

Histone locus regulation by the *Drosophila* dosage compensation adaptor protein CLAMP

Leila E. Rieder,^{1,13} Kaitlin P. Koreski,^{2,13} Kara A. Boltz,^{3,4,9} Guray Kuzu,^{5,10} Jennifer A. Urban,¹ Sarah K. Bowman,^{5,11} Anna Zeidman,^{1,12} William T. Jordan III,¹ Michael Y. Tolstorukov,⁵ William F. Marzluff,^{2,3,4,6,7} Robert J. Duronio,^{2,3,4,7,8} and Erica N. Larschan¹

¹Department of Molecular Biology, Cellular Biology, and Biochemistry, Brown University, Providence, Rhode Island 02912, USA; ²Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA; ³Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599, USA; ⁴Integrative Program for Biological and Genome Sciences, University of North Carolina, Chapel Hill, North Carolina 27599, USA; ⁵Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA; ⁶Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina 27599, USA; ⁷Department of Biology, University of North Carolina at Chapel Hill, University of North Carolina, Chapel Hill, North Carolina 27599, USA; ⁸Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, USA

The conserved histone locus body (HLB) assembles prior to zygotic gene activation early during development and concentrates factors into a nuclear domain of coordinated histone gene regulation. Although HLBs form specifically at replication-dependent histone loci, the *cis* and *trans* factors that target HLB components to histone genes remained unknown. Here we report that conserved GA repeat *cis* elements within the bidirectional *histone3–histone4* promoter direct HLB formation in *Drosophila*. In addition, the CLAMP (chromatin-linked adaptor for male-specific lethal [MSL] proteins) zinc finger protein binds these GA repeat motifs, increases chromatin accessibility, enhances histone gene transcription, and promotes HLB formation. We demonstrated previously that CLAMP also promotes the formation of another domain of coordinated gene regulation: the dosage-compensated male X chromosome. Therefore, CLAMP binding to GA repeat motifs promotes the formation of two distinct domains of coordinated gene activation located at different places in the genome.

[*Keywords:* nuclear bodies; histone locus body; CLAMP; GA repeats; GAGA factor; dosage compensation]

Supplemental material is available for this article.

Received April 27, 2017; revised version accepted July 25, 2017.

Within the complex environment of the nucleus, coordinated gene expression is facilitated by membraneless structures known as nuclear bodies (NBs). NBs are critical for the precise spatial and temporal regulation and processing of RNAs and include nucleoli, Cajal bodies, and histone locus bodies (HLBs) (Mao et al. 2011). NBs share properties and assembly mechanisms with larger nuclear domains that regulate coordinated gene expression, such as the dosage-compensated X chromosome (in mammals, the Barr body). NBs improve the efficiency and coordina-

tion of nuclear processes, such as transcription and RNA processing, by concentrating factors to promote interactions that would otherwise be stochastic (Matera et al. 2009; Mao et al. 2011; Tatomer et al. 2016). Despite their importance, our understanding of how specific NBs are formed early during development remains incomplete.

The HLB is a highly conserved NB that assembles at the replication-dependent histone genes (Liu et al. 2006), which are present in multiple clustered copies in most metazoans (Duronio and Marzluff 2017). Humans have two histone gene clusters, a major cluster on chromosome 6 and a minor cluster on chromosome 1 (Albig and Doenecke 1997; Marzluff et al. 2002), while most *Drosophila* species have a single replication-dependent

Present addresses: ⁹Department of Entomology, North Carolina State University, Raleigh, NC 27695, USA; ¹⁰Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, PA 16801, USA; ¹¹Directed Genomics, Ipswich, MA 01938, USA; ¹²Science Outreach Program, The Rockefeller University, New York, NY 10065, USA

¹³These authors contributed equally to this work.

Corresponding authors: leila_rieder@brown.edu, erica_larschan@brown.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.300855.117>.

© 2017 Rieder et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

histone gene locus. In *Drosophila melanogaster*, the histone locus resides on chromosome 2L and consists of a tandem array of ~100 copies of a 5-kb cluster containing each of the histone genes (Lifton et al. 1978; McKay et al. 2015). Coordinated expression of histone genes is necessary to maintain nucleosome subunit stoichiometry, and this requirement is reflected in the arrangement of the *Drosophila* histone genes that encode nucleosomal core proteins: In each 5-kb gene cluster, *H2A* and *H2B* share a bidirectional promoter, as do *H3* and *H4*. This same arrangement is present in other species, such as budding yeast (Smith and Murray 1983; Eriksson et al. 2012). Histone production is also tightly coordinated across the cell cycle, leading to a burst of histone mRNA production at the beginning of each S phase (Marzluff et al. 2008). Many factors involved in the cell cycle-regulated transcription and processing of histone transcripts are concentrated in the HLB (Duronio and Marzluff 2017).

A common theme for NB assembly is that a “scaffolding” protein serves as a platform to recruit other NB components. In *Drosophila*, HLB scaffolding is mediated by the multi-sex combs (Mxc) protein, the ortholog of mammalian NPAT (nuclear protein of the ataxia telangiectasia-mutated locus), a Cyclin E/Cdk2 substrate that is essential for both HLB assembly and histone gene expression (Ma et al. 2000; Zhao et al. 2000; Ye et al. 2003; White et al. 2011; Terzo et al. 2015). Early during *Drosophila* development, before the initiation of zygotic histone gene expression, Mxc assembles into a “proto-HLB” along with FLASH (FLICE-associated huge protein) (White et al. 2011; Salzler et al. 2013), a protein necessary for endonucleolytic cleavage to form mature histone mRNA (Yang et al. 2009; Burch et al. 2011; Tatomer et al. 2016). Once Mxc and FLASH assemble into a proto-HLB, other factors involved in histone mRNA biosynthesis are recruited to the HLB (White et al. 2011; Salzler et al. 2013), including the mRNA processing factor U7 snRNP (Strub and Birnstiel 1986; Mowry and Steitz 1987) and Mute (muscle wasted), a putative transcriptional repressor and homolog of the mammalian YY1-associated protein (Bulchand et al. 2010; Yang et al. 2014). These data suggest that ordered recruitment of factors contributes to HLB assembly.

How the process of scaffolding the HLB is initiated and functionally linked to regulation of the histone locus chromatin and histone gene expression is not understood. Nucleation of Mxc/FLASH proto-HLBs does not require expression of histone mRNA (Salzler et al. 2013). Thus, one possibility is that a factor expressed during early development binds DNA at or near the histone genes and initiates HLB assembly and histone gene activation, perhaps by interacting with scaffolding factors such as Mxc/NPAT. Using engineered histone transgenes, Salzler et al. (2013) determined previously that the ~300-base-pair (bp) bidirectional promoter between the *Drosophila* *H3* and *H4* genes (*H3–H4p*) is necessary and sufficient to recruit HLB factors, including Mxc, FLASH, U7 snRNP, and Mute. Although transcription from the *H3–H4p* is necessary for full recruitment of HLB factors, some Mxc and FLASH is recruited even in the absence of an active

H3–H4p (Salzler et al. 2013). In addition, once fully formed, HLBs do not require ongoing transcription for maintenance, as they are present in G1-arrested cells that do not express histone genes (Liu et al. 2006; White et al. 2007). Thus, some HLB component likely recognizes a *cis* element in the DNA at the histone locus. The scaffolding protein Mxc contains one AT-hook domain, but there is no evidence that Mxc or NPAT directly binds DNA (Wei et al. 2003; Miele et al. 2005; Terzo et al. 2015).

The *H3–H4p* is highly conserved among 12 *Drosophila* species and contains two GA repeat *cis* elements (Salzler et al. 2013). GA-rich *cis* elements have been implicated in a variety of nuclear processes in *Drosophila*, including RNA polymerase II pausing (Tsai et al. 2016), zygotic genome activation (Chen et al. 2013), three-dimensional genome organization (Quinn et al. 2014), and DNA loop formation (Eagen et al. 2017). Two known *Drosophila* zinc finger transcription factors directly interact with GA repeats. The first, the well-studied GAGA factor (GAF; *trithorax-like* [*trl*]), opens chromatin and modulates transcriptional pausing at many genes (Guertin et al. 2012; Fuda et al. 2015). The second, chromatin-linked adaptor for male-specific lethal (MSL) proteins (CLAMP), is a zinc finger DNA-binding protein that is required for male X-chromosome dosage compensation (Larschan et al. 2012). CLAMP binds throughout the genome but is enriched at evolutionarily conserved long GA repeats on the X chromosome (Kuzu et al. 2016), where it recruits the MSL complex (Larschan et al. 2012; Soruco et al. 2013). The MSL complex generates a chromosomal domain of coordinated gene activation that increases transcript levels of male X-linked genes twofold, equalizing expression between XY males and XX females (Belote and Lucchesi 1980; Hamada et al. 2005). While not historically considered a NB, the male *Drosophila* X chromosome represents a distinct domain of coordinated gene activation similar to the histone locus.

Using genetic, genomic, and biochemical approaches, we show that the conserved GA repeats within the *H3–H4p* direct HLB formation. CLAMP, but not GAF, binds to these repeats early during development, before zygotic genome activation and prior to formation of the mature HLB. CLAMP is critical for histone gene expression and opening of chromatin at the histone locus. Furthermore, tethering CLAMP to an ectopic histone locus is sufficient to recruit HLB factors. Therefore, the presence of CLAMP and the absence of GAF at GA repeats at the HLB and the male X chromosome (Soruco et al. 2013) are common properties shared by two different domains of coordinated gene activation.

