Coordinated expression of replication-dependent histone genes from multiple loci promotes histone homeostasis in *Drosophila*

Ashlesha Chaubal [a], Justin M. Waldern[b], Colin Taylor[b], Alain Laederach[b,c], William F. Marzluff [a,b,c,d], and Robert J. Duronio [a,b,c,e*]

[a]Integrative Program for Biological and Genome Sciences, [b]Department of Biology, [c]Lineberger Comprehensive Cancer Center, [d]Department of Biochemistry and Biophysics, and [e]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599;

*Corresponding author
Integrative Program for Biological and Genome Sciences 3350 Genome Sciences Building
University of North Carolina Chapel Hill, NC 27599

Running head: HLB coordination of histone expression

Abbreviations: RD, Replication dependent, HLB, Histone Locus Body, BAC, Bacterial artificial chromosome

Abstract

Production of large amounts of histone proteins during S phase is critical for proper chromatin formation and genome integrity. This process is achieved in part by the presence of multiple copies of replication dependent (RD) histone genes that occur in one or more clusters in metazoan genomes. In addition, RD histone gene clusters are associated with a specialized nuclear body, the histone locus body (HLB), which facilitates efficient transcription and 3' end-processing of RD histone mRNA. How all five RD histone genes within these clusters are coordinately regulated such that neither too few nor too many histones are produced, a process we refer to as histone homeostasis, is not understood. Here, we explored the mechanisms of coordinate regulation between multiple RD histone loci in *Drosophila melanogaster* and *Drosophila virilis*. We provide evidence for functional competition between endogenous and ectopic transgenic histone arrays located at different chromosomal locations in *D. melanogaster* that helps maintain proper histone mRNA levels. Consistent with this model, in both species we found that individual histone gene arrays can independently assemble an HLB that results in active histone transcription. Our findings suggest a role for HLB assembly in coordinating RD histone gene expression to maintain histone homeostasis.

Introduction

Nucleosomes containing an octamer of histone proteins constitute the fundamental building blocks of chromatin and regulate access to, and expression of, the information within eukaryotic genomes. Generating sufficient H2a, H2b, H3, and H4 histone proteins in the correct stoichiometric amounts to assemble nucleosomes during S phase of the cell cycle is imperative for properly packaging the newly replicated DNA and is critical for normal genome function and stability. Disruptions to this process resulting in either a deficit of histones during S-phase or an accumulation of excess, non-nucleosomal histones can have detrimental effects on cell viability. For example, depletion of H2B or H4 in yeast causes mitotic arrest and disruption of chromosome segregation (Han *et al.*, 1987; Kim *et al.*, 1988), as does reduction of maternal levels of all 4 core histone mRNAs in early *Drosophila* embryos (Sullivan *et al.*, 2001). Similarly, in human cells repression of histone expression either by knockdown of SLBP, a factor important for histone mRNA 3' end processing and translation (Wagner *et al.*, 2005), or by ectopic expression of the histone chaperone HIRA (Nelson *et al.*, 2002) (PMID: 12620223), results in S-phase arrest. Thus, S-phase cells need to rapidly produce large amounts of histones for deposition onto replicating DNA to maintain proper chromatin structure. Conversely, excess positively charged histones can bind non-specifically to nucleic acids forming aggregates or sequester histone binding proteins, thereby resulting in cytotoxicity (Singh *et al.*, 2010). Overexpression of histone genes in budding yeast causes an increased rate of chromosome loss (Meeks-Wagner and Hartwell, 1986; Au *et al.*, 2008) in spite of an active degradation system to remove excess histones (PMID: 14651846). Furthermore, excess histones increase sensitivity to DNA damage in budding yeast by interfering with the homologous recombination machinery (Liang *et al.*, 2012). In early *Xenopus* and zebrafish embryos, which store large amounts of histone proteins on chaperones, addition of excess histones delays activation of zygotic transcription (Amodeo *et al.*, 2015; Joseph *et al.*, 2017).

All these studies indicate that cells must maintain a balance of not too many or not too few histones, a process referred to as histone homeostasis. Achieving histone homeostasis likely occurs through the regulation of histone amounts at multiple steps in gene expression including transcription, pre-mRNA processing, transport of histone mRNA to the cytoplasm, mRNA translation, and protein stability (Harris *et al.*, 1991; Gunjan and Verreault, 2003; Singh *et al.*, 2009; Cook *et al.*,

2011; Eriksson *et al.*, 2012; Marzluff and Koreski, 2017; Mendiratta *et al.*, 2019; Khan *et al.*, 2022). Here we provide evidence that an additional mechanism for achieving histone homeostasis in *Drosophila* involves modulation of histone mRNA accumulation in response to differing numbers of histone genes at different genomic loci.

Commented [RJD1]: ASHLESHA: NEED TO DOUBLE CHECK THESE REFERENCES

Histones are categorized into two classes, replication-dependent (RD) or canonical histones and replication-independent histone variants (Talbert and Henikoff, 2017). RD histones comprise the bulk of histones in chromatin and their synthesis is tightly coupled to the cell cycle, only being produced during S phase, whereas variant histone expression is not coupled to the cell cycle and their location in the genome is reflective of their function (e.g. centromeric histone H3, H3.3 or H2a.Z). Eukaryotic organisms contain multiple copies of RD histone genes, and in metazoans the genes encoding the 5 RD histone proteins are clustered together at one or more loci (Lifton *et al.*, 1978; Maxson *et al.*, 1983; Marzluff *et al.*, 2002). In *Drosophila melanogaster*, there is a single RD histone locus on chromosome 2L (*HisC*) where a 5-kb unit containing one copy of each RD histone gene is tandemly repeated ~ 100 times (McKay *et al.*, 2015; Bongartz and Schloissnig, 2019). The evolutionarily conserved clustering of RD histone genes almost certainly contributes to the coordinated expression of all five histones and ensure rapid activation at the beginning of S-phase. Clusters of RD histone genes in metazoans are also associated with a phase separated nuclear body called the Histone Locus Body (HLB) (Duronio and Marzluff, 2017). The HLB is primarily organized and identified by the orthologous proteins Multi-sex-combs (Mxc) in *Drosophila* and NPAT in mammals, and contains other evolutionarily conserved factors involved only in RD histone gene transcription and pre-mRNA processing (Ye *et al.*, 2003; Dominski and Marzluff, 2007; Yang *et al.*, 2009; Bulchand *et al.*, 2010; White *et al.*, 2011). Mxc is a large (>1800aa) protein composed mostly of intrinsically disordered regions with a structured N-terminal domain that mediates multimerization and is required for HLB formation (Terzo *et al.*, 2015). HLB formation in *Drosophila* is critical for histone biosynthesis, as depletion of Mxc prevents RD histone gene expression (White *et al.*, 2011). The HLB is also important for coupling of RD histone gene expression with the cell cycle, as Cyclin E/cdk2-mediated phosphorylation of Mxc/NPAT in S-phase activates histone gene expression (Ma *et al.*, 2000; Wei *et al.*, 2003; Ye *et al.*, 2003; White *et al.*, 2007; White *et al.*, 2011; Armstrong and Spencer, 2021). Here we

explore whether HLBs play a role in coordinating expression from multiple histone genes at non-homologous loci.

To explore this question, we employed a previously established platform for engineering specific histone genotypes in *Drosophila melanogaster* (McKay *et al.*, 2015; Meers *et al.*, 2018). Removal of all ~200 copies of each endogenous RD histone gene by homozygous deletion of *HisC* is lethal, but this lethality can be rescued by just 12 or 24 copies of each RD histone gene provided by a BAC-based transgenic histone gene array (McKay *et al.*, 2015). This rescue occurs because despite the large difference in gene copy number the overall level of H2A mRNA is similar between the "24x" engineered, transgenic histone genotype and the "200x" wild-type genotype (McKay *et al.*, 2015), and the total amount of RD histone mRNA that accumulates per gene copy is modulated to achieve histone homeostasis. Moreover, when wild-type *HisC* and the homozygous 12x transgenic array were present in the same animal ("224x" genotype), the amount of H2A transcript detected from each locus changed such that the total amount of H2A mRNA was comparable to a 200x or a 24x genotype (McKay *et al.*, 2015). This result suggested a mechanism of communication among histone genes residing at different loci to achieve a precise level of H2A gene expression. In this study, we demonstrate that such regulation applies to all 5 RD histone genes. We also provide evidence for functional competition between endogenous histone genes and transgenic histone arrays that likely results from limited availability of an HLB component(s), suggesting how HLB assembly might contribute to histone homeostasis.

