Drosophila histone locus body assembly and function involves multiple interactions

Kaitlin P. Koreski[a], Leila E. Rieder[b], Lyndsey M. McLain[c], Ashlesha Chaubal[d], William F. Marzluff[a,c,d,e,f,*], and Robert J. Duronio[a,c,d,f,g,*]

[a]Curriculum in Genetics and Molecular Biology, [c]Department of Biology, [d]Integrative Program for Biological and Genome Sciences, [e]Department of Biochemistry and Biophysics, [f]Lineberger Comprehensive Cancer Center, and [g]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599; [b]Department of Biology, Emory University, Atlanta, GA 30322

ABSTRACT **The histone locus body (HLB) assembles at replication-dependent (RD) histone loci and concentrates factors required for RD histone mRNA biosynthesis. The *Drosophila melanogaster* genome has a single locus comprised of ~100 copies of a tandemly arrayed 5-kB repeat unit containing one copy of each of the 5 RD histone genes. To determine sequence elements required for *D. melanogaster* HLB formation and histone gene expression, we used transgenic gene arrays containing 12 copies of the histone repeat unit that functionally complement loss of the ~200 endogenous RD histone genes. A 12x histone gene array in which all *H3-H4* promoters were replaced with *H2a-H2b* promoters (12x^{PR}) does not form an HLB or express high levels of RD histone mRNA in the presence of the endogenous histone genes. In contrast, this same transgenic array is active in HLB assembly and RD histone gene expression in the absence of the endogenous RD histone genes and rescues the lethality caused by homozygous deletion of the RD histone locus. The HLB formed in the absence of endogenous RD histone genes on the mutant 12x array contains all known factors present in the wild-type HLB including CLAMP, which normally binds to GAGA repeats in the *H3-H4* promoter. These data suggest that multiple protein–protein and/or protein–DNA interactions contribute to HLB formation, and that the large number of endogenous RD histone gene copies sequester available factor(s) from attenuated transgenic arrays, thereby preventing HLB formation and gene expression on these arrays.**

Monitoring Editor
Tom Misteli
National Institutes of Health, NCI

Received: Mar 9, 2020
Revised: May 5, 2020
Accepted: May 7, 2020

INTRODUCTION

An important organizing principle in cells is the use of membraneless compartments to spatially and temporally regulate diverse biological processes (Mitrea and Kriwacki, 2016). Numerous membraneless compartments have been identified in both the nucleus (e.g., nucleoli, Cajal bodies, histone locus bodies) and the cytoplasm (e.g., P-bodies, stress granules, germ granules) and are collectively referred to as biomolecular condensates (Banani *et al.*, 2017). There is increasing evidence suggesting that biomolecular condensates are formed through liquid–liquid phase separation or condensation (Alberti *et al.*, 2019). This occurs when proteins and/or nucleic acids in the nucleoplasm or cytoplasm coalesce or demix into a condensed phase that often resembles liquid droplets. Large nuclear condensates that are visible under light microscopy are most often referred to as nuclear bodies (NBs) and represent an important organizing feature of the nucleus.

The histone locus body (HLB) is a conserved NB that assembles at replication-dependent (RD) histone genes and concentrates factors required for RD histone mRNA biogenesis (Duronio and Marzluff, 2017). RD histone mRNAs are the only eukaryotic mRNAs that are not polyadenylated (Marzluff and Koreski, 2017). The unique stem loop at the 3′-end of RD histone mRNAs results from a

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E20-03-0176) on May 13, 2020.

*Address correspondence to: William F. Marzluff (marzluff@med.unc.edu); Robert J. Duronio (duronio@med.unc.edu).

Abbreviations used: DWT, designer wild type; FLASH, FLICE-associated huge protein; HLB, histone locus body; HWT, histone wild type; MUTE, muscle wasted; Mxc, multi sex combs; NB, nuclear body; NPAT, nuclear protein, ataxia-telangiectasia locus; PR, promoter replacement; RD, replication-dependent.

processing reaction requiring a specialized suite of factors, some of which are constitutively localized in the HLB (Duronio and Marzluff, 2017). The HLB provides a powerful system to study how NBs form and function because it contains a well-characterized set of factors involved in producing a unique class of cell–cycle-regulated mRNAs. We previously demonstrated that concentrating factors (e.g., FLASH [FLICE-associated huge protein] and U7 snRNP) in the *Drosophila melanogaster* HLB is critical for efficient histone pre-mRNA processing (Wagner *et al.*, 2007; Tatomer *et al.*, 2016). However, a full understanding of how the HLB participates in histone mRNA biosynthesis requires knowledge of HLB assembly at the molecular level.

Prior studies of NBs have provided several important assembly concepts that are applicable to the HLB. Many NB components have an intrinsic ability to self-associate, an observation leading to two different models of NB assembly: 1) interactions among NB components occur stochastically, wherein individual factors can be recruited to the body in any order; or 2) components assemble in an ordered or hierarchical pathway, wherein the recruitment of components is predicated on prior recruitment of others (Dundr and Misteli, 2010; Rajendra *et al.*, 2010). The HLB appears to employ a hybrid version of these two possibilities. For example, genetic loss of function experiments suggest a partially ordered assembly pathway of the *Drosophila* HLB with some components being required for subsequent recruitment of others (White *et al.*, 2011; Terzo *et al.*, 2015; Tatomer *et al.*, 2016). The scaffolding protein Mxc (multi sex combs), the *Drosophila* ortholog of human NPAT (nuclear protein, ataxia-telangiectasia locus), and FLASH likely form the core of the HLB and are required for subsequent recruitment of other factors (White *et al.*, 2011). Tethering experiments in mammalian cells indicate that ectopic HLB formation also may be induced by several different HLB components, supporting a stochastic model of assembly (Shevtsov and Dundr, 2011).

The initiation event in self-organizing NB assembly is the key step in the process and is not well understood. A prevalent model postulates a "seeding" event that initiates the nucleation of critical components that form a platform for further recruitment of other components (Dundr, 2011; Shevtsov and Dundr, 2011; Altmeyer *et al.*, 2015; Falahati *et al.*, 2016; Stanek and Fox, 2017; Gomes and Shorter, 2019). In some instances, RNA is thought to help seed NB assembly, and NBs such as the nucleolus and ectopic paraspeckles can form at sites of specific transcription (Matera *et al.*, 2009; Mao *et al.*, 2011; Falahati *et al.*, 2016). Blocking transcription prevents complete HLB assembly in both zebrafish and flies (Salzler *et al.*, 2013; Heyn *et al.*, 2017; Hur *et al.*, 2020). However, the HLB is present at RD histone genes even in G1 when the genes are not active, raising the possibility that histone genes themselves participate in seeding HLB assembly (Zhao *et al.*, 1998; Liu *et al.*, 2006).

In *Drosophila*, the RD histone genes are present at a single locus with ~100 copies of a tandemly arrayed 5-kb repeat unit, each of which contains one copy of the divergently transcribed *H2a-H2b* and *H3-H4* gene pairs as well as the gene for linker histone *H1* (Lifton *et al.*, 1978; McKay *et al.*, 2015; Bongartz and Schloissnig, 2019). Using transgenes containing a wild-type or mutant derivative of a single histone repeat, we previously demonstrated that the bidirectional *H3-H4* promoter stimulated HLB assembly and transcription of the single histone repeat in salivary glands (Salzler *et al.*, 2013). We subsequently showed that the conserved GAGA repeat elements present in the *H3-H4* promoter region are targeted by the zinc-finger transcription factor CLAMP (chromatin-linked adaptor for MSL proteins), and that this interaction promotes HLB assembly (Rieder *et al.*, 2017). Thus, the *H3-H4* promoter region might act to seed HLB assembly.

In this work, we leveraged transgenic histone gene arrays to test whether the *H3-H4* promoter region is necessary for in vivo function of the RD histone locus. We found that replacement of *H3-H4* promoters with *H2a-H2b* promoters results in an attenuated transgenic histone gene array that does not function in the presence of the intact endogenous RD histone locus, but surprisingly provides full in vivo function, including normal HLB assembly and histone gene expression, when the endogenous RD histone locus is absent. These results suggest that multiple elements in the histone genes and core HLB proteins are involved in HLB assembly.

RESULTS

To study histone gene regulation and the role of histone posttranslational modifications in chromatin, we previously developed an experimental system in which BAC-based, transgenic histone gene arrays containing 12 copies of the 5-kb histone repeat unit assemble HLBs and functionally complement homozygous loss of the ~100 gene copies at the endogenous RD histone locus, *HisC* (McKay *et al.*, 2015). To study histone gene expression, we created a $12x^{HWT}$ (Histone Wild Type) transgenic construct containing a synonymous polymorphism in the *H2a* gene (i.e., mutation of a XhoI site) that allows us to distinguish transgenic *H2a* gene expression from endogenous *H2a* gene expression (McKay *et al.*, 2015). Here, we extended this design and created a "Designer Wild Type" (DWT) 5-kb histone repeat unit that has all five histone genes marked in a similar manner (Figure 1A; Supplemental Figure S1A). We also introduced restriction enzyme sites around the mRNA 3′-end processing signals, the stem loop and histone downstream element (Marzluff and Koreski, 2017), to allow us to readily manipulate those sequences (Supplemental Figure S1). We used this DWT repeat to create a $12x^{DWT}$ construct and found that a transgenic version of this array behaved identically to the original $12x^{HWT}$ array. It rescued the lethality caused by homozygous deletion of *HisC*, resulting in viable, fertile adult flies. In addition, we can maintain a stock of flies lacking the endogenous RD histone locus and containing a homozygous $12x^{DWT}$ transgene.