Results

CLAMP is recruited to the histone locus via the H3–H4p

While studying CLAMP in the context of male X-chromosome dosage compensation, we noticed distinct CLAMP puncta in the nuclei of early *D. melanogaster* embryos. Costaining revealed that these CLAMP puncta colocalized with markers of the HLB in embryos (Fig. 1A) and

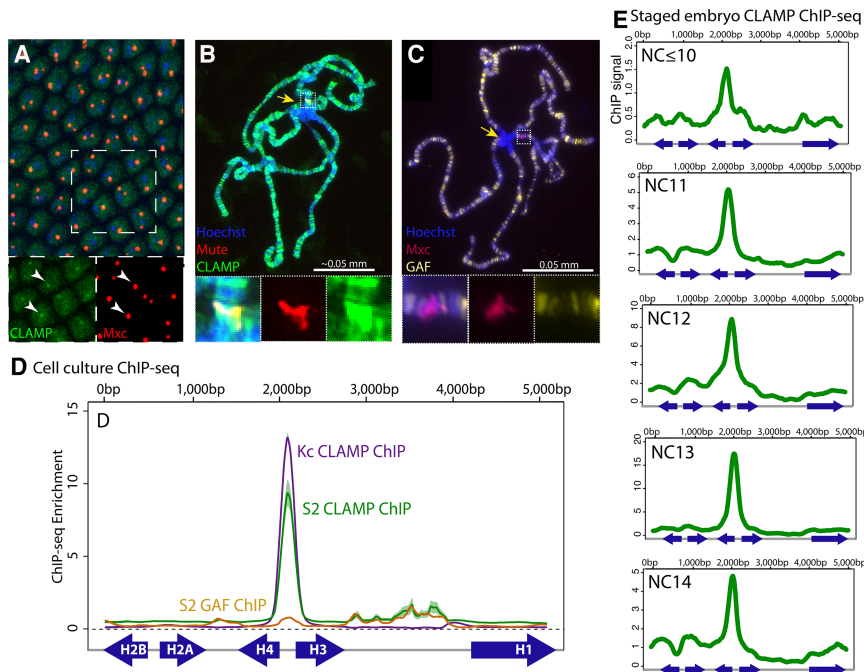


Figure 1. CLAMP colocalizes with markers of the HLB. Embryos (A) and third instar larvae salivary gland polytene chromosomes (B,C) immunostained for CLAMP (green), HLB components (Mxc and Mute; red), and GAF (yellow). (A) CLAMP forms distinct puncta in the syncytial nuclei of wild-type *Drosophila* embryos that colocalize (arrowheads) with Mxc foci. (B) In salivary gland polytene chromosomes, CLAMP colocalizes with Mute at the histone locus near the chromocenter (yellow arrow). (C) GAF, another GA-binding factor, does not colocalize with Mxc. (D) We mapped our previous CLAMP ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) data from cultured S2 (male; green) and Kc (female; purple) cells (Sorucu et al. 2013) and existing GAF ChIP-seq (yellow) from cultured S2 cells (Fuda et al. 2015) to the histone gene cluster. Shading represents 95% confidence intervals. ChIP-seq data were normalized to inputs. (E) We performed CLAMP ChIP-seq from precisely staged early embryos and mapped reads to the his-

tone gene cluster. CLAMP is present at the *H3–H4p* as early as nuclear cycle 10, before zygotic genome activation. We normalized CLAMP ChIP-seq data to ChIP input, as in D.

cultured cells (Supplemental Fig. S1A) and on the giant salivary gland polytene chromosomes of third instar larvae (Fig. 1B). A GFP-tagged full-length CLAMP also colocalized with HLB markers on salivary gland polytene chromosomes (Supplemental Fig. S1B). Both CLAMP and GAF recognize GA repeats throughout the genome, often at the same loci (Kasinathan et al. 2014; Kuzu et al. 2016). However, we found that GAF was not present at the HLB (Fig. 1C).

The endogenous *Drosophila* histone locus on chromosome 2L contains ~100 tandem copies of a 5-kb gene cluster (Lifton et al. 1978; McKay et al. 2015), each containing a single copy of the five replication-dependent histone genes (Fig. 1D). To determine the exact location of CLAMP binding within the histone locus, we mapped existing *Drosophila* cell culture CLAMP ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) data from our laboratory (Sorucu et al. 2013) and GAF ChIP-seq data (Fuda et al. 2015) to a custom genome containing a single copy of the 5-kb gene cluster (McKay et al. 2015). With this approach, the ChIP-seq signal represents an average binding profile across all ~100 gene clusters. We found that CLAMP localized precisely to the *H3–H4p* in both male S2 and female Kc cultured cells (Fig. 1D), which is the same region of the gene cluster that is minimally sufficient for recruitment of HLB components (Salzler et al. 2013). In marked contrast, GAF did not localize to the histone locus (Fig. 1C,D), suggesting that CLAMP provides a unique function at the histone locus, similar to our results on the dosage-compensated X chromosome (Sorucu et al. 2013; Kuzu et al. 2016).

A number of factors are present at the HLB constitutively throughout the cell cycle, while others are present only during S phase, when the histone genes are transcriptionally active (Duronio and Marzluff 2017). The scaffolding protein Mxc is present at the HLB throughout the cell cycle but is phosphorylated only when Cyclin E/Cdk2 is active (e.g., during S phase in cultured cells), creating a phosphoepitope recognized by the MPM-2 antibody (White et al. 2011). Therefore, to characterize whether CLAMP localization to the HLB is cell cycle-dependent, we used the MPM-2 antibody to label S-phase HLBs and the Mxc antibody to label all HLBs. Unlike the MPM-2 epitope, CLAMP was present at the HLBs in all cultured cells (Supplemental Fig. S1A), suggesting that CLAMP localizes to the HLB throughout the cell cycle.

A “proto-HLB” composed of FLASH and Mxc forms before the onset of zygotic histone gene transcription (White et al. 2011; Salzler et al. 2013). However, neither FLASH nor Mxc nor any other previously known component of the HLB has been shown to bind DNA and target the HLB to the histone locus. Our observations that CLAMP localizes to the *H3–H4p* (Fig. 1) and is present at the HLB throughout the cell cycle (Supplemental Fig. S1A) led us to hypothesize that CLAMP may be a factor that is recruited to the histone locus prior to activation of zygotic histone gene expression (i.e., by embryonic nuclear cycle 10) and therefore may be a component of the “proto-HLB.” To test this hypothesis, we performed CLAMP ChIP-seq from pools of 200–400 hand-sorted precisely staged embryos (Blythe and Wieschaus 2015). We identified pools of embryos at each of the nuclear cycles 11–14 using a PCNA-EGFP nuclear reporter. We

pooled embryos in cycle 10 and younger to obtain sufficient chromatin from such young embryos for ChIP-seq analysis. We then mapped ChIP-seq reads to the histone gene cluster. CLAMP is present at the embryonic *H3-H4p*, as observed in cultured cells (Fig. 1D), and was present at all assayed nuclear cycles, including by nuclear cycle 10 (Fig. 1E). In contrast, other HLB components (i.e., Mxc, FLASH, Mute, and U7 snRNP) are not detectable at nuclear foci prior to cycle 10 (White et al. 2007; Terzo et al. 2015). These observations demonstrate that CLAMP is present at the embryonic histone locus prior to zygotic genome activation and suggest that it is recruited to the histone locus before the mature HLB is formed (White et al. 2011; Salzler et al. 2013). Collectively, our observations led us to hypothesize that CLAMP regulates the histone locus during development by recognizing critical *cis* elements within the *H3-H4p*.

Conserved GA repeat *cis* elements in the *H3-H4p* are required for ectopic HLB formation

We showed previously that a single transgenic copy of the ~300-bp *H3-H4p* is necessary and sufficient to recruit HLB components to an ectopic locus. We identified two conserved GA repeat motifs within the *H3-H4p* as potential CLAMP-binding sites (Fig. 2A,B; Salzler et al. 2013). There is a substantial expansion of one GA repeat

in *D. melanogaster* compared with other drosophilids, including closely related species such as *Drosophila simulans* (Fig. 2B). Because we reported recently that expanded GA repeats facilitate CLAMP-mediated X-chromosome dosage compensation (Kuzu et al. 2016), we asked whether CLAMP localization to the HLB was specific to *D. melanogaster* by staining polytene chromosomes from *D. simulans* (Supplemental Fig. S2A,B) and *Drosophila virilis* (Fig. 2C,D), which diverged from *D. melanogaster* >40 million years ago (Russo et al. 1995). The genome of *D. simulans* contains a single histone locus near the chromocenter, while the genome of *D. virilis* contains two histone loci in the middles of chromosome arms (Schienman et al. 1998; Berloco et al. 2001). We found that CLAMP is present at the histone locus in both *D. simulans* and *D. virilis*. Similar to *D. melanogaster* (Fig. 1C), GAF did not colocalize with HLB factors in other species. Therefore, CLAMP localization to the histone locus is not specific to *D. melanogaster* and is not only due to the GA repeat expansion in the *D. melanogaster H3-H4p* (Fig. 2B).

Our CLAMP ChIP-seq results (Fig. 1D,E) and the sequence conservation of the *H3-H4p* (Fig. 2B) led us to hypothesize that the GA repeats may function to promote HLB formation. To identify regions in the *H3-H4p* that are important for HLB factor recruitment, we constructed four transgenes containing deletions in the ~300-bp *H3-H4p*. In three of the four constructs, either one or both

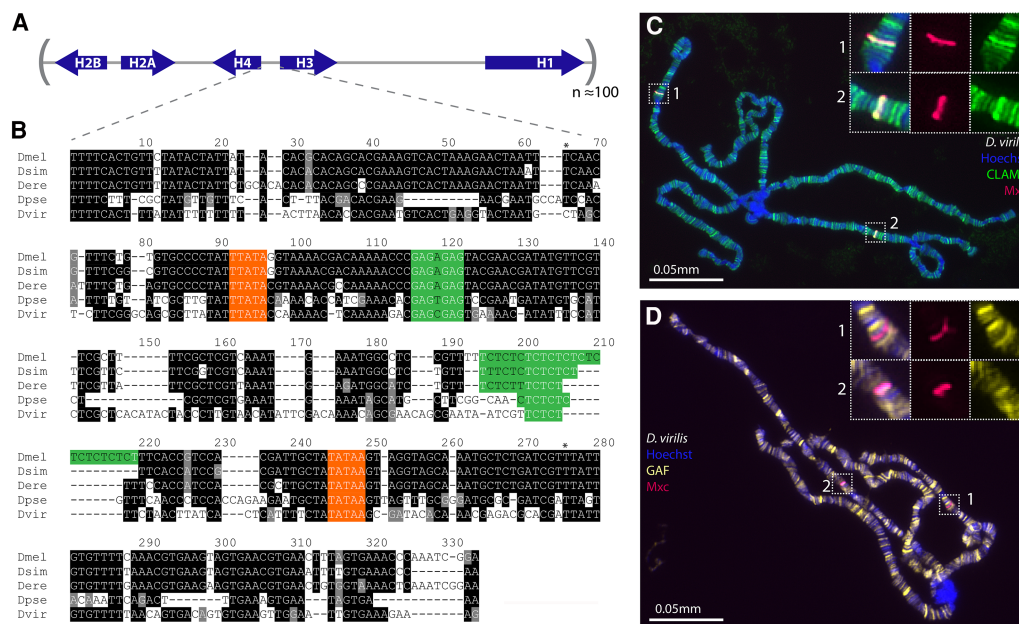


Figure 2. Localization of CLAMP at the HLB is conserved across drosophilids. (A) The five replication-dependent histone genes are clustered in a tandemly repeated array of an ~5-kb repeat unit. Approximately 100 repeat units comprise the histone locus on *D. melanogaster* chromosome 2L. (B) The ~300-bp bidirectional promoter between the *Histone3* and *Histone4* genes is highly conserved among drosophilids. (Dmel) *D. melanogaster*; (Dsim) *D. simulans*; (Dere) *Drosophila erecta*; (Dpse) *Drosophila pseudoobscura*; (Dvir) *D. virilis*. Two GA repeats (green) are highly conserved from *D. melanogaster* to *D. virilis* (~40-million-year-old divergence) (Russo et al. 1995). TATA boxes are highlighted in orange. There is a substantial expansion of one GA repeat in *D. melanogaster*. Asterisks represent transcription start sites for the *H3* and *H4* genes. (C,D) We stained polytene chromosomes from *D. virilis* (which has two histone loci) for Mxc (red), CLAMP (green), and GAF (yellow). Mxc and CLAMP are recruited to both *D. virilis* histone loci (C), and GAF is not recruited to either histone locus (D).

of the GA repeats are deleted (Supplemental Fig. S3A). We found that HLB factors were efficiently recruited only to the *H3-H4p* transgene constructs that preserve both GA repeats (Supplemental Fig. S3B–D).