Results

All five RD histone genes modulate expression to maintain histone homeostasis

In *Drosophila melanogaster*, the deficiency *Df(2L)HisC*^{ED1429} (hereafter *ΔHisC*) removes the entire endogenous histone locus on chromosome 2L and causes embryonic lethality when homozygous (Günesdogan *et al.*, 2010). This lethality can be rescued with a BAC- based transgene containing an engineered gene array with 12 tandemly repeated copies of the wild-type histone gene unit (HWT) (McKay *et al.*, 2015). The 12x^{HWT} array contains a synonymous polymorphism in the *H2A* gene resulting in loss of an *XhoI* site, enabling us to measure the amount of endogenous versus transgenic H2A transcript by restriction enzyme digestion of H2A RT-PCR

products (Figure 1A). In this study, we utilized a transgenic histone gene unit where each gene is similarly marked by the insertion or removal of a restriction site(s) without altering the protein coding sequence of any RD histone gene (Figure 1A). The transgenic array containing 12 copies of this designer wild-type (DWT) gene unit also rescues lethality caused by homozygous *ΔHisC* (Koreski *et al.*, 2020).

To test whether all RD histone genes modulate expression to compensate for differences in gene copy number, we measured zygotic RD histone mRNA amounts from 4-6hr old embryos that were either wild-type or carrying a homozygous $12x^{DWT}$ transgenic array in the presence or absence of the endogenous *HisC* locus (Supplemental Figure 1A-D). We compared RNA levels among embryos with the normal number of endogenous histone genes (wild-type, 200x), carrying two copies of $12x^{DWT}$ transgenic histone gene array (24x), or that were homozygous both for *HisC* and the $12x^{DWT}$ transgenic histone gene arrays (224x). We amplified cDNA from these three genotypes using primers that recognize each RD histone gene in both the endogenous and transgenic templates, followed by restriction digestion with specific enzymes that differentially digest endogenous versus transgenic PCR products (Figure 1A). Quantitation of band intensities from restriction digested PCR products revealed that the amount of each RD histone mRNA is similar between the 24x genotype and the 200x genotype despite ~8-fold difference in the number of histone genes (Figure 1B - D). To more accurately quantify the relative level of expression from the endogenous and transgenic histone loci we sequenced the RT-PCR products using a miSeq platform. The results are similar to the quantification of band intensities, revealing that XXXXXXXXX (Figure 1E, Supplemental Figure 1?). These data are consistent with our previously published results with the *H2A* gene (McKay *et al.*, 2015; Koreski *et al.*, 2020) and provides evidence for modulation of expression of all five RD histone genes to maintain histone homeostasis.

We next analyzed the relative mRNA contribution from endogenous versus transgenic histone genes when both were present in the same embryos. Interestingly, the $24x^{DWT}$ transgenic histone gene arrays, which when present alone can generate histone mRNA equivalent to or exceeding that made by 200 copies of the endogenous genes, contributes only ~25% of the total RD histone mRNA when present together with the endogenous *HisC* locus (Figure 1B - E). We also note that the ratio of the level of transcripts contributed by the transgenic versus endogenous

histone loci (~1:3) is still higher than expected from the ratio of number of transgenic histone genes to endogenous histone genes (1:8). These data suggest that there must be coordination between the endogenous and transgenic histone loci, which are located on different chromosomes, to maintain a particular overall amount of RD histone RNA. Thus, we have established a molecular assay to detect the relative amount of endogenous versus transgenic transcripts for each of the 5 RD histone genes and observed regulation between different histone gene loci that likely contributes to histone homeostasis.

Different histone loci compete for limiting RD histone gene expression factors

How do cells coordinate gene expression between distinct histone gene arrays at different loci to achieve a particular overall level of histone mRNA? One potential mechanism could involve competition for shared but limiting histone mRNA biosynthetic factors. In this situation, a histone gene array that is functionally attenuated via mutation may be unable to effectively compete with wild-type *HisC*, resulting in lower levels of gene expression from that array. We previously showed that a transgene carrying 12 copies of a RD histone gene unit in which each bidirectional *H3-H4* promoter is replaced by the wild-type bidirectional *H2a-H2b* promoter (Figure 2A, Supplemental Figure 1F, $12X^{PR}$ or "promoter replacement") behaves as an attenuated histone gene array: it does not from an HLB and is not well expressed in the presence of the endogenous RD histone genes at *HisC*, but in the absence of the endogenous RD histone genes the $12X^{PR}$ array forms an HLB, expresses RD histone mRNA, and can rescue the lethality caused by homozygous *ΔHisC* (Koreski *et al.*, 2020). The major difference between *H2a-H2b* and *H3-H4* bidirectional promoters is that the *H2a-H2b* region lacks the GAGA repeat present in the *H3-H4* promoter that binds the zinc finger protein CLAMP, which promotes RD histone gene expression (Rieder *et al.*, 2017). We also previously showed that the *H3-H4* promoter is important for HLB formation and expression of all the core RD histone genes salivary glands (Salzler *et al.*, 2013; Rieder *et al.*, 2017). Thus, the $12X^{PR}$ is an attenuated RD histone gene array that cannot effectively compete with the endogenous *HisC* locus but is fully functional when it is the only source of RD histone genes.

This result led us to ask whether a wild-type transgenic array with a small number of RD histone genes could also attenuate the expression of $12X^{PR}$. We

created flies carrying two different transgenic histone arrays by making a recombinant 3rd chromosome carrying the $12X^{PR}$ transgene on the left arm of chromosome 3 and a transgene with only 8 copies of the wild-type histone gene unit ($8X^{HWT}$) on the right arm of chromosome 3 (Figure 2A). Like the DWT transgene, the $12X^{PR}$ array also carries synonymous polymorphisms in each histone gene enabling us to differentiate between transcripts generated from $12X^{PR}$ versus $8X^{HWT}$ (or versus the endogenous genes at *HisC*) (Figure 2A). We determined the level of expression of all 5 RD histone genes in 3-6hr old embryos carrying a single copy of the $12X^{PR}$ array in a homozygous *ΔHisC* background (Supplemental Figure 1A, 1E). Consistent with our previous results (Koreski *et al.*, 2020), in this genotype the $12X^{PR}$ transgene produces wild-type amounts of all five RD histone genes (Figure 2B and 2C, Genotype 1 versus 2). Thus, in this genotype replacing the *H3- H4* promoter with the *H2a-H2b* promoter does not have any substantial effect on the expression of the *H3-H4* gene pair. Next, we measured the relative expression from the $12X^{PR}$ array in the presence of both *HisC* and $8X^{HWT}$ (Figure 2B, Genotype 2) or just $8X^{HWT}$ (Figure 2B, Genotype 3). Including one complement of endogenous histone genes (~100 copies at the *HisC* locus located on the CyO balancer chromosome) results in loss of expression of all five RD histone genes from the $12X^{PR}$ array (Figure 2B and 2C, Genotype 2). Thus, although the $12X^{PR}$ array carries intact wild-type *H1* and *H2a-H2b* genes with their normal promoters, the replacement of the *H3-H4* promoter with the *H2a-H2b* promoter attenuates the expression of this entire transgenic locus in the presence of ~100 copies of the endogenous histone genes (Koreski *et al.*, 2020)(Salzler *et al.*, 2013). In contrast, $12X^{PR}$ is expressed when present with one copy of the $8X^{HWT}$ transgene rather than with *HisC* (Figure 2B and 2C, Genotype 3). In this genotype, the $12X^{PR}$ and $8X^{HWT}$ transgenic arrays exhibit an ~ 70:30 relative contribution, respectively, to the total amount of histone mRNA, thereby maintaining an overall RD histone gene expression level similar to that of one copy of *HisC* (Figure 2C, Genotype 3). We conclude from this experiment that 8 copies of the wild-type RD histone gene unit do not compete with $12X^{PR}$ like ~100 copies do. We hypothesize that this competition is due to limiting amounts of a factor(s) involved in histone mRNA biosynthesis that must be distributed between different loci (Koreski *et al.*, 2020). Together, these data provide further evidence for coordination between histone loci that are present on either the same or separate chromosomes.