The H3-H4 promoter stimulates HLB formation

To test whether the *H3-H4* promoter region is necessary for histone locus function in vivo, we engineered a derivative of the $12x^{DWT}$ array in which all *H3-H4* promoters were replaced with *H2a-H2b* promoters. In this "Promoter Replacement" ($12x^{PR}$) construct, we replaced the entire 298-nt sequence between the initiation codons of the divergently transcribed *H3* and *H4* genes with the 226-nt sequence between the initiation codons of the divergently transcribed *H2a* and *H2b* genes, leaving the *H1* and *H2a* and *H2b* genes intact (Figure 1A; Supplemental Figure S1B). One difference between the two promoters is that the *H2a-H2b* region lacks the CLAMP-binding GAGA repeats present in the *H3-H4* promoter (Rieder *et al.*, 2017). Otherwise it is an intact, native bidirectional histone gene promoter. Consequently, the $12x^{PR}$ construct should retain the ability to initiate transcription from all RD histone genes. The 72-nt size difference between the promoters provides a way to unambiguously distinguish between the $12x^{PR}$ and the $12x^{DWT}$ array genotypes using PCR primers in the *H3* and *H4* coding regions in flies heterozygous (Figure 1B) or homozygous (Figure 1C) for the *HisC* deletion.

We first assessed whether the $12x^{PR}$ and $12x^{DWT}$ transgenes could support HLB formation in the presence of the endogenous histone genes in polytene chromosome spreads from 3rd instar larval salivary glands. In these polyploid cells, the genome is amplified greater than 1000-fold and sister chromatids line up in register, resulting in large chromosomes providing high-resolution cytology.

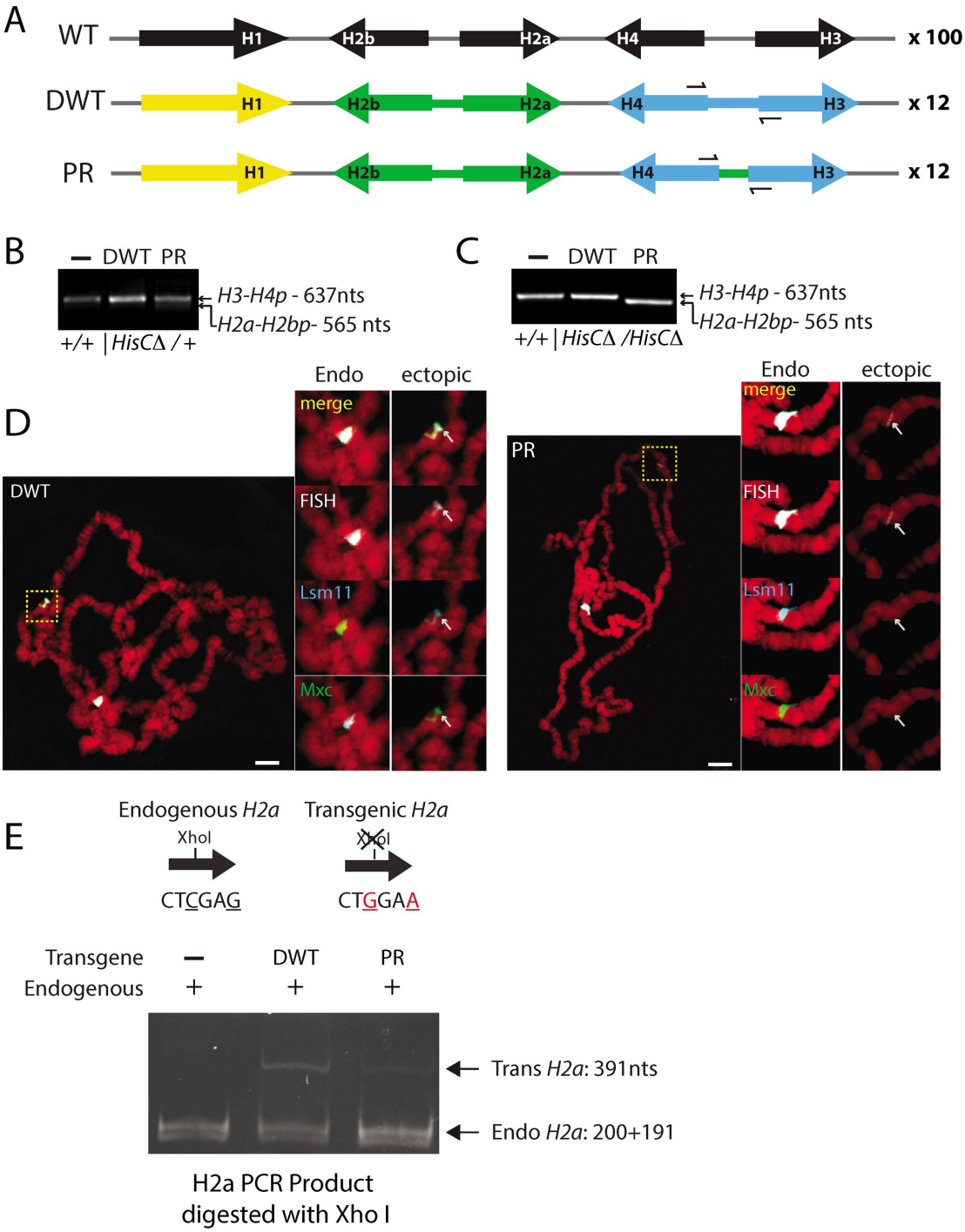

FIGURE 1: The *H3-H4* bidirectional promoter is required for ectopic HLB formation and gene expression in the presence of the endogenous RD histone genes. (A) Schematic of the endogenous *Drosophila* RD histone gene repeat, which is found in a tandem array of ~100 copies at *HisC*, and the BAC-based DWT and PR synthetic histone repeats, which arrayed to 12 copies and inserted at VK33 on chromosome 3. The colors indicate that each transgenic histone gene was engineered to be distinguishable molecularly from the endogenous histone genes (see Supplemental Figure S1). In the PR construct, the *H2a-H2b* promoters (green) replaced the *H3-H4* promoters (blue). (B, C) Primers that anneal to *H3* and *H4* genes (A) were used to amplify the promoter region from genomic DNA extracted from whole flies heterozygous (B) or embryos homozygous (C) for a deletion of the *HisC* locus and containing either the $12x^{DWT}$ or the $12x^{PR}$ transgene. The amplicon from a wild-type locus is larger (637 nts) than the PR locus (565 nts). (D) Polytene chromosome squashes from 3rd instar larval salivary glands containing both *HisC* and either the $12x^{DWT}$ or the $12x^{PR}$ transgenes hybridized with a probe to the histone locus and stained for HLB components Lsm11 and Mxc. Note that absence of an HLB at $12x^{PR}$ (arrows) when *HisC* is present (right panels). Scale bar, 10 microns. (E) Assay for specific detection of endogenous versus transgenic RD histone gene expression. Total RNA was prepared from 3rd instar larvae containing the endogenous genes and one copy of the $12x^{DWT}$ or $12x^{PR}$ transgene. RT-PCR with *H2a* primers followed by digestion with *Xho I* cleaves cDNA from the endogenous genes but not the transgenes, which was visualized using an 8% polyacrylamide gel.

We visualized the HLB by immunofluorescence using antibodies that recognize one of several HLB components. These components include Mxc, the *Drosophila* ortholog of human NPAT, which is necessary for HLB assembly and histone gene expression (White *et al.*, 2011; Terzo *et al.*, 2015); histone pre-mRNA processing factors FLASH (Yang *et al.*, 2009) and Lsm11, a component of the U7 snRNP (Liu *et al.*, 2006; Godfrey *et al.*, 2009; Burch *et al.*, 2011); and muscle wasted (Mute), an ortholog of human YARP (YY1AP-related protein 1) that is a putative repressor of histone gene expression (Bulchand *et al.*, 2010). In these preparations, we also visualized both the endogenous histone locus and the ectopic, transgenic histone genes by FISH using a probe derived from the entire histone repeat unit. We observed HLB formation at the control $12x^{DWT}$ transgenic locus but not at the $12x^{PR}$ transgenic locus (Figure 1D). These data are consistent with our previous results that sequences within the *H3-H4* promoter are necessary for HLB assembly in the presence of the endogenous histone genes (Salzler *et al.*, 2013).