To test directly whether the GA repeats are required for HLB formation, we used mutational analysis of BAC-based transgenes carrying 12 copies of the full 5-kb histone gene cluster (12xHistone^{WT}). We showed previously that the 12xHistone^{WT} transgene forms a functional ectopic HLB, recruiting all tested HLB factors in the presence of the endogenous histone locus (Fig. 3A), and rescues the lethality caused by homozygous deletion of the endogenous histone locus (McKay et al. 2015). The 12xHistone^{WT} transgene also recruits CLAMP (Fig. 3A). We therefore generated a mutant 12x array, 12x2xLacO, in which both GA repeats in the *H3-H4p* are replaced with LacO sequences (Fig. 3B; promoter sequence in the

Supplemental Material). We integrated this transgene into the *Drosophila* genome at the same site and asked whether it recruits HLB factors in the presence of the endogenous histone locus. The 12x2xLacO transgene fails to recruit CLAMP (Fig. 3B), indicating that the GA repeat sequences are necessary for CLAMP binding to the *H3-H4p* in the presence of the endogenous histone locus. The 12x2xLacO transgene also failed to recruit Mxc, Lsm10, or Mute to polytene chromosome spreads (Fig. 3B), consistent with our results from the promoter deletion transgene constructs (Supplemental Fig. S3). In contrast to the 12xHistone^{WT} array (McKay et al. 2015), the 12x2xLacO array in which the GA repeats have been replaced with LacO sequences did not express detectable amounts of *Histone2A* mRNA (Fig. 3E). Thus, the GA repeats in the *H3-H4p* are key *cis* elements that promote HLB formation and histone gene expression.

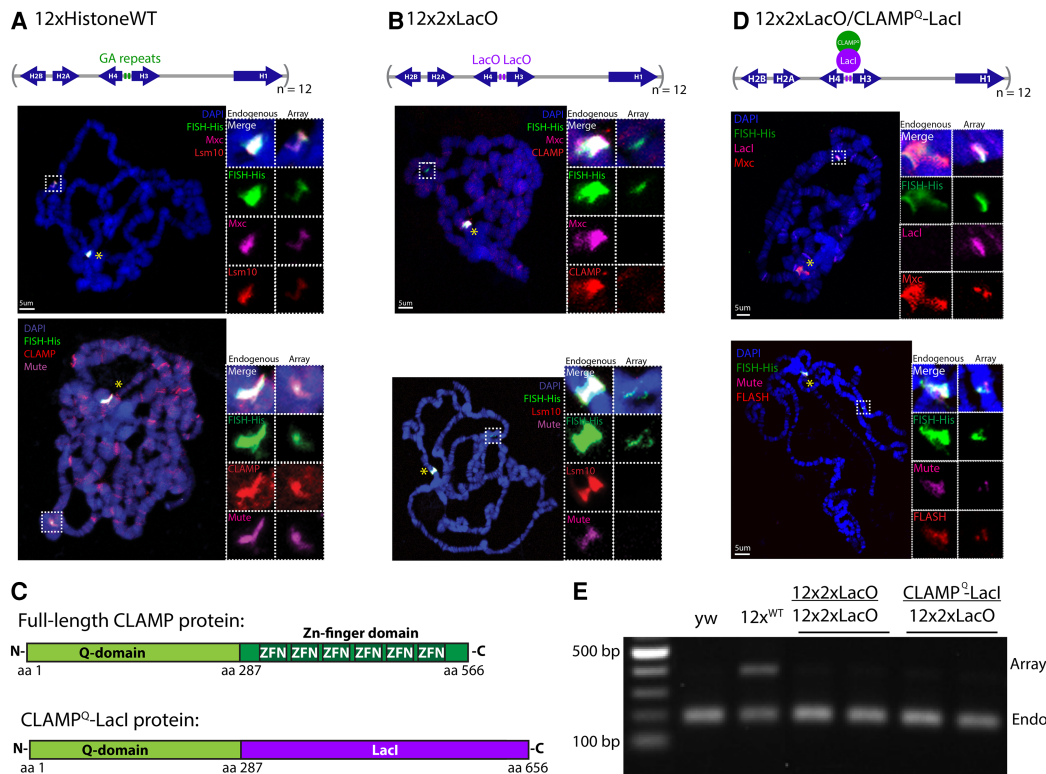


Figure 3. The GA repeats in the *H3-H4p* are required for HLB formation, and retargeted CLAMP recruits HLB factors. We integrated transgenes carrying 12 tandem arrays of the wild-type histone gene cluster (A) or a cluster in which the GA repeats are replaced by two LacO sites (B) into the *Drosophila* genome and scored for recruitment of HLB factors (Mxc, CLAMP, Mute, and Lsm10) by fluorescent in situ hybridization (FISH) against the histone gene cluster (FISH-His; green). The yellow asterisk indicates the endogenous histone locus, while the array locus is boxed. DNA is visualized by DAPI (blue). The sequences of the mutant *H3H4p* are in the Supplemental Material. (C) We designed the CLAMP^Q-LacI fusion protein to include the N-terminal glutamine-rich domain of full-length CLAMP (amino acids 1–287). We replaced the C-terminal zinc finger DNA-binding domain of CLAMP with LacI (purple). (D) We recombined the 12x2xLacO array transgene and the *clamp^Q-lacI* transgene onto the same chromosome, performed FISH-His (green), and assayed recruitment of CLAMP^Q-LacI (detected with anti-LacI antibody; pink) and the HLB components Mxc (red), Mute (pink), and FLASH (red). Yellow asterisks indicate the endogenous histone loci, while array loci are boxed. (E) Expression of *Histone2a* mRNA from endogenous histone locus and transgenic histone gene cluster arrays. We mutated a site within the transgenic *H2a* gene (array) to prevent restriction digestion of the *H2a* cDNA (McKay et al. 2015). *Yellow⁻*, *white⁻* (*yw*) larvae do not carry the transgenic locus and produce only histone transcripts from the endogenous locus (cut; Endo), while animals transgenic for the 12x^{WT} array produce both endogenous and array transcripts (uncut; array). Animals carrying the 12x2xLacO array and animals carrying the 12x2xLacO array and *clamp^Q-lacI* transgene do not express array transcripts.

CLAMP promotes recruitment of HLB components to an ectopic histone locus

We demonstrated previously that CLAMP recognizes GA-rich sequence elements and promotes recruitment of the MSL complex specifically to the male X chromosome (Soruceo et al. 2013; Kuzu et al. 2016). We therefore hypothesized that CLAMP functions similarly at the histone locus, recognizing GA repeat *cis* elements and promoting the recruitment of HLB-specific factors. We performed a tethering experiment using the 12x2xLacO transgenic array described above (Fig. 3B) and a synthetic CLAMP^Q-LacI protein in which the CLAMP zinc finger DNA-binding domain (Larschan et al. 2012) is replaced with LacI (Robinett et al. 1996). The resulting transgene, “CLAMP^Q-LacI,” contains only the N-terminal 287 amino acids of CLAMP, which includes the polyglutamine domain of the CLAMP protein but does not include the DNA-binding domain. We conducted the experiment in this way to avoid the presence of two competing DNA-binding domains on the same protein that could prevent binding to the histone locus (Fig. 3C). We expressed the CLAMP^Q-LacI transgene using a *ubiquitin* promoter and measured expression of CLAMP^Q-LacI (Supplemental Fig. S4A) by Western blot using protein from third instar larval salivary glands and *clamp^q-lacI* mRNA levels by quantitative RT-PCR (qRT-PCR) (Supplemental Fig. S4B) using RNA from whole larvae. From this analysis, we determined that the CLAMP^Q-LacI protein is expressed at higher levels than endogenous CLAMP (Supplemental Fig. S4A).

We performed polytene chromosome immunostaining on animals expressing CLAMP^Q-LacI in the presence of the transgenic 12x2xLacO histone array. We found that CLAMP^Q-LacI as well as Mxc, FLASH, and Mute colocalized with the fluorescent in situ hybridization (FISH) signal against the histone gene repeat when paired with the 12x2xLacO histone array (Fig. 3D). In contrast, in the absence of CLAMP^Q-LacI, the 12x2xLacO array does not recruit HLB factors (Fig. 3B). Therefore, CLAMP^Q-LacI binds the 12x2xLacO array and promotes the recruitment of HLB factors. CLAMP^Q-LacI also localized to several additional genomic locations (Fig. 3D), possibly through dimerization with endogenous CLAMP. However, HLB factors were recruited only to the 12x2xLacO locus after CLAMP^Q-LacI expression. Thus, like endogenous CLAMP, CLAMP^Q-LacI is not sufficient to recruit HLB factors to nonhistone loci, suggesting that the *H3-H4p* may contain other important *cis* elements and/or recruit additional critical components. Therefore, promoting HLB factor recruitment to the histone locus is a context-specific function of the ubiquitously expressed CLAMP protein, similar to its function in promoting the recruitment of the MSL complex to the dosage-compensated X chromosome.

Although multiple HLB factors were recruited to the 12x2xLacO histone array in the presence of CLAMP^Q-LacI, we did not detect any histone mRNA expressed from the array (Fig. 3E). Thus, HLB factor recruitment can be uncoupled from histone locus transcription, as

observed previously in the formation of a “proto-HLB” containing Mxc and FLASH (Salzler et al. 2013). There are several possible explanations for why CLAMP^Q-LacI can recruit HLB factors but not promote transcriptional activation. First, only low levels of HLB factors were recruited to the 12x2xLacO histone array in the presence of CLAMP^Q-LacI (Fig. 3D), compared with the levels recruited to the 12xHistone^{WT} array in the presence of endogenous wild-type CLAMP (Fig. 3A). Therefore, the stoichiometry of CLAMP^Q-LacI compared with each 12x2xLacO sequence may be insufficient to support transcription. Second, a specific conformation of the endogenous CLAMP protein or the zinc finger domain of CLAMP, which is replaced by LacI in the CLAMP^Q-LacI construct (Fig. 3C), may be required for transcriptional activation of the histone genes but not for recruitment of HLB factors. Alternatively, wild-type CLAMP at the endogenous histone locus may recruit critical limiting factors much more efficiently than ectopically localized CLAMP^Q-LacI. Together, our data indicate that the N-terminal 287 amino acids of CLAMP are sufficient for promoting the recruitment of HLB factors to an ectopic chromosomal location but not for transcriptional activation.