HLB assembly reflects competition between different RD histone loci

The basis for distributing limiting gene expression factors to multiple RD histone loci is likely rooted in the mechanism of HLB assembly, which occurs through a combination of ordered and stochastic processes (Duronio and Marzluff, 2017). Recruitment of histone mRNA biosynthetic factors to the HLB is consistent with both "seed and grow" and phase separation mechanisms of assembly (White *et al.*, 2011; Hur *et al.*, 2020). The H3-H4 promoter and/or nascent histone mRNA provides the "seed" (Rieder *et al.*, 2017; Hur *et al.*, 2020) while multimerization of Mxc (Terzo *et al.*, 2015) provides a scaffold for recruitment of other HLB components ("grow") resulting in a phase separated nuclear body that facilitates activation of histone gene expression. To explore whether HLB assembly might play a role in coordination or competition between different histone loci (e.g. by assembling these loci into a single body or multiple, distinct bodies), we stained embryos with antibodies against Mxc to visualize HLBs in the different genotypes described above. We first asked whether a single $12X^{PR}$ locus was competent for HLB assembly in diploid cells by analyzing *ΔHisC*/*ΔHisC* embryos containing either one ($12X^{PR}$/+) or two ($12X^{PR}$/$12X^{PR}$) copies of the transgene. We observed a single HLB in all (n=372) epidermal cell nuclei of germ band retracted *ΔHisC*/*ΔHisC*; $12X^{PR}$/+ embryos in which these cells are arrested in G1 phase of the cell cycle, consistent with the presence of a single, un-replicated $12X^{PR}$ transgene (Figure 3A, B). In *ΔHisC*/*ΔHisC*; $12X^{PR}$/$12X^{PR}$ blastoderm embryos most nuclei (~99%) had either one or two HLBs, which is reflective of paired versus unpaired homologous chromosomes, respectively (Supplemental Figure 2A). We obtained a similar result in control *ΔHisC*/*ΔHisC*; $12X^{DWT}$/$12X^{DWT}$ embryos (Supplemental Figure 2A). We conclude from these data that HLB formation occurs at the $12X^{PR}$ transgenic array when the endogenous RD histone genes are absent, consistent with our previous observation that $12X^{PR}$ can assemble an HLB in highly polyploid salivary gland nuclei as well as syncytial stage embryos and that $12X^{PR}$ is active for RD histone gene expression in the homozygous *ΔHisC* genotype (Figure 2) (Koreski *et al.*, 2020).

We next determined the effect on HLB assembly at $12X^{PR}$ of introducing one (*ΔHisC*/*CyO*) or two (*+/+)* copies of the wild-type *HisC* locus. We found in G1-arrested embryonic epidermal cells that most (96%, n=485) nuclei in *ΔHisC*/*CyO*;$12X^{PR}$/+ embryos contained a single HLB, while a small proportion of

nuclei (4%, n=485) contained two HLBs of differing sizes (Figure 3A, B). In +/+; $12X^{PR}/12X^{PR}$ blastoderm embryos essentially all nuclei (n=5681) contained either two HLBs that appear similarly sized (18%) or one HLB (82%) (Supplemental Figure 2B). This result is identical to that obtained in true wild-type Oregon R embryos (Figure 4A). These data suggest that an HLB does not form at $12X^{PR}$ in the presence of two copies of *HisC*, but forms with a low frequency when there is a single copy of the histone locus. One possibility is that the ~100-200 copies of each wild-type RD histone gene unit (or of the *H3-H4* promoter itself) present at *HisC* sequester a limiting factor(s) and prevent HLB components from assembling on the $12X^{PR}$ locus. We conclude that HLB assembly at $12X^{PR}$ is severely impaired by the presence of wild-type *HisC*, resulting in very low or no histone mRNA production from $12X^{PR}$ (Figure 2).

To test the competition hypothesis, we examined whether only 8 copies of wild-type histone genes could attenuate $12X^{PR}$ HLB assembly. In *ΔHisC*/*ΔHisC* embryos simultaneously carrying one copy of the $12X^{PR}$ transgene and one copy of the $8X^{HWT}$ transgene on different arms of the third chromosome, we observed that 25% (n=506) of nuclei had two HLBs in G1- arrested epidermal cells (Figure 3A, B). This result indicates that an HLB can simultaneously form at both $12X^{PR}$ and $8X^{HWT}$ transgenic loci. Furthermore, because $8X^{HWT}$ does not suppress HLB formation at, or expression from, the $12X^{PR}$ transgene, our observations suggest that $8X^{HWT}$ is not as effective as *HisC* in sequestering factors from $12X^{PR}$. At present we cannot explain why ~75% of nuclei had a single HLB in *ΔHisC*/*ΔHisC*; $12X^{PR}$, $8X^{HWT}$/+ embryos (Figure 3B), but one possibility is that these nuclei reflect HLB formation only at the $12X^{PR}$ or only at the $8X^{HWT}$ transgene.

Transcription of RD histone mRNA occurs in individual ectopic HLBs

To determine whether RD histone gene transcription is always coincident with HLB formation, we performed RNA FISH to core histone genes while simultaneously staining with anti-Mxc antibodies. As noted above, in nuclei of wild-type blastoderm embryos we observe either one (80%) or two (20%) HLBs (Figure 4A). The fraction of single HLB nuclei is the same as the fraction of paired homologous *HisC* loci previously determined using DNA in situ hybridization (Hiraoka *et al.*, 1993). Furthermore, we have observed fusion of individual HLBs (two HLBs merging into one) by live imaging embryos carrying GFP-tagged Mxc (Hur *et al.*, 2020). Thus, the

distribution of one versus two HLBs in wild-type likely results from homologous chromosome pairing in early *Drosophila* embryos. In cellular blastoderm (cycle 14) embryos that are homozygous for both *HisC* (+/+) and a $12X^{DWT}$ transgenic histone gene array ($12X^{DWT}/12X^{DWT}$), we observe a broad distribution of nuclei (n=4545) with one (26%), two (42%), three (27%), or four (5%) individual HLBs (Figure 4A). Nuclei with 4 HLBs represent the situation in which neither the homologous *HisC* loci on chromosome 2 nor the homologous $12X^{DWT}$ loci on chromosome 3 are paired. In these nuclei we observe two larger and two smaller HLBs, consistent with our previous observation that the number of histone genes at a locus determines HLB size (Hur *et al.*, 2020). Nuclei with fewer than 4 HLBs likely result from homologous chromosome pairing.

To determine whether individual HLBs are active for transcription, we hybridized +/+; $12X^{DWT}/12X^{DWT}$ embryos with a fluorescent probe set that simultaneously recognizes the four core histone RNAs (i.e. H2A, HB, H3, H4). This approach provides a highly sensitive method for detecting nascent RD histone transcripts. We found that every focus of nascent histone transcripts was associated with an HLB as assessed by Mxc staining (Figure 4B), including in those nuclei with 4 HLBs. Moreover, and as we have observed previously, these HLBs display a "core-shell" organization with nascent histone transcripts residing in the core surrounding by Mxc (Figure 4B, High_Resolution)(Kemp *et al.*, 2021). Thus, both unpaired *HisC* and unpaired $12X^{DWT}$ loci can independently support RD histone gene transcription in nuclear cycle 14 embryos. In a small number of early interphase nuclei (as assessed by nuclear morphology), we observed an Mxc focus that was not associated with a nascent histone transcript, suggesting that HLB assembly may have occurred prior to detectable RD histone gene transcription.