Mutant, single copy transgenic histone gene repeats that fail to form an HLB (Salzler *et al.*, 2013), and Mxc mutants that don't form an HLB (Terzo *et al.*, 2015), result in reduced histone mRNA levels. These data suggest that the HLB is necessary for efficient histone mRNA biosynthesis. We therefore tested whether the $12x^{PR}$ transgene could support histone gene expression in the absence of HLB assembly. To determine expression from the ectopic $12x^{DWT}$ or $12x^{PR}$ genes in the presence of the endogenous RD histone genes, we randomly prime cDNA from total RNA and then amplify a fragment of each coding region containing the restriction enzyme site that is changed in the DWT construct. By digesting the PCR fragment with the appropriate restriction enzyme and resolving the fragments by gel electrophoresis, we can determine the relative level of expression from the transgenic and endogenous RD histone genes. For example, RT-PCR of the *H2a* gene results in the same size amplification product from both the transgenic and endogenous genes, but only the product from the endogenous histone genes is sensitive to digestion with *Xho*I (Figure 1E). When assayed in the presence of the endogenous genes, we found that *H2a* was expressed from the $12x^{DWT}$ at a much higher level than from the $12x^{PR}$ transgene, even though identical promoters drove *H2a* in each transgenic array (Figure 1E; Figure 2, A, lanes 2 and 4, and B, lanes 2 and 3). These data are consistent with our previous observations using transgenes with a single histone gene repeat unit (Salzler *et al.*, 2013). Together these results demonstrate that the *H3-H4* promoter is required both for HLB formation and for histone gene expression in the presence of the endogenous RD histone genes.

A histone gene array lacking the H3-H4 promoter forms HLBs and is expressed in the absence of the endogenous genes

To assess the biological activity of the $12x^{PR}$ transgene, we determined the developmental outcome of having a $12x^{PR}$ transgene as the only zygotic source of histone mRNA. Due to large stores of maternal histone protein and mRNAs, embryos homozygous for a deletion that removes the endogenous RD histone gene array develop normally through S phase of cell cycle 14, but require zygotic RD histone gene expression for progression through S phase of cycle 15 and beyond (Gunesdogan *et al.*, 2010, 2014). Consequently, embryos lacking RD histone genes arrest at nuclear cycle 15 and do not hatch. As noted above, this embryonic lethality is rescued by a single $12x^{DWT}$ transgene, which supports development of histone deletion progeny into viable, fertile adult flies. Surprisingly, we found that embryos lacking endogenous histone genes and containing the $12x^{PR}$ transgene hatched and developed into nearly the expected number of fertile adults without any overt developmental delays, similar to embryos lacking endogenous histone genes and rescued by the $12x^{DWT}$ transgene. These data suggest that the $12x^{PR}$ transgene provides normal amounts of histone gene function when the endogenous RD histone genes are absent.

We interrogated this unexpected result more thoroughly by analyzing 1) histone gene expression, 2) HLB formation, and 3) histone pre-mRNA processing in both 5- to 7-h-old embryos and 3rd instar larvae lacking endogenous histone genes and containing either the $12x^{DWT}$ or the $12x^{PR}$ transgene. In both genotypes, all histone mRNA in either 5- to 7-h-old embryos (Figure 2A, lanes 3 and 5) or 3rd instar larvae (Figure 2B, lanes 4 and 5) was derived only from the ectopic histone gene array, indicating that the maternal histone mRNA stores had been degraded normally. We observed robust histone gene expression from the $12x^{PR}$ transgene when the endogenous histone genes are absent, in stark contrast with the low level of expression from the $12x^{PR}$ transgene when endogenous histone genes are present (Figure 2, A, lane 4, and B, lane 3).

High levels of histone gene expression are strongly correlated with the ability to form an HLB (Salzler *et al.*, 2013; Terzo *et al.*, 2015; Rieder *et al.*, 2017). Given that the $12x^{PR}$ can support histone gene expression in the absence of the endogenous genes, we assayed for HLB formation in polytene chromosome spreads from 3rd instar larval salivary glands and in embryos. We detected robust HLB formation at the $12x^{PR}$ transgenic locus with antibodies against FLASH and Lsm11 (Figure 2C), or Mxc and Mute (Figure 2D), similar to that observed at the $12x^{DWT}$ locus. Thus, the $12x^{PR}$ transgene, which lacks *H3-H4* promoter sequences, can support HLB formation and normal levels of histone gene expression in the absence of endogenous RD histone genes.

Histone mRNA from the $12x^{PR}$ transgenic locus is properly processed

An important function of the HLB is concentrating factors to promote efficient histone pre-mRNA processing (Tatomer *et al.*, 2016). Therefore, we reasoned it was possible that the attenuated $12x^{PR}$ gene array might affect other aspects of histone mRNA biosynthesis, including pre-mRNA processing. All *Drosophila* RD histone genes contain cryptic polyadenylation signals downstream of the normal processing sites that are only used when the histone pre-mRNA processing reaction is compromised, resulting in the production and accumulation of poly(A)+ histone mRNA (Lanzotti *et al.*, 2002; Godfrey *et al.*, 2009; Tatomer *et al.*, 2016). We examined pre-mRNA processing efficiency in flies rescued by the $12x^{PR}$ transgene. Our PCR-based assays to detect histone mRNA expression cannot differentiate between properly processed and misprocessed histone mRNAs. We therefore used Northern blotting with a probe against the *H3* coding region to determine whether histone pre-mRNA was efficiently processed. In contrast to a *FLASH* mutant, which expresses large amounts of poly(A)+ histone mRNA (Tatomer *et al.*, 2016), we did not detect poly(A)+ histone mRNA from early embryos (Figure 2E) or from whole 3rd instar larvae (Figure 2F) in *HisC* deletion animals rescued by the $12x^{PR}$ transgene. Note also that the levels of histone mRNA from the $12x^{PR}$ transgenes are similar to wild-type levels. These results indicate that the HLB formed on the $12x^{PR}$ array in the absence of endogenous histone genes supports efficient histone pre-mRNA processing.

HLB assembly at the $12x^{PR}$ transgenic locus occurs at the normal time during embryogenesis

The above data suggest that the $12x^{PR}$ represents an attenuated histone gene array that provides normal biological function when