CLAMP regulates histone locus chromatin and histone gene expression

Because CLAMP localizes to the histone locus and binds directly to the essential GA repeat *cis* elements, we hypothesized that CLAMP regulates histone locus chromatin and histone gene transcription. All five histone transcripts and the *clamp* transcript are expressed in the oocyte and maternally deposited in 0- to 2-h embryos. Mature HLBs are formed by cycle 11, and the zygotic histone locus is activated by zygotic genome activation (~2 h after egg lay) (White et al. 2007). To determine how CLAMP functions as the HLB forms and the histone locus becomes transcriptionally activated in the early embryo, we depleted maternally deposited CLAMP protein and mRNA by RNAi using a UAS-driven *clamp*-specific shRNA (Ni et al. 2008) and a strong GAL4 driver (maternal triple driver [MTD]) that is expressed in the female germline (Staller et al. 2013). As controls, we also analyzed matched MTD controls. We performed qRT-PCR for histone mRNA levels using primers specific to the *H3* and *H4* transcripts (Bulchand et al. 2010). CLAMP mRNA and protein are nearly completely depleted in 0- to 2-h and 2- to 4-h embryos derived from *clamp* RNAi mothers (Figs. 4A, 5A). We found that in unfertilized eggs and 0- to 2-h embryos laid by *clamp* RNAi mothers, both *H3* and *H4* transcript levels were significantly decreased (Fig. 4A), indicating that CLAMP depletion results in a reduction in the amount of histone mRNA deposited in the egg. Histone transcript levels began to recover in older embryos, likely due to zygotic genome activation of both the embryonic histone locus and the *clamp* locus. Nearly 100% (99.87%) of *clamp* RNAi embryos do not hatch, demonstrating that maternally deposited CLAMP is critical for early development.

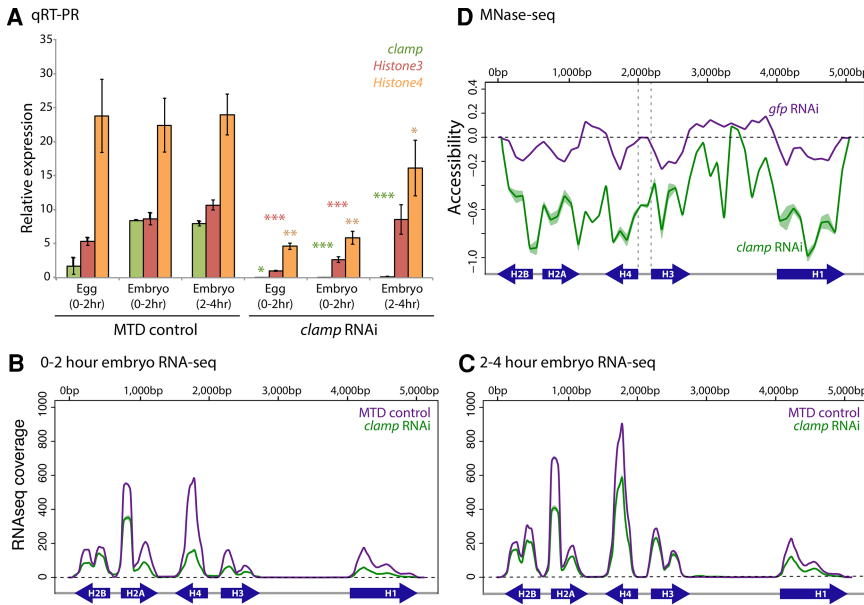


Figure 4. CLAMP regulates histone gene transcription and histone locus chromatin accessibility. (A–C) We performed qRT-PCR and mRNA sequencing (mRNA-seq) using RNA from unfertilized eggs laid by virgin mothers, 0- to 2-h fertilized embryos, and 2- to 4-h fertilized embryos laid by MTD control mothers or *clamp* RNAi mothers. (A) *clamp* (green), *Histone3* (red), and *Histone4* (orange) transcripts are significantly reduced at all stages (with the exception of *Histone3* in 2- to 4-h embryos) after *clamp* RNAi compared with MTD controls, although transcripts begin to recover in 2- to 4-h embryos after zygotic genome activation. Error bars represent \pm one standard deviation from the mean. *clamp* expression was normalized to *pka*, and *Histone3* and *Histone4* expression was normalized to *rp49*. (*) $P < 0.05$; (**) $P < 0.001$; (***) $P < 0.0001$. Western blots for CLAMP protein from 2- to 4-h embryos are shown in Figure 5A. (B) Maternal deposition of all five histone mRNAs is significantly reduced in 0-

to 2-h embryos from *clamp* RNAi mothers (green) compared with control MTD mothers (purple), as assayed by mRNA-seq. (C) *Histone* transcript levels begin to recover in 2- to 4-h embryos. See Supplemental Table S1 for raw FPKM (fragments per kilobase per million mapped fragments) data and *P*-values. (D) We analyzed MNase-seq (micrococcal nuclease [MNase] digestion followed by high-throughput sequencing) data from cultured S2 cells treated with *clamp* RNAi (green) or *gfp* RNAi (control; purple). Accessibility scores >0 were assigned to chromatin that is more open compared with the rest of the genome, while scores <0 were assigned to chromatin that is relatively closed. (B–D) Shading represents 95% confidence intervals.

We measured transcript accumulation from all histone genes by performing mRNA sequencing (mRNA-seq) from the same embryo time points and RNAi conditions that we used in our qRT-PCR assay (Fig. 4A). We found that maternal *clamp* RNAi resulted in decreased expression of all histone genes compared with MTD controls (Fig. 4B,C; Supplemental Table S1), including *H1*, *H2A*, and *H2B*, which are not adjacent to the site of CLAMP binding within the *H3–H4p* (Fig. 1D,E). To determine whether this reduction of histone transcripts might be an indirect effect of CLAMP regulating the expression of known HLB components, we compared the levels of mRNAs encoding known HLB components, including FLASH, Mxc, Mute, Lsm10, and Lsm11 (Duronio and Marzluff 2017), from 0- to 2-h and 2- to 4-h *clamp* RNAi embryos with those from control MTD embryos. The transcript levels of the known HLB factors were not affected by *clamp* RNAi in the early embryo (Supplemental Table S2). Together, these observations suggest that the effects on histone transcript levels that we observed in the early developing embryo after CLAMP depletion (Fig. 4A–C) are not due to misregulation of genes encoding other HLB components but a specific effect of CLAMP depletion on histone mRNA deposition in the egg.

We next sought to understand how CLAMP might regulate histone gene expression. To determine whether CLAMP promotes chromatin accessibility at the histone locus, we performed *clamp* RNAi in cultured male S2 cells and measured chromatin accessibility using a previously described micrococcal nuclease (MNase) digestion followed by high-throughput sequencing

(MNase-seq) titration approach developed by our group (Mieczkowski et al. 2016). Accessible chromatin is digested even under dilute MNase concentrations, while inaccessible chromatin is protected until it is subjected to higher concentrations. Therefore, sequencing libraries generated from chromatin subjected to a low concentration of MNase are enriched for accessible regions, while libraries from chromatin digested with a higher concentration of enzyme are enriched for more inaccessible regions (Mieczkowski et al. 2016).

Using our chromatin accessibility data, we mapped reads to a single copy of the histone gene array (McKay et al. 2015) and calculated chromatin accessibility as described in Mieczkowski et al. (2016). In control S2 cells subjected to *gfp* control RNAi, the most accessible regions of the histone gene cluster included promoter and intergenic regions (Fig. 4D). In cells subjected to *clamp* RNAi, the entire histone gene cluster decreased in accessibility. Overall, our data support the requirement for CLAMP recruitment to the histone locus for both promoting chromatin accessibility across the entire histone locus and activating gene expression of all replication-dependent histone genes (Fig. 4A–D).

CLAMP is specifically retained at the HLB in clamp nulls

The ability of CLAMP to promote HLB factor recruitment to the transgenic array locus (Fig. 3D) and regulate chromatin organization at the histone locus (Fig. 4D) led us to test whether removing CLAMP resulted in loss of endogenous HLB formation in *clamp*-null larvae and in

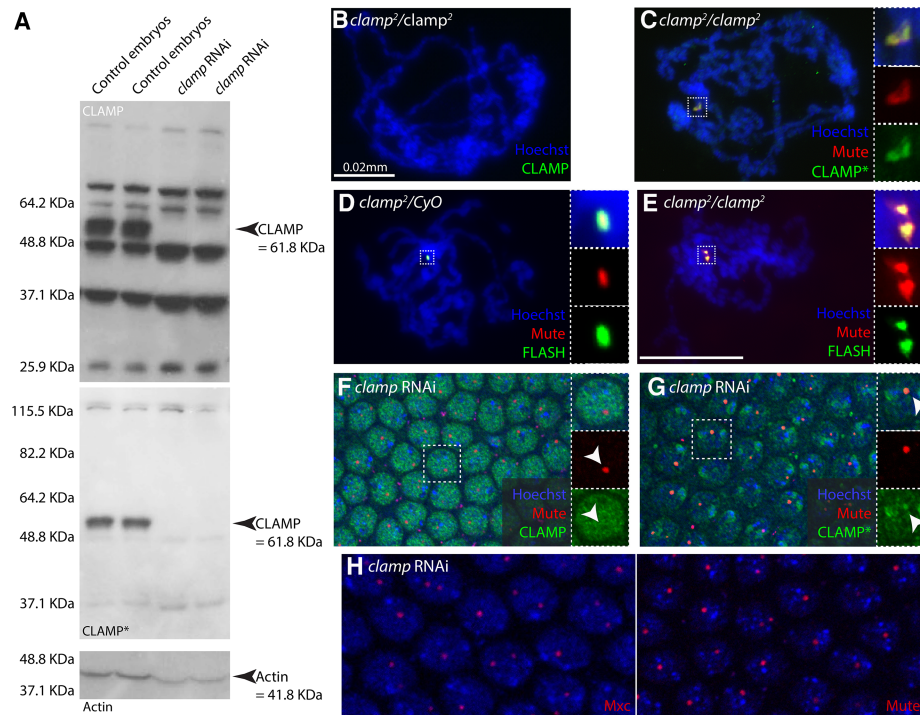


Figure 5. CLAMP remains at the histone locus in CLAMP-depleted larvae and embryos. (A) Western blots demonstrating the antibody specificity and efficacy of *clamp* RNAi in embryos (2–4 h). Control embryos (those expressing the MTD without *clamp* RNAi) have abundant CLAMP protein (61.8 kDa), while CLAMP is undetectable in embryos subjected to *clamp* RNAi. We developed two custom antibodies against CLAMP. The CLAMP antibody (top) (Larschan et al. 2012; Urban et al. 2017) cross-reacts with several bands by Western blot, while the CLAMP* antibody (bottom) has lower cross-reactivity. Note that these Westerns also demonstrate the efficacy of *clamp* RNAi in 2- to 4-h embryos. We stained for Actin as a loading control. (B–E) Larval salivary gland polytene chromosome spreads from salivary glands from homozygous (*clamp*²/*clamp*²; B,C,E) and heterozygous (*clamp*²/*CyO*; D) third instar larvae stained for CLAMP (CLAMP [B] and CLAMP* [C]; both green) and HLB components Mute (red; C–E) and FLASH (green; D,E). HLBs are formed in both *clamp*²/*CyO* controls (D) and *clamp*²/*clamp*² nulls (C,E). (F–H) *clamp*-depleted embryos immunostained for CLAMP (CLAMP [F] and CLAMP* [G]; both green) and the HLB factors Mute (red) and Mxc (red).

embryos in which *clamp* is depleted by RNAi. While almost all embryos laid by mothers with RNAi-reduced *clamp* levels are inviable, ~30% of zygotically mutant animals homozygous for the *clamp*²-null allele survive until the third instar larval stage (Urban et al. 2017). These data indicate that survival of *clamp*² larvae is likely due to maternal deposition of wild-type *clamp* mRNA or protein. We reported previously that *clamp*²-null animals produce undetectable amounts of CLAMP protein by both Western blot of salivary gland tissue and polytene chromosome immunostaining (Urban et al. 2017). However, our previously characterized affinity-purified anti-CLAMP antibody detected many cross-reacting proteins by Western blot (Fig. 5A), and we suspected that it might not be sufficiently sensitive to detect small amounts of CLAMP protein in *clamp*²-null larvae. Therefore, we developed a new affinity-purified anti-CLAMP antibody (“CLAMP*”)—produced using a different approach (see the Materials and Methods)—that is more specific and does not cross-react with other *Drosophila* proteins (Fig. 5A).