***Drosophila virilis* non-homologous RD histone loci behave similarly to engineered non-homologous *D. melanogaster* loci**

Thus far, we have used ectopic transgenic histone gene arrays to analyze expression and HLB assembly at non-homologous RD histone loci in *Drosophila melanogaster*. To examine HLB formation and histone gene transcription in a natural system carrying non-homologous histone loci, we analyzed early embryos from *Drosophila virilis*, which contains two RD histone gene clusters at different loci. In *D.*

virilis, the major RD histone gene locus (~30 repeats) is located at the cytogenetic position 25F on chromosome 2 and the minor locus (~6 repeats) is located at position 43C on chromosome 4 (Schienman *et al.*, 1998; Shiotsugu, 2002; Rieder *et al.*, 2017) (Figure 5A). Furthermore, the RD histone gene units in *D. virilis* exist as either quintets (gene units containing all five RD histone genes) or quartets (gene units containing only core RD histone genes and lacking the H1 gene) (Domier *et al.*, 1986; Schienman *et al.*, 1998).

D. virilis syncytial blastoderm embryos stained with antibodies against *D. melanogaster* Mxc exhibited a distribution of nuclei (n=1868) with one (23%), two (66%), three (10%) and four (1%) individual HLBs, similar to our engineered system in *D. melanogaster* (Figure 5B). Nuclei with four HLBs exhibit two larger and two smaller HLBs, implying that HLB formation occurs at both the major and minor locus via a mechanism like *D. melanogaster* where the number of histone genes determines HLB size at this stage of development (Hur *et al.*, 2020). Furthermore, like *D. melanogaster*, it is likely that *D. virilis* nuclei with less than four Mxc foci represent fused HLBs due to pairing of homologous chromosomes.

To test whether nascent transcription can be detected at both these non-homologous histone loci, we hybridized *D. virilis* blastoderm embryos that were stained with anti-Mxc antibodies with fluorescent probes that recognize *D. virilis* histone H4 mRNA. We found that all individual HLBs were active for transcription, including those in nuclei exhibiting 3 or 4 HLBs (Figure 5C). Thus, our data demonstrate that both the major and minor histone loci in *Drosophila virilis* independently form HLBs and express RD histone genes in the same nucleus.

Discussion

The number of RD histone genes varies widely in different species, ranging from two copies in the yeast *Saccharomyces cerevisiae* to hundreds of copies in fruit flies and sea urchins (Hentschel and Birnstiel, 1981; Maxson *et al.*, 1983). Furthermore, these genes can either be organized in highly regular tandem repeats at a single locus (e.g., *Drosophila melanogaster*) (Lifton *et al.*, 1978), randomly arrayed in multiple clusters at distinct chromosomal locations (e.g., mammals) (Marzluff *et al.*, 2002; Seal *et al.*, 2022), or distributed as small clusters throughout the genome (e.g., *Caenorhabditis elegans*) (Roberts *et al.*, 1987). Since the overall histone level needs to be tightly controlled and coupled with S phase for genomic stability and cell

survival, it is likely that coordinate expression from multiple histone loci is actively regulated to maintain histone homeostasis. In this study, we have examined the relationship between expression and HLB assembly at different RD histone loci in *Drosophila*.

We previously demonstrated that in *D. melanogaster* embryos, 24 copies of transgenic *H2A* genes generate an amount of mRNA equivalent to that made by ~200 endogenous *H2A* genes (McKay *et al.*, 2015). Here, we extended this analysis to show that the other three core RD histone genes as well as the linker histone *H1* gene display this same homeostatic regulation. Because total RNA was measured in our experiments, we cannot distinguish the relative contribution of transcriptional or post-transcriptional (e.g. mRNA half-life) mechanism to the maintenance of overall RD histone mRNA levels. Nevertheless, our data clearly show that individual histone gene arrays can be simultaneously expressed in both *D. melanogaster* and *D. virilis*, suggesting that transcriptional regulation contributes to expression homeostasis between distinct RD histone gene loci.

How are all 5 RD histone genes coordinately regulated? There are no transcription factors known to bind simultaneously to each gene and coordinately regulate them. Instead, HLB assembly factors like Mxc are required for expression of each of the 5 RD histone genes (White *et al.*, 2011), and mammalian NPAT is present at each of the active RD histone gene promoters (PMID: 31036827). We probed the relationship between HLB assembly and RD histone gene transcription using engineered BAC-based transgenic histone gene arrays, particularly the functionally attenuated $12X^{PR}$ array in which the bidirectional *H3- H4* promoter is replaced by the bidirectional *H2a-H2b* promoter in each of the 12 histone gene units. This natural *H2a-H2b* promoter is capable of driving *H3-H4* expression and providing RD histone gene function, as $12X^{PR}$ fully rescues homozygous deletion of the endogenous *HisC* RD histone gene array (Koreski *et al.*, 2020). Interestingly, in the presence of *HisC* none of the RD histone genes in $12X^{PR}$ are expressed, though three of the genes (i.e. *H1*, *H2a* and *H2b*) are unperturbed and contain their endogenous promoters. Thus, the lack of *H2a-H2b* and *H1* expression from $12X^{PR}$ in the presence of *HisC* cannot be due to the absence of a key cis element. Rather, we conclude that the lack of HLB assembly, which does not occur at $12X^{PR}$ in the presence of *HisC* but does in the absence of *HisC*, is the reason for the failure of expression. Consistent with this interpretation, $12X^{PR}$ is expressed in the presence

of *HisC* when integrated in trans with $12X^{HWT}$ at the same locus on the third chromosome (i.e., the $12X^{PR}/12X^{HWT}$ genotype), a situation that promotes assembly of a single HLB that includes both transgenes (Koreski *et al.*, 2020).

HLB assembly fails to occur at $12X^{PR}$ in the presence of *HisC* because of the absence of the *H3-H4* bidirectional promoter. We previously showed that in the presence of the endogenous genes at *HisC*, HLB components can be recruited to an ectopic RD histone locus by a single *H3-H4* gene pair or just the *H3-H4* promoter but not by the *H2a-H2b* and *H1* genes (Salzler *et al.*, 2013). In addition, GAGA repeats found only within the *H3-H4* promoter and that bind the zinc finger protein CLAMP are required for ectopic HLB assembly (Rieder *et al.*, 2017). Thus, HLB assembly nucleated by the *H3-H4* promoter provides a mechanism for how the *H3- H4* promoter can stimulate *H2a-H2b* and *H1* transcription. We hypothesize that the *H3-H4* promoters in a histone array provide a strong binding site for the recruitment of HLB components, thereby nucleating HLB assembly and facilitating expression of the entire RD histone gene array. Note that CLAMP is present in the 12XPR HLBs, but not bound to DNA, suggesting that it interacts with both DNA and an HLB factor(s) in the wild-type array (Koreski *et al.*, 2020).

We also found that unlike the *HisC* locus, which contains ~100 histone gene units, a single copy of a transgene containing an array of 8 histone gene units did not prevent HLB formation and transcription at the $12X^{PR}$ transgene. One interpretation of this result is that many histone gene units sequester the limited supply of essential HLB components from the attenuated $12X^{PR}$ transgene, and hence requires a higher concentration of Mxc to seed the HLB. In this model, the presence of only 8 histone gene units is insufficient to bind enough HLB factors to achieve this level of sequestration. In contrast, 12 wild-type histone gene units can effectively compete with *HisC*, as we observed HLB assembly and transcription at $12X^{DWT}$ in the presence of *HisC*. However, the amount of expression from $12X^{DWT}$ is modulated by the presence of *HisC*, as more RD histone mRNA is produced by $12X^{DWT}$ in the absence of *HisC* than in its presence. There is also less expression from the *HisC* locus also in the presence of the $12X^{DWT}$. Thus, these distinct loci may compete for limiting factors necessary for RD histone mRNA biosynthesis.

