome using Bowtie2 v2.2.2

(Langmead and Salzberg, 2012) with options -N 1 -k 1. Duplicated reads for generation of TDF tracks were identified and removed with sambamba 0.8.0 (Tarasov et al., 2015). TDF tracks were generated with IGVTools2 (Thorvaldsdóttir et al., 2013) and options count -z 5 -w 25 -e 250. Peak calling between IP and Input samples was performed independently for individual and pooled replicates per group using MACS 1.4.2 (Zhang et al., 2008) with default options (duplicate removal) and genome=dm. Candidate called peaks from pooled replicates were defined as final peaks if they presented overlap with peaks identified in both their respective individual replicates. For differential binding analysis between conditions, the DiffBind package version 2.10.0 (Ross-Innes et al., 2012) was used, with functions: dba.count with options *score = DBA_SCORE_RPKM_FOLD, bLog = TRUE, bScaleControl = TRUE*; dba.contrast with options *categories = DBA_CONDITION, block = DBA_REPLICATE, minMembers = 2*; and dba.analyze with options *method = DBA_DESEQ2, bFullLibrarySize = TRUE, bSubControl = FALSE*. Peaks were annotated using Homer version 2 (Heinz et al., 2010) using the built-in dm3 genome. Peak chromosomal location was plotted using htSeqTools v1.4.2 (Planet et al., 2012). Peak/Feature distribution was obtained from Homer annotated tables. Statistical assessment of peak overlaps was performed using overlap permutation tests with the regioneR package version 1.14.0 (Gel et al., 2016) and default options. Unless otherwise specified all downstream analyses were performed with R3.5.1. (Team, 2012).

6.2.9.5.3. DRIP-seq

For analysis of the retrieved reads, NGMerge (Gaspar, 2018) was used to automatically detect and remove sequencing adapters with options -a -n 18 -v. Alignment and peak calling were performed as described above for ChIP-seq analyses. Called peaks from pooled replicates were defined as final peaks (candidate Rloops) if they presented overlap with peaks identified in both their respective individual replicates. Definition of final Rloops was performed using DESeq2 with lfcShrinkage fold change estimation (type='normal') to assess level of RNH+ depletion. Final Rloops

were defined as those RNH- candidate Rloops with a DESeq2 lfcShrinkage FC RNH+/RNH- (RNH depletion) of -1.25 or less. Peak annotation, chromosomal location and statistical assessment of peak overlaps were performed as described above for ChIP-seq analyses.

6.2.9.5.4. ATAC-seq

Bioinformatics analyses of ATAC-seq data was performed according to the guidelines of Harvard FAS Informatics (https://informatics.fas.harvard.edu/atac-seq-guidelines.html). NGMerge (Gaspar, 2018) was used to automatically detect and remove sequencing adapters using options -a -n 10 -v. Adapter cleaned FastQ files were downsampled to the library size of the smallest replicate/pair file (˜21.2 million reads) using setqtk (https://github.com/lh3/seqtk). For analysis of THSS fragments, reads were aligned to dm3 with Bowtie2 using a maximum insert size of 100bp (options --very-sensitive -k 10 -X 100), and Genrich version 0.5 (https://github.com/jsh58/Genrich) was used to detect putative ATAC peaks with options -j -y -r -e chrM. Differential analysis between conditions for all putative ATAC peaks was performed with DESeq2 version 1.22.1 (Love et al., 2014) using lfcShrinkage fold change estimation with options type='normal'.

6.2.9.6. Data availability

cRNA-seq data generated in this work is GSE227997. ChIP-seq, DRIP-seq and ATAC-seq data generated in this work are GSE228142.

7. REFERENCES

Abakir A, Giles TC, Cristini A, Foster JM, Dai N, et al. 2020. N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nat Genet*. 52(1):48–55

AeRi K, Ann D. 2003. A Human Globin Enhancer Causes both Discrete and Widespread Alterations in Chromatin Structure. *Mol Cell Biol*. 23(22):8099–8109

Aguilera A. 2002. The connection between transcription and genomic instability. *EMBO J*. 21(3):195–201

Aguilera A, García-Muse T. 2012. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol Cell*. 46(2):115–24

Aguilera A, Gómez-González B. 2017. DNA–RNA hybrids: the risks of DNA breakage during transcription. *Nat Struct Mol Biol*. 24(5):439–43

Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. 2015a. HNRNPA2B1 is a mediator of m6A-dependent nuclear RNA processing events. *Cell*. 162(6):1299–1308

Alarcón CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. 2015b. N 6-methyladenosine marks primary microRNAs for processing. *Nature*. 519(7544):482–85

Alekseyenko AA, Gorchakov AA, Zee BM, Fuchs SM, Kharchenko P V, Kuroda MI. 2014. Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. *Genes Dev*. 28(13):1445–60

Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol*. 19(4):229–44

Andrés M, García-Gomis D, Ponte I, Suau P, Roque A. 2020. Histone H1 post-translational modifications: update and future perspectives. *Int J Mol Sci*. 21(16):5941

Andreyeva EN, Bernardo TJ, Kolesnikova TD, Lu X, Yarinich LA, et al. 2017. Regulatory functions and chromatin loading dynamics of linker histone H1 during endoreplication in Drosophila. *Genes Dev*. 31(6):603–16

Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. 1991. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences*. 88(22):10148–52

Ashwin SS, Maeshima K, Sasai M. 2020. Heterogeneous fluid-like movements of chromatin and their implications to transcription. *Biophys Rev*. 12(2):461–68

Attwood JT, Yung RL, Richardson BC. 2002. DNA methylation and the regulation of gene transcription. *Cell Mol Life Sci*. 59(2):241–57

Bai L, Morozov A V. 2010. Gene regulation by nucleosome positioning. *Trends in Genetics*. 26(11):476–83

Baldi S, Krebs S, Blum H, Becker PB. 2018. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nat Struct Mol Biol*. 25(9):894–901

Bayona-Feliu A, Casas-Lamesa A, Carbonell A, Climent-Cantó P, Tatarski M, et al. 2016. Histone H1: Lessons from Drosophila. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 1859(3):526–32

Bayona-Feliu A, Casas-Lamesa A, Reina O, Bernués J, Azorín F. 2017. Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. *Nat Commun*. 8(1):

Becker M, Becker A, Miyara F, Han Z, Kihara M, et al. 2005. Differential In Vivo Binding Dynamics of Somatic and Oocyte-specific Linker Histones in Oocytes and During ES Cell Nuclear Transfer. *Mol Biol Cell*. 16(8):3887–95

Bednar J, Hamiche A, Dimitrov S. 2016. H1–nucleosome interactions and their functional implications. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 1859(3):436–43

Bentley DL. 2014. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet*. 15(3):163–75

Bernués J, Izquierdo-Boulstridge A, Reina O, Castejón L, Fernández-Castañer E, et al. 2022. Lysine 27 dimethylation of Drosophila linker histone dH1 contributes to heterochromatin organization independently of H3K9 methylation. *Nucleic Acids Res*. 50(16):9212–25

Blanchette M, Green RE, Brenner SE, Rio DC. 2005. Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. *Genes Dev*. 19(11):1306–14

Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, et al. 1986. Characterization of monoclonal antibody to DNA· RNA and its application to immunodetection of hybrids. *J Immunol Methods*. 89(1):123–30

Bonet-Costa C, Vilaseca M, Diema C, Vujatovic O, Vaquero A, et al. 2012. Combined bottom-up and top-down mass spectrometry analyses of the pattern of post-translational modifications of Drosophila melanogaster linker histone H1. *J Proteomics*. 75(13):4124–38

Borah S, Wong AC, Steitz JA. 2009. Drosophila hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron. *Proceedings of the National Academy of Sciences*. 106(8):2577–82

Brangwynne CP, Tompa P, Pappu RV. 2015. Polymer physics of intracellular phase transitions. *Nat Phys*. 11(11):899–904

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic

profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 10(12):1213–18

Carbonell A, Pérez-Montero S, Climent-Cantó P, Reina O, Azorín F. 2017. The Germline Linker Histone dBigH1 and the Translational Regulator Bam Form a Repressor Loop Essential for Male Germ Stem Cell Differentiation. *Cell Rep*. 21(11):3178–89

Caridi PC, Delabaere L, Zapotoczny G, Chiolo I. 2017. And yet, it moves: nuclear and chromatin dynamics of a heterochromatic double-strand break. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 372(1731):20160291

Casas-Lamesa A. 2018. *Analysis of the contribution of linker histone H1 to the dynamics of RNA: DNA hybrids.* Universitat de Barcelona

Cerritelli SM, Crouch RJ. 2009. Ribonuclease H: the enzymes in eukaryotes. *FEBS J*. 276(6):1494–1505

Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, Gregor T. 2018. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat Genet*. 50(9):1296–1303

Chen L, Chen J-Y, Zhang X, Gu Y, Xiao R, et al. 2017. R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol Cell*. 68(4):745-757.e5

Chereji R V, Bryson TD, Henikoff S. 2019. Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol*. 20(1):198

Chujo T, Hirose T. 2017. Nuclear bodies built on architectural long noncoding RNAs: unifying principles of their construction and function. *Mol Cells*. 40(12):889

Clapier CR, Chakravarthy S, Petosa C, Fernández-Tornero C, Luger K, Müller CW. 2008. Structure of the Drosophila nucleosome core

particle highlights evolutionary constraints on the H2A-H2B histone dimer. *Proteins: Structure, Function, and Bioinformatics*. 71(1):1–7

Climent-Cantó P, Carbonell A, Tamirisa S, Henn L, Pérez-Montero S, et al. 2021. The tumour suppressor brain tumour (Brat) regulates linker histone dBigH1 expression in the Drosophila female germline and the early embryo. *Open Biol*. 11(5):200408

Climent-Cantó P, Carbonell A, Tatarski M, Reina O, Bujosa P, et al. 2020. The embryonic linker histone dBigH1 alters the functional state of active chromatin. *Nucleic Acids Res*. 48(8):4147–60

Cremer T, Cremer M. 2010. Chromosome territories. *Cold Spring Harb Perspect Biol*. 2(3):a003889

Cutter AR, Hayes JJ. 2015. A brief review of nucleosome structure. *FEBS Lett*. 589(20PartA):2914–22

Dahl C, Grønbæk K, Guldberg P. 2011. Advances in DNA methylation: 5-hydroxymethylcytosine revisited. *Clinica Chimica Acta*. 412(11):831–36

Daujat S, Zeissler U, Waldmann T, Happel N, Schneider R. 2005. HP1 Binds Specifically to Lys26-methylated Histone H1.4, whereas Simultaneous Ser27 Phosphorylation Blocks HP1 Binding*. *Journal of Biological Chemistry*. 280(45):38090–95

Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. 2002. Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution††We dedicate this paper to the memory of Max Perutz who was particularly inspirational and supportive to T.J.R. in the early stages of this study. *J Mol Biol*. 319(5):1097–1113

Di Stefano M, Stadhouders R, Farabella I, Castillo D, Serra F, et al. 2020. Transcriptional activation during cell reprogramming correlates with the formation of 3D open chromatin hubs. *Nat Commun*. 11(1):2564

Dreyfuss G, Kim VN, Kataoka N. 2002. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol*. 3(3):195–205

Drolet M, Phoenix P, Menzel R, Massé E, Liu LF, Crouch RJ. 1995. Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proceedings of the National Academy of Sciences*. 92(8):3526–30

Dubochet J, Adrian M, Chang J-J, Homo J-C, Lepault J, et al. 1988. Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys*. 21(2):129–228

Eberharter A, Becker PB. 2004. ATP-dependent nucleosome remodelling: factors and functions. *J Cell Sci*. 117(17):3707–11

Ekwall K, Nimmo ER, Javerzat JP, Borgstrom B, Egel R, et al. 1996. Mutations in the fission yeast silencing factors clr4+ and rik1+ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J Cell Sci*. 109(11):2637–48

El Hage A, French SL, Beyer AL, Tollervey D. 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev*. 24(14):1546–58

Elgin SCR, Reuter G. 2013. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol*. 5(8):a017780

Ernst J, Kellis M. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol*. 28(8):817–25

Fan Y, Sirotkin A, Russell RG, Ayala J, Skoultchi AI. 2001. Individual somatic H1 subtypes Are dispensable for mouse development even in mice lacking the H1(0) replacement subtype. *Mol Cell Biol*. 21(23):7933–43

Fernández-Justel JM, Santa-María C, Martín-Vírgala S, Ramesh S, Ferrera-Lagoa A, et al. 2022. Histone H1 regulates non-coding RNA turnover on chromatin in a m6A-dependent manner. *Cell Rep*. 40(11):111329

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, et al. 2010. Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. *Cell*. 143(2):212–24

Finch JT, Klug A. 1976. Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences*. 73(6):1897–1901

Fussner E, Strauss M, Djuric U, Li R, Ahmed K, et al. 2012. Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep*. 13(11):992–96

Fyodorov D V., Zhou BR, Skoultchi AI, Bai Y. 2018. Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol Cell Biol*. 19(3):192–206

Gan W, Guan Z, Liu J, Gui T, Shen K, et al. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev*. 25(19):2041–56

García-Muse T, Aguilera A. 2019. R Loops: From Physiological to Pathological Roles. *Cell*. 179(3):604–18

Germier T, Kocanova S, Walther N, Bancaud A, Shaban HA, et al. 2017. Real-Time Imaging of a Single Gene Reveals Transcription-Initiated Local Confinement. *Biophys J*. 113(7):1383–94

Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, et al. 2019. Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell*. 179(2):470-484.e21

Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. 2012. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Mol Cell*. 45(6):814–25

Gottesfeld JM, Garrard WT, Bagi G, Wilson RF, Bonner J. 1974. Partial Purification of the Template-Active Fraction of Chromatin: A Preliminary Report. *Proceedings of the National Academy of Sciences*. 71(6):2193–97

Grewal SIS, Jia S. 2007. Heterochromatin revisited. *Nat Rev Genet*. 8(1):35–46

Gu T, Elgin SCR. 2013. Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in Drosophila. *PLoS Genet*. 9(9):e1003780

Hale TK, Contreras A, Morrison AJ, Herrera RE. 2006. Phosphorylation of the Linker Histone H1 by CDK Regulates Its Binding to HP1α. *Mol Cell*. 22(5):693–99

Halmer L, Gruss C. 1996. Effects of Cell Cycle Dependent Histone H1 Phosphorylation on Chromatin Structure and Chromatin Replication. *Nucleic Acids Res*. 24(8):1420–27

Han SP, Tang YH, Smith R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. *Biochemical Journal*. 430(3):379–92

Hansen JC, Maeshima K, Hendzel MJ. 2021. The solid and liquid states of chromatin. *Epigenetics Chromatin*. 14(1):50

Harshman SW, Young NL, Parthun MR, Freitas MA. 2013. H1 histones: Current perspectives and challenges. *Nucleic Acids Res*. 41(21):9593–9609

Heitz E. 1928. *Das Heterochromatin Der Moose*. Bornträger

Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. 2009. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res*. 19(3):460–69

Henikoff S, Smith MM. 2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol*. 7(1):a019364

Hergeth SP, Schneider R. 2015. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep*. 16(11):1439–53

Hihara S, Pack C-G, Kaizu K, Tani T, Hanafusa T, et al. 2012. Local Nucleosome Dynamics Facilitate Chromatin Accessibility in Living Mammalian Cells. *Cell Rep*. 2(6):1645–56

Ho JWK, Jung YL, Liu T, Alver BH, Lee S, et al. 2014. Comparative analysis of metazoan chromatin organization. *Nature*. 512(7515):449–52

Holehouse AS, Das RK, Ahad JN, Richardson MOG, Pappu R V. 2017. CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys J*. 112(1):16–21

Holmes DS, Mayfield JE, Sander G, Bonner J. 1972. Chromosomal RNA: Its Properties. *Science (1979)*. 177(4043):72–74

Horn PJ, Peterson CL. 2002. Chromatin Higher Order Folding--Wrapping up Transcription. *Science (1979)*. 297(5588):1824–27

Hovemann BT, Reim I, Werner S, Katz S, Saumweber H. 2000. The protein Hrb57A of Drosophila melanogaster closely related to hnRNP K from vertebrates is present at sites active in transcription and coprecipitates with four RNA-binding proteins. *Gene*. 245(1):127–37

Huang RC, Bonner J. 1965. Histone-bound RNA, a component of native nucleohistone. *Proceedings of the National Academy of Sciences*. 54(3):960–67

Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell*. 169(2):216-228.e19

Hyman AA, Weber CA, Jülicher F. 2014. Liquid-Liquid Phase Separation in Biology. *Annu Rev Cell Dev Biol*. 30(1):39–58

Iyer LM, Abhiman S, Aravind L. 2011. Chapter 2 - Natural History of Eukaryotic DNA Methylation Systems. In *Progress in Molecular Biology and Translational Science*, Vol. 101, eds. X Cheng, RM Blumenthal, pp. 25–104. Academic Press

Izquierdo-Bouldstridge A, Bustillos A, Bonet-Costa C, Aribau-Miralbés P, García-Gomis D, et al. 2017. Histone H1 depletion triggers an interferon response in cancer cells via activation of heterochromatic repeats. *Nucleic Acids Res*. 45(20):11622–42

Izzo A, Kamieniarz K, Schneider R. 2008. The histone H1 family: Specific members, specific functions? *Biol Chem*. 389(4):333–43

Janssen A, Colmenares SU, Karpen GH. 2018. Heterochromatin: Guardian of the Genome. *Annu Rev Cell Dev Biol*. 34(1):265–88

Jensen RH, Chalkley R. 1968. Physical state of nucleohistone under physiological ionic strength. Effect of interaction with free nucleic acids. *Biochemistry*. 7(12):4388–95

Jenuwein T, Allis CD. 2001. Translating the Histone Code. *Science (1979)*. 293(5532):1074–80

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (1979)*. 337(6096):816–21

Johnson WL, Yewdell WT, Bell JC, McNulty SM, Duda Z, et al. 2017. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife*. 6:e25299

Joti Y, Hikima T, Nishino Y, Kamada F, Hihara S, et al. 2012. Chromosomes without a 30-nm chromatin fiber. *Nucleus*. 3(5):404–10

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596(7873):583–89

Kasinsky HE, Lewis JD, Dacks JB, Ausio J. 2001. Origin of H1 linker histones. *FASEB journal*. 15(1):34–42

Keenen MM, Brown D, Brennan LD, Renger R, Khoo H, et al. 2021. HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *Elife*. 10:e64563

Kharchenko P V, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, et al. 2011. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature*. 471(7339):480–85

Kimura H, Cook PR. 2001. Kinetics of Core Histones in Living Human Cells: Little Exchange of H3 and H4 and Some Rapid Exchange of H2b. *Journal of Cell Biology*. 153(7):1341–54

Kornberg RD. 1977. Structure of Chromatin. *Annu Rev Biochem*. 46(1):931–54

Kornberg RD, Lorch Y. 1999. Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome. *Cell*. 98:285–94

Kouzarides T. 2007. Chromatin Modifications and Their Function. *Cell*. 128(4):693–705

Lafontaine DLJ, Riback JA, Bascetin R, Brangwynne CP. 2021. The nucleolus as a multiphase liquid condensate. *Nat Rev Mol Cell Biol*. 22(3):165–82

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, et al. 2017. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*. 547(7662):236–40

Leicher R, Osunsade A, Chua GNL, Faulkner SC, Latham AP, et al. 2022. Single-stranded nucleic acid binding and coacervation by linker histone H1. *Nat Struct Mol Biol*. 29(5):463–71

Lever MA, Th'ng JPH, Sun X, Hendzel MJ. 2000. Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature*. 408(6814):873–76

Li X, Fu X-D. 2019. Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat Rev Genet*. 20(9):503–19

Li Y, Syed J, Sugiyama H. 2016. RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem Biol*. 23(11):1325–33

Liao Y, Smyth GK, Shi W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 41(10):

Liu LF, Wang JC. 1987. Supercoiling of the DNA template during transcription. *Proceedings of the National Academy of Sciences*. 84(20):7024–27

Loda A, Heard E. 2019. Xist RNA in action: Past, present, and future. *PLoS Genet*. 15(9):e1008333-

Lorch Y, Kornberg RD. 2017. Chromatin-remodeling for transcription. *Q Rev Biophys*. 50:e5

Lu X, Wontakal SN, Emelyanov A V., Morcillo P, Konev AY, et al. 2009. Linker histone H1 is essential for Drosophila development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. *Genes Dev*. 23(4):452–65

Lu X, Wontakal SN, Kavi H, Kim BJ, Guzzardo PM, et al. 2013. Drosophila H1 Regulates the Genetic Activity of Heterochromatin by Recruitment of Su(var)3-9. *Science (1979)*. 340(6128):78–81

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 389(6648):251–60

Luque A, Collepardo-Guevara R, Grigoryev S, Schlick T. 2014. Dynamic condensation of linker histone C-terminal domain regulates chromatin structure. *Nucleic Acids Res*. 42(12):7553–60

Madigan JP, Chotkowski HL, Glaser RL. 2002. DNA double-strand break-induced phosphorylation of Drosophila histone variant H2Av helps

prevent radiation-induced apoptosis. *Nucleic Acids Res*. 30(17):3698–3705

Maeshima K, Ide S, Babokhov M. 2019. Dynamic chromatin organization without the 30-nm fiber. *Curr Opin Cell Biol*. 58:95–104

Maeshima K, Imai R, Tamura S, Nozaki T. 2014. Chromatin as dynamic 10-nm fibers. *Chromosoma*. 123(3):225–37

Maeshima K, Matsuda T, Shindo Y, Imamura H, Tamura S, et al. 2018. A Transient Rise in Free Mg2+ Ions Released from ATP-Mg Hydrolysis Contributes to Mitotic Chromosome Condensation. *Current Biology*. 28(3):444-451.e6

Maeshima K, Tamura S, Hansen JC, Itoh Y. 2020. Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Curr Opin Cell Biol*. 64:77–89

Marsano RM, Dmitri P. 2022. Constitutive Heterochromatin in Eukaryotic Genomes: A Mine of Transposable Elements. *Cells*. 11(761):

Martire S, Banaszynski LA. 2020. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol*. 21(9):522–41

Marzluff WF, Gongidi P, Woods KR, Jin J, Maltais LJ. 2002. The Human and Mouse Replication-Dependent Histone Genes. *Genomics*. 80(5):487–98

Matunis EL, Matunis MJ, Dreyfuss G. 1992. Characterization of the major hnRNP proteins from Drosophila melanogaster. *Journal of Cell Biology*. 116(2):257–69

Matunis EL, Matunis MJ, Dreyfuss G. 1993. Association of individual hnRNP proteins and snRNPs with nascent transcripts. *Journal of Cell Biology*. 121(2):219–28

Mayor R, Izquierdo-Bouldstridge A, Millán-Ariño L, Bustillos A, Sampaio C, et al. 2015. Genome Distribution of Replication-independent

Histone H1 Variants Shows H1.0 Associated with Nucleolar Domains and H1X Associated with RNA Polymerase II-enriched Regions *. *Journal of Biological Chemistry*. 290(12):7474–91

McBryant SJ, Lu X, Hansen JC. 2010. Multifunctionality of the linker histones: an emerging role for protein-protein interactions. *Cell Res*. 20(5):519–28

McDowall AW, Smith JM, Dubochet J. 1986. Cryo-electron microscopy of vitrified chromosomes in situ. *EMBO J*. 5(6):1395–1402

McGinty RK, Tan S. 2015. Nucleosome Structure and Function. *Chem Rev*. 115(6):2255–73

McNulty SM, Sullivan LL, Sullivan BA. 2017. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev Cell*. 42(3):226-240.e6

Mimura M, Tomita S, Sugai H, Shinkai Y, Ishihara S, Kurita R. 2021. Uncharged Components of Single-Stranded DNA Modulate Liquid–Liquid Phase Separation With Cationic Linker Histone H1. *Front Cell Dev Biol*. 9:710729

Misteli T, Gunjan A, Hock R, Bustin M, Brown DT. 2000. Dynamic binding of histone H1 to chromatin in living cells. *Nature*. 408(6814):877–81

Monahan Z, Ryan VH, Janke AM, Burke KA, Rhoads SN, et al. 2017. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J*. 36(20):2951–67

Muzzopappa F, Hertzog M, Erdel F. 2021. DNA length tunes the fluidity of DNA-based condensates. *Biophys J*. 120(7):1288–1300

Nagel S, Grossbach U. 2000. Histone H1 Genes and Histone Gene Clusters in the Genus Drosophila. *J Mol Evol*. 51(3):286–98

Ni J-Q, Liu L-P, Hess D, Rietdorf J, Sun F-L. 2006. Drosophila ribosomal proteins are associated with linker histone H1 and suppress gene transcription. *Genes Dev*. 20(14):1959–73

Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R. 2001. Heterochromatin Formation in Mammalian Cells: Interaction between Histones and HP1 Proteins. *Mol Cell*. 7(4):729–39

Nozaki T, Imai R, Tanbo M, Nagashima R, Tamura S, et al. 2017. Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. *Mol Cell*. 67(2):282-293.e7

Olins DE, Olins AL. 2003. Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol*. 4(10):809–14

Osley MA, Tsukuda T, Nickoloff JA. 2007. ATP-dependent chromatin remodeling factors and DNA damage repair. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 618(1):65–80

Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O'Shea CC. 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science (1979)*. 357(6349):eaag0025

Öztürk MA, Cojocaru V, Wade RC. 2018. Toward an Ensemble View of Chromatosome Structure: A Paradigm Shift from One to Many. *Structure*. 26(8):1050–57

Paulsen RD, Soni D V, Wollman R, Hahn AT, Yee M-C, et al. 2009. A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol Cell*. 35(2):228–39

Peng JC, Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol*. 9(1):25–35

Peng JC, Karpen GH. 2009. Heterochromatic Genome Stability Requires Regulators of Histone H3 K9 Methylation. *PLoS Genet*. 5(3):e1000435-

Pérez-Montero S, Carbonell A, Azorín F. 2016. Germline-specific H1 variants: the "sexy" linker histones. *Chromosoma*. 125(1):1–13

Pérez-Montero S, Carbonell A, Morán T, Vaquero A, Azorín F. 2013. The Embryonic Linker Histone H1 Variant of Drosophila, dBigH1, Regulates Zygotic Genome Activation. *Dev Cell*. 26(6):578–90

Perišić O, Collepardo-Guevara R, Schlick T. 2010. Modeling Studies of Chromatin Fiber Structure as a Function of DNA Linker Length. *J Mol Biol*. 403(5):777–802

Perišić O, Portillo-Ledesma S, Schlick T. 2019. Sensitive effect of linker histone binding mode and subtype on chromatin condensation. *Nucleic Acids Res*. 47(10):4948–57

Peters AHFM, O'Carroll D, Scherthan H, Mechtler K, Sauer S, et al. 2001. Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. *Cell*. 107(3):323–37

Piacentini L, Fanti L, Negri R, Del Vescovo V, Fatica A, et al. 2009. Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in Drosophila. *PLoS Genet*. 5(10):e1000670

Planells J, Jordan Pla A, Jain S, Guadalupe JJ, Proux-Wera E, et al. 2023. The exosome degrades chromatin-associated RNAs genome-wide and maintains chromatin homeostasis. *bioRxiv*. 2023–24

Pohjoismäki JLO, Holmes JB, Wood SR, Yang M-Y, Yasukawa T, et al. 2010. Mammalian Mitochondrial DNA Replication Intermediates Are Essentially Duplex but Contain Extensive Tracts of RNA/DNA Hybrid. *J Mol Biol*. 397(5):1144–55

Portela A, Esteller M. 2010. Epigenetic modifications and human disease. *Nat Biotechnol*. 28(10):1057–68

Prendergast L, Reinberg D. 2021. The missing linker: emerging trends for H1 variant-specific functions. *Genes Dev*. 35(1–2):40–58

Probst AlineV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. 2010. A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. *Dev Cell*. 19(4):625–38

Protter DSW, Parker R. 2016. Principles and Properties of Stress Granules. *Trends Cell Biol*. 26(9):668–79

R Core Team. 2017. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. www.gbif.org

Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. 1993. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*. 362(6417):219–23

Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M, Cosma MP. 2015. Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. *Cell*. 160(6):1145–58