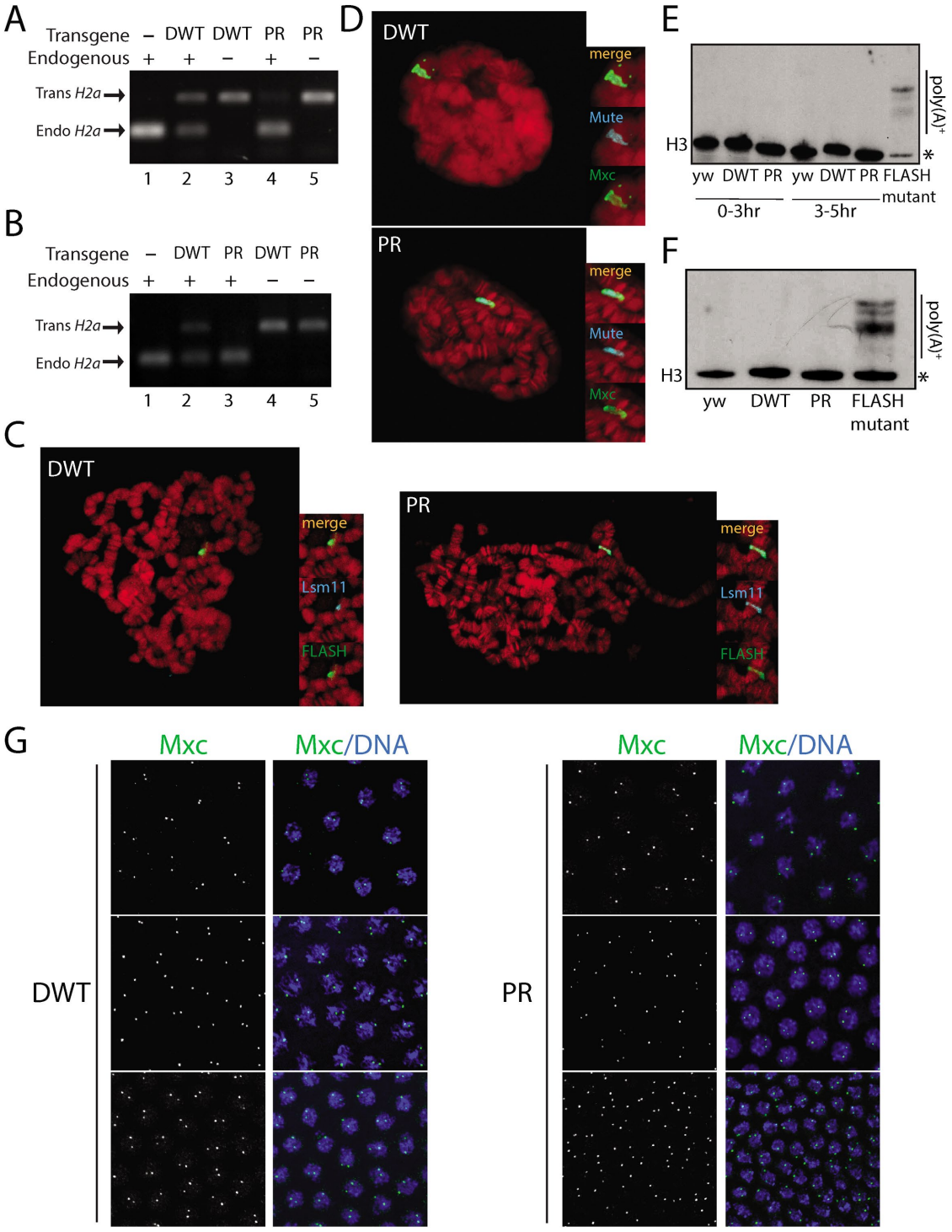

FIGURE 2: The PR transgene is expressed and forms an HLB in the absence of the endogenous RD histone genes. (A, B) RNA was isolated from 5–7 h embryos (A) or 3rd instar larvae (B) containing the $12x^{DWT}$ or $12x^{PR}$ transgene in the presence or absence of *HisC*, with wild-type embryos as a control (lane 1). RT-PCR products using *H2a* primers were digested with *Xho I* and visualized using a 1.5% agarose gel (which does not resolve the two restriction fragments from the endogenous genes as in Figure 1E). (C, D) Polytene chromosomes from *HisC* deletion homozygous 3rd instar larval salivary glands rescued by either the $12x^{DWT}$ or the $12x^{PR}$ transgene stained with antibodies against HLB components Lsm11 and FLASH (C) or Mxc and Mute (D). (E, F) Total RNA from 0–3 h or 3–5 h embryos (D) or 3rd instar larvae (E) from wild-type (yw) or HisC deletion homozygotes containing either the $12x^{DWT}$ or the $12x^{PR}$ transgene analyzed by Northern blotting using a radiolabeled *H3* coding region probe. *Normally processed H3 mRNA. A FLASH mutant that cannot recruit the processing machinery to the HLB (Tatomer *et al.*, 2016) was used as a positive control for production of poly(A)+ histone mRNA. (G) Syncytial, embryos from *HisC* deletion homozygous stocks rescued by $12x^{DWT}$ or $12x^{PR}$ transgenes and stained with Mxc to monitor HLB formation in development. Three consecutive syncytial nuclear cycles (top to bottom) were analyzed.

not in competition with the wild-type endogenous histone genes. To further explore this model, we determined if an HLB assembles on the 12x^{PR} array in the early embryo at the same time that it assembles on the 12x^{DWT} array. The HLB begins assembling in syncytial blastoderm embryos just prior to the onset of zygotic histone transcription (White *et al.*, 2007, 2011). We previously reported that a "proto-HLB" consisting of Mxc and FLASH forms in nuclear cycle 10, followed by the onset of zygotic histone gene expression and further recruitment of additional HLB components (Mute and Lsm11) in cycle 11 (White *et al.*, 2011; Salzler *et al.*, 2013). To determine if the HLB forms at the 12x^{PR} transgenic locus with normal timing, we stained syncytial blastoderm embryos lacking endogenous histone genes and rescued by a 12x^{PR} transgene with antibodies against Mxc (Figure 2G). In these experiments, HLB assembly during the syncytial blastoderm cycles was indistinguishable from that of histone deletion embryos rescued by the control 12x^{DWT} transgene. Thus, in the absence of the endogenous genes, an HLB assembles on the 12x^{PR} array at the same time in early development as it does on a wild-type array.

Note that zygotic histone transcription occurs coincident with HLB formation in syncytial blosterm cycle 11 (Terzo *et al.*, 2015; Hur *et al.*, 2020). HLBs increase in size in each cycle after they are formed (Hur *et al.*, 2020). Given how short S phase is at this stage of embryogenesis, the levels of histone mRNA formed during these rapid cell cycles are likely very small and are not necessary for development since embryos can progress through cycle 14 without any histone genes using only the maternal supply of histone proteins and mRNAs. Only starting in cycle 15 are large amounts of zygotic histone mRNAs required (Smith *et al.*, 1993; Gunesdogan *et al.*, 2010, 2014).

CLAMP is present in the 12x^{PR} HLB in the absence of the endogenous RD histone genes

The *H3-H4* promoter is highly conserved among *Drosophilids* and contains conserved GAGA repeats (Salzler *et al.*, 2013; Rieder *et al.*, 2017), which we previously showed are essential for HLB formation and expression of RD histone genes in the presence of the endogenous histone locus (Rieder *et al.*, 2017). Although in vitro these repeats can bind both the zinc-finger GA-repeat binding protein CLAMP and the major *Drosophila* GAGA repeat binding protein GAF (GAGA factor*; trithorax-like*) (Gilmour *et al.*, 1989), only CLAMP is bound to the histone locus in wild-type animals (Rieder *et al.*, 2017). The *H2a-H2b* promoter and the rest of the histone repeat unit do not contain any GAGA repeats longer than 4 nts, and CLAMP preferentially binds to longer GAGA repeats (Kuzu *et al.*, 2016). We asked whether CLAMP is recruited to the 12x^{PR} transgene in salivary gland polytene chromosomes. As with all other HLB components we tested, CLAMP is not recruited to the 12x^{PR} transgene or a 12x histone gene array in which the GAGA sequences are replaced with *lacO* binding sites (GA^{M}, Figure 3A) when the endogenous RD histone genes are present (Rieder *et al.*, 2017). Surprisingly, we found that in the absence of endogenous RD histone genes, CLAMP (Figure 3C), but not GAF (Figure 3E), is recruited to the 12x^{PR} transgenic locus with similar intensity to CLAMP recruitment to the 12x^{DWT} transgenic locus (Figure 3, B and D). Furthermore, the 12x GA^{M} array also supports *H2a* gene expression in the absence of the endogenous genes (Figure 3F). These data indicate that CLAMP can be recruited to the HLB in a histone gene array lacking GAGA repeats when the GAGA binding sites within the *H3-H4* promoter at the endogenous RD histone locus have been removed. The *H3* and *H4* genes are transcribed, indicating that CLAMP is not an essential DNA-binding transcription factor for the *H3* and *H4* genes, but likely serves as a factor that alters chromatin structure (Rieder *et al.*, 2017).

We carried out ChIP-qPCR and ChIP-seq experiments on *HisC* deletion embryos containing only the 12x^{DWT} array or only the 12x^{PR} array to determine how CLAMP is interacting with the histone gene array. In agreement with ChIP-seq results on the endogenous histone genes, which showed that CLAMP is localized precisely to the *H3-H4* promoter (Rieder *et al.*, 2017), CLAMP was bound to the *H3-H4* promoter but not to the *H2a-H2b* genes or promoter on the 12x^{DWT} array (Figure 4, A–C). In contrast, CLAMP was not enriched at any location on the 12x^{PR} array (Figure 4, A, B, and D). As in wild-type flies, in the 12x^{PR} embryos CLAMP bound the endogenous *rox1* locus on the X chromosome, a locus to which CLAMP is recruited in males for its role in X chromosome dosage compensation (Figure 4, B, E, and F) (Soruco *et al.*, 2013). Thus, CLAMP is not specifically enriched at any location in the entire histone gene repeat unit when the *H3-H4* promoter is replaced with *H2a-H2b* promoter in a 12x transgenic array, despite our observation using microscopy that CLAMP is present in the HLB at the 12x^{PR} array (Figure 3C). These data suggest that CLAMP can be recruited to RD histone genes by protein–protein interactions independently of its binding to specific DNA sequences in the histone gene repeat.

A wild-type array can activate the 12x^{PR} array when present *in trans* at the homologous locus

The results above demonstrate that the 12x^{PR} transgenic locus does not form an HLB in the presence of the endogenous histone genes unlike the wild-type 12x^{DWT} transgene, which does. We tested whether juxtaposing a wild-type histone gene array near the 12x^{PR} locus would nucleate a functional HLB and activate expression from the 12x^{PR} transgene. Our BAC-based transgenes are inserted into the genome via site-specific recombination at the same chromosomal location. We created flies in which the 12x^{PR} was placed at position VK33 on chromosome 3 *in trans* to the 12x^{HWT} transgene used in our initial studies (McKay *et al.*, 2015) and analyzed these transgenes in the presence of the endogenous RD histone genes on chromosome 2. We examined HLB formation at the ectopic VK33 location by staining whole mount salivary glands with antibodies against Mxc and FLASH. Using this approach, we detect only a single, large HLB on the endogenous RD histone genes in wild-type nuclei lacking any transgenes (Figure 5, A-1 and B). When 12x^{PR} was the only transgene present in addition to the endogenous RD histone genes, none of the nuclei contained an ectopic HLB and we only detected the endogenous HLB (Figure 5, A-2 and B), consistent with the results of staining spread salivary gland polytene chromosomes (Figure 1D). In contrast, we observed formation of a second small HLB in addition to the single large endogenous HLB in 100% of the 12x^{HWT}/12x^{PR} nuclei (Figure 5, A-5 and B), as expected since this HLB also forms when only the 12x^{HWT} array is present at this site.