We suggest that HLB assembly and competition for limiting HLB components between histone gene arrays present at different loci provides a possible mechanism for coordinating RD histone gene expression to maintain histone homeostasis. We

cannot exclude the possibility that alternative chromatin configurations between the different loci contribute to the effects we observe. In addition, *C. elegans* doesn't have NPAT, FLASH, or U7 snRNA (and thus no HLBs), and forms the histone mRNA 3' end using an RNAi type mechanism after making a polyadenylated pre-mRNA (PMID: 22863779). Thus, nematodes coordinate expression from multiple histone genes via a mechanism that does not rely on HLB formation.

During early embryonic development when homologous chromosomes begin to pair in *Drosophila*, individual HLBs associated with each homologous *HisC* locus come into proximity and fuse into a single HLB, consistent with the liquid droplet properties of a phase separated nuclear body (Hur *et al.*, 2020). Accordingly, we observed nuclei with either one or two HLBs in embryos carrying homologous histone loci in OregonR, Δ*HisC*/Δ*HisC*; $12X^{PR}/12X^{PR}$ and *ΔHisC*/*ΔHisC*; $12X^{DWT}/12X^{DWT}$ genotypes. Whether the properties of HLB fusion play a role in histone homeostasis by facilitating physical proximity of non-arrayed histone loci in the 3-dimensional nuclear space, thereby enabling efficient RD histone gene expression from multiple loci, remains to be investigated. This line of inquiry may prove to be important for better understanding the regulation of multiple histone genes in mammalian cells (Albig and Doenecke, 1997; Seal *et al.*, 2022).

Materials and Methods

Fly strains and genetic crosses

Drosophila virilis (National *Drosophila* Species Stock center # 15010-1051.118) was a gift from Dr. Daniel Matute (University of North Carolina, Department of Biology). The Bloomington Stock Center provided Oregon R (stock #25211) and *yw* (stock #6599). *ΔHisC(Df(2L)HisC^{ED1429})*,UAS-2xEYFP / CyO and *ΔHisC(Df(2L)HisC^{ED1429})*, twi-GAL4 / CyO were a gift from Alf Herzig (Max Planck Institute for Biophysical Chemistry, Molecular Developmental Biology). Other fly stocks are described in (McKay *et al.*, 2015; Koreski *et al.*, 2020). All fly stocks were maintained on standard corn medium. For gene expression analysis (Figure 2) and HLB formation (Figure 3) embryos were collected as follows: Embryos of the genotype Δ*HisC*/Δ*HisC*; $12X^{PR}$/+ and Δ*HisC*/Cyo; $12X^{PR}$/+ were obtained by crossing males of the genotype Δ*HisC*, UAS-2xEYFP/Δ*HisC*,UAS-2xEYFP; $12X^{PR}/12X^{PR}$ to females of the genotype Δ*HisC*,twi-Gal4/CyO; +/+ (Supplemental Figure1E, F). Embryos of the genotype Δ*HisC*/Δ*HisC*; $12X^{PR}$,$8X^{HWT}$/+ and Δ*HisC*/CyO; $12X^{PR}$,$8X^{HWT}$ /+ were obtained by crossing males of

the genotype Δ*HisC*,UAS- 2xEYFP/Δ*HisC*,UAS-2xEYFP; $12X^{PR}$,$8X^{HWT}$ /$12X^{PR}$,$8X^{HWT}$ to females of the genotype Δ*HisC*,twi-Gal4/CyO; +/+ (Supplemental Figure1E,F). GFP signal was used to distinguish between the Δ*HisC*/Δ*HisC* and Δ*HisC*/CyO embryonic genotypes.

Histone expression analysis

Total RNA was isolated from embryos using the Trizol reagent (Invitrogen). cDNA was synthesized with random hexamers using SuperscriptII (Invitrogen). PCR was performed using the cDNA template and gene-specific primers to each histone gene. Each reaction was performed at least three times. PCR products were digested using AflII (H1), XhoI (H2a), XbaI (H2a), NruI (H2b), Eco53KI (isoschizomer of SacI) (H3) and NgoMIV(isoschizomer of NaeI)(H4). Digested PCR products were run on an 8% polyacrylamide gel and stained with SYBR gold. Quantitation of band intensities was performed using Image Lab software. Bar plots of band intensities normalized to tubulin and relative to *yw* (control) were generated in GraphPad Prism (Dotmatics).

Embryo collection and fixation

Embryos were collected on apple juice agar plates and aged at 25°C. Embryos were dechorionated in 50% bleach and then fixed in 4% formaldehyde in PBS with 50% heptane, on a nutator for 15 min at room temperature. The lower formaldehyde layer was removed and replaced with methanol. Embryos were vigorously shaken for 30s to remove the vitellin membrane. Devitellinized embryos sink to the bottom. The heptane- methanol mixture and the embryos that did not sink were removed and replaced with fresh methanol. These embryos were then stored in methanol at -20°C to be used for immunostaining and FISH experiments.

Immunostaining

Fixed embryos were rehydrated using PBST (PBS + 0.1% TritonX-100) and blocked in Image-iT FX signal enhancer (Invitrogen) for 30m. The signal enhancer was removed and replaced with primary antibodies diluted in PBST at 4°C overnight. The embryos were then washed 3X with PBST, followed by an incubation in secondary antibodies diluted in PBST for 1h at room temperature. Embryos were then either stained with DAPI and mounted in Prolong-Gold antifade (Invitrogen) for imaging

or used further for FISH experiments.

Antibodies

Primary antibodies were guinea pig anti-Mxc (1:6000) (White *et al.*, 2011) and rabbit anti- GFP (1:1000) (Rockland #600-401-215). For FISH experiments rabbit anti-MXC was used (1:500 for *D. virilis* and 1:1000 for *D. melanogaster*) (White *et al.*, 2011). Alexa Fluor secondary antibodies were goat anti-rabbit-488 and goat anti-guinea pig-647 (at a dilution of 1:1000).

Fluorescent in situ hybridization

Custom Stellaris RNA FISH probes targeting the coding region of core histone mRNA (H2a, H2b, H3 and H4) in *Drosophila melanogaster* and those targeting the *H4* mRNA in *Drosophila virilis* were designed using the Stellaris RNA FISH Probe Designer (LGC Biosearch Technologies) and labelled with Quasar670. Embryos that were fixed and stained as stated above were incubated in 4% formaldehyde in PBS for 10 min to crosslink bound antibodies, then washed thrice in 2XSSC with 10% formamide. Wash buffer was removed and replaced with hybridization buffer (2XSSC + 10% formamide + 10% dextran sulphate). FISH probes diluted in hybridization buffer (final concentration 50nm for *D. melanogaster* and 100nm for *D. virilis*) were incubated with the embryos overnight at 37°C. Following hybridization, embryos were washed with wash buffer, stained with DAPI and mounted in Prolong-Gold anti-fade (Invitrogen) for imaging.

Confocal imaging and analysis

All images were acquired with a 63X oil immersion objective using a Zeiss LSM880 confocal microscope with Zen software. High_resolution images of HLBs (Figure 4B, bottom panel) were acquired with a 63X oil immersion objective using the Leica SP8 Lightning system at the highest resolution with LAS X software. Images were analyzed using FIJI and IMARIS(9.7.2) software. Quantitation of the number of HLBs within a nucleus (Figure 4 and 5), was performed in IMARIS as follows: Using the Surface function, nuclei within a blastoderm embryo were converted into individual surfaces by selecting DAPI as the source channel and a seed diameter of 4.5-5 microns. A quality threshold was applied to ensure that every surface generated corresponded to a nucleus. Merged or overlapping surfaces were manually deleted.