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. 2007. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*. 129(7):1311–23

Rodriguez-Campos A, Azorin F. 2007. RNA is an integral component of chromatin that contributes to its structural organization. *PLoS One*. 2(11):e1182

Rogers SL, Rogers GC. 2008. Culture of Drosophila S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy. *Nat Protoc*. 3(4):606–11

Roque A, Ponte I, Arrondo JLR, Suau P. 2008. Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. *Nucleic Acids Res*. 36(14):4719–26

Roy D, Yu K, Lieber MR. 2008. Mechanism of R-Loop Formation at Immunoglobulin Class Switch Sequences. *Mol Cell Biol*. 28(1):50–60

Saeki H, Ohsumi K, Aihara H, Ito T, Hirose S, et al. 2005. Linker histone variants control chromatin dynamics during early embryogenesis. *Proceedings of the National Academy of Sciences*. 102(16):5697–5702

Saito M, Hess D, Eglinger J, Fritsch AW, Kreysing M, et al. 2019. Acetylation of intrinsically disordered regions regulates phase separation. *Nat Chem Biol*. 15(1):51–61

Sancho M, Diani E, Beato M, Jordan A. 2008. Depletion of Human Histone H1 Variants Uncovers Specific Roles in Gene Expression and Cell Growth. *PLoS Genet*. 4(10):e1000227-

Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, et al. 2016. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol Cell*. 63(1):167–78

Scaffidi P. 2016. Histone H1 alterations in cancer. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 1859(3):533–39

Schmitz K-M, Mayer C, Postepska A, Grummt I. 2010. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev*. 24(20):2264–69

Schmitz RJ, Lewis ZA, Goll MG. 2019. DNA Methylation: Shared and Divergent Features across Eukaryotes. *Trends in Genetics*. 35(11):818–27

Schneider I. 1972. Cell lines derived from late embryonic stages of Drosophila melanogaster

Schwarz PM, Felthauser A, Fletcher TM, Hansen JC. 1996. Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. *Biochemistry*. 35(13):4009–15

Serna-Pujol N, Salinas-Pena M, Mugianesi F, Le Dily F, Marti-Renom MA, Jordan A. 2022. Coordinated changes in gene expression, H1 variant distribution and genome 3D conformation in response to H1 depletion. *Nucleic Acids Res*. 50(7):3892–3910

Shakya A, Park S, Rana N, King JT. 2020. Liquid-Liquid Phase Separation of Histone Proteins in Cells: Role in Chromatin Organization. *Biophys J*. 118(3):753–64

Shatskikh AS, Kotov AA, Adashev VE, Bazylev SS, Olenina L V. 2020. Functional significance of satellite DNAs: insights from Drosophila. *Front Cell Dev Biol*. 8:312

Simpson RT. 1978. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*. 17(25):5524–31

Simpson RT, Thoma F, Brubaker JM. 1985. Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system for study of higher order structure. *Cell*. 42(3):799–808

Singh AK, Lakhotia SC. 2012. The hnRNP A1 homolog Hrp36 is essential for normal development, female fecundity, omega speckle formation and stress tolerance in Drosophila melanogaster. *J Biosci*. 37(4):659–78

Singh AK, Lakhotia SC. 2016. The hnRNP A1 homolog Hrb87F/Hrp36 is important for telomere maintenance in Drosophila melanogaster. *Chromosoma*. 125(3):373–88

Skourti-Stathaki K, Proudfoot NJ, Gromak N. 2011. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Mol Cell*. 42(6):794–805

Sollier J, Cimprich KA. 2015. Breaking bad: R-loops and genome integrity. *Trends Cell Biol*. 25(9):514–22

Strick R, Strissel PL, Gavrilov K, Levi-Setti R. 2001. Cation–chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. *Journal of Cell Biology*. 155(6):899–910

Strickfaden H, Tolsma TO, Sharma A, Underhill DA, Hansen JC, Hendzel MJ. 2020. Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells. *Cell*. 183(7):1772-1784.e13

Strom AR, Emelyanov A V, Mir M, Fyodorov D V, Darzacq X, Karpen GH. 2017. Phase separation drives heterochromatin domain formation. *Nature*. 547(7662):241–45

Suissa Y, Kalifa Y, Dinur T, Graham P, Deshpande G, et al. 2010. Hrp48 attenuates Sxl expression to allow for proper notch expression and signaling in wing development. *Proceedings of the National Academy of Sciences*. 107(15):6930–35

Sullivan BA, Karpen GH. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol*. 11(11):1076–83

Sun Q, Csorba T, Skourti-Stathaki K, Proudfoot NJ, Dean C. 2013. R-Loop Stabilization Represses Antisense Transcription at the Arabidopsis FLC Locus. *Science (1979)*. 340(6132):619–21

Szostak E, García-Beyaert M, Guitart T, Graindorge A, Coll O, Gebauer F. 2018. Hrp48 and eIF3d contribute to msl-2 mRNA translational repression. *Nucleic Acids Res*. 46(8):4099–4113

Talasz H, Sarg B, Lindner HH. 2009. Site-specifically phosphorylated forms of H1.5 and H1.2 localized at distinct regions of the nucleus are related to different processes during the cell cycle. *Chromosoma*. 118(6):693–709

Talbert PB, Henikoff S. 2010. Histone variants — ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol*. 11(4):264–75

Thakur J, Fang H, Llagas T, Disteche CM, Henikoff S. 2019. Architectural RNA is required for heterochromatin organization. *bioRxiv*. 784835

Thakur J, Henikoff S. 2020. Architectural RNA in chromatin organization. *Biochem Soc Trans*. 48(5):1967–78

Thiriet C, Hayes JJ. 2009. Linker histone phosphorylation regulates global timing of replication origin firing. *Journal of Biological Chemistry*. 284(5):2823–29

Thomas M, White RL, Davis RW. 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proceedings of the National Academy of Sciences*. 73(7):2294–98

Thorslund T, Ripplinger A, Hoffmann S, Wild T, Uckelmann M, et al. 2015. Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. *Nature*. 527(7578):389–93

Trojer P, Reinberg D. 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell*. 28(1):1–13

Turner AL, Watson M, Wilkins OG, Cato L, Travers A, et al. 2018. Highly disordered histone H1−DNA model complexes and their condensates. *Proceedings of the National Academy of Sciences*. 115(47):11964–69

Ura K, Nightingale K, Wolffe AP. 1996. Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. *EMBO J*. 15(18):4959–69

van Holde K, Zlatanova J. 2007. Chromatin fiber structure: Where is the problem now? *Semin Cell Dev Biol*. 18(5):651–58

Velazquez Camacho O, Galan C, Swist-Rosowska K, Ching R, Gamalinda M, et al. 2017. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife*. 6:e25293

Vujatovic O, Zaragoza K, Vaquero A, Reina O, Bernués J, Azorín F. 2012. Drosophila melanogaster linker histone dH1 is required for transposon silencing and to preserve genome integrity. *Nucleic Acids Res*. 40(12):5402–14

Vyas P, Brown DT. 2012. N-and C-terminal domains determine differential nucleosomal binding geometry and affinity of linker histone isotypes H10 and H1c. *Journal of Biological Chemistry*. 287(15):11778–87

Waddington CH. 1957. The Strategy of the Genes. *Geo Allen & Unwin, London*,

Weintraub H. 1984. Histone-h1-dependent chromatin superstructures and the suppression of gene activity. *Cell*. 38(1):17–27

Werner MS, Ruthenburg AJ. 2015. Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. *Cell Rep*. 12(7):1089–98

Willcockson MA, Healton SE, Weiss CN, Bartholdy BA, Botbol Y, et al. 2021. H1 histones control the epigenetic landscape by local chromatin compaction. *Nature*. 589(7841):293–98

Wiśniewski JR, Zougman A, Krüger S, Mann M. 2007. Mass Spectrometric Mapping of Linker Histone H1 Variants Reveals Multiple Acetylations, Methylations, and Phosphorylation as Well as Differences between Cell Culture and Tissue * S. *Molecular & Cellular Proteomics*. 6(1):72–87

Woodcock CL, Frado LL, Rattner JB. 1984. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *Journal of Cell Biology*. 99(1):42–52

Woodcock CL, Skoultchi AI, Fan Y. 2006. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Research*. 14(1):17–25

Xu B, Clayton DA. 1996. RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *EMBO J*. 15(12):3135–43

Xu N, Emelyanov A V, Fyodorov D V, Skoultchi AI. 2014. Drosophila linker histone H1 coordinates STAT-dependent organization of heterochromatin and suppresses tumorigenesis caused by hyperactive JAK-STAT signaling. *Epigenetics Chromatin*. 7(1):16

Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. 2010. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. 1804(4):996–1010

Yadav T, Whitehouse I. 2016. Replication-Coupled Nucleosome Assembly and Positioning by ATP-Dependent Chromatin-Remodeling Enzymes. *Cell Rep*. 15(4):715–23

Yang X, Liu Q-L, Xu W, Zhang Y-C, Yang Y, et al. 2019. m6A promotes R-loop formation to facilitate transcription termination. *Cell Res*. 29(12):1035–38

Yano T, de Quinto SL, Matsui Y, Shevchenko A, Shevchenko A, Ephrussi A. 2004. Hrp48, a Drosophila hnRNPA/B Homolog, Binds and Regulates Translation of oskar mRNA. *Dev Cell*. 6(5):637–48

Yuhong F, Tatiana N, M M-KE, Jie Z, R MT, et al. 2003. H1 Linker Histones Are Essential for Mouse Development and Affect Nucleosome Spacing In Vivo. *Mol Cell Biol*. 23(13):4559–72

Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. *Nat Genet*. 48(11):1385–95

Zhang H, Ji X, Li P, Liu C, Lou J, et al. 2020. Liquid-liquid phase separation in biology: mechanisms, physiological functions and human diseases. *Sci China Life Sci*. 63(7):953–85

Zhang Y, Sun Z, Jia J, Du T, Zhang N, et al. 2021. Overview of histone modification. In *Histone Mutations and Cancer*, pp. 1–16. Springer

Zhou B-R, Bai Y. 2019. Chromatin structures condensed by linker histones. *Essays Biochem*. 63(1):75–87

Zhou BR, Jiang J, Feng H, Ghirlando R, Xiao TS, Bai Y. 2015. Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Mol Cell*. 59(4):628–38

Zhou J, Fan JY, Rangasamy D, Tremethick DJ. 2007. The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat Struct Mol Biol*. 14(11):1070–76

ANNEX I

Published online 27 February 2020

Nucleic Acids Research, 2020, Vol. 48, No. 8 ***4147–4160***
doi: 10.1093/nar/gkaa122

The embryonic linker histone dBigH1 alters the functional state of active chromatin

Paula Climent-Cantó[1,2], Albert Carbonell[1,2], Milos Tatarski[1,2], Oscar Reina[2], Paula Bujosa[1,2], Jofre Font-Mateu[3,4], Jordi Bernués[1,2], Miguel Beato[3,4] and Fernando Azorín[1,2,*]

[1]Institute of Molecular Biology of Barcelona, IBMB, CSIC, Baldiri Reixac, 4, 08028 Barcelona, Spain, [2]Institute for Research in Biomedicine, IRB Barcelona. The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain, [3]Centre de Regulació Genòmica (CRG). The Barcelona Institute of Science and Technology, Barcelona, Spain and [4]Universitat Pompeu Fabra (UPF), Barcelona, Spain

Received October 17, 2019; Revised January 30, 2020; Editorial Decision February 13, 2020; Accepted February 25, 2020

ABSTRACT

Linker histones H1 are principal chromatin components, whose contribution to the epigenetic regulation of chromatin structure and function is not fully understood. In metazoa, specific linker histones are expressed in the germline, with female-specific H1s being normally retained in the early-embryo. Embryonic H1s are present while the zygotic genome is transcriptionally silent and they are replaced by somatic variants upon activation, suggesting a contribution to transcriptional silencing. Here we directly address this question by ectopically expressing dBigH1 in *Drosophila* S2 cells, which lack dBigH1. We show that dBigH1 binds across chromatin, replaces somatic dH1 and reduces nucleosome repeat length (NRL). Concomitantly, dBigH1 expression down-regulates gene expression by impairing RNApol II binding and histone acetylation. These effects depend on the acidic N-terminal ED-domain of dBigH1 since a truncated form lacking this domain binds across chromatin and replaces dH1 like full-length dBigH1, but it does not affect NRL either transcription. *In vitro* reconstitution experiments using *Drosophila* preblastodermic embryo extracts corroborate these results. Altogether these results suggest that the negatively charged N-terminal tail of dBigH1 alters the functional state of active chromatin compromising transcription.

INTRODUCTION

Linker histones H1 constitute an evolutionarily conserved family of chromosomal proteins that play an important structural role in regulating chromatin compaction and higher order chromatin organization (see (1) for a recent review). In metazoan species, histones H1 usually exist as multiple variants, some of which are specifically expressed in the germline (reviewed in (2)). For instance, four of the eleven mice/human H1 isoforms are germline specific, of which three are expressed in males (H1T, HILS1 and H1T2) and one in females (H1oo). Female-specific variants usually accumulate in the oocyte and are retained during early embryogenesis (2). In comparison to most metazoa, H1 complexity in *Drosophila* is much reduced since it contains a single somatic dH1 variant (3–5), which is ubiquitously expressed throughout development, and a single germline specific variant dBigH1, which is expressed in both the female and male germlines, and it is retained in the early embryo (6,7). Embryonic H1s persist as long as the zygotic genome remains transcriptionally silent, being replaced by somatic variants when transcription begins during zygotic genome activation (ZGA) (reviewed in (2)). In *Drosophila*, dBigH1 is present during early embryogenesis until ZGA onset at cellularization (6). At this stage, dBigH1 is replaced by somatic dH1 in somatic cells, whereas it is retained in the primordial germ cells (PGC) (6), which remain transcriptionally silent.

These observations suggest that dBigH1, and embryonic H1s in general, are general transcriptional regulators that contribute to silencing. Linker histones H1 have been usually associated with transcription repression, but somatic H1s are readily detected across expressed genes (8–14). In this regard, it has been reported that somatic H1s can even enhance the synergism between transcription factors (15). In contrast, in the presence of dBigH1, chromatin appears to be transcriptionally silent, suggesting that dBigH1 enhances transcriptional silencing. Here, we analyze the mechanisms of action of dBigH1. For this purpose, we have performed ectopic expression experiments in *Drosophila* S2 cells, which lack dBigH1. These experiments confirm the contribution of dBigH1 to transcriptional silencing, iden-

*To whom correspondence should be addressed. Tel: +34 93 4034958; Fax: +34 93 4034979; Email: fambmc@ibmb.csic.es

tifying the acidic N-terminal ED-domain as responsible for its negative effect on transcription.

MATERIALS AND METHODS

Antibodies and recombinant proteins

αdBigH1 antibodies are described in (6). αdH1 antibodies were a gift from Dr J. Kadonaga and are described in (16). αRpb3 antibodies were a gift from Dr J. Zeitlinger and are described in (17). The rest of antibodies used in these experiments were commercially available: αPol IIo^{ser5} (Abcam, ab5131), αH3Kac (Millipore 06-599), αH3K36me3 (Cell Signaling, 9715), αH3 (Abcam, ab9050), αH4 (Abcam, ab10158) and αFLAG (Sigma F3165).

Recombinant His-tagged dBigH1 constructs were expressed in *Escherichia coli* BL21-LysS cells using pET30b(+) expression vectors (Novagen), and purified using Ni-NTA columns (BioRad) by conventional methods.

Stable S2 cell lines and transgenic flies

Stable S2 cell line expressing dBigH1::FLAG under the control of a Cu^{2+}-inducible promoter is described in (7). Stable lines expressing dBigH1ΔED::FLAG and dBigH1ΔNTD::FLAG were generated identically as the dBigH1::FLAG-expressing line.

w1118 and *daughterless(da)*-GAL4 flies were obtained from the Bloomington Stock Center. Transgenic flies carrying UAS-*dBigH1* construct are described in (6). Transgenic flies carrying UAS-*dBigH1ΔCTD,* UAS-*dBigH1ΔED* and UAS-*dBigH1ΔNTD* were obtained by specific site-directed integration of the appropriate constructs into ZH-51D att line. All *Drosophila* stocks were maintained at 25°C on standard media.

Expression of dBigH1 constructs in S2 cells and salivary glands

In S2 cells, expression of the dBigH1::FLAG, dBigH1ΔED::FLAG, dBigH1ΔNTD::FLAG constructs was induced with 1mM $CuSO_4$ for 24 h. To determine the extent of binding to chromatin of the different constructs, crosslinked chromatin was prepared as for ChIP (see below), diluted in PLB and analyzed by WB using αdBigH1 (1:10,000), αFLAG (1:2500), αH3 (1:2500) and αH4 (1:5000) antibodies. For quantitative analysis, WBs were digitalized using a GS-800 Calibrated laser densitometer (BioRad) and analyzed using Fiji software (18). From these analyses we estimated that chromatin abundance of the full-length dBigH1::FLAG, and the truncated dBigH1ΔED::FLAG and dBigH1ΔNTD::FLAG forms was similar (Figure 1B and Supplementary Figure S7A). To determine the actual abundance of the various dBigH1 constructs, total histones were prepared by acid extraction of purified nuclei in 0.2N HCl overnight in a rotating wheel at 4°C. Acid extracts were analyzed by Coomassie staining and WB using αdBigH1 (1:10,000) and αdH1 (1:20,000) antibodies (Figure 1C). Quantitative analysis of Coomassie stained gels using Fiji software, showed that dBigH1ΔED::FLAG accounted for ∼23% of total linker histones. The levels of full-length dBigH1::FLAG could not be directly determined since it overlapped with an unrelated species present in control non-induced cells. Similarly, dBigH1ΔNTD::FLAG expression levels could neither be determined directly since it overlapped with dH1 (Supplementary Figure S7B). However, considering that chromatin abundance of the full-length dBigH1::FLAG and the truncated dBigH1ΔNTD::FLAG forms was similar to that of dBigH1ΔED::FLAG (Figure 1B and Supplementary Figure S7A), we estimated that all three forms were expressed to a similar extent, accounting for ∼20–25% of total linker histones.

For expression in salivary glands homozygous flies carrying the various dBigH1 constructs described above were crossed to homozygous *da*-GAL4 flies at 25°C.

ChIP experiments

For ChIP experiments, 10^8 cells were collected and fixed with 1.8% of formaldehyde for 10 min at room temperature by gently mixing. Cross-linking was stopped adding glycine to a final concentration of 125mM. After 5 min, cells were spun down for 2 min at 1500g and washed with 5 ml of PBS. Then, cells were resuspended in 10 ml of 10 mM HEPES pH 7.9, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100 and incubated for 10 min in a wheel at 4°C. Cells were spun down again and resuspended in 10 ml of 10 mM HEPES pH 7.9, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100, incubated for 10 min in a rotating wheel at 4°C and spun again for 2 min at 1500g. Later, cells were lysed in 5 ml of TE (10 mM Tris–HCl pH 8, 1 mM EDTA) with 1% SDS. Chromatin was washed three times with 5 ml of TE, resuspended in TE, 1 mM PMSF, 0.1% SDS and sonicated in a Bioruptor Pico sonifier (Diagenode) to obtain fragments of 200–500 bp. After sonication, lysates were adjusted to 1% Triton X-100, 0.1% sodium deoxycholate (DOC), 140 mM NaCl, incubated for 10 min in a rotating wheel at 4°C and chromatin was recovered by centrifugation. For each experiment, 400 μl of chromatin was used for the immunoprecipitation (IP) and 40 μl for the input sample. IPs were carried out in RIPA buffer (140 mM NaCl, 10 mM Tris–HCl pH 8, 1 mM EDTA, 1% Triton X-100, 0.01% SDS, 0.1% DOC). A pre-clearing step was carried out in a rotating wheel for 1 h at 4°C with 30 μl of 50% (v/v) protein A-Sepharose CL4B beads (GE Healthcare, 17-0780-01) previously blocked with RIPA–1%BSA. Then, the corresponding antibodies were added and incubated overnight at 4°C in a rotating wheel. IPs were performed by adding 40 μl of 50% (v/v) protein A-Sepharose CL4B beads previously blocked with RIPA–1%BSA and incubating in a rotating wheel for 3 h at 4°C. Beads were washed five times for 5 min in 1 ml of RIPA buffer, once for 5 min in 250 mM LiCl, 10 mM Tris–HCl pH 8, 1 mM EDTA, 0.5% NP-40, 0.5% DOC and twice for 5 min in TE. Beads were then resuspended in 40 μl of TE, DNase-free RNaseA was added at 0.25 μg/ml to the IPs and input samples, and incubated for 30 min at 37°C. To purify the DNA, samples were adjusted to 1% SDS, 0.1 M $NaHCO_3$, 0.2 mg/ml Proteinase K and incubated overnight at 65°C. DNA was purified by phenol-chloroform extraction. Antibodies used were αdBigH1 (1 μl), αH1 (1 μl), αRpb3 (3.5 μg), αPol IIo^{ser5} (2 μg), αH3K36me3 (2 μg) and αH3Kac (2 μg).

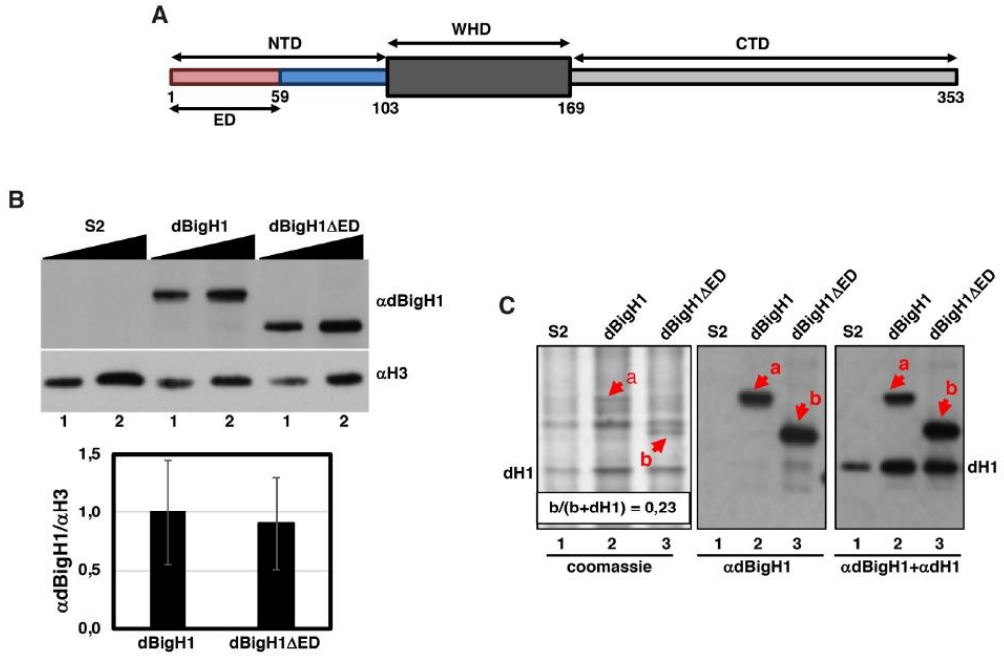


Figure 1. Expression of dBigH1 constructs used in these experiments. (**A**) Schematic representation of the domain organization of dBigH1. NTD, N-terminal domain; WHD, winged helix domain; CTD, C-terminal domain; ED, acidic domain. Numbers indicate aa position in the primary sequence. (**B**) The abundance of the indicated dBigH1 constructs in crosslinked chromatin prepared from control S2 cells and cells expressing full-length dBigH1::FLAG and the truncated dBigH1ΔED::FLAG form was determined by WB with αdBigH1 antibodies. Increasing amounts of chromatin are analyzed (lanes 1 and 2). αH3 antibodies were used for normalization. Quantitative analysis of the results is shown in the bottom ($N = 3$; error bars are SD; two tailed t-test, P-value = 0.247). (**C**) Total acid extracts prepared from control S2 cells (lane 1) and cells expressing full-length dBigH1::FLAG (lane 2) and the truncated dBigH1ΔED::FLAG (lane 3) forms were analyzed in Coomassie-blue stained gels (left) and by WB with αdBigH1 (center) and αdBigH1+αdH1 (right) of the gel in the left. Bands corresponding to dBigH1::FLAG (a) and the truncated dBigH1ΔED::FLAG (b) forms are indicated. The band corresponding to dH1 is also indicated. Note that dBigH1::FLAG overlaps with an unrelated species present in control S2 cells. The proportion of dBigH1ΔED::FLAG (band b) respect to total linker histones (b+dH1) is indicated.

For ChIP-seq experiments, all experiments were done in duplicate. Libraries were prepared from 3 to 10 ng of DNA using the NEBNext Ultra II DNA Library Prep kit for Illumina (NEB, E7645S) according to manufacturer's instructions. Libraries were then subjected to PCR amplification (10 cycles) and cluster generation for subsequent sequencing. For ChIP-qPCR experiments, IPs were analyzed according to the ΔΔCt method. Primers used in these experiments are described in Supplementary Table S1.

Expression profiling and RT-qPCR analysis

For expression analysis, total RNA was isolated from 10^7 cells using the RNeasy Mini Kit (Qiagen) and cDNAs were generated from 25 ng of total RNA and subjected to PCR amplification (17 cycles) using the TransPlex® Complete Whole Transcriptome Amplification Kit (Sigma, WTA2) according to manufacturer's instructions. 8 μg of amplified cDNAs were fragmented and labeled using the GeneChip Human Mapping 250K Nsp Assay kit (Affymetrix, 900766) according to manufacturer's instructions. Labeled cDNAs were then hybridized in GeneChip *Drosophila* Genome arrays 2.0 from Affymetrix (Thermo Fisher, 900531) for 16 h at 45°C. After incubation, the arrays were processed in the GeneChip Fluidics Station 450 from Affymetrix and scanned in GSC3000 System from Affymetrix.

For RT-qPCR analysis, total RNA was isolated from 10^7 cells using the RNeasy Mini Kit (Qiagen). For cDNA synthesis, 1 μg of total RNA was used and qPCRs were run in triplicate in at least 4 independent experiments. Expression data were normalized to Rpl32 and analyzed using ΔΔCt method. Primers used in these experiments are described in Supplementary Table S1.

Bioinformatics and biostatistics analyses

Quality control of raw ChIP-Seq data was assessed with the FastQC tool version 0.11 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Then, FastQ files were aligned against the dm3 genome using Bowtie 0.12.5 (19), allowing for two mismatches (-n 2) and reporting only unique hit alignments (-m 1). Afterwards, potential PCR over-amplification artifacts were assessed and removed using sambamba 0.5.1 (20). Number of uniquely

aligned reads were checked to be within the ENCODE recommended guidelines (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3431496/). A summary of these ChIP-seq quality control data is presented in Supplementary QC Data. Binary tracks for samples in TDF format for visual assessment of sample signal were generated with IGVTools 2 (21). In order to determine whole genome distribution, peak calling for IP vs their respective input sample was performed with MACS 1.4.2 (22), setting the option –g dm to account for *Drosophila melanogaster* genome size, and leaving the rest of options as default. Unless otherwise specified, all downstream analyses were performed using R3.4.4 and Bioconductor (23). Additional assessment of immunoprecipitation, group separation and biological replicate correlation was performed via generation of Gini/SSD coverage inequality indexes, PCA-like MDS plots and coverage correlation heatmaps using htSeqTools 1.26.0 (24). Lorenz curves for assessing coverage distribution in the analyzed samples were generated with the ineq package version 0.2–13 using genome bins of 1000bp. Per-sample gene level quantification in $\log_2$RPKM and $\log_2$RPKM(IP/Input) ratios were generated with the countOverlaps function from the GenomicRanges package version 1.30.3, and using the UCSC *Drosophila melanogaster* dm3 annotation. Raw metagene profiles were generated from normalized coverage ratio values with the regionsCoverage function from the htSeqTools package. Whole-genome epigenetic profile Multidimensional Scaling maps were generated with the chroGPS package version 2.0 (25,26). The pausing index (PI) of total RNApol II (Rpb3) was calculated as described elsewhere (27,28). In brief, Rpb3 occupancy (average $\log_2$RPKM(IP/input) ratio between replicates) was computed at the TSS-region (defined as TSS ± 250 bp) and the CDS (defined as TSS + 250 bp to TTS – 250 bp) for each gene longer than 1 kb and, then, the PI was determined as the TSS/CDS ratio of Rpb3 occupancy.