We measured histone gene expression from the 12x^{PR} transgene in these genotypes. In the 12x^{HWT} array the histone *H2a* gene, and not the other histone genes, is marked with a restriction site change (McKay *et al.*, 2015), whereas the 12x^{PR} array has all five histone genes marked with a restriction site change. Therefore, to specifically detect histone gene expression from the 12x^{PR} transgene, we digested embryonic *H3* RT-PCR products with *Sac* I, whose recognition site is missing from the 12x^{PR} *H3* gene but is present in both the endogenous and the 12x^{HWT} *H3* genes (Figure 5C). Strikingly, the 12x^{PR} transgene supports stronger *H3* gene expression in the presence compared with the absence of the endogenous histone genes when present *in trans* with the 12x^{HWT} transgene (Figure 5C, lanes 1 and 2). These data demonstrate that the wild-type histone sequences in the 12x^{HWT} were able to activate the 12x^{PR} transgenic locus in the presence of the endogenous histone genes, likely by nucleating

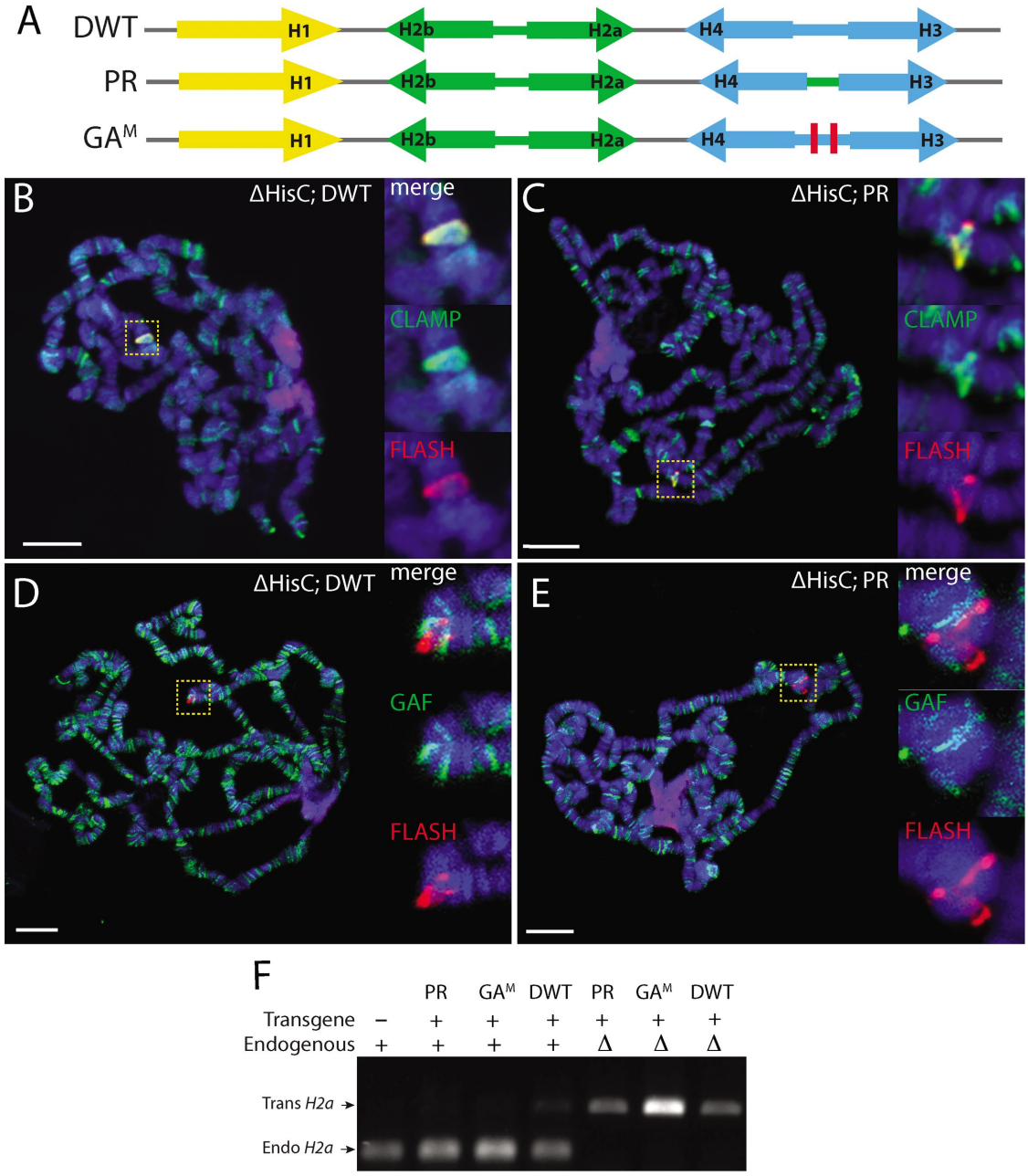

FIGURE 3: CLAMP but not GAF is recruited to HLBs that form in the absence of GA repeats. (A) Schematic of the WT, PR, and GA mutant (GA^M) synthetic histone repeats. For GA^M, GA sequences in the *H3-H4* promoter were mutated to *lacO* sites or scrambled (Rieder *et al.*, 2017). (B–E) Polytene chromosome squashes from 3rd instar larval salivary glands lacking endogenous histone genes and rescued by either a $12x^{DWT}$ transgene (B, D) or a $12x^{PR}$ transgene (C, E) were stained with antibodies against FLASH and CLAMP (B, C) or FLASH and GAF (D, E). Scale bar, 10 microns. (F) RT-PCR analysis with *H2a* primers performed on cDNA from 5–7 h embryos of indicated genotypes and visualized with ethidium bromide on a 0.8% agarose gel.

ectopic HLB formation that encompasses the paired homologous chromosomes. This result is similar to transvection, a phenomenon in which transcription of a gene can be activated by an enhancer located *in trans* through pairing of homologous chromosomes (Duncan, 2002; Fukaya and Levine, 2017).

Transgenes with a single copy of the wild-type repeat partially activate a 12x array *in trans*

Since 12 copies of a wild-type histone repeat at the same integration site-stimulated expression from the $12x^{PR}$ transgene, we tested whether a single copy could do the same using two different approaches; one where the single copy is present *in trans* and the other where it is present *in cis*. For the first approach, we inserted a 1x HWT transgene at the VK33 integration site. In these strains we detected an ectopic HLB in 90% of the salivary gland nuclei in the presence of the endogenous genes (Figure 5B), similar to what we observed previously (Salzler *et al.*, 2013). When this 1x HWT transgene was placed *in trans* with the $12x^{PR}$ array, we detected an ectopic HLB in 94% of the nuclei examined (Figure 5, A-6 and B). We tested whether there was any activation of expression of the

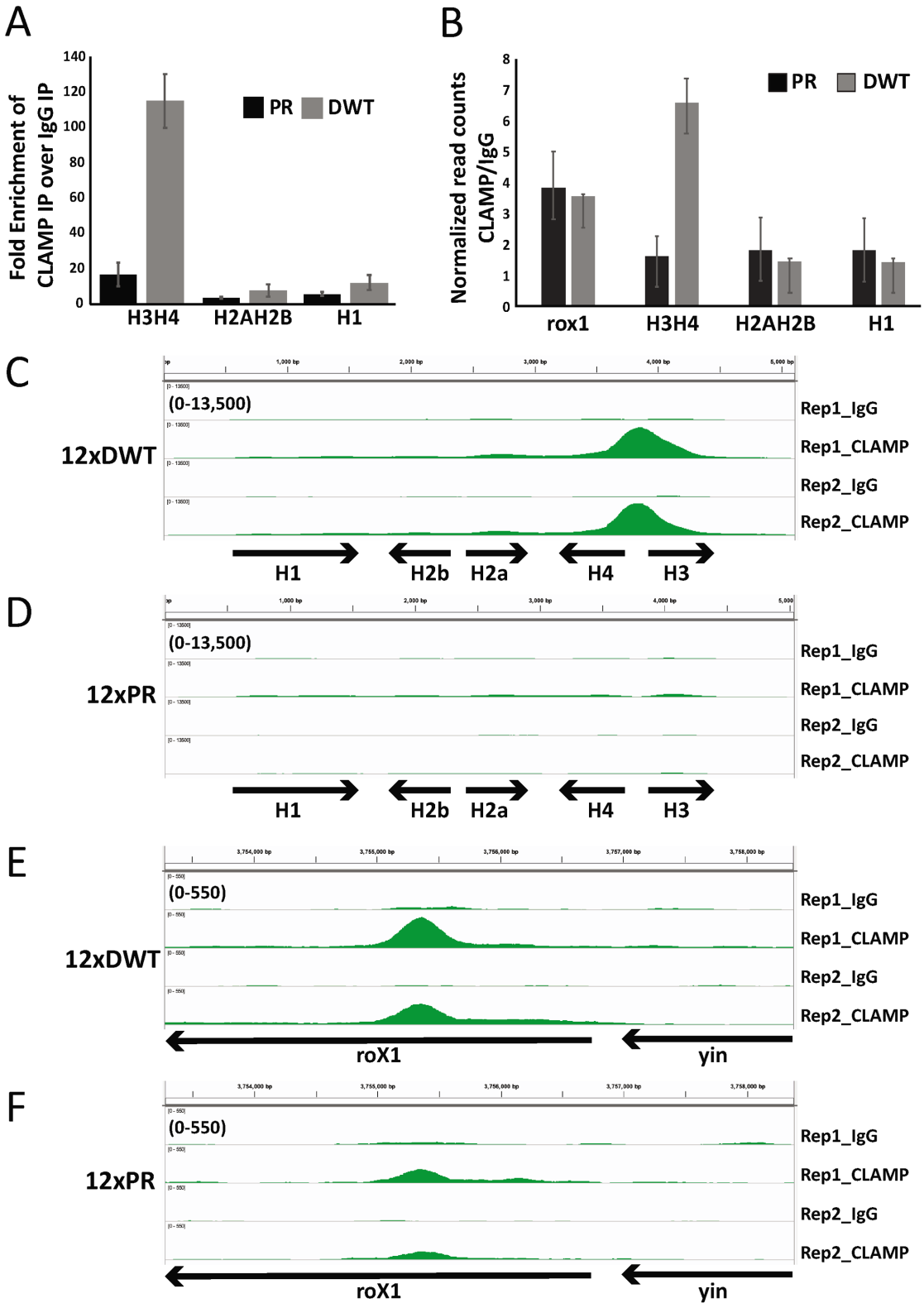

FIGURE 4: CLAMP does not directly bind to a sequence in the 12x^{PR} transgene. (A) Two- to four-h-old *HisC* homozygous deletion embryos rescued by homozygous 12x^{DWT} or homozygous 12x^{PR} transgenes were analyzed by CLAMP ChIP-qPCR using primers recognizing the indicated promoters. (B) Biological replicates of 2- to 4-h-old embryos of the same genotypes were analyzed by ChIP-seq. Reads were quantified within a defined window containing the indicated genes (see *Materials and Methods*). (C–F) Genome browser views of reads from the same CLAMP ChIP-seq data that map to the histone gene repeat (C, D) or the *roX1* locus (E, F), an X-linked control locus for CLAMP binding to the genome. Note that >12× more reads mapping to *HisC* relative to *roX1* suggest that there is strong binding of CLAMP to all histone repeats in the 12x^{DWT} the array. Note also that neither the entire *roX1*, but only the region of *roX1* that binds CLAMP, or *yin* genes are shown in the browser window.