Using the Spots function, HLBs within nuclei were converted into spots by selecting the Mxc signal as the source channel and a seed diameter of 0.4-0.6 microns. The quality threshold was applied to ensure every spot generated was aligned with an Mxc focus. Spots that did not align with Mxc foci were manually deleted. Finally, HLBs within nuclei were counted using the "Split (spots) into surface objects" extension. The data generated was exported into excel sheets. Bar plots were created in GraphPad Prism. Despite our best effort to manually delete merged or overlapping surfaces and background staining spots that may be considered HLBs by the algorithm, a low level of error was observed (e.g. a nucleus with one HLB that overlaps with another nucleus with two HLBs sometimes can be considered as one surface with three spots. We have included these data in the bar plots. Taking into consideration a high n value, the low level of technical error does not affect the interpretation of our data.

Amplicon RNA-sequencing and Bioinformatic quantification of histone cDNA

Histone mRNAs were reverse transcribed using XXX RT with random priming and amplified using the PCR primers listed below. Illumina adaptors and sample specific bar codes were subsequently added in two consecutive rounds of PCR amplification (XX and YY cycles respectively). The libraries were prepared using the small RNA protocol as described previously in Smola et al., 2015 (PMID 26426499). Following library preparation and quality control, the samples were loaded on an Illumina MiSeq subjected to paired-end sequencing using a 600-cycle kit. The reads were demultiplexed using Illumina BaseSpace and the fastq files analyzed.

Commented [AL2]: Bill, is this correct?

Commented [AL3]: Bill did we list the correct primers below for amplicon amplification?

Commented [JW4]: Ashlesha did these steps, and I'm not sure how many cycles she ended up doing for PCR-1 and PCR-2. It can vary based on how much sample you have to begin with.

Since the mutant and wild-type histone genes differ only by several nucleotides, we opted to use an exact match criterion to quantify relative expression in our sequencing data. We identified unique 25 nucleotide sequences in each of the histone mRNAs and used an exact match regular expression to count reads. The unique sequences are provided in supplementary table XX.

We counted both exact matches and reverse complement matches in both R1 and R2, and used the raw read counts from the read with the higher quality scores as defined by the MiSeq Illumina sequencer. We then computed relative ratios of read

counts and report these in Table YYY. Raw fastq files were uploaded to the Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra), under project ID PRJN

Commented [AL5]: Bill or Bob, make sure to reference this correctly.

Commented [LA6]: Colin please update this.

PCR Primers

H1_forward 5'-GTCTGATTCTGCAGTTGCAACG-3' H1_reverse 5'-TCCAGTTTTCTTGGCATCC-3'

H2a_forward 5'-GGCCATGTCTGGACGTGGAAAAGGT-3'

H2a_reverse 5'-GGCCTTAGGCCTTCTTCTCGGTCTT-3' H2b_forward 5'-CTAGTGGAAAGGCAGCCA-3'

H2b_reverse 5'-GAGCTGGTGTACTTGGTGA-3'

H3_forward 5'-GCTACTAAGGCCGCTCG-3'

H3_reverse 5'-GGCATTATGGTGACACGC-3'

H4_forward 5'-GCC AAA TCC GTA GAG GGT-3'

H4_reverse 5'-GGTCGTGGTAAAGGAGGCA-3'

α-tubulin_forward 5'-GGCAGTTCGAACGTATACGC-3'

α-tubulin_reverse 5'-GACCACAGTGGGTTCCAGAT-3'

attB 5'- AGTGTGTCGCTGTCGAGATG-3'

attP 5'-CCTTCACGTTTTCCCAGGT-3'

Lamp1_forward 5'- CCTGTGTTATATAAACCCGTGATA-3'

Lamp1_reverse 5'- CTAACGAACGTAAGCGACAC-3'

Pry4_forward 5'- CAATCATATCGCTGTCTCACTCA-3'

PR_verification_forward 5'-CGATGACGCTTGGCGCCAC-3'

PR_verification_reverse 5'-CCACCAGTCGATTTGCGAGCAG-3'

Acknowledgements

We thank Mark Geisler for developing the *D. melanogaster* core histone FISH probes, Sierra Cole for helping develop the PCR assay, and Tony Perdue, the

Department of Biology Imaging core and James Kemp for help with confocal microscopy. This work was supported by GM058921 from the National Institutes of Health. Additional support for this work was provided by NHLBI R01 HL111527 and NIGMS R35 GM140844 to A.L.

Figure Legends

Figure 1. All five RD histone genes display homeostatic control of histone mRNA levels. (A) Schematic representation of the endogenous RD histone gene unit of Drosophila melanogaster and the transgenic histone gene unit (Designer Wild-Type, DWT). The endogenous histone gene unit is tandemly arrayed ~100 times at the HisC locus on chromo- some 2L, resulting in ~200 copies of RD histone genes in a diploid fly. The transgenic array consists of 12 repeating gene units inserted on chromosome 3L, making a total of 24 RD histone gene copies in a homozygous genotype. Each histone gene in the transgenic array is designed to be molecularly distinguishable from the endogenous counterpart through insertion or removal of specific restriction sites as indicated. (B-D) Polyacrylamide gels of restriction digested RT-PCR products from 3-6 hour old embryos for (B) the *H2A-H2B* gene pair, (C) the *H3-H4* gene pair and (D) the linker histone, H1. mRNA level for each gene was measured in three genotypes: yw control (“200”, lanes 1), *ΔHisC* / *ΔHisC*; $12X^{DWT}/12X^{DWT}$ (“24”, lanes 2) and +/+; $12X^{DWT}/12X^{DWT}$ (“200, 24”, lanes 3). Bar plots of band intensity normalized to tubulin and relative to *yw* from three biological replicates. Values indicate mean and error bars indicate SD.

Figure 2. The 12X promoter replacement transgenic array is outcompeted by ~100 endogenous histone genes but not by an 8X wild-type histone array. (A) Schematic representation of transgenic RD histone gene arrays inserted on chromosome 3. The promoter replacement (PR) array inserted on chromosome 3L has 12 histone gene units in which the H3-H4 promoter (blue rectangle) is replaced by the H2A-H2B promoter (yellow rectangle). The Histone Wild-Type (HWT) array inserted on chromosome 3R consists of 8 repeating wild-type RD histone gene units. (B) Polyacrylamide gels of restriction digested RT-PCR products. mRNA level for each RD histone gene was measured in three genotypes: *ΔHisC* / *ΔHisC*; 12XPR/+ (lanes 1), *ΔHisC* /CyO; 12XPR,8XHWT/+ (lanes 2) and *ΔHisC* / *ΔHisC*; 12XPR, 8XHWT /+

(lanes 3). Asterisk indicates low molecular weight restriction digested product(~50bp). (C) Bar plots of band intensity normalized to undigested PCR products from three biological replicates. Values indicate mean and error bars indicate SD.

Figure 3. HLB assembly at the promoter replacement array is impaired by the presence of endogenous histone genes. (A) G1-arrested epidermal cells stained with antibodies against Mxc in germband retracted embryos from three different genotypes; *ΔHisC* / *ΔHisC*; 12XPR/12XPR (top panel), *ΔHisC*/CyO;12XPR/+ (middle panel) and *ΔHisC* / *ΔHisC*; 12XPR, 8XHWT/+ (bottom panel). Schematics on the right represent HLB formation at the histone loci (endogenous HisC and transgenic) in each respective genotype. Possible fusion of HLBs at non-homologous loci located on the same chromosome is also depicted (bottom panel schematic). Red circles represent nuclei with two HLBs. Scale bar, 5 microns. (B) Bar plots represent the number of HLBs detected in each nucleus in embryos from each genotype. “n” indicates the number of nuclei analyzed for each genotype. Values over bars show percentage of nuclei.