Affymetrix data from Drosophila2 arrays for dBigH1-expressing and control mock-induced cells were normalized with R and Bioconductor using RMA background correction, quantile normalization and RMA summarization to obtain probeset expression estimates (29). This method of data normalization is based on the assumption that there is not a general change in gene expression between the analyzed experimental conditions, with most genes suffering no big expression changes. In our case, raw intensity values before RMA did not present a global shift in expression in a consistent manner, except for a weak global decrease detected in one of the replicates (Supplementary QC Data). Next, we used limma (29) to determine differentially expressed probesets in dBigH1-expressing vs control cells, using a Benjamini-Hochberg adjusted $P < 0.1$ and a $|FC| > 1.5$. The GSEA pre-ranked algorithm (30) was used to identify significantly enriched and depleted KEGG pathways as provided by the org.Dm.eg.db package (November 2014), using all probesets in the *Drosophila* genome ranked by mean $\log_2$FC, summarized at gene level using the annotation from the drosophila2.db package version 2.8.1.

ChIP-seq and arrays data are deposited at NCBI GEO (GSE127227 and GSE103292).

NRL determination

For NRL determination, 2×10^7 cells of each condition were grown and treated with 1 mM $CuSO_4$ for 24 h. The cells were collected washed twice with PBS and fixed with 1.1% of formaldehyde for 10 min at room temperature by gently mixing. Cross-linking was stopped adding glycine to a final concentration of 125 mM. After 5 min, cells were spun down for 8 min at 1000g at 4°C and washed twice with 10 ml of cold PBS before pelleting and flash freezing in liquid nitrogen. For MNase digestion, the pellet was resuspended in PBS, 0.1% Triton X-100 (PBS-TX). The digestion was performed at 37°C in a volume of 400 μl PBS–TX with 4×10^6 cells per digestion time. $CaCl_2$ was adjusted to 1 mM, cells were pre-warmed and, after adding 0.8 U of MNase (Sigma-Aldrich), incubated for increasing time from 30 s to 5 min. Digestion was stopped in ice by adding EDTA and EGTA to a final concentration of 20 mM each. To purify the DNA, samples were adjusted to 10 mM Tris HCl pH 8, 0.4% SDS, 0.4 mg/ml Proteinase K and incubated overnight at 65°C. DNA was extracted by phenol-chloroform and precipitated with ethanol. Later the samples were treated with RNase A for 30 min at 37°C, ran in 2% agarose gels at 90 V for 5 h and stained with ethidium bromide. Images were analyzed using Fiji software and the size of each fragment was determined from the maximum of the corresponding size distribution using MW markers. To determine the apparent NRL, the size of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes, was plotted against the number of nucleosomes and the apparent NRL calculated from the slope of the corresponding regression curve. To compare NRLs of the different conditions, samples showing similar extent of digestion were analyzed.

In vitro chromatin reconstitution and transcription experiments

Drosophila embryo extract DREX was prepared in Exb50 buffer (10 mM HEPES pH 7.6, 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA pH 8, 10 mM β-glicerophosphate, 10% glycerol) from preblastodermic embryos, as described in (31). To prepare dBigH1-depleted DREX, αBigH1 antibodies were coupled to protein A magnetic beads and incubated with DREX for five rounds of 45 min each at 4°C. The extent of depletion was determined by WB (Supplementary Figure S4B). Chromatin reconstitution experiments were performed as described in (31), using 600 ng of a pAc5.1-V5-His plasmid (Invitrogen) carrying an EGFP reporter gene under the control of the *Drosophila actin5C* promoter. After reconstitution, chromatin was precipitated with 15 mM of $MgCl_2$ and resuspended in Exb50. The extent of chromatin assembly was determined by MNase digestion as described in (31). For *in vitro* transcription assays, 12 μl of the reconstituted chromatin were incubated for 60 min at 30°C with 10 μl Buffer-C_{90} (20 mM Tris–HCl pH 7.8, 1 mM EDTA, 1 mM DTT, 10% glycerol, 90 mM NaCl), 20 μl Rxn mix (20 mM Tris pH 8.3, 5 mM $MgCl_2$, 3 mM DTT, 25 mM rNTP mix, 6.25% PEG 8000) and 14 μl of HeLa nuclear extract (CILBiotech) as described in (15). Transcription was stopped by incubation at 39°C

for 15 min in stop-mix (200 mM NaCl, 20 mM EDTA, 1% SDS, 12 μl glycogen, 0.12 μg/μl Proteinase K). Then, RNA was purified using the RNeasy Mini Kit (Qiagen) and cDNA was prepared using the Transcriptor First Strand cDNA Synthesis Kit (Roche) with oligo $(dT)_{18}$ primers. The amount of GFP mRNA was quantified by RT-qPCR using primers against GFP (Supplementary Table S1) and normalized with respect to the total amount of chromatin template used in the assay determined by qPCR with the same primers before transcription. For each experimental condition, three independent biological replicates assembled in the same dBigH1-depleted DREX (Supplementary Figure S4B) were analyzed.

Immunostaining experiments in salivary glands

For immunostaining, salivary glands from L3 larvae were dissected in Cohen solution (10 mM $MgCl_2$, 25 mM glycerol-3-phosphate, 3 mM $CaCl_2$, 10 mM KH_2PO_4, 0.5% NP-40, 30 mM KCl, 160 mM sucrose) and fixed in 0.74% formaldehyde in PBS for 2 min. Glands were then incubated for 3 min and squashed in 45% acetic acid, 0.62% formaldehyde. Preparations were washed for 5 min in PBS-T (PBS, 0.05% Tween 20) and incubated with αdBigH1 (1:400) antibodies in PBS-T, 1%BSA overnight at 4°C. Preparations were washed three times for 5 min in PBS-T and incubated with the secondary antibody for 2 h at room temperature. Slides were mounted in Mowiol (Calbiochem-Novabiochem) containing 0.2 ng/μl DAPI (Sigma) and visualized on a Nikon Eclipse E1000 microscope.

RESULTS

Ectopically expressed dBigH1 binds across chromatin and down-regulates gene expression

To ectopically express dBigH1 in *Drosophila* S2 cells we used a stable cell line expressing a dBigH1::FLAG construct under the control of a Cu^{2+}-inducible promoter (Figure 1) (7). Upon induction, dBigH1 accounted for ∼23% of total linker histones (Figure 1B and C) (see Materials and Methods). ChIP-seq analyses using αdBigH1 antibodies showed that ectopically expressed dBigH1 bound broadly across chromatin since similar genomic inequality Lorenz' curves were observed in the immunoprecipitated (IP) and input samples (Figure 2A). ChIP-qPCR experiments confirmed these results since dBigH1 was detected at multiple randomly selected genomic regions, including repetitive DNA elements (Figure 2B). We observed that, though moderately, the genomic distribution of ectopically expressed dBigH1 in S2 cells positively correlated with those observed in embryos (6) (Pearson's correlation coefficient, $r = 0.239$) and testes (7) (Pearson's correlation coefficient, $r = 0.331$) (Supplementary Figure S1). Interestingly, in S2 cells, deposition of ectopically expressed dBigH1 strongly correlated with dH1 content in control non-induced cells since genes containing high dBigH1 levels in induced cells corresponded to genes with high dH1 content in non-induced cells (Pearson's correlation coefficient, $r = 0.796$) (Figure 2C). Along genes, linker histones H1 are usually depleted in the region around the transcription start-site (TSS) of highly expressed genes and their occupancy increases progressively along the transcription unit (CDS) (8–14) (see also Supplementary Figure S9B). On the other hand, no such depletion is detected in low expressed genes that, overall, have higher H1 content than highly expressed genes (8–14) (see also Supplementary Figure S9B). The dBigH1 coverage profile showed similar features (Figure 2D), though, in comparison to dH1 (Supplementary Figure S9B), the depletion at TSS in highly expressed genes, as well as the differences in dBigH1 content between high and low expressed genes, were less pronounced. We also performed mock ChIP-seq experiments with αdBigH1 antibodies in control non-induced cells to assess the background level of non-specificity. Although, like in induced cells, genomic inequality of the mock IP samples was similar to the inputs (Supplementary Figure S2A), the background αdBigH1 coverage profile along genes was markedly different (Supplementary Figure S2B). ChIP-qPCR experiments confirmed these results since, in comparison to the signal observed in induced cells, background αdBigH1 signal in control non-induced cells was equally negligible at both TSS and CDS of several genes (Supplementary Figure S3A). ChIP-qPCR experiments also showed that, at TSS of highly expressed genes, the occupancy of dBigH1 was higher in comparison to dH1 and similar to the occupancy at CDS (Supplementary Figure S3B, left). Instead, in low expressed genes, dH1 and dBigH1 showed similar high relative occupancy at TSS (Supplementary Figure S3B, right). ChIP-qPCR experiments also confirmed that dBigH1 occupancy was similar in high and low expressed genes (Supplementary Figure S3C, center), while dH1 was generally more abundant in low expressed genes (Supplementary Figure S3C, left).

Next, we performed expression-profiling analyses to determine the effects of dBigH1 on gene expression. Upon dBigH1 expression, we detected 630 differentially expressed (DE) genes at fold-change $|FC|>1.5$ (Supplementary Table S2), accounting for ∼5% of the genes present in the array. Considering all genes, the proportion of up-regulated and down-regulated genes was roughly the same (9669 versus 9283 probesets, respectively) (Figure 3A). However, within the subset of DE genes, ∼68% were down-regulated (Figure 3A) and, overall, a global down-regulation was observed at $|FC|>1.5$ (Figure 3B). RT-qPCR experiments confirmed down-regulation of six randomly selected DE genes showing $FC < -1.5$ (Figure 3D, left) and, though weakly, the extent of down-regulation of the differentially down-regulated genes significantly correlated with their dBigH1 content (Pearson's correlation coefficient, $r = 0.11$; P-value = 0.0227) (Figure 3C). Regarding differentially up-regulated genes, RT-qPCR experiments confirmed up-regulation of three out of four genes tested (Figure 3D, right). However, up-regulation did not significantly correlate with dBigH1 content (Pearson's correlation coefficient, $r = 0.09$; P-value = 0.1737) (Figure 3C). Furthermore, permutation test analysis showed that the extent of up-regulation was significantly lower than the extent of down-regulation (permutation test; P-value = 0.007, $B = 5000$).

In vitro chromatin reconstitution experiments using *Drosophila* preblastodermic extracts (DREX) confirmed down-regulation induced by dBigH1. DREX has been extensively used in chromatin reconstitution experiments

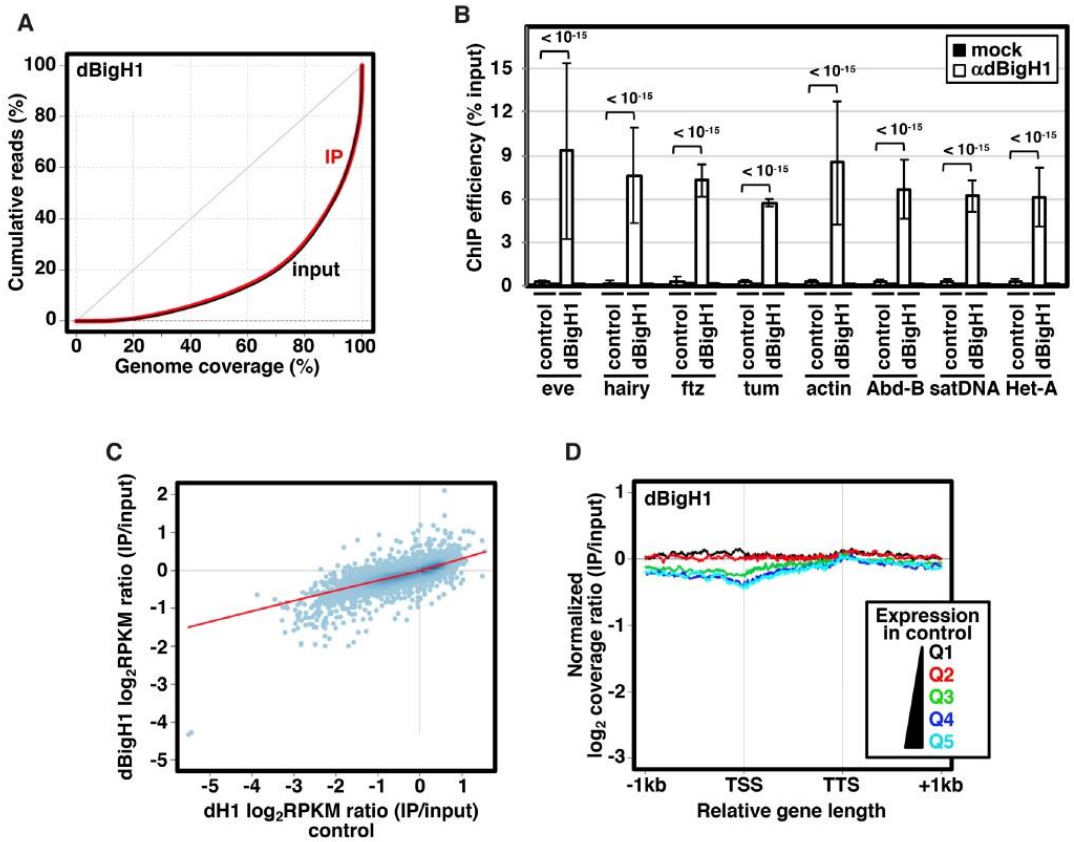


Figure 2. Ectopic dBigH1 expression in S2 cells results in its binding across chromatin. (**A**) Lorenz curves showing coverage inequality for IP (red) and input samples (black) of dBigH1 ChIP-seq analyses in induced cells. (**B**) ChIP-qPCR analyses with αdBigH1 antibodies and preimmune serum (mock) at the indicated genomic regions in control and dBigH1-expressing cells ($N = 2$; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted P-values are indicated). (**C**) Correlation of gene level dBigH1 content (log_2RPKM IP/Input) in dBigH1-expressing cells and dH1 content (log_2RPKM IP/Input) in control cells. (**D**) The normalized log_2 coverage ratio IP/input of dBigH1 in dBigH1-expressing cells is presented along an idealized gene-length ± 1 kb for genes longer than 1 kb categorized according to their expression quantile in control cells (Q1-lowest to Q5-highest). TSS, transcription start site; TTS, transcription termination site.

(32,33). DREX is enriched in dBigH1 (34) and chromatin reconstituted in DREX contains dBigH1 (Supplementary Figure S4A). To assess the effect of dBigH1 on transcription, a plasmid construct carrying a GFP-reporter was subjected to chromatin reconstitution in full DREX or after dBigH1 depletion using αdBigH1 antibodies (Supplementary Figure S4B), and the resulting reconstituted minichromosomes were used as template in *in vitro* transcription assays using HeLa extracts. MNase digestion showed that dBigH1-depleted DREX was as competent as full DREX for chromatin reconstitution (Supplementary Figure S4C). However, the GFP-reporter was transcribed to a significantly higher extent when the chromatin template was reconstituted in dBigH1-depleted DREX than in full or control mock-depleted DREX (Figure 3E). Furthermore, the addition of bacterially expressed recombinant dBigH1 during reconstitution with dBigH1-depleted DREX strongly reduced GFP transcription (Figure 3E).

We observed that, in comparison to non-DE or up-regulated genes down-regulated genes exhibited higher expression (Figure 3F) and stronger RNApol II pausing in control cells (Figure 3G). Gene ontology analyses showed that down-regulated genes mainly associated with KEGG pathways related to metabolic processes and housekeeping functions, while up-regulated ones correlate with more diverse functional pathways (Supplementary Figure S5A). We also generated whole-genome gene level epigenetic profile maps using Multidimensional Scaling with the chroGPS package (25,26). In these maps, all genes in the genome are represented in a two-dimensional space according to similarity between their epigenetic state and three main domains can be identified corresponding to active, Polycomb (PC)

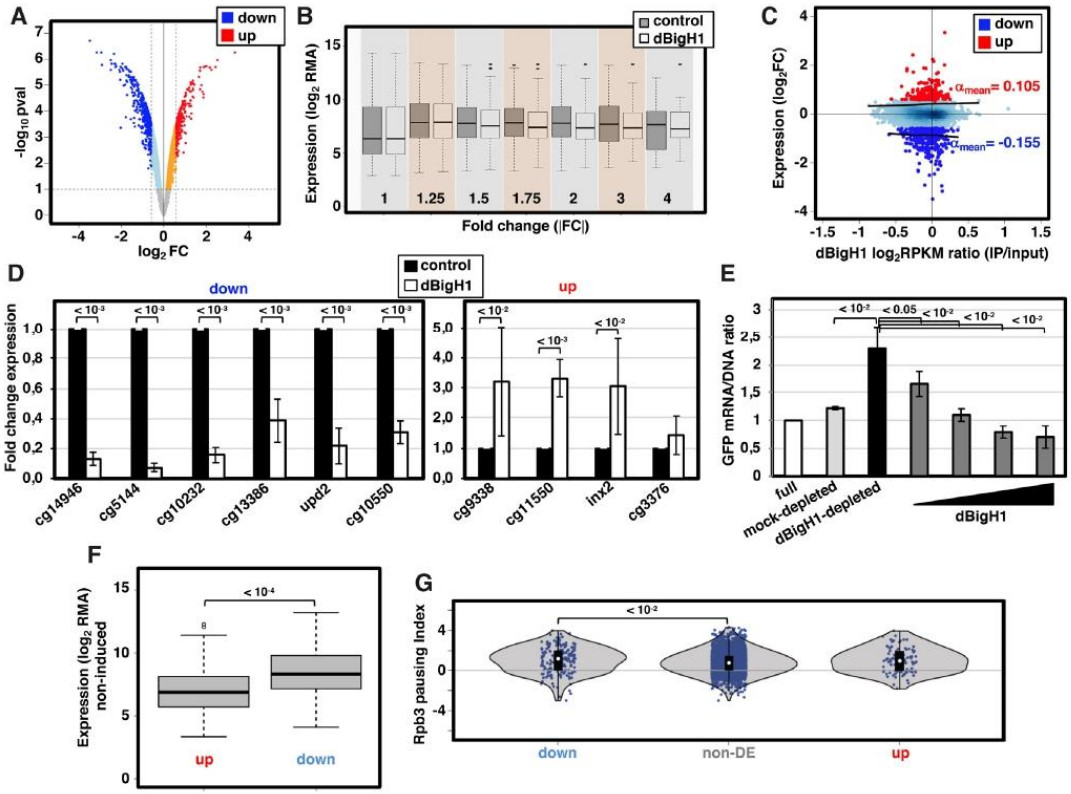


Figure 3. dBigH1 down-regulates gene expression. (**A**) Volcano plot showing the change in expression of each individual gene in cells expressing dBigH1 in comparison to control mock-induced cells. Differentially down-regulated and up-regulated genes (|FC| > 1.5; Benjamini-Hochberg adjusted *P*-value < 0.1) are indicated in blue and red, respectively. (**B**) The normalized gene expression (log_2RMA) in dBigH1-expressing cells and control mock-induced cells is presented for all genes classified according to their absolute fold-change (|FC|) expression upon dBigH1 expression. (**C**) Correlation of gene level dBigH1 content (log_2RPKM IP/input) and expression FC upon dBigH1 expression (log_2FC). Genes differentially up-regulated (red) and down-regulated (blue) are indicated. Non-DE genes are also indicated (light blue). Black lines are lowess regression fits for up-regulated and down-regulated genes; the average slopes (α_{mean}) are indicated. (**D**) The fold change expression with respect to control mock-induced cells of six randomly selected down-regulated genes (left) and four randomly selected up-regulated genes (right) was determined by RT-qPCR in dBigH1-expressing cells (white) ($N = 7$; error bars are SD; two-tailed t-test, *P*-values are indicated). (**E**) The relative GFP mRNA/DNA ratios of a chromatin template carrying a GFP-reporter gene assembled *in vitro* in full (white), mock-depleted (light grey) and dBigH1-depleted DREX in the absence (black) and presence of increasing amounts of bacterially-expressed dBigH1 added during assembly (dark grey) ($N = 3$; error bars are SD; two-tailed *t*-test, *P*-values are indicated). (**F**) The expression (log_2RMA) in control cells is presented for up-regulated and down-regulated genes (Wilcoxon rank sum test, *P*-value is indicated). (**G**) The pausing index of total RNApol II (Rpb3) in control cells is presented for down-regulated, non-DE and up-regulated genes (Wilcoxon rank sum test, *P*-values are indicated).

and HP1a chromatin (25). Down-regulated genes scattered throughout the map (Supplementary Figure S5B). Instead, we observed a lower proportion of up-regulated genes in the PC-chromatin domain (Supplementary Figure S5B).

The acidic ED-domain of dBigH1 is required for down-regulation

dBigH1 has the characteristic tripartite domain organization of linker histones, in which a central winged-helix domain (WHD) is flanked by unstructured N-terminal (NTD) and C-terminal (CTD) domains (35–37) (Figure 1A). The WHD and CTD of dBigH1 are relatively well-conserved with respect to somatic dH1 (6). In contrast, the NTD of dBigH1 contains an extra domain enriched in acidic E and D residues (6) (Figure 1A). As generally observed in linker histones (38,39), the positively charged CTD is required for dBigH1 binding to chromatin since, opposite to full length dBigH1, a truncated form missing the CTD did not bind to polytene chromosomes when ectopically expressed in salivary glands (Supplementary Figures S6A–C). On the other hand, truncated forms missing the acidic ED-domain or the entire NTD bound polytene chromosomes (Supplementary Figures S6D and E).

Next, we addressed the contribution of the extra ED-domain of dBigH1 to gene expression. For this purpose, we used a stable S2 cell line expressing a truncated dBigH1ΔED form, which was expressed to a similar level

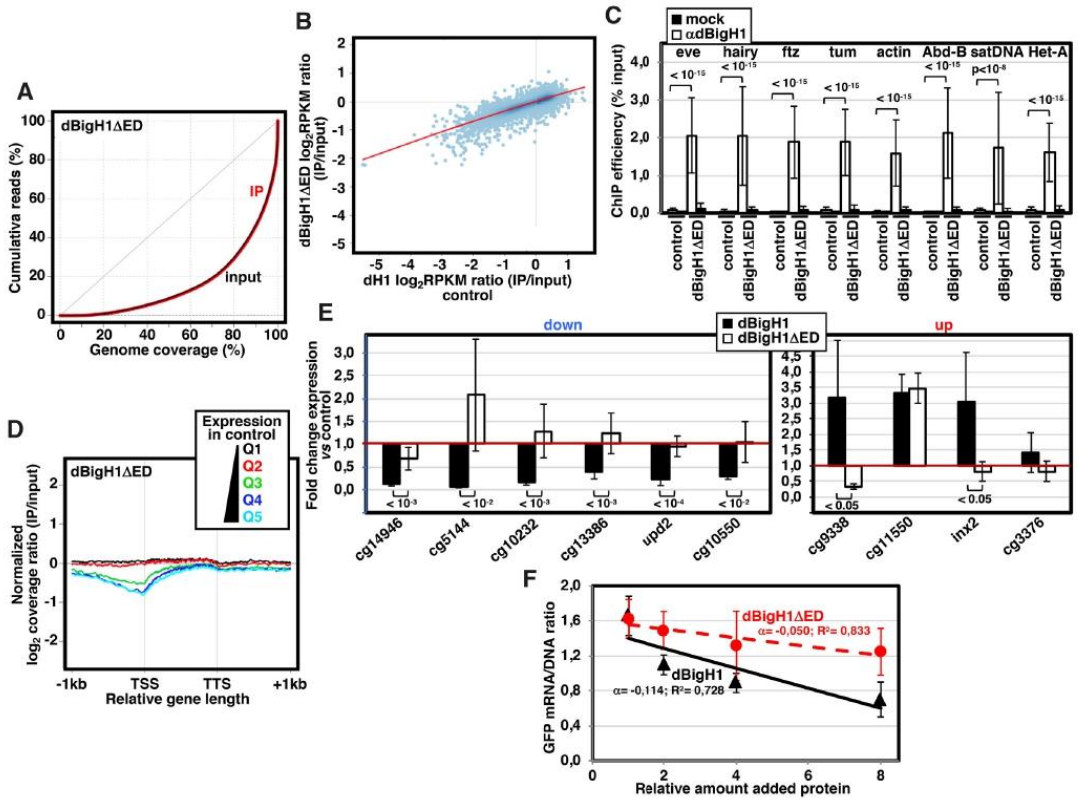


Figure 4. The acidic N-terminal ED-domain of dBigH1 is required for down-regulation. (**A**) Lorenz curves showing coverage inequality for IP (red) and input samples (black) of dBigH1ΔED ChIP-seq analyses in induced cells. (**B**) Correlation of gene level dBigH1ΔED content (log_2RPKM IP/Input) in dBigH1ΔED-expressing cells and dH1 content (log_2RPKM IP/Input) in control cells. (**C**) ChIP-qPCR analyses with αdBigH1 antibodies and preimmune serum (mock) at the indicated genomic regions in control and dBigH1ΔED-expressing cells ($N = 2$; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted *P*-values are indicated). (**D**) The normalized log_2 coverage ratio IP/input of dBigH1ΔED in dBigH1ΔED-expressing cells is presented along an idealized gene-length±1kb for genes longer than 1kb categorized according to their expression quantile in control cells (Q1-lowest to Q5-highest). TSS, transcription start site; TTS, transcription termination site. (**E**) The fold change expression with respect to control mock-induced cells of six randomly selected down-regulated genes (left) and four randomly selected up-regulated genes (right) was determined by RT-qPCR in dBigH1ΔED-expressing cells (white) and compared to dBigH1-expressing cells (black) ($N \geq 4$; error bars are SD; two-tailed *t*-test, *P*-values are indicated). (**F**) Inhibition curves showing the relative GFP mRNA/DNA ratio of a chromatin template carrying a GFP-reporter assembled *in vitro* in dBigH1-depleted DREX upon the addition of increasing amounts of bacterially-expressed dBigH1 (black) and dBigH1ΔED (red). The correlation coefficients (R^2) and slopes (α) are indicated (error bars are SD; $N = 3$).

than full-length dBigH1 (Figure 1B). ChIP-seq analyses showed that, similar to full-length dBigH1, dBigH1ΔED was binding across chromatin (Figure 4A) and its deposition correlated with dH1 content in control non-induced cells (Pearson's correlation coefficient, r = 0.863) (Figure 4B). ChIP-qPCR experiments confirmed binding of dBigH1ΔED across chromatin (Figure 4C). However, at the gene level, the coverage profile of dBigH1ΔED showed a deeper depletion at TSS of highly expressed genes in comparison to full-length dBigH1 (Figure 4D). Regarding the effects on gene expression, we observed that deletion of the ED-domain in dBigH1ΔED abolished the down-regulation observed with full-length dBigH1 (Figure 4E, left). Furthermore, addition of dBigH1ΔED during *in vitro* reconstitution experiments using dBigH1-depleted DREX inhibited GFP expression to a weaker extent than addition of full-length dBigH1, as judged from the slopes of the inhibition curves obtained upon the addition of doubling amounts of the corresponding proteins (Figure 4F). Altogether these results suggest that the acidic ED-domain is required for down-regulation induced by dBigH1. The effects on up-regulated genes were less consistent (Figure 4E, right). Deletion of the entire NTD caused similar effects (Supplementary Figure S7).

dBigH1 interferes with RNApol II binding

Next, we performed ChIP-seq experiments with αRpb3 antibodies to determine the effects of dBigH1 binding on the genomic distribution of RNApol II. We observed that, in