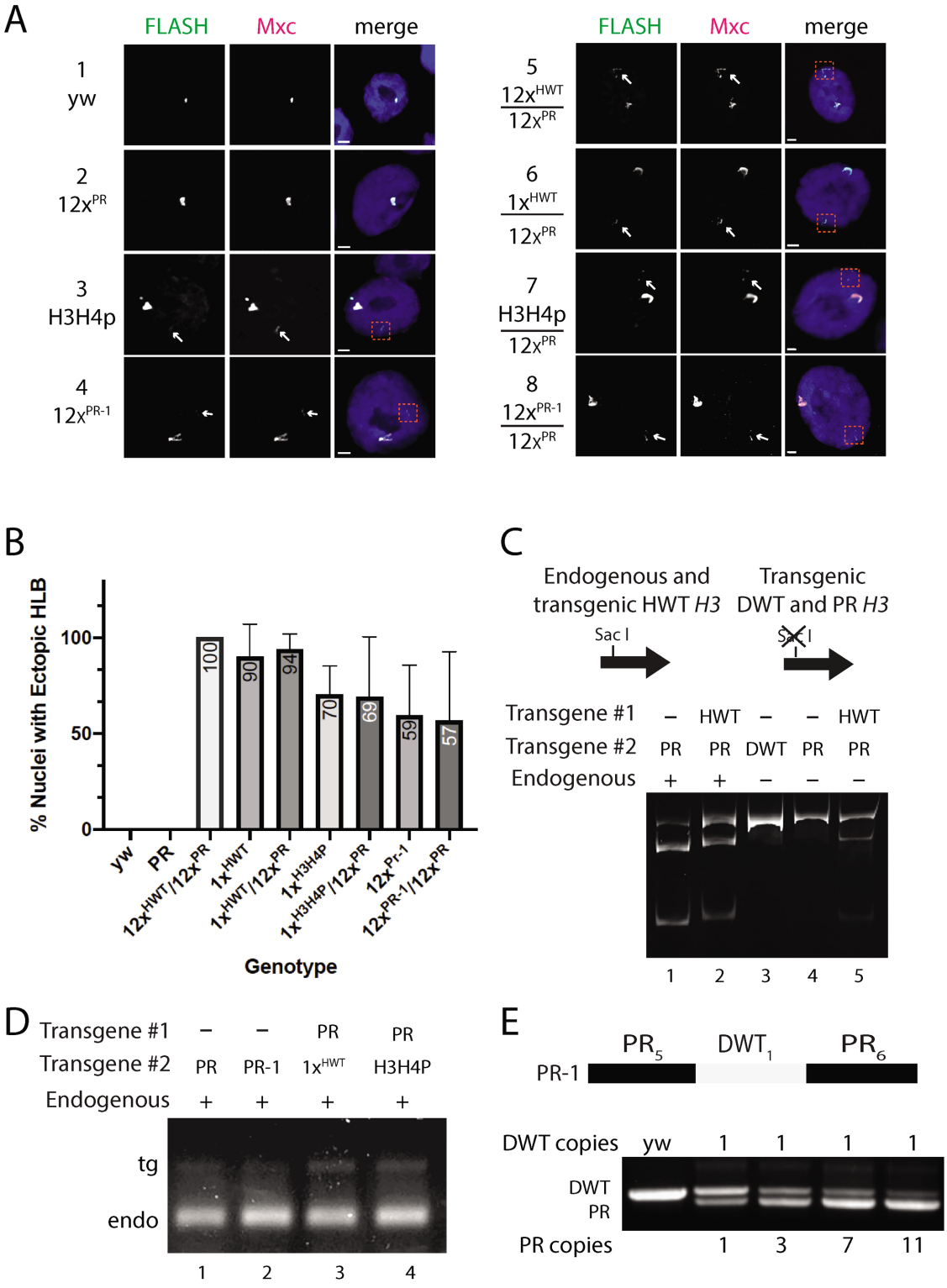

FIGURE 5: Activation of PR transgenes by wild-type histone gene arrays. (A-1–A-8) Intact salivary gland nuclei from 3rd instar larvae of the indicated genotypes, all of which contained *HisC* and stained for the HLB components Mxc and FLASH. Ectopic HLBs are indicated by arrows and dashed boxes and were present in all genotypes except the $12x^{PR}$. Scale bar, 10 microns. (B) Quantification of the percentage of ectopic HLBs in the salivary gland nuclei from panel A. (C) Schematic of the assay used to distinguish $12x^{PR}$ from $12x^{HWT}$ cDNA. RT-PCR amplification of *H3* mRNA from 5- to 7-h-old embryos of the indicated genotypes, followed by digestion with *Sac*I and visualized on an 8% polyacrylamide gel. (D) RT-PCR amplification of *H2a* mRNA from larvae of indicated genotypes, followed by digestion with *Xho*I and visualization on an agarose gel. (E) Top: schematic of the BAC-based PR-1 histone array containing 11 PR and 1 DWT repeat unit. Bottom: analysis of BAC constructs containing WT and PR repeats using PCR as in Figure 1, B and C.

$12x^{PR}$ array when present *in trans* with 1x HWT using our RT-PCR assay. There was a modest activation of the expression $12x^{PR}$ array, but much weaker than when the $12x^{HWT}$ was *in trans* with the $12x^{PR}$ as measured by the ratio of the bands (Figure 5D, lanes 1 and 3).

We next tested if including a single wild-type repeat in the $12x^{PR}$ array could stimulate HLB formation and histone mRNA expression from that array in the presence of the endogenous genes. We created a 12x array ($12x^{PR\text{-}1}$) containing one wild-type histone repeat unit in the center of 11 PR repeat units (Figure 5E). Like $12x^{PR}$, the $12x^{PR\text{-}1}$ transgene rescued a deletion of the endogenous histone genes, resulting in viable and fertile adults and indicating that it is likely fully active when present as the only source of RD histone genes. We then examined HLB formation in intact salivary glands from animals containing both the $12x^{PR\text{-}1}$ transgene and the endogenous RD histone genes. In this genotype, we detected HLB formation at the ectopic $12x^{PR\text{-}1}$ transgenic locus in 59% of the nuclei (Figure 5, A-4 and B), compared with 0% of nuclei from the $12x^{PR}$. We also analyzed expression from this array in the presence of the endogenous RD histone genes and observed only modest activation of histone mRNA expression (Figure 5D, lane 2) and less than the level of expression when the 1x WT was *in trans* to the 12x array (Figure 5D, lane 3). Thus, a single histone repeat unit can stimulate HLB formation, but activates expression of neighboring genes better when placed *in trans* (i.e., on the other homolog) rather than *in cis* (i.e., in the same 12x array).

We next tested a transgene containing only the *H3-H4* promoter with no coding or 3′-end sequences, which we previously demonstrated could form an HLB in salivary glands when inserted into a ectopic site (Salzler *et al*., 2013). Ectopic HLBs were detected in 70% of the salivary glands when only the *H3-H4p* promoter transgene was present and 69% when it was placed *in trans* to the $12x^{PR}$ array (Figure 5, A-3, A-7, and B). We assayed the expression of the *H3* gene from the $12x^{PR}$ when the *H3-H4p* transgene was *in trans* to the $12x^{PR}$ array. We detected low level expression, similar to the expression detected when the 1x WT transgene was present opposite the $12x^{PR}$ array (Figure 5D, lane 4). Collectively, these data demonstrate that the presence of an HLB nucleating sequence *in trans* to $12x^{PR}$ can induce formation of an ectopic HLB and histone gene expression in the presence of the endogenous histone genes. Further, these data emphasize that the *H3-H4* promoter is a critical element in promoting ectopic HLB formation in the presence of the endogenous histone genes, consistent with our previous observations (Salzler *et al*., 2013; Rieder *et al*., 2017).