Using both anti-CLAMP antibodies, we analyzed CLAMP localization and HLB formation on *clamp*²-null polytene chromosomes from larval salivary glands. In addition, we analyzed early embryos after RNAi depletion

of maternally deposited CLAMP protein and mRNA (Figs. 4A, 5A). While these embryos die, it is still possible to analyze CLAMP localization and HLB formation during the early syncytial stages (0–2 h) of embryonic development. Using the CLAMP antibody, we observed loss of CLAMP from all loci on polytene chromosomes from *clamp*² nulls, as observed previously (Fig. 5B). There were no CLAMP foci in early embryos after CLAMP RNAi, although Mute foci were readily detected (Fig. 5F). However, using the CLAMP* antibody, we detected a modest amount of CLAMP immunostaining, which colocalizes specifically and solely with HLB factors on *clamp*²-null mutant polytenes (Fig. 5C) and in embryos after *clamp* RNAi (Fig. 5G). Therefore, the CLAMP* antibody is more sensitive than our previously published CLAMP antibody for immunofluorescence. Multiple HLB factors continue to colocalize on *clamp*²-null polytene chromosomes, similar to control chromosomes where abundant CLAMP is present (Fig. 5D,E), and in *clamp*-depleted embryos (Fig. 5H). Because zygotic CLAMP protein is not produced in *clamp*² homozygous animals (Urban et al. 2017), we conclude that maternally deposited CLAMP from *clamp*² heterozygous mothers persists at the histone locus throughout development of

the larval salivary gland. Furthermore, the strong MTD does not deplete all maternally deposited *clamp* transcript or protein, and therefore a small amount of CLAMP remains at embryonic HLBs. Some of the ~200 GA repeat CLAMP-binding sites within the histone gene array are likely capable of recruiting the small amount of maternal CLAMP that remains in *clamp*²-null third instar larvae and early *clamp*-depleted embryos.

GAF localizes to the HLB when CLAMP is depleted

Since only trace amounts of maternally deposited CLAMP are present in HLBs in *clamp*²-null larvae and embryos from *clamp* RNAi mothers (Fig. 5C,G), we tested whether CLAMP depletion results in novel recruitment of the GAF protein to some *H3-H4*ps. Consistent with this possibility, we found that GAF colocalized with HLB markers on polytene chromosomes from *clamp*² homozygous larvae (Fig. 6B) and in syncytial embryos after MTD-driven *clamp* RNAi (Fig. 6D) even though GAF was absent from the HLBs when CLAMP is present (Figs. 1C, 6A,C). Previous in vitro experiments suggested that GAF is able to bind the GA-rich *cis* elements in the *H3-H4*p (Gilmour et al. 1989). However, consistent with our current observations, later in vivo studies suggested that GAF is not present at the histone locus (O'Brien et al. 1995; Bhat et al. 1996). Our results suggest that the 61-kDa CLAMP protein (Urban et al. 2017) is the likely ~66-kDa protein bound to the *H3-H4*p in vivo that was identified >25 years ago (Gilmour et al. 1989; Weber and Gilmour 1995).

However, in agreement with previous in vitro observations (Weber and Gilmour 1995), GAF localizes to the histone locus when CLAMP is depleted (Fig. 6B,D), suggesting that it can bind to the *H3-H4*p when there is not sufficient CLAMP to occupy all GA repeat-binding sites. There are several models that could explain the relation-

ship between CLAMP and GAF at the HLB. For example, residual amounts of CLAMP (Fig. 5) may be sufficient to open the chromatin and initiate HLB formation but not stimulate transcription. In this situation, although GAF can now bind unoccupied GA repeats, it is not able to stimulate histone gene transcription to the level that CLAMP does (Fig. 4A–C). A second possibility is that GAF can open chromatin to promote HLB factor recruitment but cannot stimulate transcription at the histone locus to the same level as CLAMP because it cannot interact with the same cofactors as CLAMP. We discuss possible functional relationships between GAF and CLAMP further in the Discussion.

Discussion

NBs provide a means to coordinate gene expression at specific sites in the nucleus. For example, the HLB coordinates histone gene expression to maintain correct nucleosomal stoichiometry. Important open questions include the following: (1) How does the HLB form specifically at the histone locus? (2) How are the histone genes coordinately regulated? Here, we show that two GA repeat arrays in the *H3-H4*p direct HLB assembly and that a GA repeat-binding factor, CLAMP, localizes to this region early in development, where it regulates the histone locus by promoting enhanced chromatin accessibility and expression of all five histone genes.

Regulation of the histone locus

The mature *Drosophila* HLB is formed by cycle 11 during early embryogenesis, prior to widespread zygotic genome activation (White et al. 2011). Activation of histone genes is critical for organismal development and viability (Saget et al. 1998; Godfrey et al. 2006, 2009; White et al. 2007; Gunesdogan et al. 2010). CLAMP is present at the histone locus when the proto-HLB containing Mxc and FLASH is first detectable (Fig. 1E), prior to zygotic histone gene activation (White et al. 2007). As a DNA-binding transcription factor that directly recognizes the GA repeats that promote HLB formation in the presence of the endogenous locus (Fig. 1D,E), CLAMP represents a potential pioneer factor involved in promoting HLB formation and activation of histone gene transcription. Furthermore, CLAMP facilitates chromatin opening across the entire histone locus repeat (Fig. 4D), suggesting that it could recruit additional factors, such as general transcription factors or cofactors necessary for histone gene regulation.

It is likely that CLAMP affects chromatin accessibility at the histone locus through recruitment of a chromatin-remodeling factor. For example, the similar GAF protein recruits the NURF301 chromatin remodeler (Tsukiyama and Wu 1995). In addition, CLAMP may control histone gene transcription (Fig. 4A–C) through modulating chromatin accessibility, or changes in accessibility may be caused by CLAMP directly regulating histone transcription.

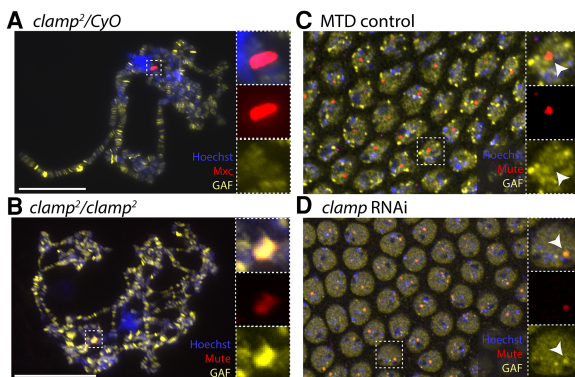


Figure 6. GAF localizes to the HLB when CLAMP is depleted. GAF (yellow) does not localize to the HLB (Mxc or Mute; red) in polytene chromosomes from *clamp*²/*CyO* heterozygous larvae (A) or in syncytial embryos laid by MTD control mothers (C). However, GAF colocalizes with HLB markers in polytene chromosomes from *clamp*² homozygote larvae (B) and in nuclei from syncytial embryos laid by MTD-GAL4/*UAS-clamp* shRNA mothers (D).

Retargeting the polyglutamine domain of CLAMP to a synthetic histone array transgene rescues recruitment of Mxc (Fig. 3E), the core HLB scaffolding protein (Terzo et al. 2015), and several other HLB components. However, CLAMP^Q-LacI is not sufficient to activate transcription (Fig. 3E). It is possible that transcriptional activation requires additional domains of CLAMP that were removed in our CLAMP^Q-LacI fusion protein, a particular conformation of the wild-type CLAMP protein that is not attainable by CLAMP^Q-LacI, or higher CLAMP occupancy than the LacO/LacI system allows. Overall, we demonstrate that a reduction in wild-type CLAMP levels leads to decreased chromatin accessibility and histone gene expression at the endogenous histone locus. Furthermore, artificially tethering CLAMP^Q-LacI to a synthetic histone gene array recruits HLB-specific factors but does not stimulate histone gene expression.

Interestingly, previous work demonstrated that additional HLB components, including Mute and U7 snRNP, are recruited to the HLB only when transcription is initiated (Salzler et al. 2013) and are present at the mature HLB by nuclear cycle 11, when the histone genes are expressed (White et al. 2007, 2011). However, in the present study, we demonstrate that recruitment of Mute can be uncoupled from histone locus transcription (Fig. 3E), perhaps capturing an intermediate stage in HLB formation.

Conservation of HLB formation

Factors that regulate histone gene expression are well conserved, but current evidence suggests that *cis* elements that target these factors are less conserved across species. For example, critical components of the HLB that do not directly interact with DNA are conserved across metazoans, including Mxc/NPAT, Mute/YY1-associated protein, FLASH, U7 snRNP, and Coilin (Duronio and Marzluff 2017). In contrast, CLAMP is highly conserved across *Diptera* (Kuzu et al. 2016), but we did not identify a CLAMP ortholog outside of insects. CLAMP may function together with an early acting *Drosophila*-specific transcription factor, such as the master zygotic genome activator Zelda, which, although conserved among drosophilids, also has no identified mammalian ortholog (Paris et al. 2013).

The mammalian histone genes are more dispersed and there are no repeated units, making it possible that mammalian *cis* elements may be within poorly conserved noncoding sequences. Therefore, the plasticity of the components that link highly conserved HLB factors to the histone locus—and the flexibility of their low-complexity domains—may allow for evolutionary variation among critical *cis*-acting sequences that specify HLB formation. CLAMP binding to GA repeats likely provides this function in *Drosophila*, while other DNA-binding proteins and *cis* elements may nucleate the HLB in other species.