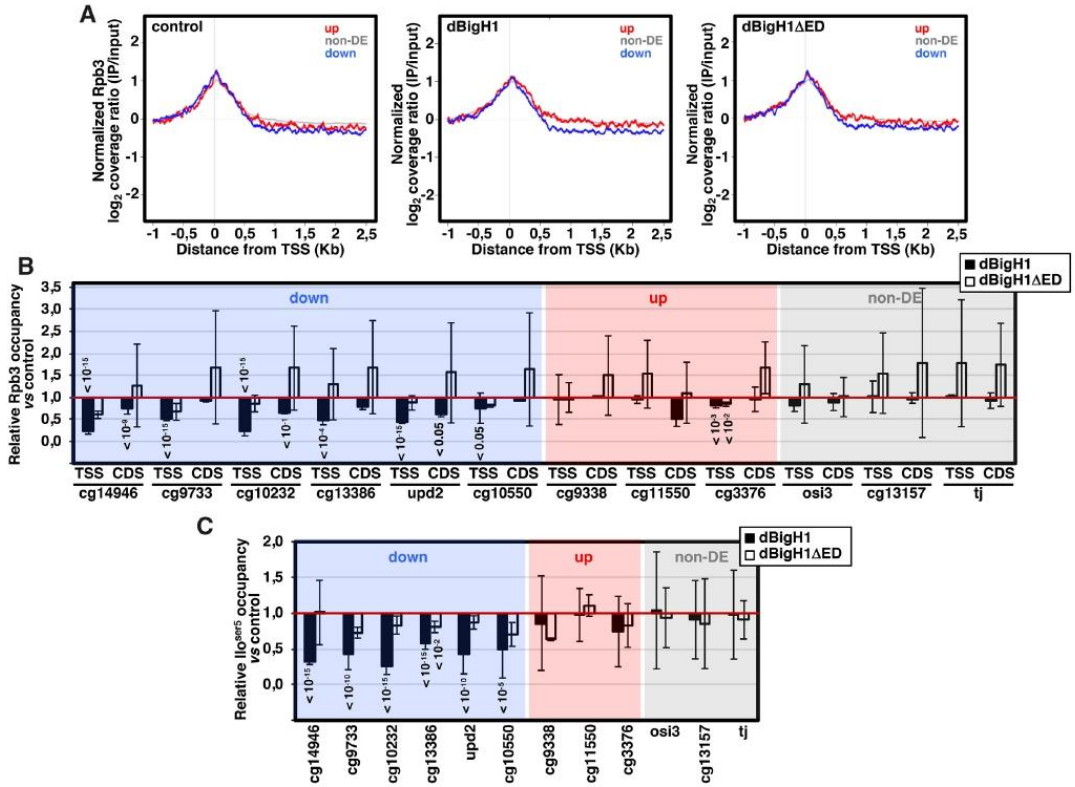


Figure 5. The acidic N-terminal ED-domain of dBigH1 perturbs RNApol II binding. (**A**) The normalized $\log_2$ coverage ratio IP/input of Rpb3 in control mock-induced cells (left), dBigH1-expressing cells (center) and dBigH1ΔED-expressing cells (right) are presented as a function of the distance to TSS for up-regulated genes (red), down-regulated genes (blue) and non-DE genes (gray). (**B**) The relative Rpb3 occupancy with respect to control mock-induced cells was determined by ChIP-qPCR at the indicated genomic regions in dBigH1-expressing cells (black) and dBigH1ΔED-expressing cells (white) ($N = 2$; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted *P*-values with respect to mock-induced cells are indicated). (**C**) As in B but for the promoter-proximal RNApol IIo^{ser5} form at TSS of the indicated genes (N = 3 (dBigH1), 2 (dBigH1ΔED); error bars are SD; Mixed linear model Benjamini–Hochberg adjusted *P*-values with respect to mock-induced cells are indicated).

comparison to non-DE and up-regulated genes, RNApol II coverage along down-regulated genes was reduced upon dBigH1 expression (Figure 5A). Actually, in comparison to control mock-induced cells, total RNApol II content of down-regulated genes tended to decrease upon dBigH1 expression (Supplementary Figure S8A). These defects depended on the acidic ED-domain since they were not observed when the truncated dBigH1ΔED form was expressed (Figure 5A and Supplementary Figure S8A). ChIP-qPCR experiments confirmed these results since, upon dBigH1 expression, several randomly selected down-regulated genes showed reduced Rpb3 occupancy at TSS and/or CDS, whereas expression of dBigH1ΔED showed no such general reduction (Figure 5B). Similar results were obtained when occupancy at TSS of the promoter-proximal active RNApol IIo^{ser5} form was analyzed (Figure 5C). In the down-regulated genes, RNApol II occupancy was generally more reduced at TSS than at CDS (Figure 5B), suggesting an effect on RNApol II pausing. In this regard, ChIP-seq analysis showed that the pausing index (PI) of down-regulated genes tended to decrease upon dBigH1 expression (Supplementary Figure S8B). On the other hand, dBigH1 expression did not affect RNApol II occupancy in non-DE and up-regulated genes (Figure 5B and C, and Supplementary Figure S8A).

dBigH1 replaces somatic dH1 and perturbs the pattern of histone modifications

WB analysis showed that the levels of chromatin-bound dH1 decreased upon expression of full-length dBigH1 or the truncated dBigH1ΔED form (Figure 6A). This reduction occurred broadly across chromatin since ChIP-seq analyses performed with αdH1 showed similar Lorenz' inequality curves for the IP samples obtained from control cells and cells expressing dBigH1 or dBigH1ΔED (Supplementary Figure S9A). Consistently, ChIP-qPCR experiments showed that dH1 occupancy tended to de-

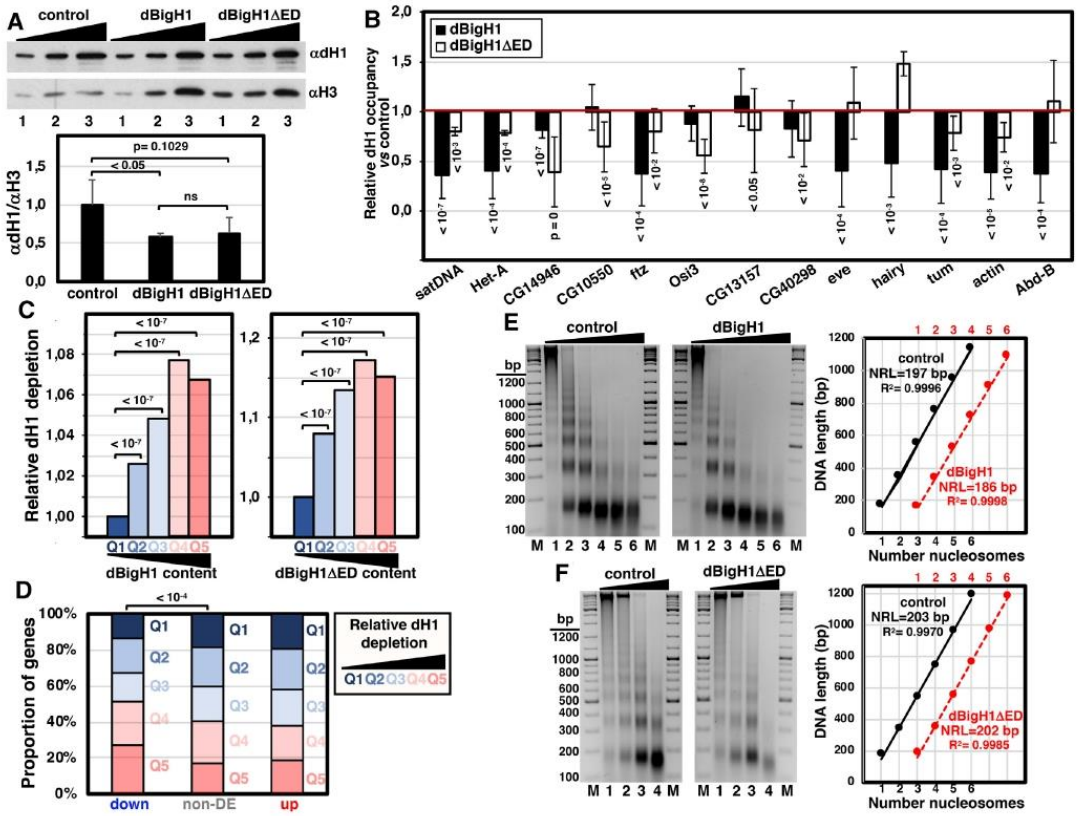


Figure 6. dBigH1 replaces dH1. (**A**) WB analysis with αdH1 antibodies of crosslinked chromatin prepared from control mock-induced cells and cells expressing full-length dBigH1 and the truncated dBigH1ΔED forms. Increasing amounts of chromatin are analyzed (lanes 1–3). αH3 antibodies were used for normalization. Quantitative analysis of the results is shown in the bottom (N = 3; errors bars are SD; two tailed *t*-test, *P*-values are indicated). (**B**) The relative dH1 occupancy with respect to control mock-induced cells was determined by ChIP-qPCR at the indicated genomic regions in dBigH1-expressing cells (black) and dBigH1ΔED-expressing cells (white) ($N = 3$; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted *P*-values with respect to mock-induced cells are indicated). (**C**) The average extent of dH1 depletion is presented for all genes categorized according to their dBigH1 (left) and dBigH1ΔED (right) content quantile in dBigH1-expressing and dBigH1ΔED-expressing cells (Q1-lowest to Q5-highest), normalized respect to Q1 (Wilcoxon rank-sum test, Benjamini–Hochberg adjusted *P*-values are indicated). (**D**) The proportion of genes categorized according to their relative dH1 depletion (Q1-lowest to Q5-highest) is presented for down-regulated, up-regulated and non-DE genes (Chi-square test, *P*-values are indicated). (**E**) MNase digestion for increasing time (lanes 1–6) of nuclei obtained from control mock-induced cells (left) and dBigH1-expressing cells (right). Lanes M correspond to MW markers (the sizes in bp are indicated). Quantitative analysis of the results is shown in the right where the size in bp of fragments containing increasing number of nucleosomes, from mono- to hexanucleosomes, are plotted against the number of nucleosomes for samples showing equivalent extent of digestion (lanes 2 in the left). The correlation coefficients (R^2) and slopes, which correspond to the apparent NRL, are indicated. (**F**) As in E but for control mock-induced cells (left) and dBigH1ΔED-expressing cells (right).

crease at multiple genomic locations, including repetitive DNA elements (Figure 6B). Analysis of ChIP-seq data showed that the extent of dH1 depletion increased as a function of increasing dBigH1/dBigH1ΔED content. For these analyses, we grouped genes into five quantiles according to their dBigH1/dBigH1ΔED content and determined for each quantile the average dH1 content fold-change upon dBigH1/dBigH1ΔED expression. We observed that genes having high dBigH1/dBigH1ΔED content showed higher dH1 depletion relative to genes with low dBigH1/dBigH1ΔED content (Figure 6C). Furthermore, we observed that, in comparison to non-DE and up-regulated genes, down-regulated ones were enriched in genes showing high dH1 depletion upon dBigH1 expression (Figure 6D). Altogether these results suggest that both dBigH1 and dBigH1ΔED replace somatic dH1 to a similar extent. In this regard, in comparison to control mock-induced cells, we observed that dBigH1 expression reduced the overall nucleosome repeat length (NRL) by ∼10 bp (Figure 6E; see also Supplementary Figure S10A), while it was not altered upon dBigH1ΔED expression (Figure 6F; see also Supplementary Figure S10B). In addition, at the gene level, the relative dH1 depletion at TSS of highly expressed genes was reduced in dBigH1-expressing cells, in

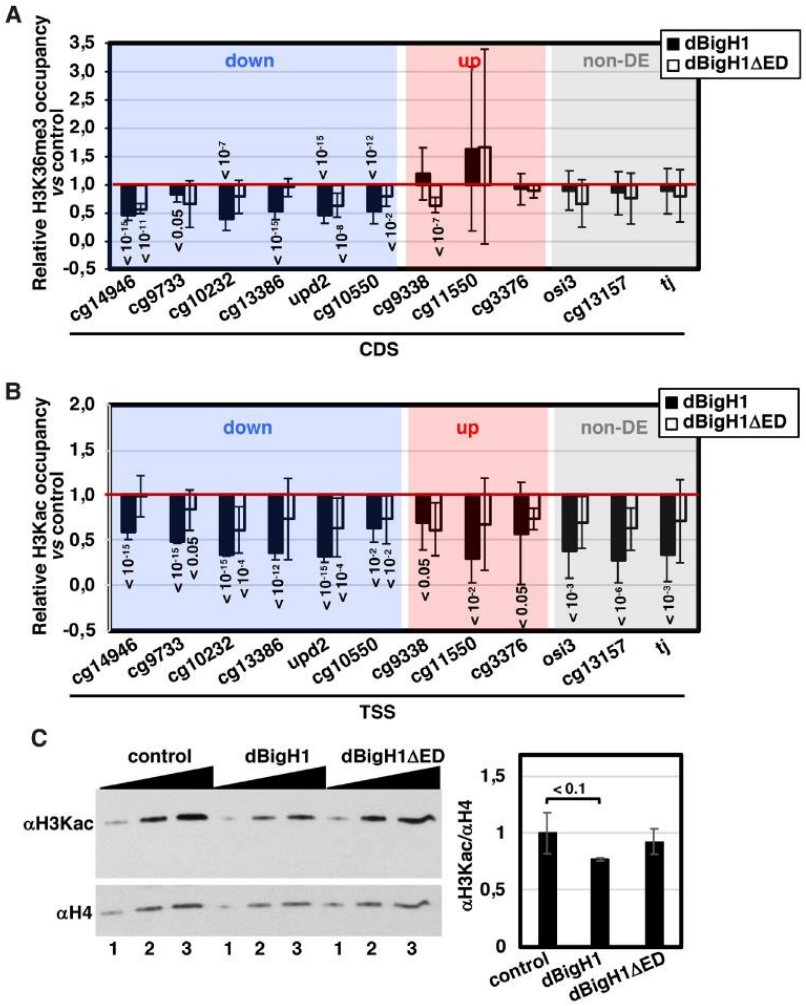


Figure 7. dBigH1 perturbs the pattern of histone modifications. (**A**) The relative H3K36me3 occupancy with respect to control mock-induced cells was determined by ChIP-qPCR at CDS of the indicated genes in dBigH1-expressing cells (black) and dBigH1ΔED-expressing cells (white) (N = 3; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted *P*-values with respect to mock-induced cells are indicated). (**B**) The relative H3Kac occupancy with respect to control mock-induced cells was determined by ChIP-qPCR at TSS of the indicated genes in dBigH1-expressing cells (black) and dBigH1ΔED-expressing cells (white) (*N* = 2; error bars are SD; Mixed linear model Benjamini-Hochberg adjusted *P*-values with respect to mock-induced cells are indicated). (**C**) In the left, WB analysis with αH3Kac antibodies of crosslinked chromatin prepared from control mock-induced cells and cells expressing full-length dBigH1 and the truncated dBigH1ΔED forms. Increasing amounts of chromatin are analyzed (lanes 1–3). αH4 antibodies were used for normalization. Quantitative analysis of the results is shown in the right (*N* = 3; error bars are SD; two tailed *t*-test, *P*-values are indicated).

comparison to control and dBigH1ΔED-expressing cells (Supplementary Figure S9B).

Transcriptional activity correlates with specific patterns of histone modifications (reviewed in (40)). In particular, transcriptionally active genes are enriched in H3K36me3 along the CDS, a modification associated with transcription elongation. In this regard, we observed that dBigH1 expression reduced H3K36me3 levels at CDS of down-regulated genes (Figure 7A). This effect correlated well with the effect on transcription since it was not observed in non-DE and up-regulated genes, and it was much weaker when dBigH1ΔED was expressed (Figure 7A). Histone K-acetylation is also a hallmark of active genes that generally occurs at promoters and correlates with transcription initiation. Similar to the effect on H3K36me3, we observed that expression of dBigH1 reduced H3 K-acetylation (H3Kac) levels at promoters of down-regulated genes (Figure 7B). However, in contrast to the effect on H3K36me3, this ef-

fect was also observed at promoters of non-DE and up-regulated genes (Figure 7B). Furthermore, WB analyses showed that total H3Kac levels decreased upon dBigH1 expression (Figure 7C). Altogether these results suggest that dBigH1 expression impaired H3Kac globally across chromatin. This effect mostly depended on the acidic ED-domain since dBigH1ΔED expression did not significantly decrease total H3Kac levels (Figure 7C) and, in comparison to full-length dBigH1, it had a weaker effect on H3Kac levels at promoters (Figure 7B).

DISCUSSION

Here, we have addressed the mechanism of dBigH1 action in transcription regulation using ectopic dBigH1 expression experiments in S2 cells. Though weakly, the genomic distribution of ectopically expressed dBigH1 positively correlates with those observed in embryos and testes, where dBigH1 is naturally expressed, suggesting that the mechanisms governing dBigH1 deposition might be partially conserved in S2 cells. Our results suggest that binding of dBigH1 negatively affects transcription. Upon dBigH1 expression, more than two-thirds of the DE-genes were down-regulated. This effect was probably underestimated since, though only in one replicate, we observed a global decrease in gene expression that, considering the methodology used for normalization, could hamper identification of differentially down-regulated genes. This down-regulation occurred at the transcriptional level, as down-regulated genes showed reduced RNApol II content. Conversely, RNApol II content of up-regulated genes was not increased upon dBigH1 expression, suggesting that the observed up-regulation was not transcriptional. Moreover, *in vitro* experiments showed that dBigH1 inhibited transcription of a chromatin template. Consistent with the negative effect on transcription, dBigH1 expression specifically decreased H3K36me3 levels at CDS of down-regulated genes.

Our results show that dBigH1 replaces dH1. In our experiments, dBigH1 binding to chromatin was likely taking place in the absence of DNA replication, as dBigH1 induction was sustained for 24h and, during this time, cell density did not increase noticeably. Whether dBigH1 deposition involves active dH1 replacement remains to be determined. *In vitro,* incubation of purified nuclei with DREX results in binding of dBigH1 to chromatin without dH1 displacement (32), suggesting that the replacement observed in S2 cells responds to an active process. Along the same lines, we observed that dBigH1 was preferentially deposited at regions enriched in dH1. Replacement of somatic H1s by embryonic H1s has been reported in nuclear transfer experiments (41–45) and NAP-1 has been shown to be involved in both B4/H1M deposition and somatic H1s removal in *Xenopus* (46,47). Further work is required to determine the mechanisms regulating dBigH1 deposition.

The acidic ED-domain of dBigH1 is required to inhibit transcription since expression of the truncated dBigH1ΔED form, which also replaced somatic dH1, did not down-regulate gene expression either affected RNApol II loading or H3Kac levels. The presence of the negatively charged acidic ED-domain in dBigH1 is peculiar as histones are highly positively charged. It is possible that, due to the negative charge of the ED-domain, the structural organization of chromatin is compromised in the presence of dBigH1. Actually, the overall NRL changed upon dBigH1 expression, but not when dBigH1ΔED was expressed. Interestingly, although the ED-domain of dBigH1 is not conserved outside of the *Drosophila* genus (6), embryonic H1s are generally more acidic than somatic ones (reviewed in (2)). In this regard, it was shown that both the *Xenopus* B4/H1M and the mammalian H1oo embryonic linker histones alter chromatin organization and dynamics (41,44,48–53). An altered chromatin organization would perturb access to chromatin and/or functioning of chromatin remodelers/modifiers and transcription factors that, ultimately, would affect RNApol II loading and transcription. In fact, regardless of the actual transcriptional outcome, dBigH1 expression globally affected H3Kac. In contrast, we reported earlier that incubation of purified nuclei with DREX, which also results in dBigH1 binding, increased H3Kac levels (32). However, it is important to note that the increase in H3Kac levels observed in this case was independent of dBigH1 binding (32).

It might be argued that the down-regulation observed upon dBigH1 expression is a consequence of increased global linker histones content. However, similar or even higher levels of expression of the truncated dBigH1ΔED and dBigH1ΔNTD forms did not affect transcription. Moreover, binding of dBigH1 is compensated by removal of dH1, thus total linker histones content is not greatly increased.

dBigH1 binding affected expression of a relatively small subset of genes. This may reflect the fact that in our experimental setup, dBigH1 accounted for only 20–25% of total linker histones. Thus, from this point of view, affected genes appear to correspond to a subset of genes more sensitive to dBigH1 levels. In this regard, we observed that down-regulated genes had strong RNApol II pausing, which tended to decrease upon dBigH1 expression. In addition, though dBigH1 binding reduced H3Kac globally, only the down-regulated genes were transcriptionally affected. Interestingly, impairing RNApol II pausing generally down-regulates gene expression (27,54,55), while reduced H3Kac levels preferentially affects expression of highly paused genes (56). These observations suggest that the higher sensitivity to dBigH1 expression of down-regulated genes is likely due to the way their transcription is regulated.

In summary, we have presented here direct evidence supporting that the acidic N-terminal tail of the embryonic dBigH1 linker histone of *Drosophila* compromises transcription by altering the functional epigenetic state of active chromatin. Other embryonic H1s might share similar properties since they are generally more acidic than their somatic counterparts.

DATA AVAILABILITY

GEO IDs: GSE127227 and GSE10329 In house scripts are provided upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are thankful to Drs J. Kadonaga and J. Zeitlinger for antibodies. P.C.-C. and P.B. acknowledge receipt of a 'Severo Ochoa' FPI fellowship from MINECO, and MT of a 'La Caixa' fellowship.

FUNDING

MINECO [BFU2015-65082-P, PGC2018-094538-B-100]; Generalitat de Catalunya [SGR2014-204, SGR2017-475]; European Community FEDER program (to F.A.); MEC 'Centro de Excelencia Severo Ochoa 2013–2017' [SEV-2012-0208]; MINECO [SAF2016-75006-P to M.B.]; 'Centre de Referència en Biotecnologia' of the Generalitat de Catalunya. Funding for open access charge: MINECO.
Conflict of interest statement. None declared.

REFERENCES

1. Fyodorov,D.V., Zhou,B.R., Skoultchi,A.I. and Bai,Y. (2018) Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* **19**, 192–206.
2. Pérez-Montero,S., Carbonell,A. and Azorín,F. (2016) Germline-specific H1 variants: the "sexy" linker histones. *Chromosoma*, **125**, 1–13.
3. Mariño-Ramírez,L., Hsu,B., Baxenavis,A.D. and Landsman,D. (2006) The histone database: a comprehensive resource for histones and histone fold-containing proteins. *Proteins*, **62**, 838–842.
4. Nagel,S. and Grossbach,U. (2000) Histone H1 genes and histone clusters in the genus *Drosophila*. *J. Mol. Evol.*, **51**, 286–298.
5. Bayona-Feliu,A., Casas-Lamesa,A., Carbonell,A., Climent-Cantó,P., Tatarski,M., Pérez-Montero,S., Azorín,F. and Bernués,J. (2016) Histone H1: lessons from *Drosophila*. *Biochim. Biophys. Acta*, **1859**, 526–532.
6. Pérez-Montero,S., Carbonell,A., Morán,T., Vaquero,A. and Azorín,F. (2013) The embryonic linker histone H1 variant of *Drosophila*, dBigH1, regulates zygotic genome activation. *Dev. Cell*, **26**, 578–590.
7. Carbonell,A., Pérez-Montero,S., Climent-Cantó,P., Reina,O. and Azorín,F. (2017) The germline linker histone dBigH1 and the translational regulator Bam form a repressor loop essential for male germ stem cell differentiation. *Cell Rep.*, **21**, 3178–3189.
8. Braunschweig,U., Hogan,G.J., Pagie,L. and van Steensel,B. (2009) Histone H1 binding is inhibited by histone variant H3.3. *EMBO J.*, **28**, 3635–3645.
9. Izzo,A., Kamieniarz-Gdula,K., Ramírez,F., Noureen,N., Kind,J., Manke,T., van Steensel,B. and Schneider,R. (2013) The genomic landscape of the somatic linker histone subtypes H1.1 to H1.5 in human cells. *Cell Rep.*, **3**, 2142–2154.
10. Li,J.Y., Patterson,M., Mikkola,H.K., Lowry,W.E. and Kurdistani,S.K. (2012) Dynamic distribution of linker histone H1.5 in cellular differentiation. *PLoS Genet.*, **8**, e1002879.
11. Cao,K., Lailler,N., Zhang,Y., Kumar,A., Uppal,K., Liu,Z., Lee,E.K., Wu,H., Medrzycki,M., Pan,C. *et al.* (2013) High-resolution mapping of h1 linker histone variants in embryonic stem cells. *PLoS Genet.*, **9**, e1003417.
12. Millán-Ariño,L., Islam,A.B., Izquierdo-Bouldstridge,A., Mayor,R., Terme,J.M., Luque,N., Sancho,M., López-Bigas,N. and Jordan,A. (2014) Mapping of six somatic linker histone H1 variants in human breast cancer cells uncovers specific features of H1.2. *Nucleic Acids Res.*, **42**, 4474–4493.
13. Nalabothula,N., McVicker,G., Maiorano,J., Martin,R., Pritchard,J.K. and Fondufe-Mittendorf,Y.N. (2014) The chromatin architectural proteins HMGD1 and H1 bind reciprocally and have opposite effects on chromatin structure and gene regulation. *BMC Genomics*, **15**, 92.
14. Mayor,R., Izquierdo-Bouldstridge,A., Millán-Ariño,L., Bustillos,A., Sampaio,C., Luque,N. and Jordan,A. (2015) Genome distribution of replication-independent histone H1 variants shows H1.0 associated with nucleolar domains and H1X associated with RNA polymerase II-enriched regions. *J. Biol. Chem.*, **290**, 7474–7491.
15. Koop,R., Croce,Di and Beato,M. (2003) Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. *EMBO J.*, **22**, 588–599.
16. Bayona-Feliu,A., Casas-Lamesa,A., Reina,O., Bernués,J. and Azorín,F. (2017) Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. *Nat. Commun.*, **8**, 283.
17. Shao,W. and Zeitlinger,J. (2017) Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.*, **49**, 1045–1051.
18. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
19. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
20. Tarasov,A., Vilella,A.J., Cuppen,E., Nijman,I.J. and Prins,P. (2015) Sambamba: fast processing of NGS alignment formats. *Bioinformatics*, **31**, 2032–2034.
21. Thorvaldsdóttir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.*, **14**, 178–192.
22. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
23. Gentleman,R.C., Carey,V.J., Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
24. Planet,E., Attolini,C.S.-O., Reina,O., Flores,O. and Rossell,D. (2012) htSeqTools: high-throughput sequencing quality control, processing and visualization in R. *Bioinformatics*, **28**, 589–590.
25. Font-Burgada,J., Reina,O., Rossell,D. and Azorín,F. (2014) chroGPS, a global chromatin positioning system for the functional analysis and visualization of the epigenome. *Nucleic Acids Res.*, **42**, 2126–2137.
26. Reina,O., Azorin,F. and Stephan-Otto Attolini,C. (2019) chroGPS2: differential analysis of epigenome maps in R. bioRxiv doi: https://doi.org/10.1101/425892, 14 May 2019, preprint: not peer reviewed.
27. Kessler,R., Tisserand,J., Font-Burgada,J., Reina,O., Coch,L., Attolini,C.S., Garcia-Bassets,I. and Azorin,F. (2015) dDsk2 regulates H2Bub1 and RNA polymerase II pausing at dHP1c complex target genes. *Nat. Commun.*, **6**, 7049.
28. Muse,G.W., Gilchrist,D.A., Nechaev,S., Shah,R., Parker,J.S., Grissom,S.F., Zeitlinger,J. and Adelman,K. (2007) RNA polymerase is poised for activation across the genome. *Nat. Genet.*, **39**, 1507–1511.
29. Carvalho,B.S. and Irizarry,R.A. (2010) A framework for oligonucleotide microarray preprocessing. *Bioinformatics*, **26**, 2363–2367.
30. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. U.S.A.*, **102**, 15545–15550.
31. Bonte,E. and Becker,P.B. (1999) Preparation of chromatin assembly extracts from preblastoderm *Drosophila* embryos. *Methods Mol. Biol.*, **119**, 187–194.
32. Becker,P.B., Tsukiyama,T. and Wu,C. (1994) Chromatin assembly extracts from *Drosophila* embryos. *Methods Cell Biol.*, **44**, 207–223.
33. Becker,P.B. and Wu,C. (1992) Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Mol. Cell Biol.*, **12**, 2241–2249.
34. Šatović,E., Font-Mateu,J., Carbonell,A., Beato,M. and Azorín,F. (2018) Chromatin remodeling in *Drosophila* preblastodermic embryo extract. *Sci. Rep.*, **8**, 10927.
35. Chapman,G.E., Hartman,P.G. and Bradbury,E.M. (1976) Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The isolation of the globular and non-globular regions of the histone H1 molecule. *Eur. J. Biochem.*, **61**, 69–75.