DISCUSSION

In this study, we used our histone gene replacement platform to analyze the cis acting elements within the *Drosophila* histone repeat unit that are necessary for HLB formation and histone gene expression. Previously we showed using a single, transgenic histone gene repeat unit that the promoter region of the divergently transcribed *H3-H4* gene pair is capable of stimulating HLB formation (Salzler *et al*., 2013). We subsequently further mapped this functionality using a 12x gene array to conserved GAGA repeats in this region that are targeted by the CLAMP protein (Rieder *et al*., 2017). Here, we present evidence that a $12x^{PR}$ histone gene array devoid of the *H3-H4* promoter and lacking any CLAMP binding elements cannot assemble an HLB in the presence of the ~100 RD histone gene copies at the endogenous locus (*HisC*). However, the $12x^{PR}$ array surprisingly can rescue homozygous deletion of *HisC* and fully support the entire *Drosophila* life cycle. In the *HisC* deletion background, the $12x^{PR}$ array assembles an HLB and expresses the same amount of properly processed histone mRNAs as the endogenous genes or as a $12x^{DWT}$ wild-type array. Below we discuss the implications of these observations on HLB assembly and organization.

The RD histone locus stimulates HLB formation in Drosophila

Biomolecular condensates form via a seeding event that promotes a high concentration of factors at a discrete location, leading to recruitment of additional factors that ultimately result in a structure that can be observed by light microscopy (Gomes and Shorter, 2019). A number of putative seeding events for biomolecular condensates have been described (Dellaire *et al*., 2006; Dundr, 2011; Mao *et al*., 2011; Shevtsov and Dundr, 2011; White *et al*., 2011), but in many cases the precise mechanism of seeding is not known. Nucleic acids, particularly RNA, have been proposed to seed different NBs. Both the nucleolus and the HLB are associated with specific genomic loci (Mao *et al*., 2011; Shevtsov and Dundr, 2011; Salzler *et al*., 2013), and it is likely that the DNA (or chromatin) and/or nascent RNA at the locus participates in the seeding event. The activation of zygotic transcription of rRNA leads to the precise spatiotemporal formation of the nucleolus in *Drosophila* embryos (Falahati *et al*., 2016). In the absence of rDNA, *Drosophila* nucleolar components still form high concentration assemblies, but these are smaller, more numerous, and do not form at the same time in the early embryo as the wild-type nucleolus.

Drosophila HLB components also stochastically assemble smaller and more unstable foci in embryos lacking the RD histone locus (White *et al*., 2007; Salzler *et al*., 2013; Hur *et al*., 2020), suggesting that HLBs and the nucleolus form similar seeding events. Indeed, the dynamics of HLB assembly in single early *Drosophila* embryos display properties consistent with liquid–liquid phase transition seeded by *HisC* (Hur *et al*., 2020). Blocking transcription in the early embryo prevents normal HLB growth (Hur *et al*., 2020), and a defective *H3-H4* promoter (with mutated TATA boxes) does not support HLB formation in the context of a single copy histone gene repeat in salivary glands (Salzler *et al*., 2013). These data suggest that active transcription is essential for forming a complete HLB.

It is important to note that HLBs assemble and persist in nonproliferating *Drosophila* tissues that do not express histone mRNAs (Liu *et al*., 2006; White *et al*., 2007) and are also present in G0/G1 mammalian cells (Ma *et al*., 2000; Zhao *et al*., 2000). Histone gene expression is activated as a result of phosphorylation of Mxc/NPAT by Cyclin E/Cdk2, resulting in changes in the HLB that promote histone gene transcription and pre-mRNA processing (Wei *et al*., 2003; Ye *et al*., 2003; White *et al*., 2011). We propose that in early embryonic development the histone locus DNA and/or chromatin seeds HLB assembly in *Drosophila*, with the *H3-H4* promoter region being particularly important. We further propose that subsequent transcriptional activation of histone genes then drives HLB growth and maturation.

Formation of an HLB on a transgenic RD histone gene array requires that this array compete effectively with the endogenous *HisC* locus for recruitment of HLB components. This is the situation with $12x^{HWT}$ and $12X^{DWT}$ arrays, which form HLBs in the presence of *HisC*. These results also indicate that there are no other elements within *HisC* that are necessary for HLB formation. Because the $12x^{PR}$ array does not form an HLB in the presence of *HisC*, but it does so in the absence of *HisC*, we hypothesize that the endogenous RD histone gene array sequesters critical HLB components, likely including Mxc and CLAMP, thereby preventing HLB assembly at the transgenic locus. By removing the *H3-H4* promoter from the transgene, we removed an element that provided additional interactions with HLB components, notably CLAMP, weakening the overall ability of the locus to stably nucleate an HLB.

HLB assembly involves multiple interactions among components

Interactions among multivalent proteins, or multivalent protein–nucleic acid interactions, are driving forces in the assembly of biomolecular condensates (Shin and Brangwynne, 2017). Mxc is likely the critical factor that together with histone genes seeds *Drosophila* HLB formation and activates histone gene expression. Mxc is a large (~1800 aa) protein that oligomerizes in vivo and likely provides a scaffold for multivalent interactions (Terzo *et al.*, 2015). A C-terminal truncation mutant of Mxc that fails to recruit histone pre-mRNA processing factors still forms an HLB and activates histone gene expression at sufficient levels to complete development, underscoring the multivalent nature of Mxc (White *et al.*, 2011; Landais *et al.*, 2014; Tatomer *et al.*, 2016).

Surprisingly, the HLB that assembles on the $12x^{PR}$ array in the absence of *HisC* contains CLAMP, even though we have removed all of the known CLAMP binding sites from the histone repeat. Although CLAMP may bind another sequence in the $12x^{PR}$ array, no other favorable GAGA repeats are present, and we were unable to detect CLAMP bound to any other location in the histone array by ChIP-qPCR and ChIP-seq experiments. More likely, CLAMP interacts with other HLB components, possibly Mxc or the Mxc-FLASH complex, providing multivalent contacts between CLAMP and other HLB components. Deleting the GAGA sequences from the *H3-H4* promoter did not affect transcription of the *H3* or *H4* genes in the absence of *HisC*, suggesting that CLAMP's major function is to promote HLB assembly and not to act as a canonical DNA binding transcription factor. Supporting this interpretation is the observation that another, more abundant transcription factor that binds to GAGA repeats, GAF, is not found at the HLB unless CLAMP is absent (Rieder *et al.*, 2017), consistent with CLAMP's critical interactions with both the GAGA repeats and the HLB factors in seeding the HLB.

Because the $12x^{PR}$ array is capable of assembling a completely functional HLB in the absence of *HisC*, the *H3-H4* promoter is not absolutely essential for HLB formation. One possibility is that there are multiple pathways for assembling functional HLBs. Previous work suggests that not all seeding events are equivalent in their ability to assemble biomolecular condensates. In artificial systems, changes in scaffold stoichiometry, which can stem from changes in valence, alter the recruitment of components (Banani *et al.*, 2016). Further, mathematical modeling has revealed that scaffolds can nucleate distinct complexes when at different concentrations, and that this can qualitatively alter the transcriptional output (Yang and Hlavacek, 2011). Additionally, P-bodies can form in multiple ways through different protein–protein or protein–nucleic acid interactions, with different interactions predominating under different conditions (Rao and Parker, 2017). Therefore, different nucleators of the HLB (i.e., the *H3-H4* promoter or other sequences in the locus) may result in similar but not identical outcomes. Collectively our results suggest that HLB formation results from the contribution of many molecular interactions, and the loss of any single one may be overcome by other multivalent interactions within the body.

MATERIALS AND METHODS

Culture condition and fly strains

Original fly strains and crosses were used as in McKay *et al.* (2015). Stocks were maintained on standard corn medium. Viability studies were performed as in Penke *et al.* (2016).

Transgenic histone gene array construction

Construction of the 5-kb histone repeat designed for this study was performed using the HiFi DNA Assembly system from New England Biolabs. PCR amplification of fragments from existing histone repeats, in addition to in vitro synthesized gblocks (IDT) containing the desired nucleotide changes, was employed for the building blocks of the reaction. Manufacturer's protocol was followed with slight modification to the incubation time of the reaction. Engineered 5-kb histone repeats were then arrayed to 12 copies in pMultiBac (McKay *et al.*, 2015; Meers *et al.*, 2018). All histone gene arrays were inserted into the VK33 attP site on chromosome 3L (65B2) (Venken *et al.*, 2006).

Northern analysis

Northern analysis was performed using a 7 M 6% urea acrylamide gel to resolve histone mRNAs; ~1 µg of RNA from embryos or larvae was used with a radio-labeled probe to the coding region of H3 as in Lanzotti *et al.* (2002).

Histone expression analysis

Total RNA was prepared in Trizol and cDNA synthesized with random hexamers using Superscript II (Invitrogen), according to the manufacturer's instructions. RT-RCR was performed using gene-specific primers to H2a (McKay *et al.* 2015) and H3. Each reaction was performed at least three times with similar results. PCR products were digested using *Xho*I (H2a) or *Sac*I (H3). Digested amplicons were run on an 8% acrylamide gel or a 1.5% agarose gel.

Immunofluorescence

We used primary antibodies at the following concentrations: rabbit anti-CLAMP (1:1000), guinea pig anti-Mxc (1:2000), guinea pig anti-Mute (1:5000), rabbit anti-FLASH (1:2000), rabbit anti-Lsm10 (1:1000), mouse anti-MPM-2 (1:100; Millipore), rabbit anti-GAF (1:1000), and mouse anti-LacI (1:1000; Millipore). We used Alexa Fluor secondary antibodies (Thermo Fisher Scientific) at a concentration of 1:1000. In situ probes were detected using 15 µg/ml streptavidin-DyLight-488 (Vector Laboratories). Salivary gland dissections and squashes were performed as in Tatomer *et al.* (2016). Images were acquired using a Zeiss Lsm710 with ZEN DUO software. Images were analyzed using ImageJ. Ectopic HLBs in embryos were quantified as Salzler *et al.* (2013).