A single transcription factor can mediate multiple domains of coordinated gene activation

We originally identified CLAMP as an essential factor (Urban et al. 2017) that is required to nucleate the forma-

tion of a domain of coordinated gene activation during male X-chromosome dosage compensation (Larschan et al. 2012; Soruco et al. 2013). GA repeat expansion and elevated GA repeat density on the X chromosome evolved to increase CLAMP occupancy and promote MSL recognition (Kuzu et al. 2016). Here we show that CLAMP is also recruited to the histone locus via GA repeats in the *H3-H4p* (Figs. 1D,E, 2B), where it promotes the formation of another domain of coordinated gene activation. In contrast to CLAMP, GAF is not required for male X-chromosome dosage compensation (Greenberg et al. 2004) and is dramatically depleted on the X chromosome compared with autosomes (Soruco et al. 2013). GAF is also not present at the histone locus in wild-type situations (Fig. 1C, D). Therefore, we identified the enrichment of CLAMP relative to GAF as a common feature of two different chromatin domains of enhanced chromatin accessibility and active transcription.

Although CLAMP regulates two very different active chromatin domains to which different specific cofactors are recruited, CLAMP is also a transcription factor that binds to thousands of loci throughout the genome. The mechanism by which a single protein fulfills multiple context-specific roles within the genome is not well understood in any system. It is likely that synergy with additional cofactors, such as the MSL complex on the dosage-compensated X chromosome (Soruco et al. 2013) and Mxc at the histone locus, are critical for this process.

In contrast to the histone locus, where HLB-specific factors are conserved and *cis* sequences are not, the MSL dosage compensation complex components are very poorly conserved even within drosophilids, compared with CLAMP. However, X-enriched GA-rich *cis* elements are well conserved across insect species (Kuzu et al. 2016). Because CLAMP is less conserved across species compared with HLB-specific factors, it is possible that it provides the plasticity required to couple conserved regulators to rapidly evolving *cis* elements at the histone locus. In contrast, on the X chromosome, CLAMP functions as a relatively well-conserved factor (compared with the poorly conserved MSL complex) that couples conserved *cis* elements to diverged cofactors. It is possible that the ancient evolutionary role of CLAMP is as a GA-binding transcription factor, while the context-specific roles of CLAMP at the histone locus and dosage-compensated X chromosome evolved more recently, after GA-rich *cis* elements became enriched at these locations to promote domains of coordinated gene activation. However, significant further study is required to fully understand how a single protein promotes the formation of multiple active chromatin domains.

The relationship between GA-binding factors at the histone locus

GAF is a ubiquitously expressed transcription factor that binds GA repeats genome-wide and functions as a transcriptional regulator (Horard et al. 2000; Fuda et al. 2015), chromatin modifier (Tsukiyama and Wu 1995), and insulator (Ohtsuki and Levine 1998), depending on

the context. Although GAF is excluded from the histone locus in wild-type situations, it localizes to the histone locus when CLAMP is depleted (Fig. 6).

There are several possible models for the relationship between CLAMP and GAF at the HLB: (1) CLAMP and GAF could be partially redundant such that GAF can compensate for some functions normally accomplished by CLAMP, including opening chromatin to minimally recruit HLB factors. However, GAF is not sufficient to fully rescue histone transcription levels when CLAMP is depleted (Fig. 4A–C). (2) CLAMP and GAF could have an antagonistic relationship at the histone locus such that a major function for CLAMP could be to exclude GAF. For example, GAF mislocalization to the histone locus may close chromatin and repress histone gene expression through its well-documented relationship with the repressive Polycomb complex (Horard et al. 2000). (3) Both CLAMP and GAF contain low-complexity glutamine-rich domains (Tariq et al. 2013) that are thought to modulate transcriptional activity (Gemayel et al. 2015) and are also frequently found in a class of proteins thought to undergo biophysical phase transitions to drive NB formation (Zhu and Brangwynne 2015). Therefore, the relative occupancy of CLAMP compared with GAF within a domain may promote or hinder NB formation. (4) GAF binds to unoccupied histone locus GA repeats opportunistically and does not function redundantly or antagonistically with CLAMP at the histone locus. In this model, CLAMP depletion drives the changes that we observed in histone locus regulation (Fig. 4).

Since we were not able to completely remove CLAMP even in the salivary glands of *clamp*-null mutant larvae, we conclude that CLAMP regulates the histone locus (Fig. 5C,G) but cannot conclude that CLAMP is essential for endogenous HLB formation. In addition, in agreement with previous studies on GAF autoregulation (Bernues et al. 2007), we found that the GAF-encoding mRNA *trithorax-like* is resistant to depletion by embryonic RNAi using the shRNA system. Therefore, further experimentation is required to determine the functional relationship between CLAMP and GAF at the HLB.

Here, we determined that the same critical *cis* elements (GA repeats) and *trans*-acting factor (CLAMP) and the absence of a second factor that binds similar *cis* elements (GAF) promote the formation of two different domains of coordinated gene activation. Furthermore, both at the histone locus and on the dosage-compensated male X chromosome, CLAMP is required to promote a domain of enhanced chromatin accessibility that extends well beyond its binding sites. Overall, our study provides new insight into the common regulatory mechanisms shared by two different domains of coordinated gene activation.

Materials and methods

Drosophila strains

We used the MTD (Bloomington, #31777) and a stock expressing a shRNA against *clamp* (Bloomington, #57008) made by the Transgenic RNAi Project (TRiP). For the *H3–H4p* deletion exper-

iments, we inserted promoter sequences into the pMulti-BAC vector containing a single histone repeat unit (McKay et al. 2015) and inserted these transgenes into site 86Fb on chromosome 3 using ϕ C31-mediated integration (Bestgene) (Groth et al. 2004). The full sequences of engineered *H3–H4p* deletion sequences are in the Supplemental Material. For the LacO array experiments, we synthesized *H3–H4ps* (Genescript) and used restriction digest cloning to insert the promoter containing LacO sequences in place of the wild-type promoter in the single histone repeat unit. We built an array of 12 histone repeat units in pMulti-BAC for each transgenic promoter and integrated each into site VK33 on chromosome 3 using ϕ C31-mediated integration (Model Systems Injections). The full sequences of engineered *H3–H4p* with LacO sequences are in the Supplemental Material. We inserted CLAMP^Q-LacI and LacI into the pUbi-GFP (gift from Mark Peifer), in which we swapped LacI for GFP using the LacI-HP1a vector (gift from Lori Wallrath). We amplified the CLAMP polyglutamine domain for Gibson assembly (New England Biolabs) using the primers F (5'-TAGGTCCTG TTCATTGAATGGAAGACCTTACCAAAAAC-3') and R (5'-GTTACTGGTTTCACCATAGCCACAATTTGCTGAAG-3'). We drove transcription of both CLAMP^Q-LacI and LacI genes using the ubiquitin promoter and integrated these transgenes into site VK20 on chromosome 3 using ϕ C31-mediated integration (Genetivision). To make GFP-CLAMP, we cloned *clamp* cDNA into a vector containing the *ubiquitin* promoter (pUbi-GFP; gift from Mark Peifer) and integrated the transgene into site VK33 on chromosome 3.

Promoter alignment

We obtained promoter sequences from *D. simulans* (DNA Data Bank of Japan [DDBJ] accession no. AB055959) (Tsunemoto and Matsuo 2001), *Drosophila erecta* (DDBJ accession no. AB073634) (Kakita et al. 2003), *Drosophila pseudoobscura* (DDBJ accession no. AB249651) (Nakashima et al. 2016), and *D. virilis*. We aligned sequences using T-Coffee (Notredame et al. 2000) and formatted the alignment using BoxShade.

FISH and immunofluorescence

We used primary antibodies at the following concentrations: rabbit anti-CLAMP (1:1000; Novus/SDIX) (Larschan et al. 2012), rabbit anti-CLAMP* (1:1000; custom antibody generated by our laboratory through a contract to Abcam; both anti-CLAMP antibodies were raised against the same N-terminal amino acids, CLAMP#22–121), guinea pig anti-Mxc (1:2000) (White et al. 2011), guinea pig anti-Mute (1:5000) (Bulchand et al. 2010), rabbit anti-C terminus FLASH (1:2000) (Yang et al. 2009), rabbit anti-Lsm10 (1:1000), mouse anti-MPM-2 (1:100; Millipore), rabbit anti-GAF (1:1000; gift from Giacomo Cavalli), mouse anti-LacI (1:1000; Millipore), and chicken anti-GFP (1:400; Life Technologies). We used Alexa fluor secondary antibodies (Thermo Fisher Scientific) at a concentration of 1:1000. We detected in situ probes using 15 μ g/mL streptavidin-DyLight-488 (Vector Laboratories). To make the FISH probe, we made a PCR product that spanned all five histone genes using a wild-type histone repeat in pMulti-BAC as the template (primers F [AAAGGAGGTTGGTAGGCAGC] and R [ACGCTAGCGCTTATCTGCA]) (McKay et al. 2015). We made biotinylated FISH probes by nick translation using the purified PCR product: 1 μ g of purified PCR product was incubated for 2 h at 15°C in a total of 50 μ L containing 1 \times DNA Pol I buffer (Fisher Optizyme); 0.05 mM each dCTP, dATP, and dGTP; 0.05 mM biotin-11-dUTP (Thermo Scientific); 10 mM 2-mercaptoethanol; 0.004 U of DNaseI (Fisher Optizyme); and 10 U of

DNA Pol I (Fisher Optizyme). The reaction was purified on a PCR purification column (Thermo Scientific) and diluted in hybridization buffer (2× SSC, 10% dextran sulfate, 50% formamide, 0.8 mg/mL salmon sperm DNA) to a final volume of 220 μ L. We performed FISH according to Grimaud et al. (2005) except that we added hybridization mixture with the probe to the slide before heating. We added a coverslip, sealed it with rubber cement, and heated the slide for 2 min on a 91°C heat block.

We obtained embryos (mixed sex) by mating virgin females, aged 3–4 d, of either genotype (1) homozygous MTD or (2) MTD crossed to Bloomington #57008 for *clamp* RNAi with *w¹¹¹⁸* males. Embryos were fixed in 3.7% formaldehyde in PBS with an equal volume of n-heptane for 20 min, immunostained using the above antibody concentrations, mounted using Prolong Diamond anti-fade reagent (Thermo Fisher), and imaged on a Zeiss laser scanning 510 or 800 confocal microscope equipped with a 63×/1.4 oil immersion plan apochromat objective and Zen software.

We performed polytene chromosome squashes from salivary glands of mixed sex larvae. We passed glands through fix 1 (1.5% formaldehyde, 1% Triton X-100, in 1× PBS) for 1 min, fix 2 (1.5% formaldehyde, 50% glacial acetic acid) for 2 min, and 1:2:3 solution (ratio of lactic acid:water:glacial acetic acid) for 5 min prior to squashing and spreading. Slides were immunostained using the antibody concentrations above and mounted using Prolong Diamond anti-fade reagent (Thermo Fisher), and spreads were imaged on a Zeiss Imager.M1 using a 40×/0.75 plan neofluar objective and AxioVision software.