36. Allan,J., Hartman,P.G., Crane-Robinson,C. and Aviles,F.X. (1980) The structure of histone H1 and its location in chromatin. *Nature*, **288**, 675–679.
37. Ramakrishnan,V., Finch,J.T., Graziano,V., Lee,P.L. and Sweet,R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219–223.
38. Th'ng,J.P., Sung,R., Ye,M. and Hendzel,M.J. (2005) H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J. Biol. Chem.*, **280**, 27809–27814.
39. Hendzel,M.J., Lever,M.A., Crawford,E. and Th'ng,J.P. (2004) The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo. *J. Biol. Chem.*, **279**, 20028–20034.
40. Schones,D.E. and Zhao,K. (2008) Genome-wide approaches to studying chromatin modifications. *Nat. Rev. Genet.*, **9**, 179–191.
41. Jullien,J., Astrand,C., Halley-Stott,R.P., Garrett,N. and Gurdon,J.B. (2010) Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. *Proc. Natl Acad. Sci. U.S.A.*, **107**, 5483–5488.
42. Gao,S., Chung,Y.G., Parseghian,M.H., King,G.J., Adashi,E.Y. and Latham,K.E. (2004) Rapid H1 linker histone transitions following fertilization or somatic cell nuclear transfer: evidence for a uniform developmental program in mice. *Dev. Biol.*, **266**, 62–75.
43. Jullien,J., Miyamoto,K., Pasque,V., Allen,G.E., Bradshaw,C.R., Garrett,N.J., Halley-Stott,R.P., Kimura,H., Ohsumi,K. and Gurdon,J.B. (2014) Hierarchical molecular events driven by oocyte-specific factors lead to rapid and extensive reprogramming. *Mol. Cell*, **55**, 524–536.
44. Teranishi,T., Tanaka,M., Kimoto,S., Ono,Y., Miyakoshi,K., Kono,T. and Yoshimura,Y. (2004) Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. *Dev. Biol.*, **266**, 76–86.
45. Yun,Y., Zhao,G.M., Wu,S.J., Li,W. and Lei,A.M. (2012) Replacement of H1 linker histone during bovine somatic cell nuclear transfer. *Theriogenology*, **78**, 1371–1380.
46. Kepert,J.F., Mazurkiewicz,J., Heuvelman,G.L., Tóth,K.F. and Rippe,K. (2005) NAP1 modulates binding of linker histone H1 to chromatin and induces an extended chromatin fiber conformation. *J. Biol. Chem.*, **280**, 34063–34072.
47. Shintomi,K., Iwabuchi,M., Saeki,H., Ura,K., Kishimoto,T. and Ohsumi,K. (2005) Nucleosome assembly protein-1 is a linker histone chaperone in *Xenopus* eggs. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 8210–8215.
48. Nightingale,K., Dimitrov,S., Reeves,R. and Wolffe,A.P. (1996) Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. *EMBO J.*, **15**, 548–561.
49. Saeki,H., Ohsumi,K., Aihara,H., Ito,T., Hirose,S., Ura,K. and Kaneda,Y. (2005) Linker histone variants control chromatin dynamics during early embryogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5697–5702.
50. Ura,K., Nightingale,K. and Wolffe,A.P. (1996) Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. *EMBO J.*, **15**, 4959–4969.
51. Becker,M., Becker,A., Miyara,F., Han,Z., Kihara,M., Brown,D.T., Hager,G.L., Latham,K., Adashi,E.Y. and Misteli,T. (2005) Differential *in vivo* binding dynamics of somatic and oocyte-specific linker histones in oocytes and during ES cell nuclear transfer. *Mol. Biol. Cell*, **16**, 3887–3895.
52. Godde,J.S. and Ura,K. (2009) Dynamic alterations of linker histone variants during development. *Int. J. Dev. Biol.*, **53**, 215–224.
53. Meshorer,E., Yellajoshula,D., George,E., Scambler,P.J., Brown,D.T. and Misteli,T. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell*, **10**, 105–116.
54. Gilchrist,D.A., Nechaev,S., Lee,C., Ghosh,S.K., Collins,J.B., Li,L., Gilmour,D.S. and Adelman,K. (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.*, **22**, 1921–1933.
55. Wang,X., Hang,S., Prazak,L. and Gergen,J.P. (2010) NELF potentiates gene transcription in the *Drosophila* embryo. *PLoS One*, **5**, e11498.
56. Boija,A., Mahat,D.B., Zare,A., Holmqvist,P.H., Philip,P., Meyers,D.J., Cole,P.A., Lis,J.T., Stenberg,P. and Mannervik,M. (2017) CBP regulates recruitment and release of promoter-proximal RNA Polymerase II. *Mol. Cell*, **68**, 491–503.

ANNEX II

Linker histone H1 regulates homeostasis of heterochromatin associated cRNAs

Paula Bujosa[1,2], Oscar Reina[2], Adrià Caballé[2], Anna Casas-Lamesa[1,2], Ana Silvina Nacht[3], Guillermo P. Vicent[1,3], Jordi Bernués[1,2,*] and Fernando Azorín[1,2,*].

[1]Institute of Molecular Biology of Barcelona, CSIC. Baldiri Reixac, 4. 08028 Barcelona. Spain.

[2]Institute for Research in Biomedicine, IRB Barcelona. The Barcelona Institute of Science and Technology. Baldiri Reixac, 10. 08028 Barcelona. Spain.

[3]Centre de Regulació Genòmica (CRG). The Barcelona Institute of Science and Technology. Barcelona. Spain.

Running title: H1 regulates chromosomal RNAs

Key words: Chromosomal RNAs/ Histone H1/ hrp36/ hrp48/ RNPs/ R-loops/ Chromatin/ Epigenetics/ *Drosophila*

*Co-corresponding authors:
Dr. J. Bernués (jbmbmc@ibmb.csic.es)
Dr. F. Azorín (fambmc@ibmb.csic.es)

ABSTRACT

Chromosomal RNAs (cRNAs) are a poorly understood fraction of cellular RNAs that co-purify with chromatin. Here we show that, in *Drosophila*, cRNAs constitute a heterogeneous group of RNA species that cover ~28% of the genome. Intriguingly, we found that cRNAs are highly enriched in heterochromatic transcripts. Our results show that heterochromatic cRNAs interact with the hnRNP A/B proteins hrp36 and hrp48 to assemble into RNP particles. We also show that depletion of linker histone dH1, a major component of chromatin, impairs assembly of hrp36 and hrp48 onto heterochromatic cRNAs. Concomitantly, impaired cRNAs assembly induces the accumulation of heterochromatic cRNAs and the formation of unscheduled RNA::DNA hybrids (R-loops). Linker histones H1 are known to regulate chromatin structure and compaction and, indeed, we show that dH1 depletion perturbs chromatin organization, reducing nucleosome occupancy and increasing accessibility of heterochromatin. These perturbations facilitate annealing of cRNAs to the DNA template, enhancing R-loops formation and cRNAs retention at heterochromatin. Altogether, these results unveil the unexpected contribution of linker histones to RNPs assembly and homeostasis of cRNAs.

INTRODUCTION

The association of RNAs with chromatin has long been appreciated. Chromatin associated RNAs, originally called cRNAs (Bonner and Widholm, 1967; Holmes et al., 1972; Huang and Bonner, 1965), account for 2%-5% of total nucleic acids in chromatin (Rodríguez-Campos and Azorín, 2007) and they have been proposed to play important regulatory roles, from enhancer function to heterochromatin and centromere integrity, and, in general, in the maintenance of a native chromatin organization (reviewed in (Li and Fu, 2019; Thakur and Henikoff, 2020)). Chromosomal cRNAs belong to different classes. Though a fraction corresponds to nascent pre-mRNAs, most of them are non-coding. In recent years, much attention has been given to CheRNAs (Gayen and Kalantry, 2017; Sun et al., 2020; Werner and Ruthenburg, 2015; Werner et al., 2017; Zhang et al., 2021; Zhang et al., 2022), which constitute a special class of chromatin

associated long non-coding RNAs (lncRNAs) that share strong similarities with enhancer RNAs (eRNAs). However, our current knowledge of the full complexity of chromosomal cRNAs is still incomplete. Here, we show that, in *Drosophila* S2 cells, an important fraction of cRNAs corresponds to heterochromatic transcripts. Current models invoke that transcription by RNApol II is involved in tethering cRNAs to chromatin (Gayen and Kalantry, 2017; Werner and Ruthenburg, 2015; Werner et al., 2017; Zhang et al., 2021). RNA transcripts are generally assembled into ribonuclear particles (RNPs) through the interaction with heterogeneous nuclear ribonucleoproteins (hnRNPs). hnRNPs constitute a large group of RNA binding proteins (reviewed in (Dreyfuss et al., 2002)). From these, the hnRNPA/B family is a group of abundant proteins that are well evolutionarily conserved and play important regulatory functions in RNA splicing, export and localization (Dreyfuss et al., 2002). *Drosophila* encodes five hnRNP A/B proteins, two of which (hrp36/Hrb87F/hnRNP-A and hrp48/Hrb27C/DAZAP) have been shown to localize across chromatin and, in particular, at heterochromatin (Matunis et al., 1993; Matunis et al., 1992; Piacentini et al., 2009). It is anticipated that chromosomal cRNAs assemble into RNPs and, indeed, we show here that hrp36 and hrp48 are involved in RNP assembly of cRNAs and, in particular, of heterochromatic cRNAs.

RNP assembly occurs co-transcriptionally on the chromatin template (Dreyfuss et al., 2002). However, how chromatin itself contributes to RNP assembly is not well understood. Here, we show that linker histone H1, a major component of chromatin, plays an important role in RNP assembly of cRNAs. Linker histones H1 are abundant chromosomal proteins that bind nucleosome core particles at the DNA entry/exit site and interact with internucleosomal linker DNA (reviewed in (Fyodorov et al., 2018)). In comparison to core histones, histones H1 are less well evolutionarily conserved and, in general, metazoan species encode for multiple variants that play partially redundant functions (Fyodorov et al., 2018). However, histones H1 complexity in *Drosophila* is much reduced since it encodes for a single somatic dH1 variant and a second germline specific dBigH1 isoform (reviewed in (Bayona-Feliu et al., 2016; Pérez-Montero et al., 2016)). Histones H1 stabilize nucleosomes and are required for folding of

the nucleofilament into higher-order structures, thus regulating chromatin compaction (Fyodorov et al., 2018). In particular, histones H1 play a crucial role in maintenance of heterochromatin integrity, compaction and silencing (Izquierdo-Bouldstridge et al., 2017; Lu et al., 2009; Lu et al., 2013; Vujatovic et al., 2012). In metazoan species, histones H1 are essential for genome stability and cell/organismal viability (Bayona-Feliu et al., 2017; Fan et al., 2003; Fan et al., 2001; Lu et al., 2009; Sirotkin et al., 1995; Vujatovic et al., 2012). Notably, histones H1s have been reported to regulate dynamics of RNA::DNA hybrids (R-loops) (Almeida et al., 2018; Bayona-Feliu et al., 2017), whose unscheduled formation is a frequent source of genomic instability (reviewed in (García-Muse and Aguilera, 2019)). In this regard, it has been shown that, in *Drosophila*, somatic dH1 prevents the accumulation of deleterious RNA::DNA hybrids (R-loops) in heterochromatin (Bayona-Feliu et al., 2017), preserving genome integrity and cell viability. However, the mechanisms by which histones H1 regulate R-loops dynamics are not well understood.

Here, we report a novel and unexpected contribution of histones H1 to RNPs assembly and homeostasis of cRNAs. We show that *Drosophila* dH1 is required for hrp36 and hrp48 assembly onto cRNAs. We also show that impaired assembly of cRNAs results in their increased retention in chromatin, specially of heterochromatic cRNAs, and promotes R-loops formation. Moreover, loss of dH1 profoundly perturbs the structural organization of chromatin, reducing nucleosome occupancy and increasing heterochromatin accessibility. Altogether our results suggest a model by which histones H1 prevent the deleterious accumulation of R-loops in heterochromatin by, on one hand, promoting RNPs assembly of cRNAs and, on the other hand, stabilizing heterochromatin structure and organization.

RESULTS

The hnRNP proteins hrp36 and hrp48 associate with heterochromatic cRNAs

To characterize chromosomal cRNAs in *Drosophila* S2 cells, we performed strand-specific RNA-seq experiments of total RNA extracted from

extensively purified chromatin (Werner and Ruthenburg, 2015). *De novo* transcriptome assembly detected ~14,000 cRNA species that covered ~28% of the *Drosophila* genome, being distributed across all chromosomes and in both strands (**Figure 1A**). Most cRNAs matched genic regions and many were sense (**Figure 1B**), suggesting that they corresponded to transcribing pre-mRNAs. Additionally, we also detected antisense and intergenic transcripts, while long non-coding RNAs (putative CheRNAs) accounted only for a minor proportion of total cRNAs (**Figure 1B**). Intriguingly, we observed that ~34% of cRNA transcripts mapped to the assembled heterochromatic regions of chromosomes 2LHet, 2RHet, 3LHet, 3RHet, YHet and XHet, and the unassembled repetitive heterochromatic elements of the artificial scaffolds U and Uextra (**Figure 1A**). These heterochromatic transcripts include various classes of transposable elements, satellite DNAs and simple-repeated DNA elements (**Supplementary Table S1**). Heterochromatic regions are generally poorly transcribed. However, in the cRNA fraction, heterochromatic transcripts were remarkably enriched, with ~25% of them corresponding to the highest decile of expression (**Figure 1C**, **right**). In contrast, the distribution by expression quantiles of the euchromatic cRNAs was rather uniform (**Figure 1C**, **left**). Altogether these results indicate that cRNAs constitute a heterogeneous group of RNA species that is enriched in heterochromatic transcripts.

Little is known about packaging and assembly of cRNAs into RNPs. In this regard, as previously reported by others (Matunis et al., 1993; Matunis et al., 1992; Piacentini et al., 2009), we detected the association of the *Drosophila* hnRNP A/B proteins hrp36 and hrp48 with polytene chromosomes (**Figure 1D**, **top**). hrp36 and hrp48 were detected at the heterochromatic chromocenter (**Supplementary Figure S1A**) and at DAPI interbands along the euchromatic chromosome arms (**Supplementary Figure S1B**), showing extensive overlapping. In addition, Western blot (WB) analysis of crosslinked chromatin from S2 cells confirmed the association of hrp36 and hrp48 with chromatin (**Figure 1E**). Notably, treatment with RNase A abolished the association of hrp36 and hrp48 with polytene chromosomes, both at the heterochromatic chromocenter and across the euchromatic chromosome arms (**Figure 1D**,

bottom). Along the same lines, WB analysis of crosslinked S2 chromatin showed that the levels of chromatin-associated hrp36 and hrp48 were drastically reduced upon RNase A treatment (**Figure 1E**).

To further analyze the association of hrp36 and hrp48 with chromatin, we performed ChIP-seq experiments in S2 cells. We detected a similar number of enriched regions, of which >80% were common to both hrp36 and hrp48 (**Figure 2B**), which is consistent with the high overlap observed in polytene chromosomes. Though hrp36 and hrp48 were detected across the genome, they were preferentially enriched at heterochromatin (**Figure 2A**). Furthermore, according to the *Drosophila* nine chromatin states model (Kharchenko et al., 2011), ~37% of the common hrp36 and hrp48 enriched regions that could be assigned to any particular chromatin state mapped to heterochromatin (state 7) and this enrichment was statistically significant (permutation test, p-value: 0.022) (**Figure 2C**). In addition, we observed that ~27% of the enriched regions mapped to silent/intergenic chromatin (state 9) (**Figure 2C**). Notably, hrp36 and hrp48 enriched regions strongly overlapped with cRNAs (permutation test, p-value< 0.001) (**Figure 2D**) and, in comparison to non-enriched regions, cRNAs abundance was significantly higher at common hrp36 and hrp48 enriched regions (permutation test, p-value< 0.001) (**Figure 2E**).

Altogether these results indicate that hrp36 and hrp48 associate with cRNAs, particularly with those of heterochromatic origin, likely participating in their assembly into RNPs.

dH1 depletion impairs RNPs assembly of cRNAs and induces their accumulation in heterochromatin

We observed that depletion of linker histone dH1 strongly reduced the association of hrp36 and hrp48 with chromatin. In comparison to control flies, hrp36 and hrp48 were barely detected along chromosome arms of polytene chromosomes from dH1-depleted $dH1^{RNAi}$ flies (**Figure 3A**), while they were strongly reduced at the heterochromatic chromocenter (**Figure 3B** and **Supplementary Figure S2A**). No such effects were observed in control mock depleted GFP^{RNAi} flies (**Supplementary Figures S2B** and **S2C**). Along the same lines, WB analysis of crosslinked S2 chromatin showed significantly reduced

levels of chromatin-associated hrp36 and hrp48 in dH1-depleted cells treated with dsRNA against dH1 (dsRNA^{dH1}) in comparison to control cells treated with dsRNA against LacZ (dsRNA^{LacZ}) (**Figure 3C**). In these experiments, dH1 levels were reduced to ~50% of the native levels, causing a similar reduction of the levels of chromatin-bound hrp36 and hrp48 (**Figure 3C**). Noteworthy, dH1 depletion did not affect hrp36 and hrp48 mRNA levels (**Figure 3D**). ChIP-seq experiments confirmed these results since, in comparison to control dsRNA^{LacZ} cells, the intensity of hrp36 and hrp48 enriched regions decreased in dH1-depleted cells (**Figures 3E** and **3F**). These results indicate that depletion of linker histone dH1 results in decreased chromatin association of hrp36 and hrp48 and, hence, impaired assembly of cRNAs into RNPs.

Next, we performed cRNA-seq experiments in dH1-depleted dsRNA^{dH1} and control dsRNA^{lacZ} cells. We observed a similar number of cRNAs species in dH1-depleted and control cells, showing strong overlapping (**Figure 4A**). In addition, considering differentially expressed cRNAs (p-value < 0.05), we observed similar proportions of up- and down-regulated cRNAs upon dH1 depletion (**Figure 4B**). In contrast, we found that a high proportion (78%) of differentially expressed heterochromatic cRNAs were up-regulated in dH1-depleted cells and this enrichment was statistically significant (permutation test, p-value < 0.001) (**Figure 4C**), suggesting that heterochromatic cRNAs accumulate upon dH1 depletion. Similarly, intergenic cRNAs were also enriched in dH1-depleted cells (permutation test, p-value < 0.001) (**Supplementary Figure S3A**) and, although the magnitude of the change was low and not statistically significant (permutation test, p-value = 0.110), lncRNAs had a tendency to increase (**Supplementary Figure S3B**). As for the rest of cRNAs classes, genic antisense did not show a clear trend (**Supplementary Figure S3C**), while, intriguingly, genic sense cRNAs were down-regulated upon dH1 depletion (permutation test, p-value < 0.001) (**Supplementary Figure S3D**). Along the same lines, we observed that cRNAs mapping to active chromatin states were down-regulated in dH1-depleted cells (permutation test, p-values < 0.001) (**Supplementary Figures S3E-S3G**). Interestingly, we observed that, in dH1-depleted cells, intron retention (IR) increased in a very high proportion (91%)

of genic cRNAs showing differential IR (**Figure 4G**), suggesting that dH1-depletion impairs co-transcriptional splicing of cRNAs.

Impaired assembly of heterochromatic cRNAs enhances R-loops formation

It was shown earlier that depletion of linker histone dH1 induces the accumulation of R-loops in heterochromatin (Bayona-Feliu et al., 2017). Next, considering that dH1 depletion disrupts assembly of heterochromatic cRNAs into RNPs, we posit that impaired assembly of heterochromatic cRNAs might facilitate the formation of R-loops in heterochromatin. For this purpose, we analyzed the effects of directly impairing RNPs assembly through depletion of hrp36 or hrp48. First, we performed cRNA-seq experiments with hrp36-depleted ($dsRNA^{hrp36}$) and hrp48-depleted ($dsRNA^{hrp48}$) cells. The total number of cRNAs detected in hrp36- and hrp48-depleted cells was similar to that detected in dH1-depleted cells and they strongly overlapped (**Figure 4D**). In addition, like dH1 depletion, depletion of hrp48 induced a strong accumulation of heterochromatic cRNAs (permutation test, p-value < 0.001) (**Figure 4E**). This effect was not observed upon depletion of hrp36 (**Figure 4F**). We also observed that, unlike dH1 depletion, depletion of hrp36 and hrp48 did not significantly altered the frequency of IR (**Figures 4H** and **4I**).

Next, we addressed R-loops formation upon hrp36 or hrp48 depletion. Immunostaining with S9.6 antibodies, which detect R-loops, showed increased S9.6 signal at the heterochromatic chromocenter of polytene chromosomes from hrp36-depleted ($hrp36^{RNAi}$) and hrp48-depleted ($hrp48^{RNAi}$) flies (**Figure 5A**), which is similar to what was previously reported in dH1-depleted flies (Bayona-Feliu et al., 2017). To further analyze the effects of hrp36 and hrp48 depletion on R-loops formation, we performed DRIP-seq experiments in hrp36- and hrp48-depleted cells, and in control $dsRNA^{lacZ}$ cells. In these experiments, we called R-loops enriched regions to those genomic regions significantly enriched after immunoprecipitation with S9.6 antibodies that, upon treatment with RNase H before immunoprecipitation, decreased coverage by FC < -1.25. It was expected that R-loops regions will be enriched in cRNAs and, indeed, R-loops detected in hrp36- and hrp48-depleted cells strongly overlapped with cRNAs (permutation test, p-values < 0.001) (**Figure 5B**). In addition, R-loops were detected in ~50%

of the hrp36 and hrp48 enriched regions (**Figure 5C**) and this overlap was statistically significant (permutation test, p-values < 0.001). We observed that, in comparison to control dsRNA^{lacZ} cells, the number of R-loops regions increased by ~2-fold in hrp48-depleted cells (**Figure 5D**, **top**). Notably, the vast majority (99%) of R-loops detected only in hrp48-depleted cells mapped at heterochromatic regions (**Figure 5E**, **left**). Similarly, R-loops common to both control and hrp48-depleted cells preferentially mapped at heterochromatin (**Figure 5E**, **center**). In contrast, a large proportion (74%) of those detected only in control dsRNA^{lacZ} cells mapped at euchromatic regions (**Figure 5E**, **right**). Moreover, Volcano plot analysis showed strongly increased R-loops intensity in hrp48-depleted cells in comparison to control dsRNA^{lacZ} cells (**Figure 5F**). Depletion of hrp36 had similar effects. In comparison to control dsRNA^{lacZ} cells, the total number of R-loops was not increased in hrp36-depleted cells (**Figure 5D, bottom**). However, R-loops detected in hrp36-depleted cells strongly accumulate at heterochromatic regions (**Figure 5G**, **left** and **center**), while euchromatic R-loops were lost upon hrp36 depletion (**Figure 5G**, **right**), which is similar to what was observed upon hrp48 depletion (**Figure 5E**). In addition, like in hrp48-depleted cells, R-loops intensity also increased upon hrp36 depletion (**Figure 5H**).

These results confirm that impairing RNPs assembly of heterochromatic cRNAs through depletion of hrp36 and hrp48 enhances the formation of R-loops in heterochromatin. However, like on the accumulation of heterochromatic cRNAs (**Figures 4E** and **4F**), depletion of hrp48 had a stronger effect on R-loops formation in comparison to hrp36 depletion. In fact, only 20% of R-loops detected upon hrp48 depletion were also detected in hrp36-depleted cells, while, *vice versa*, 55% of those detected in hrp36-depleted cells were also found in hrp48-depleted cells (**Supplementary Figure S4A**). We also observed that R-loops specific of hrp48-depleted and hrp36-depleted cells, as well as those common to both, were enriched at heterochromatic regions, though this enrichment was higher for R-loops specific of hrp48-depleted cells (**Supplementary Figure S4B**). Overall, these results suggest that, despite their strong colocalization (**Figure 2B** and **Supplementary Figure S1**), depletion of hrp36 and hrp48 show non-

equivalent effects regarding accumulation of heterochromatic cRNAs and R-loops formation. These differences are unlikely due to differences in depletion efficiency since hrp36 and hrp48 levels were reduced to a similar extent (**Supplementary Figure S4D)**. In fact, IF experiments performed in *hrp36*^{RNAi} and *hrp48*^{RNAi} depleted flies showed that the association of hrp36 and hrp48 with polytene chromosomes are largely independent of each other (**Supplementary Figure S4C**), and WB analysis of crosslinked S2 chromatin confirmed these results (**Supplementary Figure S4D**).

Altogether these results suggest that the effects of dH1 depletion on R-loops accumulation in heterochromatin is, at least in part, mediated by its disrupting effect on RNPs assembly of heterochromatic cRNAs.

dH1 depletion reduces nucleosome occupancy and increases accessibility of heterochromatin

Next, we performed ATAC-seq experiments to determine the effects of dH1 depletion on the structural organization of the chromatin template. We performed paired-end sequencing to be able to separately analyze transposase hypersensitive sites (THSS) and fragments consistent with the sizes of nucleosomes by categorizing the reads by insert size (<120 bp *vs* 120-250 bp respectively). We observed that dH1 depletion resulted in an increased proportion of THSS fragments, while the proportion of nucleosomal fragments decreased (**Figure 6A**), indicating that, overall, chromatin accessibility increases upon dH1 depletion. THSS peaks detected in dH1-depleted and control dsRNA^{lacZ} cells strongly overlapped (**Figure 6B**). We found that THSS peaks detected in dH1-depleted cells preferentially mapped to heterochromatic regions (**Figure 6C**, **right** and **center**), while the peaks detected only in control dsRNA^{lacZ} cells were mostly euchromatic (**Figure 6C**, **left**). We also found that, at heterochromatic regions, the intensity of THSS peaks increased upon dH1 depletion (**Figure 6D**). These results suggest that accessibility of heterochromatin increases upon dH1 depletion.

In contrast, we observed that the intensity of THSS peaks lying at euchromatic regions apparently decreased upon dH1 depletion (**Figures 6D**). Rather than reduced accessibility, decreased intensity of euchromatic THSS

peaks reflects that, in comparison to control dsRNA^{lacZ} cells, THSS ATAC-seq signal spreads over wider regions in dH1-depleted cells. This was best shown when THSS signal was analyzed at the transcription-start sites (TSSs) that are nucleosome-free in active genes and, hence, widely accessible (Buenrostro et al., 2013). In fact, TSSs account for a large proportion of euchromatic THSSs (**Supplementary Figure S5**). THSS ATAC-seq signal at TSSs was constrained to a narrow 100bp long region in control dsRNA^{lacZ} cells (**Figure 6E**, **left**). Instead, in dH1-depleted cells, it spread over >500bp especially downstream of the TSS (**Figure 6E**, **right**). This widespread ATAC-seq signal suggests that dH1 depletion perturbs nucleosomal organization. Analysis of the distribution of nucleosomal ATAC-seq fragments confirmed these results. Two main nucleosomal fragments of 151bp and 138bp were detected (**Figures 6A** and **7A**). Similar nucleosomal fragments were previously detected in MNase-seq experiments in S2 cells (Chereji et al., 2019). We observed that, in comparison to control dsRNA^{lacZ} cells, the nucleosomal fragments were less intense in dH1-depleted cells (**Figure 7A, top**) and accounted for a lower proportion of total fragments (**Figure 7B**), indicating that nucleosome occupancy was decreased. Moreover, determination of the average dyad density showed that the distribution of nucleosomes around TSSs became fuzzier in dH1-depleted cells and, in particular, the position of the +1 nucleosome was less well defined (**Figure 7A, bottom**), which is in good agreement with the observed spreading of the THSS ATAC-seq signal downstream of the TSS (**Figure 6E**). This perturbation is likely general, not restricted to TSSs, since similar results were observed when the distribution of nucleosomes was analyzed relative to hrp36/hrp48 enriched regions (**Figures 7C** and **7D**). In this case, in addition to the two nucleosomal fragments of 151bp and 138bp, a subnucleosomal fragment of 125bp was also detected (**Figures 7C, top**), which was much weaker at TSSs sites. Like at TSSs, the intensity of nucleosomal fragments at hrp36/hrp48 sites decreased upon dH1 depletion (**Figures 7C, top** and **7D**). At these sites, it is expected that nucleosomes would not be well positioned as around TSSs. Consequently, in this case, average dyad density plots are less informative, showing only weak differences upon dH1 depletion (**Figure 7C, bottom**). Altogether these results

suggest that dH1 depletion perturbs chromatin organization, reducing nucleosome occupancy. These defects are observed at euchromatic TSSs as well as at hrp36/48 sites that, as shown above (**Figure 2A**), preferentially locate in heterochromatin.