Figure 4. Transcription occurs at individual HLBs formed at ectopic histone gene arrays. (A) Syncytial nuclear cycle 14 embryos stained with antibodies against Mxc. Top panel shows Oregon R wild type embryos, and bottom panel shows embryos carrying a homozygous wild-type transgenic histone gene array (12XDWT) in the pres- ence of HisC. Colored circles represent nuclei with 1-4 HLBs (1:red, 2:yellow, 3:blue, 4:orange). Bar plots represent the number of HLBs detected in each nucleus, quanti- fied using IMARIS imaging software (See Methods). “n” indicates the number of nuclei analyzed for each genotype. Values over bars show percentage of nuclei. Schematics on the right represent HLB formation at the endogenous (blue chromosomes) and transgenic (green chromosomes) histone loci present in each respective genotype. Scale bar, 5 microns. (B) Syncytial blastoderm embryos from the genotypes indicated above, simultaneously stained for Mxc (left panel) and hybridized with fluorescent probes detecting RD core histone RNAs (i.e. H2A, HB, H3, H4) (middle panel). The panel on the right shows a merge of HLBs and nascent

RNA transcripts. The bottom panel shows high-resolution images of HLBs obtained using the Leica SP8 Lightning system where the pinhole was set at 0.6 Airy units for increased resolution. Dashed square represents zoomed-in images of a single nucleus shown in smaller panels on the right. Scale bar, 5 microns.

Figure 5. HLB formation and transcription can occur independently at both the major and minor histone loci in Drosophila virilis. (A) Schematic representation of the Drosophila virilis histone loci on chromosome 2 (major locus) and on chromosome 4 (minor locus). (B) D. virilis syncytial blastoderm embryos stained with antibodies raised against D. melanogaster Mxc. Colored circles represent nuclei with 1-4 HLBs (1:red, 2:yellow, 3:blue, 4:orange). Bar plots represent the number of HLBs detected in each nucleus, quantified using IMARIS imaging software. “n” indicates the number of nuclei analyzed. Values over bars show percentage of nuclei. (C) D. virilis syncytial blastoderm embryos simultaneously stained for Mxc (left panel) and hybridized with fluorescent probes detecting D. virilis H4 histone mRNA (middle panel). Panel on the right shows a merge of HLBs and nascent H4 mRNA transcripts. Dashed square represents zoomed-in images of a single nucleus shown in smaller panels on the right. Scale bar, 5 microns.

Supplemental Figure 1. (related to Figure 1, 2 and 3). Genotype verification of Drosophila stocks carrying transgenic histone gene arrays. Schematic representation of PCR design used to detect (A) HisC deletion (B) Insertion of a trans- genic histone gene array. For (C) and (D) PCR was performed using genomic DNA from three genotypes: yw control (“200”, lanes 1), *ΔHisC* / *ΔHisC*; 12XDWT/12XDWT (“24”, lanes 2) and +/+; 12XDWT/12XDWT (“200, 24”, lanes 3). (C) Poly-acrylamide gel of PCR products demonstrating the presence and absence of the HisC locus. In the presence of HisC (lane 1 and 3), a PCR product is generated by the Lamp1 forward (L_F) and Lamp1 reverse (L_R) primers resulting in a product of ~748 bp. In a HisC deletion genotype (lane 2), a portion of the Lamp1 gene is deleted and therefore L_F cannot bind. Instead, a PCR product is generated using Pry4 forward (P_F) and L_R, resulting in a product of ~600 bp. (D) Polyacrylamide gel of PCR products demonstrating formation of an attR site due to

attB-attP recombination for the insertion of 12XDWT at VK33. A PCR product is observed only in genotypes carrying a transgenic histone gene array (lane 2 and 3). For (E) and (F) PCR was performed using genomic DNA from following genotypes: OregonR control (lane 1), *ΔHisC* / *ΔHisC*; 12XDWT/12XDWT control (lane 2), *ΔHisC* /CyO; +/+ female parent (lane 3), *ΔHisC* / *ΔHisC*; 12XPR/12XPR male parent (lane 4), *ΔHisC* /CyO; +/+ female parent (lane 5) and *ΔHisC* / *ΔHisC*; 12XPR, 8XHWT /12XPR, 8XHWT male parent (lane 6). (E) Polyacrylamide gel of PCR products demonstrating the presence and absence of the HisC locus. Two bands are observed in lanes 3 and 5 because these genotypes are heterozygous for the HisC deletion (*ΔHisC* /CyO). (F) Polyacrylamide gel of PCR products detecting the presence of either the H3-H4 promoter (298 bp) or the H2a-H2b promoter (226 bp) between the H3-H4 gene pair. The 72 bp difference between the two bidirectional promoters allows us to differentiate between genotypes that are either wild-type (lane 1, 3 and 5), carrying wildtype transgenic arrays (lane 2, 12XDWT, lane 6 8XHWT) or carrying the promoter replacement array (lane 4 and 6). (G) Table representing histone gene PCR products and the respective restriction digestion products of endogenous and trans- genic histone genes used for the analyses shown in Figure 1B-D and Figure 2B (sizes in basepairs (bp).

Supplemental Figure 2. (related to Figure 3 and 4). HLB formation can occur at the promoter replacement array only in the absence of HisC. Mxc staining of syncytial nuclear cycle 14 embryos that are either (A) carrying homozygous transgenic arrays (12XDWT, top panel and 12XPR, bottom panel) in a histone deletion background or (B) carrying homozygous promoter replacement transgenic array (12XPR) in presence of HisC. Colored circles represent nuclei with 1 / 2 HLBs (1:red, 2:yellow). Bar plots represent the number of HLBs detected in each nucleus. Mxc foci in nuclei were counted using the IMARIS imaging software (See Methods). n indicates the number of nuclei analyzed for each genotype. Values over bars show percentage of nuclei. Schematics on the right represent HLB formation at the histone loci (endogenous and transgenic) present in each respective genotype. Scale bar, 5 microns.

References

Albig, W., and Doenecke, D. (1997). The human histone gene cluster at the D6S105 locus. Hum Genet *101*, 284-294.

Amodeo, A.A., Jukam, D., Straight, A.F., and Skotheim, J.M. (2015). Histone titration against the genome sets the DNA-to-cytoplasm threshold for the Xenopus midblastula transition. Proc Natl Acad Sci U S A *112*, E1086-1095.

Armstrong, C., and Spencer, S.L. (2021). Replication-dependent histone biosynthesis is coupled to cell-cycle commitment. Proc Natl Acad Sci U S A *118*.

Au, W.C., Crisp, M.J., DeLuca, S.Z., Rando, O.J., and Basrai, M.A. (2008). Altered dosage and mislocalization of histone H3 and Cse4p lead to chromosome loss in Saccharomyces cerevisiae. Genetics *179*, 263-275.

Bongartz, P., and Schloissnig, S. (2019). Deep repeat resolution-the assembly of the Drosophila Histone Complex. Nucleic Acids Res *47*, e18.

Bulchand, S., Menon, S.D., George, S.E., and Chia, W. (2010). Muscle wasted: a novel component of the Drosophila histone locus body required for muscle integrity. J Cell Sci *123*, 2697-2707.

Cook, A.J., Gurard-Levin, Z.A., Vassias, I., and Almouzni, G. (2011). A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain. Mol Cell *44*, 918-927.

Domier, L.L., Rivard, J.J., Sabatini, L.M., and Blumenfeld, M. (1986). Drosophila virilis histone gene clusters lacking H1 coding segments. J Mol Evol *23*, 149-158.

Dominski, Z., and Marzluff, W.F. (2007). Formation of the 3' end of histone mRNA: getting closer to the end. Gene *396*, 373-390.

Duronio, R.J., and Marzluff, W.F. (2017). Coordinating cell cycle-regulated histone gene expression through assembly and function of the Histone Locus Body. RNA Biol *14*, 726-738.

Eriksson, P.R., Ganguli, D., Nagarajavel, V., and Clark, D.J. (2012). Regulation of histone gene expression in budding yeast. Genetics *191*, 7-20.

Günesdogan, U., Jäckle, H., and Herzig, A. (2010). A genetic system to assess in vivo the functions of histones and histone modifications in higher eukaryotes. EMBO Rep *11*, 772-776.