Western blotting

We conducted Western blotting as in Urban et al. (2017). We collected 2- to 4-h embryos of the relevant genotypes (at least 150 per sample) on grape juice agar plates and washed them briefly with 1× PBS in a cell strainer basket. We dechlorinated them for 2 min in 50% bleach and then washed them with several milliliters of PBS before transferring them to lysis buffer (50 mM Tris-HCl at pH 8.0, 150 mM NaCl, 1% SDS, 0.5× protease inhibitor). For salivary glands, we dissected glands from third instar larvae ($n = 10$ per sample) in cold PBS and froze samples in liquid nitrogen. We extracted total protein from samples by homogenizing the samples in cold lysis buffer using a small pestle. We cleared the samples by centrifuging at 14,000g for 10 min at room temperature. To blot for CLAMP and Actin, we ran 20 μ g of total protein on a Novex 4%–12% Tris-glycine precast gradient gel (Life Technologies). We transferred proteins to PVDF membranes using the iBlot transfer system (Thermo Fisher Scientific) and probed the membranes for CLAMP (rabbit anti-CLAMP, 1:1000; rabbit anti-CLAMP*, 1:1000) and Actin (mouse anti-Actin, 1:400,000; Millipore) using the Western Breeze kit, following the manufacturer's protocol (Thermo Fisher Scientific).

MNase-seq

We maintained S2 cells in standard Schneider's medium (Gibco) with 10% heat-inactivated fetal bovine serum (Thermo Fisher Scientific). We performed RNAi as in Soruco et al. (2013). We performed and analyzed MNase-seq data as in Mieczkowski et al. (2016). We mapped reads to the custom histone locus genome (McKay et al. 2015) using Bowtie aligner with the parameters “-M 5 -k 1 -I 50 -X 500 --solexa-quals --best --chunkmbs 256” (Langmead et al. 2009). We identified genomic positions with abnormally high numbers of mapped reads (Z -score = 7) and discarded tags mapped to such positions. We computed read frequencies in 100-bp nonoverlapping bins and normalized for the library

size. We calculated MNase accessibility (MACC) values for each bin by fitting linear regression on the normalized read frequencies computed for each titration point (1.5-, 6.25-, 25-, and 100-U MNase concentrations). We used log scale for the MNase concentrations in the fitting procedure. We applied the GC content correction to obtain the final accessibility scores (MACC values). The chromatin accessibility data are available at NCBI Gene Expression Omnibus (GEO) with series number GSE99894.

Quantitative real-time PCR

We conducted qRT-PCR as described in Urban et al. (2017) using the embryo RNA obtained for mRNA-seq (below) as well as RNA extracted from unfertilized oocytes laid by unmated mothers and collected 0–2 h after egg lay. We used four biological replicates for each genotype and time point. Primers for histone transcripts *H3* and *H4* and the normalization gene *rp49* are listed in Bulchand et al. (2010), while *clamp* and *pka* primers can be found in Urban et al. (2017). We normalized histone transcript abundance against *rp49* and *clamp* transcript abundance against *pka*. We analyzed data using a Student's *t*-test, comparing transcript abundance between *clamp* RNAi embryos or oocytes and matched MTD control embryos or oocytes.

Embryo mRNA-seq

We used embryo RNA collected for qRT-PCR (above). As in Wood et al. (2016), we used 100 ng of total RNA as input for the Ovation Universal RNA-seq kit with *Drosophila* rRNA depletion module (NuGEN). We sequenced libraries on an Illumina HiSeq 2500 in 1 × 50-bp mode. We used at least four individually isolated biological replicates for each time point and RNAi condition. We mapped reads using TopHat version 2.0.13 with default parameters (Trapnell et al. 2009) and counted fragments mapping to histone gene exons (see Supplemental Table S1 for all FPKM [fragments per kilobase per million mapped fragments] values for histone genes). See Supplemental Table S2 for a list of significantly affected genes. The mRNA-seq data are available at NCBI GEO with series number GSE102922.

Staged embryo ChIP-seq

To obtain female embryos, we mated +; SD72/*CyO* females to 19-3, *yw*, *Rsp[s]-B[s]/Dp(2:y)CB25-4*, *y+*, *Rsp[s]B[s]*; *SPSD/CyO* males (both kind gifts from Cynthia Staber) to obtain +/*Dp(2:y)CB25-4*, *y+*, *Rsp[s]B[s]*; *SPSD/SD72* males, which we then mated to *yw*; attP2[*PCNA-EGFP*] females (kind gift from Shelby Blythe). We performed 0- to 4-h timed lays and collected and fixed embryos according to Blythe and Wieschaus (2015). We then hand-sorted embryos using a Zeiss Discovery.V8 microscope under GFP excitation using an X-CITE 120Q stereo light source. We pooled 200 (NC 11–14) to 400 (NC < 11) embryos and performed ChIP as in Blythe and Wieschaus (2015) using 3 μ L of rabbit anti-CLAMP antibody per sample. We synthesized libraries using the NEBNext ChIP-seq kit (New England Biosystems) and sequenced libraries on an Illumina HiSeq 2500 in 2 × 100-bp mode. We mapped CLAMP ChIP-seq reads to the custom histone locus genome (McKay et al. 2015), allowing only unique alignments by using Bowtie aligner (Langmead et al. 2009).

Analysis of H2a expression from the ectopic array

We isolated total RNA from larvae ($n = 10$) of the indicated genotypes by flash freezing samples in liquid nitrogen and

homogenizing them with a steel bead using a Retsch MM300 TissueLyser Mixer Mill. We then performed phenol/chloroform (Invitrogen) total RNA extractions. We reverse-transcribed 1 µg of total RNA using the SuperScript II kit (Invitrogen) using random primers according to the manufacturer's instructions. We analyzed expression from the transgenic array *Histone2a* genes via PCR and restriction digestion as in McKay et al. (2015).

Acknowledgments

We thank Victoria Cabot for initial observations; Shelby Blythe for the PCNA-EGFP fly stock and early embryo ChIP-seq protocol; Cindi Staber for SD and *Rsp* fly stocks; members of the Anne Hart and Casey Dunn laboratories for use of their fluorescent dissecting microscopes; Judith Bender, Arjun Ray, Jason Wood, and Steve Criscione for advice on experimental analyses that did not end up in the final manuscript; Daniel McKay for the custom histone array genome; Deirdre Tatomer for several published fly stocks; and Geoff Williams for microscopy advice. We thank Joseph Gall for an anti-Coilin antibody, Giacomo Cavalli for the anti-GAF antibody, Jacob Bliss for experimental conversations, and Brian Jones and Judith Bender for critical comments on the manuscript. We thank the *Drosophila* Genomics Resource Center, the Transgenic RNAi Project (TRiP), and the Bloomington Stock Center for fly lines. Finally, we thank anonymous reviewers for their constructive feedback. L.E.R. is supported by National Institutes of Health (NIH) grant F32 GM109663, and K.A.B. was supported by NIH training grant T32 CA009156. All experiments were supported by NIH grants GM098461 (E.N.L.) and GM058921 (W.F.M. and R.J.D.), American Cancer Society Research Scholar grant 123682-RSG-13-040-01-DMC (E.N.L.), and a Pew Biomedical Scholars program grant (E.N.L.). L.E.R., K.P.K., K.A.B., E.N.L., R.J.D., and W.F.M. conceived the study. L.E.R., K.P.K., K.A.B., J.A.U., S.B., and A.Z. performed the investigations. L.E.R., G.K., and M.Y.T. performed the formal analysis. L.E.R. wrote the original draft of the manuscript, and L.E.R., E.N.L., R.J.D., and W.F.M. reviewed and edited the manuscript. L.E.R., K.P.K., and K.A.B. performed visualization. M.Y.T. and W.T.J. supervised the study. L.E.R., K.A.B., E.N.L., R.J.D., and W.F.M. acquired the funding.

References

Albig W, Doenecke D. 1997. The human histone gene cluster at the D6S105 locus. *Hum Genet* **101**: 284–294.

Belote JM, Lucchesi JC. 1980. Control of X chromosome transcription by the maleless gene in *Drosophila*. *Nature* **285**: 573–575.

Berlaco M, Fanti L, Breiling A, Orlando V, Pimpinelli S. 2001. The maternal effect gene, abnormal oocyte (abo), of *Drosophila melanogaster* encodes a specific negative regulator of histones. *Proc Natl Acad Sci* **98**: 12126–12131.

Bernues J, Pineyro D, Kosoy A. 2007. General, negative feedback mechanism for regulation of Trithorax-like gene expression in vivo: new roles for GAGA factor in flies. *Nucleic Acids Res* **35**: 7150–7159.

Bhat KM, Farkas G, Karch F, Gyurkovics H, Gausz J, Schedl P. 1996. The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division. *Development* **122**: 1113–1124.

Blythe SA, Wieschaus EF. 2015. Zygotic genome activation triggers the DNA replication checkpoint at the midblastula transition. *Cell* **160**: 1169–1181.

Bulchand S, Menon SD, George SE, Chia W. 2010. Muscle wasted: a novel component of the *Drosophila* histone locus body required for muscle integrity. *J Cell Sci* **123**: 2697–2707.

Burch BD, Godfrey AC, Gasdaska PY, Salzler HR, Duronio RJ, Marzluff WF, Dominski Z. 2011. Interaction between FLASH and Lsm11 is essential for histone pre-mRNA processing in vivo in *Drosophila*. *RNA* **17**: 1132–1147.

Chen K, Johnston J, Shao W, Meier S, Staber C, Zeitlinger J. 2013. A global change in RNA polymerase II pausing during the *Drosophila* midblastula transition. *Elife* **2**: e00861.

Duronio RJ, Marzluff WF. 2017. Coordinating cell cycle-regulated histone gene expression through assembly and function of the histone locus body. *RNA Biol* **14**: 726–738.

Eagen KP, Aiden EL, Kornberg RD. 2017. Polycomb-mediated chromatin loops revealed by a sub-kilobase resolution chromatin interaction map. bioRxiv doi: 10.1101/099804.

Eriksson PR, Ganguli D, Nagarajavel V, Clark DJ. 2012. Regulation of histone gene expression in budding yeast. *Genetics* **191**: 7–20.

Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT. 2015. GAGA factor maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters. *PLoS Genet* **11**: e1005108.

Gemayel R, Chavali S, Pougach K, Legendre M, Zhu B, Boeynaems S, van der Zande E, Gevaert K, Rousseau F, Schymkowitz J, et al. 2015. Variable glutamine-rich repeats modulate transcription factor activity. *Mol Cell* **59**: 615–627.

Gilmour DS, Thomas GH, Elgin SC. 1989. *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science* **245**: 1487–1490.