DISCUSSION

Here we report on the contribution of linker histones H1 to homeostasis of chromosomal cRNAs, a fraction of cellular RNAs that remains poorly characterized. We show that, in *Drosophila* S2 cells, cRNAs constitute a heterogeneous group of RNA species. We found that, despite cRNAs account for only 2%-5% of total nucleic acids in chromatin (Rodríguez-Campos and Azorín, 2007), they cover up to 28% of the *Drosophila* genome, suggesting that their association with chromatin is highly dynamic. In fact, many cRNAs match at euchromatic genic regions and are sense, suggesting that they correspond to transcribing pre-mRNAs that stay associated with chromatin while being co-transcriptionally processed before leaving the chromatin template (Li and Fu, 2019). However, most remarkably, we found that cRNAs are highly enriched in transcripts matching heterochromatic regions that, normally, are poorly transcribed, suggesting that heterochromatic cRNAs associate tightly with heterochromatin likely playing important functions. As a matter of fact, heterochromatic transcripts have been detected in fission yeast, *Drosophila* and mammals (Alekseyenko et al., 2023; Cohen et al., 1973; Gaubatz and Cutler, 1990; Grewal and Jia, 2007; Halic and Moazed, 2010; Johnson et al., 2017; Lehnertz et al., 2003; Lu and Gilbert;, 2007; Rudert et al., 1995; Schoeftner and Blasco, 2008) and they have been implicated in heterochromatin establishment, maintenance and compaction through the recruitment of heterochromatin associated proteins such as HP1 and SAF-B (Grewal and Jia, 2007; Huo et al., 2020; Maison et al., 2011; Muchardt et al., 2002; Santenard et al., 2010).

Previous studies reported the association of several hnRNPs with chromatin (Matunis et al., 1993; Matunis et al., 1992; Piacentini et al., 2009). Amongst them, hrp36 and hrp48 were shown to localize at heterochromatin (Matunis et al., 1993; Matunis et al., 1992; Piacentini et al., 2009). Here we have

shown that chromatin binding of hrp36 and hrp48 depends on RNA, and that they strongly co-localize with cRNAs, especially with those of heterochromatic origin. These results suggest that, on one hand, hrp36 and hrp48 interact with cRNAs, being involved in their assembly into RNP particles, and, on the other hand, that cRNAs are mostly acting on *cis*, over the genomic regions where they originate from. Additional evidence for cRNAs acting on *cis* comes from DRIP-seq analyses showing strong co-localization of R-loops with cRNAs. hrp36 and hrp48 belong to the hnRNP A/B family, a group of highly conserved hnRNP proteins that play important functions in the regulation of RNA splicing, export and localization (Dreyfuss et al., 2002). Our results extend these functions to the assembly and regulation of chromatin associated cRNAs and, specially, of heterochromatic cRNAs. RNPs are unusually large and dynamic, containing numerous hnRNP proteins (Dreyfuss et al., 2002). Thus, hrp36 and hrp48 are unlikely the only hnRNPs involved in cRNAs assembly. In fact, the number of cRNAs species largely exceeds that of hrp36 and hrp48 enriched regions. In this regard, it is important to notice that hnRNPs have RNA-binding specificity (Dreyfuss et al., 2002). Thus, it is possible that different sets of hnRNP proteins are involved in the assembly of different cRNAs classes. Along these lines, we observed that, despite largely sharing the same targets, hrp36 and hrp48 associate with chromatin independently of each other. Thus, relative abundance of hrp36 and hrp48 might be different in different cRNAs, accounting for the observed distinct effects of their depletion. From this point of view, our results suggest that hrp48 has stronger effects on the behavior of heterochromatic cRNAs, while hrp36 acts more broadly and has milder effects. Further work will provide a full catalogue of hnRNPs associated with cRNAs, their specificity and functional contribution to cRNAs metabolism.

Our results show that binding of hrp36 and hrp48 to chromatin depends on linker histone dH1 that, interestingly, has been shown to co-immunoprecipitate with both hrp36 and hrp48 (Ni et al., 2006). dH1 depletion strongly reduces chromatin association of hrp36 and hrp48, implying that assembly of cRNAs into RNPs is impaired, and, concomitantly, we observed the accumulation of heterochromatic cRNAs. In this regard, it was shown that dH1 depletion strongly

upregulates expression of heterochromatic repetitive elements (Lu et al., 2013; Vujatovic et al., 2012), suggesting that relief of heterochromatin silencing is a major determinant of the accumulation of heterochromatic cRNAs. A similar situation is likely in mammalian cells since, as in *Drosophila*, reduced H1 levels increases expression of heterochromatic elements (Izquierdo-Bouldstridge et al., 2017). dH1 depletion also increases intergenic cRNAs and, though not significantly, long non-coding cRNAs tend to increase. Interestingly, despite the effects on heterochromatic cRNAs were not directly assessed, it was reported that reduced H1 levels in mouse ES cells increase chromatin abundance of some long non-coding CheRNAs (Fernández-Justel et al., 2022).

It has been shown that, in *Drosophila*, a major effect of dH1 depletion is the accumulation of RNA::DNA hybrids (R-loops) in heterochromatin, which causes DNA damage, genomic instability and cell death (Bayona-Feliu et al., 2017). Results reported here suggest that, at least in part, these defects are the consequence of reduced deposition of hrp36 and hrp48 since direct depletion of hrp36 and hrp48 induce a similar accumulation of R-loops in heterochromatin. Coating of cRNAs by hnRNPs likely acts as a barrier that hinders annealing to the DNA template and, thus, prevents the formation of R-loops. Along the same lines, chromatin stability and compaction also counteract R-loops formation (Bayona-Feliu and Aguilera, 2021; Chen et al., 2017; Sanz et al., 2016). In this regard, linker histones H1 have been shown to stabilize nucleosomes and regulate chromatin compaction both *in vitro* and *in vivo* (Fan et al., 2005; Fyodorov et al., 2018; Hashimoto et al., 2010; Lu et al., 2009; Sancho et al., 2008). Our results confirm that dH1 depletion perturbs chromatin organization, reducing nucleosome occupancy and increasing heterochromatin accessibility. These chromatin defects likely facilitate annealing of cRNAs and R-loops formation. Altogether these results suggest a model by which loss of linker histones promotes R-loops accumulation in heterochromatin through a dual mechanism. On one hand, reduced H1 levels upregulate expression of heterochromatic cRNAs and impair their assembly into RNPs, while, on the other hand, destabilize heterochromatin structure and increase its accessibility (**Figure**

8). In its turn, the formation of R-loops enhances cRNAs retention in heterochromatin.

Altogether these observations show that dH1 depletion affects multiple aspects of heterochromatin organization, from increased expression, cRNAs accumulation and R-loop formation to reduced nucleosome occupancy and increased accessibility. Intriguingly, in comparison to euchromatin, heterochromatin appears to be more sensitive to reduced histone H1 levels. Noteworthy, H1 content is higher in heterochromatin than in active euchromatin (Braunschweig et al., 2009; Climent-Cantó et al., 2020; Kharchenko et al., 2011; Lu et al., 2009) and, in general, experiments addressing the effects of reduced H1 levels achieved depletions not higher than 50% of the native H1 levels. It is possible that this reduction, while being sufficient to impact heterochromatin organization, might not affect euchromatic sites to the same extent. However, our results show that, actually, genic sense cRNAs and, in general, cRNAs mapped to active chromatin states decrease upon dH1 depletion. Moreover, upon hrp36 and hrp48 depletion, R-loops decrease at euchromatic regions since a large fraction of them (~70%) are detected in control cells but not in depleted cells. Thus, it is possible that factors associated with active euchromatin, such as helicases and/or nucleases, might counteract cRNAs retention and R-loop formation, preventing their accumulation (García-Muse and Aguilera, 2019). It is also possible that, rather than reduced H1 levels by themselves, the effects on heterochromatin reflect the contribution of specific post-translational modifications. In this regard, it has been shown that, in *Drosophila*, dH1K27me2 is largely constrained to heterochromatin and its reduction induces R-loops formation (Bernués et al., 2022).

In summary, results reported here show that depletion of *Drosophila* linker histone dH1 alters homeostasis of chromosomal cRNAs, impairing their assembly into RNPs and promoting their accumulation in heterochromatin and the formation of unscheduled R-loops. The mechanisms by which dH1 regulates cRNAs assembly and homeostasis remain to be determined. In this regard, in mouse ES cells, it has been reported that reduced H1 levels affects m6A modification (Fernández-Justel et al., 2022), an RNA modification that has been

shown to impact many aspects of RNA metabolism, including RNP assembly (Alarcón et al., 2015a; Alarcón et al., 2015b; Liu et al., 2015), while, on the other hand, regulates R-loops dynamics (Abakir et al., 2020; Yang et al., 2019). Notably, m6A has also been shown to regulate RNA splicing (Haussmann et al., 2016; Xiao et al., 2016) and our results suggest that dH1 depletion impairs splicing of cRNAs. Other RNA modifications might have similar effects. Further work will determine the pattern of modification of cRNAs in general and, in particular, of heterochromatic cRNAs, and the extent to which loss of linker histones alters cRNAs modification.

MATERIALS AND METHODS

Fly stocks and genetic analysis

hrp36^{RNAi} V51759 (II) and *hrp48*^{RNAi} V101555 (II) stocks were from VDRC (Vienna). *dH1*^{RNAi} was from NIG-FLY (31617R-3) and is described in (Bayona-Feliu et al., 2017). GFP^{RNAi} is described in (Font-Burgada et al., 2008). GAL4 driver stocks were *w*; *Nubb*GAL4 and *yw*; *Nubb*Gal4; UAS-*Dicer2* (kindly provided by Dr. Casali). H2AvD1::GFP stock was kindly provided by Dr. Saint.

For H1 depletion experiments *yw*; *Nubb*Gal4; UAS-*Dicer2* and *dH1*^{RNAi}; Sb/TM3 flies were crossed at 25°C as described before (Bayona-Feliu et al., 2017). For hrp36 and hrp48 depletion experiments *w*; *Nubb*GAL4 flies and either *hrp36*^{RNAi} (V51759) or *hrp48*^{RNAi} (V101555) flies were crossed at 29°C.

Culture of *Drosophila* S2 cells and depletion experiments

Drosophila S2 cells were cultured at 25ºC in Schneider's Insect Medium (L0207-500, Biowest) supplemented with 10% heat-inactivated FBS (10270, Gibco) and 1% penicillin-streptomycin (15140-122, Gibco). For RNAi-mediated knockdown experiments, dsRNAs against the coding regions of dH1, hrp36 and hrp48 were prepared by *in vitro* transcription using MEGAscript T7 kit (AM1334, Thermo Fisher Scientific) from appropriate PCR templates produced as described in (Rogers and Rogers, 2008) with the primers summarized in **Supplementary Table S2**. For depletion, 1.5×10^7 cells were resuspended in Serum-Free medium with 50 μg of dsRNA and seeded in a T-75 flask (Corning). After 1 hour at room temperature, medium was adjusted to 10% FBS and cells

were collected for downstream processing after three days. For dH1 knockdown, a second dose of 50μg of dsRNA was applied for three additional days before harvesting of the cells. As control, cells were subjected in parallel to mock depletion with dsRNA against the bacterial LacZ gene (see **Supplementary Table S2** for primers**)**.

Antibodies

Rabbit polyclonal αdH1 was kindly provided by Dr Kadonaga and is described in (Vujatovic et al., 2012) [WB: 1:10000]. Mouse monoclonal αhrp36 was kindly provided by Dr Saumweber (Hovemann et al., 2000) [WB: 1:400; ChIP: 4 μg; IF: 1:200]. Rabbit polyclonal αhrp48 was kindly provided by Dr Gebauer (Szostak et al., 2018) [WB: 1:5000; ChIP: 6 μL; IF 1:200]. Mouse monoclonal S9.6 is described in (Boguslawski et al., 1986) [DRIP: 10 μg; IF 1:500]. Rat αHP1a is described in (Font-Burgada et al., 2008) [IF 1:200]. The rest of antibodies were commercially available: rabbit polyclonal αH4 (Abcam (ab10158) [WB: 1:1000]), rabbit polyclonal αGFP (Life technologies (A11122) [IF: 1:500]) and mouse αGFP (Roche (11814 460 001) [IF: 1:50]).

Immunofluorescence analysis of polytene chromosomes

Polytene chromosomes from 3rd instar larvae were prepared and immunostained as described in (Bayona-Feliu et al., 2017). Briefly, 3rd instar salivary glands were manually dissected and incubated in Cohen's solution for 10 min at room temperature. Then, after fixation in PBS containing 0.74% *p*-formaldehyde for 2 min, and in 45% acetic acid, 0.74% *p*-formaldehyde for 3 min, glands were transferred to a cover slip, put a slide on and squashed. Chromosomes were transferred to the slide after immersion in liquid nitrogen and immunostained. RNase A treatment of polytene chromosomes was performed as above with the only modification that RNase A was added at 100 μg/mL in Cohen's solution. Images were recorded with a Leica SPE confocal microscope. When image quantification was required, salivary glands from hrp36- and hrp48-depleted larvae were mixed with control salivary glands from larvae of flies carrying a H2Av::GFP transgene before squashing and processed together to minimize experimental variability. To differentiate polytene chromosomes from control and depleted larvae, samples were immunostained with αGFP antibodies.

Image quantification was performed using FIJI software using sum slices and an *ad-hoc* macro kindly provided by Anna Lladó (Advanced Digital Microscopy Facility, IRB Barcelona).

cRNAs preparation, sequencing and analysis

cRNAs were prepared essentially as described in (Werner and Ruthenburg, 2015). Briefly, 10^8 cells were collected, washed twice with PBS and resuspended in 400 μL of Buffer A (10mM HEPES pH 7.9, 10mM KCl, 1.5mM $MgCl_2$, 0.34M Sucrose, 10% Glycerol). Triton X-100 was added to a final concentration of 0.1% and incubated for 8 min on ice. Nuclei were collected by sucrose cushion centrifugation (24% sucrose in lysis Buffer A), rinsed with ice-cold PBS + 1mM EDTA, and resuspended in 400 μL of ice-cold Glycerol Buffer (20mM Tris pH 7.9, 75mM NaCl, 0.5mM EDTA, 50% glycerol). Nuclei were lysed by adding 400 μL of ice-cold Lysis Buffer B (10mM HEPES pH 7.9, 7.5mM $MgCl_2$, 3mM EDTA, 0.2mM EGTA, 0.3M NaCl, 1M urea, 1% NP40), shaken vigorously and kept on ice for 10 min, with periodic shaking. Chromatin was sedimented by centrifugation at 14000g for 2 min at 4ºC, rinsed twice with cold PBS + 1mM EDTA, and resuspended in 150 μL PBS + 1mM EDTA. Then, RNAs were purified using RNAzol RT adding 400 μL per sample and 100 μL of chloroform, shaking the samples vigorously, incubating them on ice 3 min and centrifuging for 5 min at maximum speed at 4ºC. Subsequently, the aqueous phase was recovered and purified with the RNeasy Mini Kit (74104, Quiagen) following manufacturer's instructions.

For NGS, two independent biological replicates were obtained for each condition. Before library preparation, samples were subjected to rRNAs depletion with Ribo-Zero Plus rRNA Depletion Kit (Illumina). Libraries were generated using NEBNext Ultra II Directional RNA Library Prep Kit (Illumina) and sequenced 2x150 paired-end at the CNAG (Barcelona). For analysis of the retrieved reads, Illumina adapters were removed using cutadapt (v1.18) (Martin, 2011). Trimmed reads were aligned to *Drosophila* rRNA sequences (Quast et al., 2013) using Bowtie2 (Langmead and Salzberg, 2012) with parameters q --local -N 1 -k 1. rRNA unaligned reads were then aligned to dm3 genome using tophat2 (v2.1.1) (Kim et al., 2013) with options --library-type fr-firststrand -g 1. Aligned reads were

separated by strand (Crick vs Watson strands) using samtools (v1.3) (Li et al., 2009). Independently for the two strands, reads were merged by condition and assembled into *de novo* transcripts using Cufflinks (Mv2.2.1) (Roberts et al., 2011). Counts per transcript were computed with the R package Rsubread (v2.8.2) (Liao et al., 2013), function featureCounts with default parameters. For every condition, low abundance (<20 reads) transcripts were removed. Overlap plots (Venn diagrams) between conditions and chromosome location plots were represented for the transcriptomes obtained at this point. For differential expression analysis, the consensus transcriptomes between control and experimental conditions were considered. These transcriptomes were further curated: transcripts that overlapped partially with an annotated gene (±100bp from start/end coordinates) in the RefGene database were split, taking both the split transcripts located outside genic regions and all genic regions defined in RefGene. Reads per transcript were recalculated with featureCounts, keeping as final transcripts those with 10 or more reads. Transcriptomes were annotated with Homer (Heinz et al., 2010) leading to 5 categories of transcripts: genic-sense defined in RefGene database; genic-antisense were transcripts that overlap with genes on the complementary strand; intergenic and lncRNAs were defined by Homer; and UTRs/Promoter-TSS were defined by those transcripts that were not labeled as intergenic by Homer but where not annotated in the RefGene database. These included mostly less curated genic regions. Repetitive elements compositions of heterochromatin transcripts and heterochromatin aligned reads were obtained using repeatMasker (v4.1.2) (Smit et al.). Differential expression analysis was performed using DESeq2 (v1.34) (Love et al., 2014). To test for statistical significance, for any subset of transcripts to be analyzed, we considered the difference between the number of differentially up- and down-regulated transcripts (with p-value< 0.05 in the differential expression results) and the null distribution was obtained by re-calculating this difference for 1,000 instances of transcriptome subsets resampled from the whole transcriptome using the same number of transcripts as in each subset. When intron retention was determined, for genic regions annotated in the refGene database, gene-exon coverage was calculated with featureCounts (Liao et al., 2013) with options

GTF.featureType="exon", GTF.attrType = "gene_id", similarly gene-transcript coverage was obtained with featureCounts with options GTF.featureType="transcript", GTF.attrType = "gene_id", allowMultiOverlap = TRUE. Intron retention differences were assessed with DESeq2 (v1.34) (Love et al., 2014) by comparing exon coverage *vs* transcript coverage.

ChIP-seq analyses

ChIP experiments were performed essentially as described in (Climent-Cantó et al., 2020). Briefly, 10^8 cells were collected and fixed with 1.8% of formaldehyde for 10 min at room temperature. Then, cells were washed and lysed in ChIP wash A (10mM HEPES pH 7.9, 10mM EDTA, 0.5mM EGTA, 0.25% Triton X-100), followed by ChIP wash B (10mM HEPES pH 7.9, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.01% Triton X-100). Chromatin was pelleted, resuspended in TE (10mM Tris–HCl pH 8.0, 1mM EDTA) with 1% SDS, washed three times in TE and resuspended in TE, 1mM PMSF, 0.1% SDS. Chromatin was sonicated in a Bioruptor Twin Sonication System (Diagenode) to obtain fragments of 200–500 bp. After sonication, lysates were adjusted to 1% Triton X-100, 0.1% sodium deoxycholate (DOC), 140mM NaCl. For immunoprecipitation (IP), 400 μL of chromatin were used and 40 μL were saved as input sample. IPs were carried out in RIPA buffer (140mM NaCl, 10 mM Tris–HCl pH 8.0, 1mM EDTA, 1% Triton X-100, 0.01% SDS, 0.1% DOC). A pre-clearing step was carried out with 30 μL of 50% (v/v) protein A-Sepharose CL4B beads (GE Healthcare, 17-0780-01) previously blocked with RIPA, 1% BSA. Then, the corresponding antibodies were added and incubated over night at 4°C in a rotating wheel. IPs were performed by adding 40 μL of 50% (v/v) protein A-Sepharose CL4B beads previously blocked with RIPA, 1% BSA and incubated in a rotating wheel for 3 hours at 4°C. Beads were washed with RIPA buffer, once with 250mM LiCl, 10mM Tris–HCl pH 8.0, 1mM EDTA, 0.5% NP-40, 0.5% DOC and twice with TE. Then beads were treated with DNase-free RNase A. To purify the DNA, samples were adjusted to 1% SDS, 0.1 M $NaHCO_3$, 0.2 mg/mL Proteinase K and incubated over night at 65°C. DNA was purified by phenol-chloroform extraction.

For NGS, two independent biological replicates were obtained for each condition and libraries were generated using NEBNext Ultra II DNA Library Prep

Kit (Illumina) and sequenced by 1 x 50 single reads at the CNAG (Barcelona). FastQ files were aligned to the dm3 genome using Bowtie2 v2.2.2 (Langmead and Salzberg, 2012) with options -N 1 -k 1. Duplicated reads for generation of TDF tracks were identified and removed with sambamba 0.8.0 (Tarasov et al., 2015). TDF tracks were generated with IGVTools2 (Thorvaldsdóttir et al., 2013) and options count -z 5 -w 25 -e 250. Peak calling between IP and Input samples was performed independently for individual and pooled replicates per group using MACS 1.4.2 (Zhang et al., 2008) with default options (duplicate removal) and genome=dm. Candidate called peaks from pooled replicates were defined as final peaks if they presented overlap with peaks identified in both their respective individual replicates. For differential binding analysis between conditions, the DiffBind package version 2.10.0 (Ross-Innes et al., 2012) was used, with functions: dba.count with options *score = DBA_SCORE_RPKM_FOLD, bLog = TRUE, bScaleControl = TRUE*; dba.contrast with options *categories = DBA_CONDITION, block = DBA_REPLICATE, minMembers = 2*; and dba.analyze with options *method = DBA_DESEQ2, bFullLibrarySize = TRUE, bSubControl = FALSE*. Peaks were annotated using Homer version 2 (Heinz et al., 2010) using the built-in dm3 genome. Peak chromosomal location was plotted using htSeqTools v1.4.2 (Planet et al., 2012). Peak/Feature distribution was obtained from Homer annotated tables. Statistical assessment of peak overlaps was performed using overlap permutation tests with the regioneR package version 1.14.0 (Gel et al., 2016) and default options. Unless otherwise specified all downstream analyses were performed with R3.5.1. (R-Core-Team, 2023).

DRIP-seq analyses

For DRIP-seq experiments, 10^7 cells were collected, washed and lysed in Lysis Buffer DRIP (10mM Tris-HCl pH 7.4, 10mM EDTA, 10mM NaCl, 0.5% SDS, 0.5 μg/μL proteinase K), and genomic DNA was purified by phenol-chloroform extraction followed by ethanol precipitation. DNA was resuspended in 50 μL of ultrapure water and the concentration was measured by Qubit. 50 μg of sample were adjusted to TE and sonicated in a Bioruptor Twin Sonication System (Diagenode) to obtain fragments of 400-500 bp. Then, each sample was separated in two Eppendorf's tubes containing 5μg of genomic DNA in RNaseH

Reaction Buffer (50mM Tris-HCl pH 8.0, 75mM KCl, 3mM $MgCl_2$, 10mM DTT). One of the replicates was untreated and the other one was treated adding 6U RNaseH (18021-071, Invitrogen) as a control for the specificity of the signal. Both were incubated at 37ºC overnight. Then, Binding Buffer DRIP (10mM Phosphate buffer pH 7.3, 140mM NaCl, 0.05% TritonX-100) and 10μg of mouse monoclonal S9.6 antibody was added to each sample and incubated in a rotatory wheel at 4ºC overnight. The following day, preblocked Dynabeads Protein G (10003D, Thermo Fisher Scientific) (with Binding Buffer DRIP, 0.5% BSA) were added to each sample and incubated in a rotatory wheel at 4ºC for 2 h. Beads were washed with Binding Buffer DRIP. To purify the DNA, samples were eluted with three sequential washes of DRIP Elution Buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 0.5% SDS). An input sample was prepared with 10% of initial sample in DRIP Elution Buffer. Then, 0.2 mg/mL Proteinase K was added and incubated at 55ºC for 45 min. DNA was purified by phenol-chloroform extraction and ethanol precipitation.

For NGS, two independent biological replicates were obtained for each condition and libraries were generated using NEBNext Ultra II DNA Library Prep Kit (Illumina) and sequenced by 2 x 150 paired-end reads on a NovaSeq 6000 (Illumina) at the CNAG Barcelona. For analysis of the retrieved reads, NGMerge (Gaspar, 2018) was used to automatically detect and remove sequencing adapters with options -a -n 18 -v. Alignment of paired-end reads and peak calling were performed as described above for ChIP-seq analyses. Called peaks from pooled replicates were defined as final peaks (candidate Rloops) if they presented overlap with peaks identified in both their respective individual replicates. Definition of final Rloops was performed using DESeq2 with lfcShrinkage fold change estimation (type='normal') to assess level of RNH+ depletion. Final Rloops were defined as those RNH- candidate Rloops with a DESeq2 lfcShrinkage FC RNH+/RNH- (RNH depletion) of -1.25 or less. Peak annotation, chromosomal location and statistical assessment of peak overlaps were performed as described above for ChIP-seq analyses.

ATAC-seq analyses

ATAC-seq experiments were performed as described in (Buenrostro et al., 2013). Briefly, 10^5 nuclei were resuspended in 50 μL Tn5 transposase mixture

(2.5 µL Tn5, 25 µL 2XTD buffer, 22.5 µL RNase Free Water), and incubated at 37°C for 30 min. After incubation, DNA was purified using the MinElute PCR Purification Kit (QIAGEN). For NGS, two independent biological replicates were obtained for each condition. For library preparation, PCR cycles were determined by qPCR. Following PCR amplification, size selection of the library was performed using Ampure beads (Agencourt RNAClean XP Beads, S01307). Libraries were sequenced by 2 x 50 paired-end reads at CRG Barcelona. Bioinformatics analyses of ATAC-seq data was performed according to the guidelines of Harvard FAS Informatics (https://informatics.fas.harvard.edu/atac-seq-guidelines.html). NGMerge (Gaspar, 2018) was used to automatically detect and remove sequencing adapters using options -a -n 10 -v. Adapter cleaned FastQ files were downsampled to the library size of the smallest replicate/pair file (˜21.2 million reads) using setqtk (https://github.com/lh3/seqtk). Fragment size and distributions of ATAC-seq samples were analyzed using plot2DO (https://github.com/rchereji/plot2DO) including the scaffolds 2LHet, 2RHet, 3LHet, 3RHet, YHet and XHet, U and Uextra. For analysis of THSS fragments, reads were aligned to dm3 with Bowtie2 using a maximum insert size of 100bp (options --very-sensitive -k 10 -X 100), and Genrich version 0.5 (https://github.com/jsh58/Genrich) was used to detect putative ATAC peaks with options -j -y -r -e chrM. Differential analysis between conditions for all putative ATAC peaks was performed with DESeq2 version 1.22.1 (Love et al., 2014) using lfcShrinkage fold change estimation with options type='normal'.

Western Blot (WB) analyses

WB analyses were performed by standard procedures using total cell extracts obtained by resuspending cells in PLB (25mM Tris-HCl pH 6.8, 4.35% Glycerol, 1% SDS) or crosslinked chromatin obtained as described above for ChIP-seq experiments. When the effect of RNase A treatment was determined by WB, cells were fixed and permeabilized with ChIP wash A and ChIP wash B as described above, and resuspended in 4.5 mL of TE. Then, the treated samples were supplemented with 100μg/mL RNaseA, and both control and treated samples were incubated at 37ºC for 30 min. After incubation, samples were lysed with 1% SDS and chromatin was washed, sonicated, adjusted to 1% Triton X-

100, 0.1% sodium deoxycholate (DOC), 140mM NaCl as described above, and supplemented with PLB. For quantification, autoradiographic films were digitalized using a GS-800 Calibrated Laser Densitometer (Bio-Rad) and quantified using ImageJ Gel Analyzer plugin. Signal from each lane was normalized to the corresponding loading control. Two to three lanes were quantified per biological sample, and at least three independent biological replicates were quantified per experiment. Two-tailed paired Student's t-test was performed using GraphPad software.