Gunjan, A., and Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in S. cerevisiae. Cell *115*, 537-549.

Han, M., Chang, M., Kim, U.J., and Grunstein, M. (1987). Histone H2B repression

causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. Cell *48*, 589-597.

Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D., and Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol Cell Biol *11*, 2416-2424.

Hentschel, C.C., and Birnstiel, M.L. (1981). The organization and expression of histone gene families. Cell *25*, 301-313.

Hiraoka, Y., Dernburg, A.F., Parmelee, S.J., Rykowski, M.C., Agard, D.A., and Sedat, J.W. (1993). The onset of homologous chromosome pairing during Drosophila melanogaster embryogenesis. J Cell Biol *120*, 591-600.

Hur, W., Kemp, J.P., Jr., Tarzia, M., Deneke, V.E., Marzluff, W.F., Duronio, R.J., and Di Talia, S. (2020). CDK-Regulated Phase Separation Seeded by Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. Dev Cell *54*, 379-394.e376.

Joseph, S.R., Pálfy, M., Hilbert, L., Kumar, M., Karschau, J., Zaburdaev, V., Shevchenko, A., and Vastenhouw, N.L. (2017). Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. Elife *6*.

Kemp, J.P., Jr., Yang, X.C., Dominski, Z., Marzluff, W.F., and Duronio, R.J. (2021). Superresolution light microscopy of the Drosophila histone locus body reveals a core-shell organization associated with expression of replication-dependent histone genes. Mol Biol Cell *32*, 942-955.

Khan, S.U., Khan, M.U., Kalsoom, F., Khan, M.I., Gao, S., Unar, A., Zubair, M., and Bilal, M. (2022). Mechanisms of gene regulation by histone degradation in adaptation of yeast: an overview of recent advances. Arch Microbiol *204*, 287.

Kim, U.J., Han, M., Kayne, P., and Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. Embo j *7*, 2211-2219.

Koreski, K.P., Rieder, L.E., McLain, L.M., Chaubal, A., Marzluff, W.F., and Duronio, R.J. (2020). Drosophila histone locus body assembly and function involves multiple interactions. Mol Biol Cell *31*, 1525-1537.

Liang, D., Burkhart, S.L., Singh, R.K., Kabbaj, M.H., and Gunjan, A. (2012). Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by

the homologous recombination pathway. Nucleic Acids Res *40*, 9604-9620.

Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. Cold Spring Harb Symp Quant Biol *42 Pt 2*, 1047-1051.

Ma, T., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T., and Harper, J.W. (2000). Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. Genes Dev *14*, 2298-2313.

Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The human and mouse replication-dependent histone genes. Genomics *80*, 487-498.

Marzluff, W.F., and Koreski, K.P. (2017). Birth and Death of Histone mRNAs. Trends Genet *33*, 745-759.

Maxson, R., Cohn, R., Kedes, L., and Mohun, T. (1983). Expression and organization of histone genes. Annu Rev Genet *17*, 239-277.

McKay, D.J., Klusza, S., Penke, T.J., Meers, M.P., Curry, K.P., McDaniel, S.L., Malek, P.Y., Cooper, S.W., Tatomer, D.C., Lieb, J.D., Strahl, B.D., Duronio, R.J., and Matera, A.G. (2015). Interrogating the function of metazoan histones using engineered gene clusters. Dev Cell *32*, 373-386.

Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell *44*, 43-52.

Meers, M.P., Leatham-Jensen, M., Penke, T.J.R., McKay, D.J., Duronio, R.J., and Matera, A.G. (2018). An Animal Model for Genetic Analysis of Multi-Gene Families: Cloning and Transgenesis of Large Tandemly Repeated Histone Gene Clusters. Methods Mol Biol *1832*, 309-325.

Mendiratta, S., Gatto, A., and Almouzni, G. (2019). Histone supply: Multitiered regulation ensures chromatin dynamics throughout the cell cycle. J Cell Biol *218*, 39-54.

Nelson, D.M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G.D., Yen, T.J., Harper, J.W., and Adams, P.D. (2002). Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity. Mol Cell Biol *22*, 7459-7472.

Rieder, L.E., Koreski, K.P., Boltz, K.A., Kuzu, G., Urban, J.A., Bowman, S.K., Zeidman, A., Jordan, W.T., 3rd, Tolstorukov, M.Y., Marzluff, W.F., Duronio, R.J., and Larschan, E.N. (2017). Histone locus regulation by the Drosophila dosage

compensation adaptor protein CLAMP. Genes Dev *31*, 1494-1508.

Roberts, S.B., Sanicola, M., Emmons, S.W., and Childs, G. (1987). Molecular characterization of the histone gene family of Caenorhabditis elegans. J Mol Biol *196*, 27-38.

Salzler, H.R., Tatomer, D.C., Malek, P.Y., McDaniel, S.L., Orlando, A.N., Marzluff, W.F., and Duronio, R.J. (2013). A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. Dev Cell *24*, 623-634.

Schienman, J.E., Lozovskaya, E.R., and Strausbaugh, L.D. (1998). Drosophila virilis has atypical kinds and arrangements of histone repeats. Chromosoma *107*, 529-539.

Seal, R.L., Denny, P., Bruford, E.A., Gribkova, A.K., Landsman, D., Marzluff, W.F., McAndrews, M., Panchenko, A.R., Shaytan, A.K., and Talbert, P.B. (2022). A standardized nomenclature for mammalian histone genes. Epigenetics Chromatin *15*, 34.

Shiotsugu, J.S. (2002). Molecular evolution of repetitive sequences: The histone genes of Drosophila virilis. Ph.D., University of Connecticut, United States -- Connecticut.

Singh, R.K., Kabbaj, M.H., Paik, J., and Gunjan, A. (2009). Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. Nat Cell Biol *11*, 925-933.

Singh, R.K., Liang, D., Gajjalaiahvari, U.R., Kabbaj, M.H., Paik, J., and Gunjan, A. (2010). Excess histone levels mediate cytotoxicity via multiple mechanisms. Cell Cycle *9*, 4236-4244.

Sullivan, E., Santiago, C., Parker, E.D., Dominski, Z., Yang, X., Lanzotti, D.J., Ingledue, T.C., Marzluff, W.F., and Duronio, R.J. (2001). Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. Genes Dev *15*, 173-187.

Talbert, P.B., and Henikoff, S. (2017). Histone variants on the move: substrates for chromatin dynamics. Nat Rev Mol Cell Biol *18*, 115-126.

Terzo, E.A., Lyons, S.M., Poulton, J.S., Temple, B.R., Marzluff, W.F., and Duronio, R.J. (2015). Distinct self-interaction domains promote Multi Sex Combs accumulation in and formation of the Drosophila histone locus body. Mol Biol Cell *26*, 1559-1574.

Wagner, E.J., Berkow, A., and Marzluff, W.F. (2005). Expression of an RNAi-resistant SLBP restores proper S-phase progression. Biochem Soc Trans *33*, 471-473.

Wei, Y., Jin, J., and Harper, J.W. (2003). The cyclin E/Cdk2 substrate and Cajal body component p220(NPAT) activates histone transcription through a novel LisH-like domain. Mol Cell Biol *23*, 3669-3680.

White, A.E., Burch, B.D., Yang, X.C., Gasdaska, P.Y., Dominski, Z., Marzluff, W.F., and Duronio, R.J. (2011). Drosophila histone locus bodies form by hierarchical recruitment of components. J Cell Biol *193*, 677-694.

White, A.E., Leslie, M.E., Calvi, B.R., Marzluff, W.F., and Duronio, R.J. (2007). Developmental and cell cycle regulation of the Drosophila histone locus body. Mol Biol Cell *18*, 2491-2502.

Yang, X.C., Burch, B.D., Yan, Y., Marzluff, W.F., and Dominski, Z. (2009). FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. Mol Cell *36*, 267-278.

Ye, X., Wei, Y., Nalepa, G., and Harper, J.W. (2003). The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. Mol Cell Biol *23*, 8586-8600.