Godfrey AC, Kupsco JM, Burch BD, Zimmerman RM, Dominski Z, Marzluff WF, Duronio RJ. 2006. U7 snRNA mutations in *Drosophila* block histone pre-mRNA processing and disrupt oogenesis. *RNA* **12**: 396–409.

Godfrey AC, White AE, Tatomer DC, Marzluff WF, Duronio RJ. 2009. The *Drosophila* U7 snRNP proteins Lsm10 and Lsm11 are required for histone pre-mRNA processing and play an essential role in development. *RNA* **15**: 1661–1672.

Greenberg AJ, Yanowitz JL, Schedl P. 2004. The *Drosophila* GAGA factor is required for dosage compensation in males and for the formation of the male-specific-lethal complex chromatin entry site at 12DE. *Genetics* **166**: 279–289.

Grimaud C, Bantignies F, Cavalli G. 2005. Fluorescent in situ hybridization combined with immunostaining on polytene chromosomes (Prot 4). <https://www.epigenesys.eu/en/resources/protocols/fluorescence-microscopy/183-fluorescent-in-situ-hybridization-combined-with-immunostaining-on-polytene-chromosomes>.

Groth AC, Fish M, Nusse R, Calos MP. 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* **166**: 1775–1782.

Guertin MJ, Martins AL, Siepel A, Lis JT. 2012. Accurate prediction of inducible transcription factor binding intensities in vivo. *PLoS Genet* **8**: e1002610.

Gunesdogan U, Jackle H, 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.

Hamada FN, Park PJ, Gordadze PR, Kuroda MI. 2005. Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev* **19**: 2289–2294.

Horard B, Tatout C, Poux S, Pirrotta V. 2000. Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol Cell Biol* **20**: 3187–3197.

- Kakita M, Shimizu T, Emoto M, Nagai M, Takeguchi M, Hosono Y, Kume N, Ozawa T, Ueda M, Bhuiyan MS, et al. 2003. Divergence and heterogeneity of the histone gene repeating units in the *Drosophila melanogaster* species subgroup. *Genes Genet Syst* **78**: 383–389.
- Kasinathan S, Orsi GA, Zentner GE, Ahmad K, Henikoff S. 2014. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat Methods* **11**: 203–209.
- Kuzu G, Kaye EG, Chery J, Siggers T, Yang L, Dobson JR, Boor S, Bliss J, Liu W, Jogle G, et al. 2016. Expansion of GA dinucleotide repeats increases the density of CLAMP binding sites on the X-chromosome to promote *Drosophila* dosage compensation. *PLoS Genet* **12**: e1006120.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Larschan E, Soruco MM, Lee OK, Peng S, Bishop E, Chery J, Goebel K, Feng J, Park PJ, Kuroda MI. 2012. Identification of chromatin-associated regulators of MSL complex targeting in *Drosophila* dosage compensation. *PLoS Genet* **8**: e1002830.
- Lifton RP, Goldberg ML, Karp RW, Hogness DS. 1978. The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harb Symp Quant Biol* **42**: 1047–1051.
- Liu JL, Murphy C, Buszczak M, Clatterbuck S, Goodman R, Gall JG. 2006. The *Drosophila melanogaster* Cajal body. *J Cell Biol* **172**: 875–884.
- Ma T, Van Tine BA, Wei Y, Garrett MD, Nelson D, Adams PD, Wang J, Qin J, Chow LT, Harper JW. 2000. Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev* **14**: 2298–2313.
- Mao YS, Zhang B, Spector DL. 2011. Biogenesis and function of nuclear bodies. *Trends Genet* **27**: 295–306.
- Marzluff WF, Gongidi P, Woods KR, Jin J, Maltais LJ. 2002. The human and mouse replication-dependent histone genes. *Genomics* **80**: 487–498.
- Marzluff WF, Wagner EJ, Duronio RJ. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* **9**: 843–854.
- Matera AG, Izaguirre-Sierra M, Praveen K, Rajendra TK. 2009. Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev Cell* **17**: 639–647.
- McKay DJ, Klusza S, Penke TJ, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW, Tatomer DC, Lieb JD, et al. 2015. Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell* **32**: 373–386.
- Mieczkowski J, Cook A, Bowman SK, Mueller B, Alver BH, Kundu S, Deaton AM, Urban JA, Larschan E, Park PJ, et al. 2016. MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat Commun* **7**: 11485.
- Miele A, Braastad CD, Holmes WF, Mitra P, Medina R, Xie R, Zaidi SK, Ye X, Wei Y, Harper JW, et al. 2005. HiNF-P directly links the cyclin E/CDK2/p220NPAT pathway to histone H4 gene regulation at the G1/S phase cell cycle transition. *Mol Cell Biol* **25**: 6140–6153.
- Mowry KL, Steitz JA. 1987. Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. *Science* **238**: 1682–1687.
- Nakashima Y, Higashiyama A, Ushimaru A, Nagoda N, Matsuo Y. 2016. Evolution of GC content in the histone gene repeating units from *Drosophila lutescens*, *D. takahashii* and *D. pseudoobscura*. *Genes Genet Syst* **91**: 27–36.
- Ni JQ, Markstein M, Binari R, Pfeiffer B, Liu LP, Villalta C, Booker M, Perkins L, Perrimon N. 2008. Vector and parameters for targeted transgenic RNA interference in *Drosophila melanogaster*. *Nat Methods* **5**: 49–51.
- Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**: 205–217.
- O'Brien T, Wilkins RC, Giardina C, Lis JT. 1995. Distribution of GAGA protein on *Drosophila* genes in vivo. *Genes Dev* **9**: 1098–1110.
- Ohtsuki S, Levine M. 1998. GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. *Genes Dev* **12**: 3325–3330.
- Paris M, Kaplan T, Li XY, Villalta JE, Lott SE, Eisen MB. 2013. Extensive divergence of transcription factor binding in *Drosophila* embryos with highly conserved gene expression. *PLoS Genet* **9**: e1003748.
- Quinn JJ, Ilik IA, Qu K, Georgiev P, Chu C, Akhtar A, Chang HY. 2014. Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat Biotechnol* **32**: 933–940.
- Robinett CC, Straight A, Li G, Willhelm C, Sudlow G, Murray A, Belmont AS. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* **135**: 1685–1700.
- Russo CA, Takezaki N, Nei M. 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol* **12**: 391–404.
- Saget O, Forquignon F, Santamaria P, Randsholt NB. 1998. Needs and targets for the multi sex combs gene product in *Drosophila melanogaster*. *Genetics* **149**: 1823–1838.
- Salzler HR, Tatomer DC, Malek PY, McDaniel SL, Orlando AN, Marzluff WF, Duronio RJ. 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 JE, Lozovskaya ER, Strausbaugh LD. 1998. *Drosophila virilis* has atypical kinds and arrangements of histone repeats. *Chromosoma* **107**: 529–539.
- Smith MM, Murray K. 1983. Yeast H3 and H4 histone messenger RNAs are transcribed from two non-allelic gene sets. *J Mol Biol* **169**: 641–661.
- Soruco MM, Chery J, Bishop EP, Siggers T, Tolstorukov MY, Leydon AR, Sugden AU, Goebel K, Feng J, Xia P, et al. 2013. The CLAMP protein links the MSL complex to the X chromosome during *Drosophila* dosage compensation. *Genes Dev* **27**: 1551–1556.
- Staller MV, Yan D, Randklev S, Bragdon MD, Wunderlich ZB, Tao R, Perkins LA, Depace AH, Perrimon N. 2013. Depleting gene activities in early *Drosophila* embryos with the 'maternal-Gal4-shRNA' system. *Genetics* **193**: 51–61.
- Strub K, Birnstiel ML. 1986. Genetic complementation in the *Xenopus* oocyte: co-expression of sea urchin histone and U7 RNAs restores 3' processing of H3 pre-mRNA in the oocyte. *EMBO J* **5**: 1675–1682.
- Tariq M, Wegrzyn R, Anwar S, Bukau B, Paro R. 2013. *Drosophila* GAGA factor polyglutamine domains exhibit prion-like behavior. *BMC Genomics* **14**: 374.
- Tatomer DC, Terzo E, Curry KP, Salzler H, Sabath I, Zapotoczny G, McKay DJ, Dominski Z, Marzluff WF, Duronio RJ. 2016. Concentrating pre-mRNA processing factors in the histone locus body facilitates efficient histone mRNA biogenesis. *J Cell Biol* **213**: 557–570.
- Terzo EA, Lyons SM, Poulton JS, Temple BR, Marzluff WF, Duronio RJ. 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.
- Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**: 1105–1111.
- Tsai SY, Chang YL, Swamy KB, Chiang RL, Huang DH. 2016. GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. *Epigenetics Chromatin* **9**: 32.
- Tsukiyama T, Wu C. 1995. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell* **83**: 1011–1020.
- Tsunemoto K, Matsuo Y. 2001. Molecular evolutionary analysis of a histone gene repeating unit from *Drosophila* simulans. *Genes Genet Syst* **76**: 355–361.
- Urban JA, Doherty CA, Jordan WT III, Bliss JE, Feng J, Soruco MM, Rieder LE, Tsiarli MA, Larschan EN. 2017. The essential *Drosophila* CLAMP protein differentially regulates non-coding roX RNAs in male and females. *Chromosome Res* **25**: 101–113.
- Weber JA, Gilmour DS. 1995. Genomic footprinting of the hsp70 and histone H3 promoters in *Drosophila* embryos reveals novel protein-DNA interactions. *Nucleic Acids Res* **23**: 3327–3334.
- Wei Y, Jin J, Harper JW. 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 AE, Leslie ME, Calvi BR, Marzluff WF, Duronio RJ. 2007. Developmental and cell cycle regulation of the *Drosophila* histone locus body. *Mol Biol Cell* **18**: 2491–2502.
- White AE, Burch BD, Yang XC, Gasdaska PY, Dominski Z, Marzluff WF, Duronio RJ. 2011. *Drosophila* histone locus bodies form by hierarchical recruitment of components. *J Cell Biol* **193**: 677–694.
- Wood JG, Jones BC, Jiang N, Chang C, Hosier S, Wickremesinghe P, Garcia M, Hartnett DA, Burhenn L, Neretti N, et al. 2016. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in *Drosophila*. *Proc Natl Acad Sci* **113**: 11277–11282.
- Yang XC, Burch BD, Yan Y, Marzluff WF, 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.
- Yang XC, Sabath I, Kunduru L, van Wijnen AJ, Marzluff WF, Dominski Z. 2014. A conserved interaction that is essential for the biogenesis of histone locus bodies. *J Biol Chem* **289**: 33767–33782.
- Ye X, Wei Y, Nalepa G, Harper JW. 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.
- Zhao J, Kennedy BK, Lawrence BD, Barbie DA, Matera AG, Fletcher JA, Harlow E. 2000. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev* **14**: 2283–2297.
- Zhu L, Brangwynne CP. 2015. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol* **34**: 23–30.