RT-qPCR analyses

RT-qPCR analyses were performed by standard procedures. Briefly, cDNA was synthesized using Transcription First Strand cDNA Synthesis Kit (Roche) and qPCR was performed using SYBR Green I Master (Roche). Relative expression levels were calculated by ΔΔCt method and normalized respect to Rpl32. **Supplementary Table S3** summarizes the primers used in these experiments. Two-tailed paired Student's t-test was performed using GraphPad software.

DATA AVAILABILITY

cRNA-seq data generated in this work is GSE227997. ChIP-seq, DRIP-seq and ATAC-seq data generated in this work are GSE228142.

ACKNOWLEDGEMENTS

We are thankful to Dr. Casali and Dr. Saint for fly stocks, and to Dr Kadonaga, Dr Saumweber and Dr Gebauer for antibodies. We also thankful to the IRB Functional Genomics, Advance Digital Microscopy, and Biostatistics and Bioinformatics core facilities. This work was financed by grants PGC2018-094538-B-100 and PID2021-123303NB-I00 from MICIN/AEI 10.13039/501100011033 and "FEDER, una manera de hacer Europa", and of the Generalitat de Catalunya (SGR2017-475). P.B acknowledges receipt of an FPI fellowship from MICIN/AEI 10.13039/501100011033 and "FEDER, una manera de hacer Europa".

REFERENCES

Abakir, A., Giles, T.C., Cristini, A., Foster, J.M., Dai, N., Starczak, M., Rubio-Roldan, A., Li, M., Eleftheriou, M., Crutchley, J.*, et al.* (2020). N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. Nat Genet *52*, 48-55.

Alarcón, C.R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S.F. (2015a). HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. Cell *162*, 1299-1308.

Alarcón, C.R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S.F. (2015b). N6-methyladenosine marks primary microRNAs for processing. Nature *519*, 482-485.

Alekseyenko, A.A., Gorchakov, A.A., Zee, B.M., Fuchs, S.M., Kharchenko, P.V., and Kuroda, M.I. (2023). Heterochromatin-associated interactions of *Drosophila* HP1a with dADD1, HIPP1, and repetitive RNAs. Genes Dev *28*, 1445-1460.

Almeida, R., Fernández-Justel, J.M., Santa-María, C., Cadoret, J.C., Cano-Aroca, L., Lombraña, R., Herranz, G., Agresti, A., and Gómez, M. (2018). Chromatin conformation regulates the coordination between DNA replication and transcription. Nat Commun *9*, 1590.

Bayona-Feliu, A., and Aguilera, A. (2021). The role of chromatin at transcription-replication conflicts as a genome safeguard. Biochem Soc Trans *49*, 2727-2736.

Bayona-Feliu, A., Casas-Lamesa, A., Carbonell, A., Climent-Cantó, P., Tatarski, M., Pérez-Montero, S., Azorín, F., and Bernués, J. (2016). Histone H1: Lessons from *Drosophila*. Biochim Biophys Acta *1859*, 526-532.

Bayona-Feliu, A., Casas-Lamesa, A., Reina, O., Bernués, J., and Azorín, F. (2017). Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. Nat Commun *8*, 283.

Bernués, J., Izquierdo-Boulstridge, A., Reina, O., Castejón, L., Fernández-Castañer, E., Leal, N., Guerrero-Pepinosa, N., Bonet-Costa, C., Vujatovic, O., Climent-Cantó, P.*, et al.* (2022). Lysine 27 dimethylation of Drosophila linker histone dH1 contributes to heterochromatin organization independently of H3K9 methylation. Nucleic Acids Res *50*, 9212-9225.

Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., and Carrico, R.J. (1986). Characterization of monoclonal antibody to DNA·RNA and its application to immunodetection of hybrids. J Immunol Methods *89*, 123-130.

Bonner, J., and Widholm, J. (1967). Molecular complementarity between nuclear DNA and organ-specific chromosomal RNA. Proc Natl Acad Sci USA. *57*, 1379-1385.

Braunschweig, U., Hogan, G.J., Pagie, L., and van Steensel, B. (2009). Histone H1 binding is inhibited by histone variant H3.3. EMBO J. *28*, 3635-3645.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213–1218.

Chen, L., Chen, J.Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H.*, et al.* (2017). R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters. Mol Cell *68*, 745-757.

Chereji, R.V., Bryson, T.D., and Henikoff, S. (2019). Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol *20*, 198.

Climent-Cantó, P., Carbonell, A., Tatarski, M., Reina, O., Bujosa, P., Font-Mateu, J., Bernués, J., Beato, M., and Azorín, F. (2020). The embryonic linker histone dBigH1 alters the functional state of active chromatin. Nucleic Acids Res *48*, 4147-4160.

Cohen, A.K., Huh, T.Y., and Helleiner, C.W. (1973). Transcription of satellite DNA in mouse L-cells. Can J Biochem *51*, 529-532.

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nat. Rev. Mol. Cell Biol. *3*, 195–205.

Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing *in vivo*. Mol Cell Biol *23*, 4559-4572.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell *123*, 1199-1212.

Fan, Y., Sirotkin, A., Russell, R.G., and Skoultchi, A.I. (2001). Individual somatic H1 subtypes are dispensable for mouse development even in mice lacking the H1(0) replacement subtype. Mol Cell Biol *21*, 7933-7943.

Fernández-Justel, J.M., Santa-María, C., Martín-Vírgala, S., Ramesh, S., Ferrera-Lagoa, A., Salinas-Pena, M., Isoler-Alcaraz, J., Maslon, M.M., Jordan, A., Cáceres, J.F.*, et al.* (2022). Histone H1 regulates non-coding RNA turnover on chromatin in a m6A-dependent manner. Cell Rep *40*, 111329.

Font-Burgada, J., Rossell, D., Auer, H., and Azorín, F. (2008). *Drosophila* HP1c isoform interacts with zinc-finger proteins WOC and relative-of-WOC to regulate gene expression. Genes Dev *22*, 3007–3023.

Fyodorov, D.V., Zhou, B.R., Skoultchi, A.I., and Bai, Y. (2018). Emerging roles of linker histones in regulating chromatin structure and function. Nat Rev Mol Cell Biol *19*, 192-206.

García-Muse, T., and Aguilera, A. (2019). R Loops: From physiological to pathological roles. Cell *179*, 604-618.

Gaspar, J.M. (2018). NGmerge: merging paired-end reads via novel empirically-derived models of sequencing errors. BMC Bioinformatics *19*, 536.

Gaubatz, J.W., and Cutler, R.G. (1990). Mouse satellite DNA is transcribed in senescent cardiac muscle. J Biol Chem *265*, 17753-17758.

Gayen, S., and Kalantry, S. (2017). Chromatin-enriched lncRNAs: a novel class of enhancer RNAs. Nat Struct Mol Biol *24*, 556-557.

Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics *32*, 289-291.

Grewal, S.I., and Jia, J. (2007). Heterochromatin revisited. Nat Rev Genet *8*, 35-46.

Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504-516.

Hashimoto, H., Takami, Y., Sonoda, E., Iwasaki, T., Iwano, H., Tachibana, M., Takeda, S., Nakayama, T., Kimura, H., and Shinkai, Y. (2010). Histone H1 null vertebrade cells exhibit altered nucleosome architecture. Nucleic Acids Res. *38*, 3533-3545.

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., and Soller, M. (2016). m6A potentiates Sxl alternative pre-mRNA splicing for robust *Drosophila* sex determination. Nature *540*, 301-304.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Holmes, D.S., Mayfield, J.E., Sander, G., and Bonner, J. (1972). Chromosomal RNA: its properties. Science *177*, 72-74.

Hovemann, B.T., Reim, I., Werner, S., Katz, S., and Saumweber, H. (2000). The protein Hrb57A of *Drosophila melanogaster* closely related to hnRNP K from vertebrates is present at sites active in transcription and coprecipitates with four RNA-binding proteins. Gene *245*, 127–137.

Huang, R.C., and Bonner, J. (1965). Histone-bound RNA, a component of native nucleohistone. Proc Natl Acad Sci USA. *54*, 960-967.

Huo, X., Ji, L., Zhang, Y., Lv, P., Cao, X., Wang, Q., Yan, Z., Dong, S., Du, D., Zhang, F.*, et al.* (2020). The nuclear matrix protein SAFB cooperates with major satellite RNAs to stabilize heterochromatin architecture partially through phase separation. Mol Cell *77*, 368-383.

Izquierdo-Bouldstridge, A., Bustillos, A., Bonet-Costa, C., Aribau-Miralbés, P., García-Gomis, D., Dabad, M., Esteve-Codina, A., Pascual-Reguant, L., Peiró, S., Esteller, M.*, et al.* (2017). Histone H1 depletion triggers an interferon response in cancer cells via activation of heterochromatic repeats. Nucleic Acids Res *45*, 11622-11642.

Johnson, W.L., Yewdell, W.T., Bell, J.C., McNulty, S.M., Duda, Z., O'Neill, R.J., Sullivan, B.A., and Straight, A.F. (2017). RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. Elife *6*, e25299.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T.*, et al.* (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. Nature *471*, 480-485.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). 1000 Genome project data processing subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, X., and Fu, X. (2019). Chromatin-associated RNAs as facilitators of functional genomic interactions. Nat Rev Genet *29*, 503-519.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res *41*, e108.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature *518*, 560-564.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Lu, J., and Gilbert;, D.M. (2007). Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. J Cell Biol *179*, 411-421.

Lu, X., Wontakal, S.N., Emelyanov, A.V., Morcillo, P., Konev, A.Y., Fyodorov, D.V., and Skoultchi, A.I. (2009). Linker histone H1 is essential for *Drosophila* development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. Genes Dev *23*, 452-465.

Lu, X., Wontakal, S.N., Kavi, H., Kim, B.J., Guzzardo, P.M., Emelyanov, A.V., Xu, N., Hannon, G.J., Zavadil, J., Fyodorov, D.V., *et al.* (2013). *Drosophila* H1 regulates the genetic activity of heterochromatin by recruitment of Su(var)3-9. Science *340*, 78-81.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A.V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J.P., *et al.* (2011). SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. Nat Genet *43*, 220-227.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *1*, https://doi.org/10.14806/ej.14817.14801.14200.

Matunis, E.L., Matunis, M.J., and Dreyfuss, G. (1993). Association of individual hnRNP proteins and snRNPs with nascent transcripts. J Cell Biol *121*, 219-228.

Matunis, M.J., Matunis, E.L., and Dreyfuss, G. (1992). Isolation of hnRNP Complexes from *Drosophila melanogaster*. J Cell Biol *116*, 245-255.

Muchardt, C., Guilleme, M., Seeler, J.S., Trouche, D., Dejean, A., and Yaniv, M. (2002). Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. EMBO Rep. 3, *3*, 975–981.

Ni, J.Q., Liu, L.P., Hess, D., Rietdorf, J., and Sun, F.L. (2006). *Drosophila* ribosomal proteins are associated with linker histone H1 and suppress gene transcription. Genes Dev *20*, 1959-1973.

Pérez-Montero, S., Carbonell, A., and Azorín, F. (2016). Germline-specific H1 variants: the "sexy" linker histones. Chromosoma *125*, 1-13.

Piacentini, L., Fanti, L., Negri, R., Del Vescovo, V., Fatica, A., Altieri, F., and Pimpinelli, S. (2009). Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. PLoS Genet *5*, e1000670.

Planet, E., Attolini, C.S., Reina, O., Flores, O., and Rossell, D. (2012). htSeqTools: high-throughput sequencing quality control, processing and visualization in R. Bioinformatics *28*, 589-590.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res *41*, D590-D596.

R-Core-Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011). Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics *27*, 2325-2329.

Rodríguez-Campos, A., and Azorín, F. (2007). RNA is an integral component of chromatin that contributes to its structural organization. PLoS ONE *2*, e1182.

Rogers, S.L., and Rogers, G.C. (2008). Culture of *Drosophila* S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy. Nat Protoc *3*, 606-611.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R.*, et al.* (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389-393.

Rudert, F., Bronner, S., Garnier, J.M., and P Dollé, D. (1995). Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. Mamm Genome *6*, 76-83.

Sancho, M., Diani, E., Beato, M., and Jordan, A. (2008). Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. PLoS Genet. *4*, e1000227.

Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J., and Torres-Padilla, M.E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat Cell Biol *12*, 853-862.

Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X., and Chédin, F. (2016). Prevalent, dynamic, and conserved R-Loop structures associate with specific epigenomic signatures in mammals. Mol Cell *63*, 167-178.

Schoeftner, S., and Blasco, M.A. (2008). Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat Cell Biol *10*, 228-236.

Sirotkin, A.M., Edelmann, W., Cheng, G., Klein-Szanto, A., Kucherlapati, R., and Skoultchi, A.I. (1995). Mice develop normally without the H1(0) linker histone. Proc. Natl. Acad. Sci. USA *92*, 6434-6438.

Smit, A.F.A., Hubley, R., and Green, P. RepeatMasker Open-4.0. 2013-2015 <http://www.repeatmasker.org>.

Sun, X., Wang, Z., Hall, J.M., Perez-Cervantes, C., Ruthenburg, A.J., Moskowitz, I.P., Gribskov, M., and Yang, X.H. (2020). Chromatin-enriched RNAs mark active and repressive cis-regulation: An analysis of nuclear RNA-seq. PLoS Comput Biol *16*, e1007119.

Szostak, E., García-Beyaert, M., Guitart, T., Graindorge, A., Coll, O., and Gebauer, F. (2018). Hrp48 and eIF3d contribute to msl-2 mRNA translational repression. Nucleic Acids Res *46*, 4099–4113.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032-2034.

Thakur, J., and Henikoff, S. (2020). Architectural RNA in chromatin organization. Biochem Soc Trans *48*, 1967–1978.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.

Vujatovic, O., Zaragoza, K., Vaquero, A., Reina, O., Bernués, J., and Azorín, F. (2012). *Drosophila melanogaster* linker histone dH1 is required for transposon silencing and to preserve genome integrity. Nucleic Acids Res *40*, 5402-5414.

Werner, M.S., and Ruthenburg, A.J. (2015). Nuclear fractionation reveals thousands of chromatin-tethered noncoding RNAs adjacent to active genes. Cell Rep *12*, 1089-1098.

Werner, M.S., Sullivan, M.A., Shah, R.N., Nadadur, R.D., Grzybowski, A.T., Galat, V., Moskowitz, I.P., and Ruthenburg, A.J. (2017). Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription. Nat Struct Mol Biol *24*, 596-603.

Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H.Y., Li, A., Ping, X.L., Lai, W.Y., *et al.* (2016). Nuclear m(6)A eeader YTHDC1 regulates mRNA Splicing. Mol Cell *61*, 507-519.

Yang, X., Liu, Q.L., Xu, W., Zhang, Y.C., Yang, Y., Ju, L.F., Chen, J., Chen, Y.S., Li, K., Ren, J., *et al.* (2019). m6A promotes R-loop formation to facilitate transcription termination. Cell Res *29*, 1035-1038.

Zhang, J., Ding, T., and Zhang, H. (2021). Insight into chromatin-enriched RNA: A key chromatin regulator in tumors. Front Cell Dev Biol *9*, 649605.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y.C., Zhou, Y.F., Cheng, Y., Huang, J.H., Lian, J.P., Yang, L., He, R.R., Lei, M.Q., Liu, Y.W., Yuan, C., *et al.* (2022). Genome-wide analysis and functional annotation of chromatin-enriched noncoding RNAs in rice during somatic cell regeneration. Genome Biol *23*, 28.

FIGURE LEGENDS

Figure 1. cRNAs are enriched in heterochromatic transcripts. (**A**) The chromosomal distribution of cRNAs along *Drosophila* chromosomes is shown for both the Watson and Crick strands. chr2L and chr2R, and chr3L and chr3R correspond to chromosome 2 and 3 left and right arms respect to the position of the centromere, respectively. chr4 and chrX are oriented with the centromere to the right. chr2LHet, chr2RHet, chr3LHet, chr3RHet, chrXHet and chrYHet correspond to assembled pericentromeric heterochromatin regions of the indicated chromosomes. chrU and chrUextra correspond to unassembled highly repetitive heterochromatic scaffolds. Euchromatic and heterochromatic scaffolds are shown in blue and red, respectively. On the bottom, pie chart showing the proportion of cRNAs mapping to euchromatic and heterochromatic scaffolds. (**B**) The proportion of cRNAs matching lncRNAs, intergenic, genic sense and antisense, and regulatory (UTRs, promoter, TSS) regions are shown. (**C**) The distribution by quantile of RPKM counts is shown for euchromatic and heterochromatic cRNAs. (**D**) Immunostainings of polytene chromosomes with αhrp36 (in yellow) and αhrp48 (in magenta) antibodies before (top) and after (bottom) treatment with RNase A. Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter (arrowheads). DNA was stained with DAPI. Scale bars correspond to 20μm. (**C**) Western blot (WB) analysis with αhrp36 (left) and αhrp48 (right) antibodies of increasing amounts (lanes 1-3) of crosslinked chromatin prepared from S2 cells treated or not with RNase A after crosslinking. αH4 antibodies were used for loading control. Quantification of the results is shown on the bottom. Results are the average of 3 independent experiments. Error bars are standard deviation. (p-values: *< 0.05, **< 0.01; two-tailed paired Student's test).

Figure 2. **ChIP-seq analyses show the association of hrp36 and hrp48 with cRNAs**. (**A**) On the top, chromosomal distributions of hrp36 (left) and hrp48 (right) ChIP-seq peaks. Chromosome definition and orientation are as in **Figure 1A**. Euchromatic and heterochromatic scaffolds are shown in blue and red, respectively. On the bottom, pie charts showing the proportion of hrp36 (left) and

hrp48 (right) ChIP-seq peaks mapping to euchromatic and heterochromatic scaffolds. (**B**) Venn diagram showing the overlap between hrp36 and hrp48 ChIP-seq peaks. (**C**) The proportion of common hrp36 and hrp48 ChIP-seq peaks assigned to each of the nine chromatin epigenetic states according to (Kharchenko et al., 2011). (**D**) Pie charts showing the proportion of hrp36 (top) and hrp48 (bottom) ChIP-seq peaks overlapping (green) or not (red) with cRNAs. (**E**) The abundance of cRNAs at regions overlapping (green) or not (red) with common hrp36 and hrp48 ChIP-seq peaks is shown as the frequency of RPKM counts ranked by percentile.

Figure 3. **dH1 depletion impairs assembly of hrp36 and hrp48 onto cRNAs**. (**A**) Immunostainings of polytene chromosomes from control w^{1118} (top) and dH1-depleted *dH1*^{RNAi} (bottom) flies with αhrp36 (in yellow) and αhrp48 (in magenta) antibodies. Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter (arrowheads). DNA was stained with DAPI. Scale bars correspond to 20 μm. (**B**) Box plots of the average grey intensity of fluorescence at the chromocenter after immunostaining with αhrp36 (left) and αhrp48 (right) antibodies is shown for control w^{1118} and dH1-depleted *dH1*^{RNAi} flies (N > 11; p-values: ****< 0.0001, *****< 0.00001, Kruskal-Wallis test). (See also **Supplementary Figure S2**). (**C**) WB analysis with αdH1 (top), αhrp36 (center) and αhrp48 (bottom) antibodies of increasing amounts (lanes 1-4) of crosslinked chromatin prepared from control $dsRNA^{LacZ}$ (left) and dH1-depleted $dsRNA^{dH1}$ (right) cells. αH4 antibodies were used for loading control. Quantification of the results is shown on the right. Results are the average of 5 independent experiments. Error bars are standard deviation. (p-values: *< 0.05; two-tailed paired Student's test). (**D**) RT-qPCR experiments showing the mRNA levels of hrp36 (left) and hrp48 (right) relative to Rpl32 in control $dsRNA^{LacZ}$ (grey) and dH1-depleted $dsRNA^{dH1}$ (black) cells. Results are the average of 3 independent experiments. Error bars are standard deviation. (p-values> 0.05; two-tailed Student's test). (**E**) On the left, Volcano plot showing the change in intensity of hrp36 ChIP-seq peaks in dH1-depleted $dsRNA^{dH1}$ cells in comparison to control $dsRNA^{LacZ}$ cells. Down- and up-regulated peaks are shown in blue and

red, respectively. On the right, pie chart showing the proportion of differentially down- and up-regulated peaks (p-value< 0.05). (**F**) As in **E** but for hrp48 ChIP-seq peaks.

Figure 4. **Impaired RNPs assembly results in the accumulation of heterochromatic cRNAs**. (**A**) Venn diagram showing the overlap between cRNAs detected in control dsRNA^{LacZ} and in dH1-depleted dsRNA^{dH1} cells. (**B**) On the left, Volcano plot showing the change in intensity of all cRNAs detected in dH1-depleted dsRNA^{dH1} cells in comparison to control dsRNA^{LacZ}. Down- and up-regulated transcripts are shown in blue and red, respectively. On the right, pie chart showing the proportion of differentially down- and up-regulated cRNAs (p-value < 0.05). (**C**) As in **B**, but for heterochromatic cRNAs. (**D**) Venn diagram showing the overlap between cRNAs detected in dH1-depleted dsRNA^{dH1}, hrp36-depleted dsRNA^{hrp36} and hrp48-depleted dsRNA^{hrp48} cells. (**E** and **F**) As in **B**, but for heterochromatic cRNAs detected in hrp48-depleted dsRNA^{hrp48} and hrp36-depleted dsRNA^{hrp36} cells, respectively. (**G-I**) On the left, Volcano plots showing the frequency of intron retention (IR) of genic cRNAs detected in dH1-depleted dsRNA^{dH1} (**G**), hrp36-depleted dsRNA^{hrp36} (**H**) and hrp48-depleted dsRNA^{hrp48} cells (**I**) respect to control dsRNA^{LacZ}. Decreased and increased IR are shown in blue and red, respectively. On the right, pie charts of the proportion of cRNAs showing differentially increased and decreased IR (p-value < 0.05).

Figure 5. **Depletion of hrp36 and hrp48 induce R-loops accumulation in heterochromatin**. (**A**) Immunostainings with S.9.6 antibodies to detect R-loops (in red) of polytene chromosomes from control w^{1118} (top), hrp48-depleted *hrp48^{RNAi}* (center) and hrp36-depleted *hrp36^{RNAi}* (bottom) flies. Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter (arrowheads). DNA was stained with DAPI. Scale bars correspond to 20μm. (**B**) Pie charts showing the proportions of R-loops detected by DRIP-seq in hrp48-depleted dsRNA^{hrp48} (right) and hrp36-depleted dsRNA^{hrp36} (left) cells overlapping (green) or not (red) with cRNAs. (**C**) Pie charts showing the proportions of hrp36 and hrp48 ChIP-seq peaks overlapping (green) or not (red)

with R-loops detected by DRIP-seq in hrp48-depleted dsRNA^{hrp48} (right) and hrp36-depleted dsRNA^{hrp36} (left). (**D**) Venn diagrams showing the overlap between cRNAs detected in hrp48-depleted dsRNA^{hrp48} (top) and hrp36-depleted dsRNA^{hrp36} (bottom) with those detected in control dsRNA^{LacZ} cells. (**E**) Pie charts showing the proportions of R-loops mapping at euchromatin (blue) and heterochromatin (red) detected only in hrp48-depleted dsRNA^{hrp48} cells (left), only in control dsRNA^{LacZ} cells (right), and both in hrp48-depleted dsRNA^{hrp48} and in control dsRNA^{LacZ} cells (center). (**F**) On the left, Volcano plot showing the change in intensity of R-loops detected in hrp48-depleted dsRNA^{hrp48} cells in comparison to control dsRNA^{LacZ} cells. Down- and up-regulated R-loops are shown in blue and red, respectively. On the right, pie chart showing the proportion of differentially down- and up-regulated R-loops (p-value< 0.05). (**G**) As in **E**, but for R-loops detected in hrp36-depleted dsRNA^{hrp36} cells. (**H**) As in **F**, but for R-loops detected in hrp36-depleted dsRNA^{hrp36} cells.

Figure 6. dH1 depletion increases accessibility of heterochromatin. (**A**) On the left, heat maps showing the size (bp) distributions of ATAC-seq fragments in control dsRNA^{LacZ} and dH1-depleted dsRNA^{dH1} cells. On the right, difference heat map of the intensities of the ATAC-seq fragments in dH1-depleted dsRNA^{dH1} cells in comparison to control dsRNA^{LacZ} cells as a function of fragment size (bp). Transposase hypersensitive (THSS) and nucleosomal fragments are indicated. Arrowheads indicate the two major nucleosomal fragments detected. (**B**) Venn diagram showing the overlap between THSS fragments detected in control dsRNA^{LacZ} and dH1-depleted dsRNA^{dH1} cells. (**C**) On the top, chromosomal distributions of THSS fragments detected only in control dsRNA^{LacZ} cells (left), only in dH1-depleted dsRNA^{dH1} cells (right), and both in control dsRNA^{LacZ} and dH1-depleted dsRNA^{dH1} cells (center). Chromosome definition and orientation are as in **Figure 1A**. Euchromatic and heterochromatic scaffolds are shown in blue and red, respectively. On the bottom, pie charts showing the proportion of THSS fragments mapping to euchromatic and heterochromatic scaffolds. (**D**) Box plot showing the average change in intensity of THSS fragments in dH1-depleted dsRNA^{dH1} cells respect to control dsRNA^{LacZ} cells for each chromosome and as

a whole (all). Euchromatic and heterochromatic chromosomes are shown in blue and red, respectively. (**E**) On the top, 2D occupancy plots of THSS coverage around the TSSs in control dsRNA^{LacZ} (left) and dH1-depleted dsRNA^{dH1} cells (right). On the bottom, plots of average coverage around the TSSs. The distance from the TSS is indicated.

Figure 7. dH1 depletion reduces nucleosome occupancy. (**A**) On the top, 2D occupancy plots of coverage for nucleosomal fragments around the TSSs in control dsRNA^{LacZ} (left) and dH1-depleted dsRNA^{dH1} cells (right). Arrows indicate the two major nucleosomal bands of 151bp and 138bp detected. The asterisk indicates a subnucleosomal fragment of 125bp. On the bottom, plots of average dyad density around the TSSs. The distance from the TSS and the position of the +1 nucleosome are indicated. (**B**) The proportion of nucleosomal fragments in **A** is shown as a function of fragment size in control dsRNA^{LacZ} (top) and dH1-depleted dsRNA^{dH1} cells (bottom). Arrows indicate the two major nucleosomal fragments of 151bp and 138bp detected. The asterisk indicates a subnucleosomal fragment of 125bp. (**C**) As in **A**, but around the center of common hrp36 and hrp48 ChIP-seq peaks. (**D**) As in **B**, but around the center of common hrp36 and hrp48 ChIP-seq peaks.

Figure 8. Model for the contribution of linker histones H1 to the accumulation of cRNAs and the formation of R-loops in heterochromatin. In normal conditions (top), heterochromatin shows a regular nucleosomal organization, heterochromatic elements are poorly transcribed, and heterochromatic cRNAs are scarce and packed into RNP particles by hnRNPs. Upon H1 knockdown (bottom), nucleosome occupancy is reduced, expression of heterochromatic elements is up-regulated, and assembly of heterochromatic cRNAs into RNPs is impaired. These defects favor annealing of heterochromatic cRNAs to the DNA template and, thus, the formation of R-loops that, in its turn, enhances cRNAs retention in heterochromatin.

Supplementary Figure S1. Co-localization of hrp36 and hrp48. Immunostainings of polytene chromosomes with αhrp36 (in yellow) and αhrp48 (in magenta) antibodies are shown at the heterochromatic chromocenter (**A**) and at a representative euchromatic region (**B**). Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter in **A**. DNA was stained with DAPI. Scale bars correspond to 5 μm.

Supplementary Figure S2. **dH1 depletion reduces hrp36 and hrp48 in heterochromatin**. (**A**) Immunostainings of polytene chromosomes from control H2AvD1::GFP (top) and dH1-depleted *dH1*^{RNAi} (bottom) flies with αhrp36 (in yellow) (left) and αhrp48 (in magenta) (right) antibodies. Immunostaining with αHP1a (in green) antibodies is shown to identify the heterochromatic chromocenter. Immunostaining with αGFP (in blue) is shown to identify control H2AvD1::GFP chromosomes. DNA was stained with DAPI. Scale bars correspond to 5μm. (See **Figure 3B** for quantification of the results). (**B**) On the left, immunostainings of polytene chromosomes from control H2AvD1::GFP (top) and mock depleted *GFP*^{RNAi} (bottom) flies with αhrp36 (in yellow). Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter. Immunostaining with αGFP (in blue) is shown to identify control H2AvD1::GFP chromosomes. DNA was stained with DAPI. Scale bars correspond to 5μm. On the right, box plot showing the average grey intensity of αhrp36 fluorescence at the chromocenter of control H2AvD1::GFP and mock depleted *GFP*^{RNAi} flies ($N > 15$; p-values > 0.05, Kruskal-Wallis test). (**C**) As in **B**, but for with αhrp48 antibodies (in magenta) ($N > 10$; p-values > 0.05, Kruskal-Wallis test).

Supplementary Figure S3. (**A-D**) On the top, Volcano plots showing the change in intensity in dH1-depleted dsRNA^{dH1} cells respect to control dsRNA^{LacZ} cells of intergenic (**A**), lncRNAs (**B**), genic antisense (**C**) and genic sense (**D**) cRNAs. Down- and up-regulated transcripts are shown in blue and red, respectively. On the bottom, pie charts showing the proportion of differentially down- and up-regulated cRNAs (p-value < 0.05). (**E-G**) As in **A-D**, but for cRNAs mapping to

active chromatin states (Kharchenko et al., 2011): promoter-TSS (**E**), regulatory regions (**F**) and transcription elongation (**G**).

Supplementary Figure S4. hrp36 and hrp48 independently associate with chromatin. (**A**) Venn diagram showing the overlap between R-loops detected by DRIP-seq in hrp48-depleted dsRNA^{hrp48} and hrp36-depleted dsRNA^{hrp36} cells. (**B**) Pie charts showing the proportions of R-loops mapping at euchromatin (blue) and heterochromatin (red) detected only in hrp48-depleted dsRNA^{hrp48} cells (left), only in hrp36-depleted dsRNA^{hrp36} cells (right), and in both hrp48-depleted dsRNA^{hrp48} and hrp36-depleted dsRNA^{hrp36} cells (center). (**C**) Immunostainings of polytene chromosomes from control w^{1118} (top), hrp36-depleted *hrp36^{RNAi}* (center) and hrp48-depleted *hrp48^{RNAi}* flies with αhrp36 (in yellow) and αhrp48 (in magenta) antibodies. Immunostaining with αHP1a antibodies (in green) is shown to identify the heterochromatic chromocenter (arrowheads). DNA was stained with DAPI. Scale bars correspond to 20μm. (**D**) WB analysis with αhrp36 (top) and αhrp48 (bottom) antibodies of increasing amounts (lanes 1-3) of crosslinked chromatin prepared from control dsRNA^{LacZ} (left), hrp36-depleted dsRNA^{hrp36} (center) and hrp48-depleted dsRNA^{hrp48} (right) cells. αH4 antibodies were used for loading control. Quantification of the results is shown on the bottom. Results are the average of 3 independent experiments. Error bars are standard deviation. (p-values: **< 0.01; two-tailed paired Student's test).

Supplementary Figure S5. The proportion of euchromatic THSS fragments mapping at the indicated genomic regions is shown.

SUPPLEMENTARY TABLES

Table S1. Heterochromatic cRNAs matching repetitive DNA elements.
The proportion of heterochromatic cRNAs matching different classes of repetitive DNA elements and of the corresponding reads are shown.

	% of transcripts	% of reads
Total repeated elements	72.57%	86.81%
Retroelements	42.75%	82.89%
LINES	17.66%	4.35%
LTR-elements	25.09%	78.54%
DNA transposons	2.35%	0.63%
Satellite DNAs	19.74%	0.11%
Simple repeats	1.34%	0.20%
Other	6.39%	2.98%

Table S2. Primers used to synthesize PCR templates for dsRNA production.
The T7 RNA polymerase promoter is underlined.

Name	Sequence (5'-3')
T7LacZ-Fw	TAATACGACTCACTATAGGGATGACCATGATTACGCCAAGC
T7LacZ-Rv	TAATACGACTCACTATAGGGCAATTTCCATTCGCCATTCAG
T7H1-Fw	TAATACGACTCACTATAGGGGATGTCTGATTCTGCAGTTGC
T7H1-Rv	TAATACGACTCACTATAGGGGGGCTTCGACTTTATGATTCC
T7hrp36-Fw	TAATACGACTCACTATAGGGAGACCAACGGAAACTACGACGAT
T7hrp36-Rv	TAATACGACTCACTATAGGGAGATGCTTGGCAATAGCCTTCTT
T7hrp48-Fw	TAATACGACTCACTATAGGGAGAACGAGAGGGGCAAACTTTTT
T7hrp48-Rv	TAATACGACTCACTATAGGGAGAGCGGGACTTCTTCTTCTCCT

Table S3. Primers for RT-qPCR

Name	Sequence (5'-3')
Rpl32-Fw	GCTAAGCTGTCGCACAAATG
Rpl32-Rv	ACCAAGGAACTTCTTGAATCCG
hrp36_all_isoforms_Fw	GATGGCCTGAAGGCTCACTT
hrp36_all_isoforms_Rv	GCATTCTGCGCATTGTCGAT
hrp48_all_isoforms_Fw	ACATGCCACCTAACTCTGCC
hrp48_all_isoforms_Rv	GCCGTAGTCGTACTCAGAGC

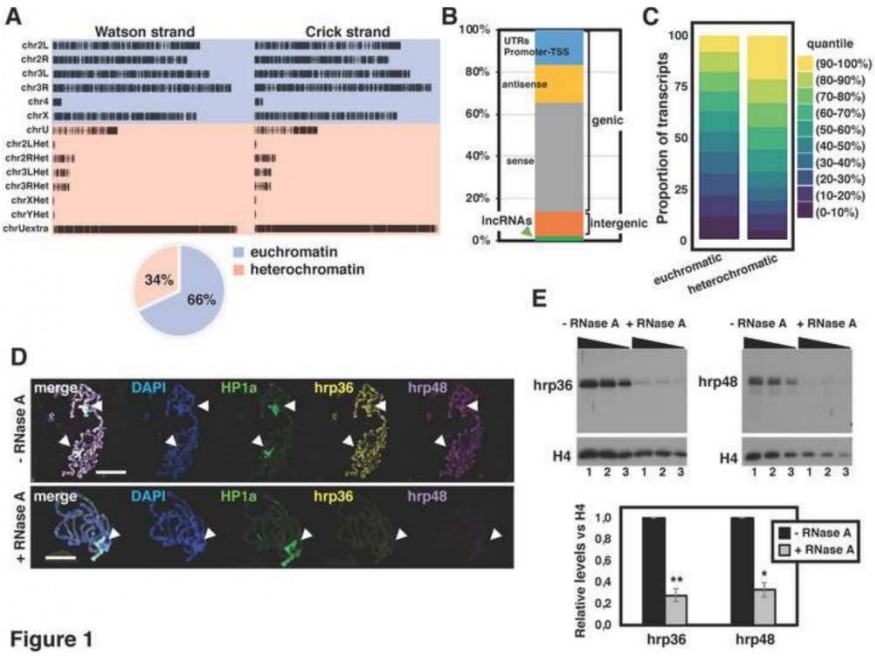


Figure 1

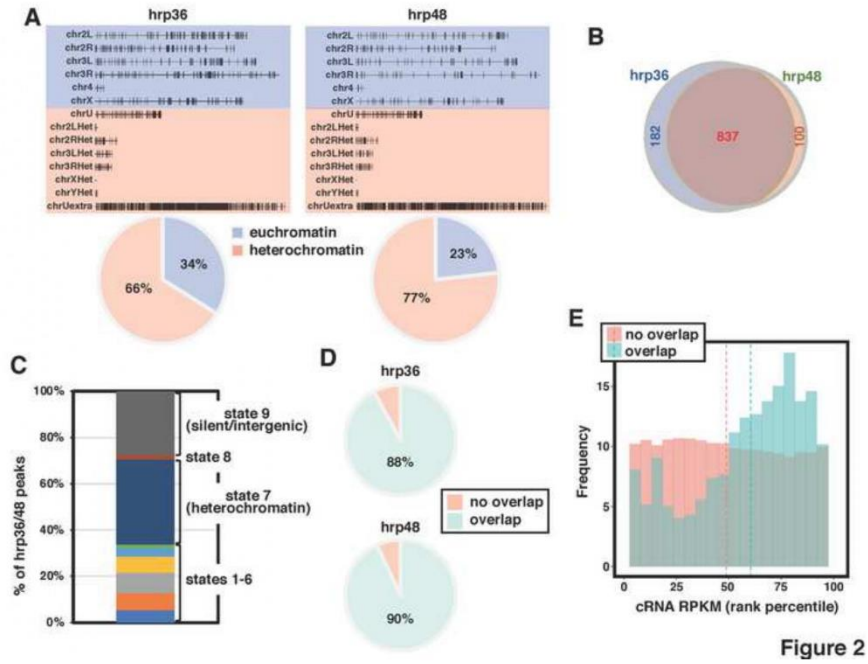


Figure 2

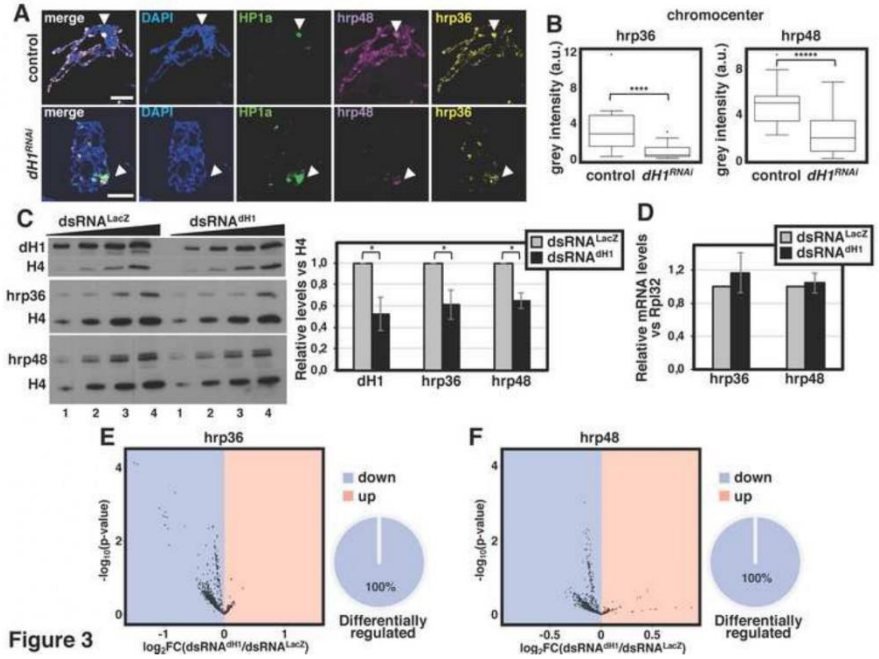


Figure 3

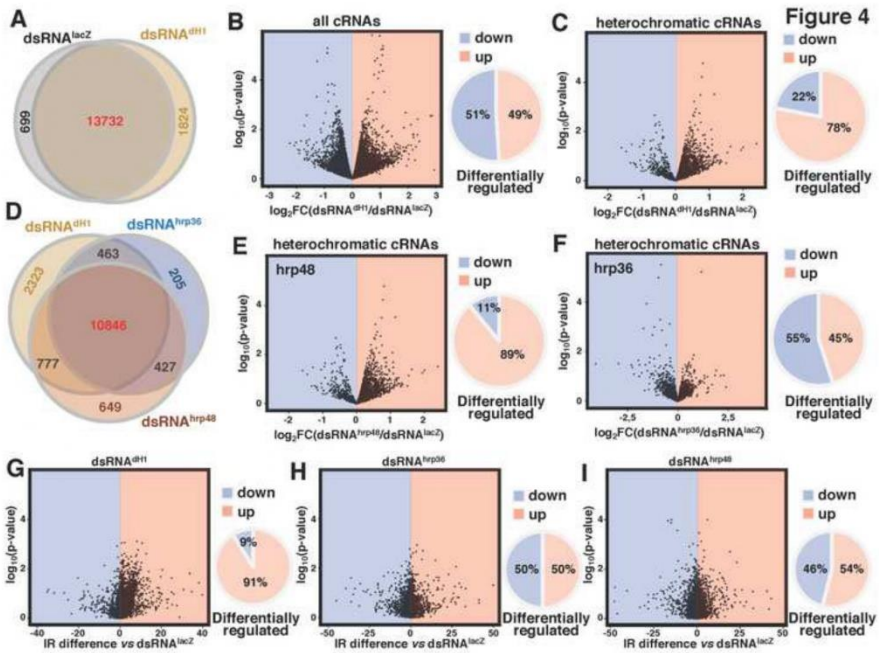


Figure 4

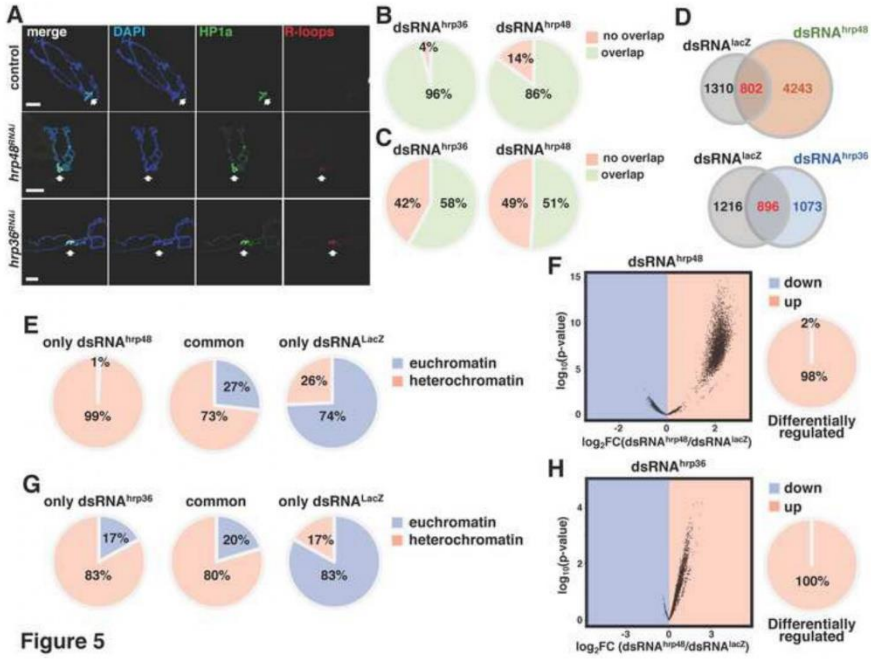


Figure 5

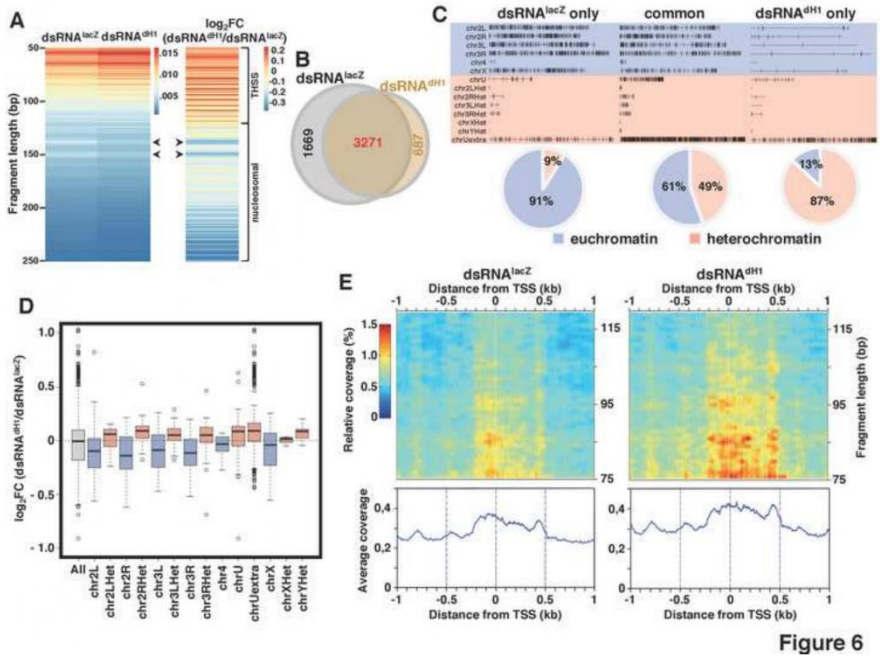


Figure 6

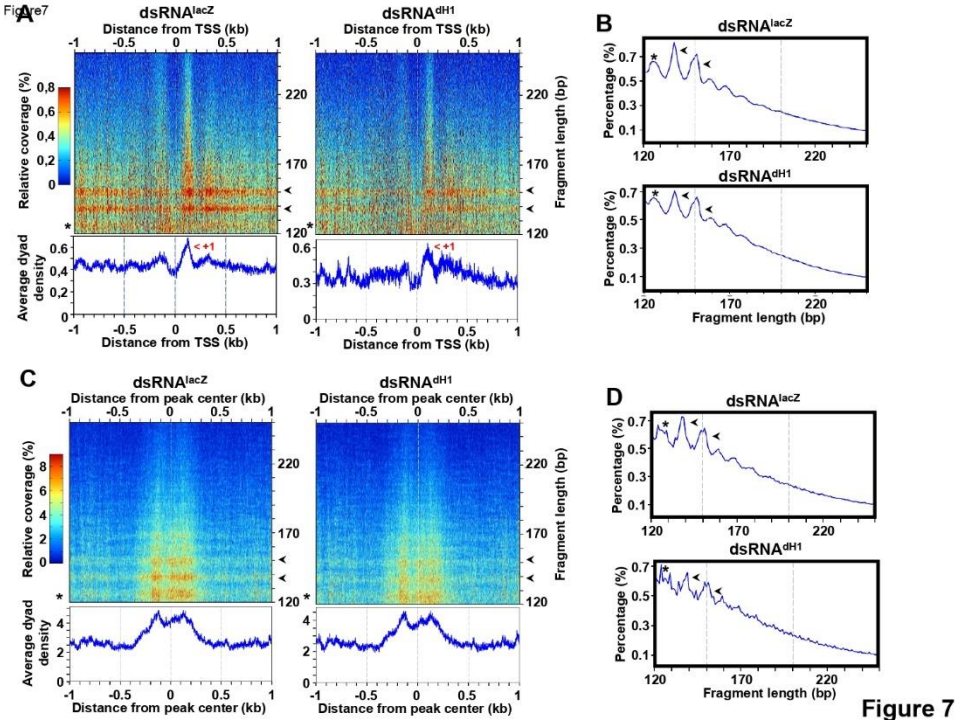


Figure 7

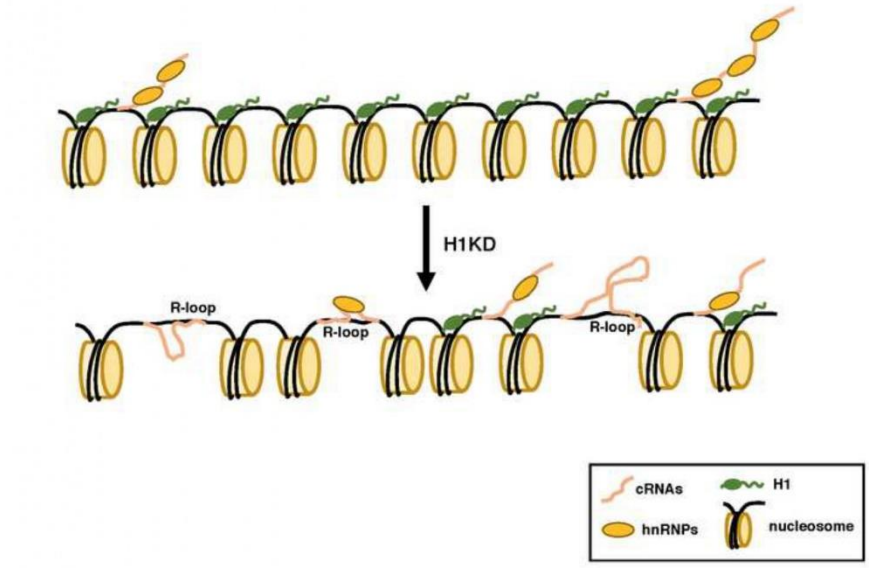


Figure 8

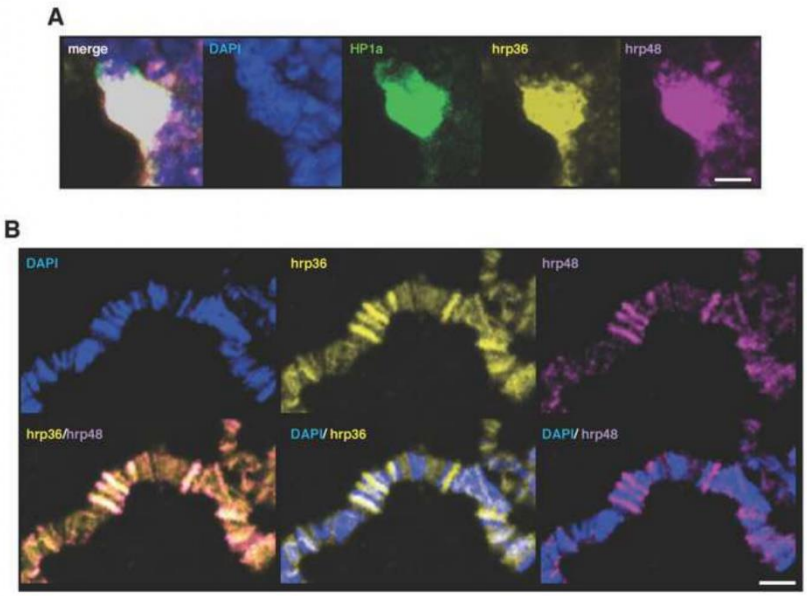


Figure S1

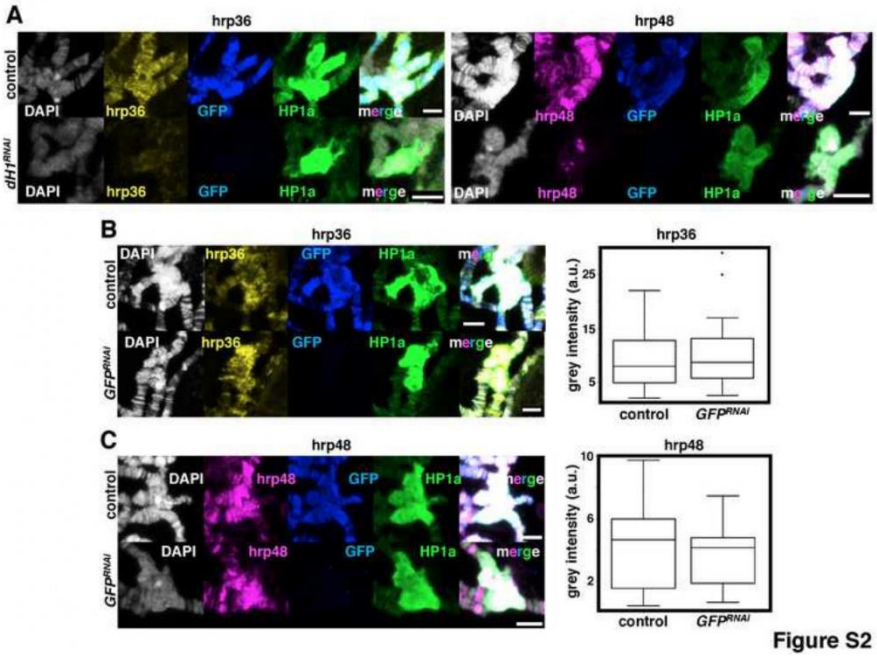


Figure S2

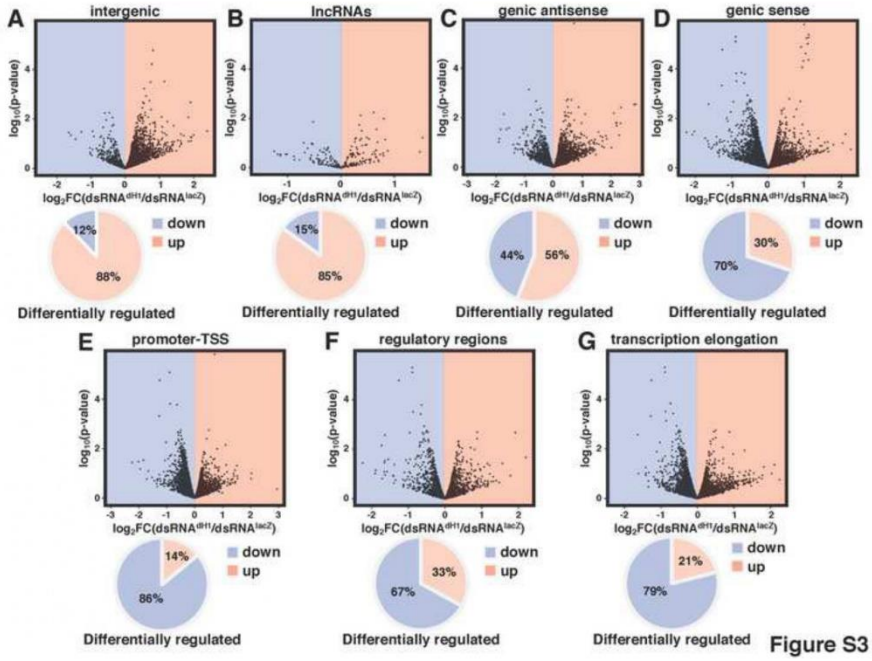


Figure S3

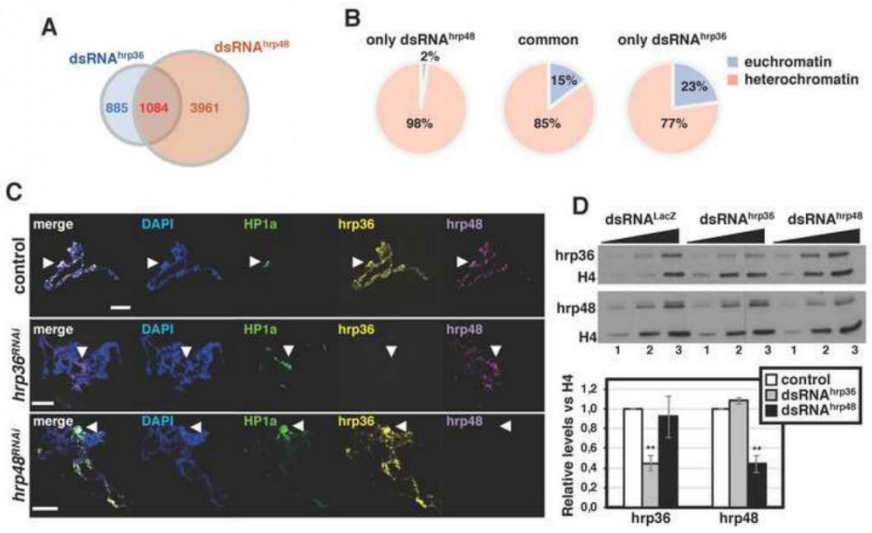


Figure S4

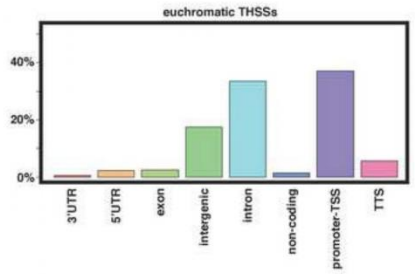

euchromatic THSSs
40%
20%
0%
3'UTR
5'UTR
exon
intergenic
intron
non-coding
promoter-TSS
TTS

Figure S5