Fluorescent in situ hybridization

FISH probes were made from a PCR product that spanned the entire wild-type histone repeat (primers AAAGGAGGTTGGTAGGCAGC and ACGCTAGCGCTTTATCTGCA). Biotinylated FISH probes were made with nick translation by incubating 1 µg of purified PCR product for 2 h at 15°C in a total of 50 µl containing 1× DNAPolI buffer (Fisher Optizyme); 0.05 mM each of 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 DNAPol 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, and 0.8 mg/ml salmon sperm DNA) to a final volume of 220 µl. FISH probes were diluted in hybridization mixture and added to slides containing salivary gland polytene chromosome spreads before heating. We added a coverslip, sealed it with rubber cement, and heated the slide for 2 min on a 91°C heat block. Slides were placed in a humid box and incubated at 37°C overnight. Immunostaining was then performed by incubating the slides in primary antibody overnight at 4°C in a humid box.

Chromatin immunoprecipitation

ChIP was performed according to Blythe and Wieschaus (2015) with the following particulars: 2- to 4-h-old embryos (~200 embryos for each of three biological replicates) from *HisC/HisC*; $12x^{DWT}/12x^{DWT}$

or $12x^{PR}/12x^{PR}$ strains were collected, fixed, and stored at −80°C until further processing. Immunoprecipitation of chromatin preparations was performed using rabbit anti-CLAMP antibody (Rieder *et al*., 2017) and a polyclonal rabbit IgG (Millipore-Sigma, 12-370). qPCR was performed as described in Urban *et al*. (2017). Three biological replicates and two technical replicates were used for each genotype. We synthesized libraries using 4 µl of inputs and 36 µl of each ChIP sample (CLAMP or IgG) according to Blythe and Wieschaus (2015). We made libraries using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs), pooled 12 samples per lane, and sequenced libraries on an Illumina HiSeq 2500 in 2 × 150-bp mode (Genewiz). We normalized CLAMP IP target abundance against IgG IP target abundance. We analyzed data using a Student's *t* test, comparing target abundance between the DWT and PR lines.

Analysis and quantification of ChIPseq data

The 150-bp paired-end sequencing reads were quality checked using Fastqc (version 0.11.8). ILLUMINA adapter sequences were trimmed from the reads using the BBDuk tool in the BBtools package (BBmap version 38.82). Due to the presence of contaminating reads from yeast, mouse, and human genomes, the Bbsplit tool was used to obtain reads that mapped to the *Drosophila* genome (BBmap version 38.82). Custom reference genomes were created by removing all RD histone gene repeat sequences from genome version dm3 and by adding back a single 5-kb histone gene repeat unit with either DWT or PR sequences (McKay *et al*., 2015). The reads were mapped to the custom genomes using Bowtie2 (version 2.3.4.1) using default parameters (Langmead and Salzberg, 2012; Langmead *et al*., 2019). We used MACS2 (version 2.1.2) to identify peaks with the following parameters: -g dm –keep-dup all -B (Zhang *et al*., 2008). The resulting bedGraph files were visualized using Integrative Genomics Viewer. For quantitation analysis, reads overlapping with specified coordinates, chrX:3755103-3755605 (roX1), DWT:3575-4077, PR:3575-4077 (H3-H4), DWT:1797-2299, PR:1797-2299 (H2A-H2B), DWT:1811-2313, PR:1811-2313 (H1), were counted using SAMtools (version 1.10) using -c -F 260 options (Li *et al*., 2009). The reads for specified regions were normalized to total reads and then normalized to input. ChIP-seq data are deposited with NCBI Gene Expression Omnibus (GEO) under accession number GSE150297.

ACKNOWLEDGMENTS

We thank Erica Larschan's lab at Brown University for assistance with ChIP experiments and John Atwater for technical assistance. This work was supported by National Institutes of Health grant R01GM058921 to R.J.D and W.F.M and F32GM109663, K99HD092625, and R00HD092625 to L.E.R.

REFERENCES

Alberti S, Gladfelter A, Mittag T (2019). Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. Cell 176, 419–434.

Altmeyer M, Neelsen KJ, Teloni F, Pozdnyakova I, Pellegrino S, Grofte M, Rask MB, Streicher W, Jungmichel S, Nielsen ML, Lukas J (2015). Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). Nat Commun 6, 8088.

Banani SF, Lee HO, Hyman AA, Rosen MK (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18, 285–298.

Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, Rosen MK (2016). Compositional control of phase-separated cellular bodies. Cell 166, 1–13.

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

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

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.

Dellaire G, Eskiw CH, Dehghani H, Ching RW, Bazett-Jones DP (2006). Mitotic accumulations of PML protein contribute to the re-establishment of PML nuclear bodies in G1. J Cell Sci 119, 1034–1042.

Duncan IW (2002). Transvection effects in Drosophila. Annu Rev Genet 36, 521–556.

Dundr M (2011). Seed and grow: a two-step model for nuclear body biogenesis. J Cell Biol 193, 605–606.

Dundr M, Misteli T (2010). Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol 2, a000711.

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.

Falahati H, Pelham-Webb B, Blythe S, Wieschaus E (2016). Nucleation by rRNA dictates the precision of nucleolus assembly. Curr Biol 26, 277–285.

Fukaya T, Levine M (2017). Transvection. Curr Biol 27, R1047–R1049.

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, 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.

Gomes E, Shorter J (2019). The molecular language of membraneless organelles. J Biol Chem 294, 7115–7127.

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.

Gunesdogan U, Jackle H, Herzig A (2014). Histone supply regulates S phase timing and cell cycle progression. Elife 3, e02443.

Heyn P, Salmonowicz H, Rodenfels J, Neugebauer KM (2017). Activation of transcription enforces the formation of distinct nuclear bodies in zebrafish embryos. RNA Biol 14, 752–760.

Hur W, Kemp JP Jr, Tarzia M, Deneke VE, Marzluff WF, Duronio RJ, Di Talia S (2020). CDK-regulated phase separation seeded by histone genes ensures precise growth and function of Histone Locus Bodies. Dev Cell, https://doi.org/10.1016/j.devcel.2020.06.003.

Kuzu G, Kaye EG, Chery J, Siggers T, Yang L, Dobson JR, Boor S, Bliss J, Liu W, Jogl 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.

Landais S, D'Alterio C, Jones DL (2014). Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in Drosophila melanogaster. Cell Rep 7, 859–870.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359.

Langmead B, Wilks C, Antonescu V, Charles R (2019). Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics 35, 421–432.

Lanzotti DJ, Kaygun H, Yang X, Duronio RJ, Marzluff WF (2002). Developmental control of histone mRNA and dSLBP synthesis during Drosophila embryogenesis and the role of dSLBP in histone mRNA 3′ end processing in vivo. Mol Cell Biol 22, 2267–2282.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

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, Sunwoo H, Zhang B, Spector DL (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol 13, 95–101.
Marzluff WF, Koreski KP (2017). Birth and death of histone mRNAs. Trends Genet 33, 745–759.
Matera AG, Izaguire-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.
Meers MP, Leatham-Jensen M, Penke TJR, McKay DJ, Duronio RJ, Matera AG (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.
Mitrea DM, Kriwacki RW (2016). Phase separation in biology; functional organization of a higher order. Cell Commun Signal 14, 1.
Penke TJ, McKay DJ, Strahl BD, Matera AG, Duronio RJ (2016). Direct interrogation of the role of H3K9 in metazoan heterochromatin function. Genes Dev 30, 1866–1880.
Rajendra TK, Praveen K, Matera AG (2010). Genetic analysis of nuclear bodies: from nondeterministic chaos to deterministic order. Cold Spring Harb Symp Quant Biol 75, 365–374.
Rao BS, Parker R (2017). Numerous interactions act redundantly to assemble a tunable size of P bodies in Saccharomyces cerevisiae. Proc Natl Acad Sci USA 114, E9569–E9578.
Rieder LE, Koreski KP, Boltz KA, Kuzu G, Urban JA, Bowman SK, Zeidman A, Jordan WT 3rd, Tolstorukov MY, Marzluff WF, *et al*. (2017). Histone locus regulation by the Drosophila dosage compensation adaptor protein CLAMP. Genes Dev 31, 1494–1508.
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.
Shevtsov SP, Dundr M (2011). Nucleation of nuclear bodies by RNA. Nat Cell Biol 13, 167–173.
Shin Y, Brangwynne CP (2017). Liquid phase condensation in cell physiology and disease. Science 357, 1253.
Smith AV, King JA, Orr-Weaver T (1993). Identification of genomic regions required for DNA replication during Drosophila embryogenesis. Genetics 135, 817–829.
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.
Stanek D, Fox AH (2017). Nuclear bodies: news insights into structure and function. Curr Opin Cell Biol 46, 94–101.
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.
Urban JA, Doherty CA, Jordan WT 3rd, 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.
Venken KJ, He Y, Hoskins RA, Bellen HJ (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science 314, 1747–1751.
Wagner EJ, Burch BD, Godfrey AC, Salzler HR, Duronio RJ, Marzluff WF (2007). A genome-wide RNA interference screen reveals that variant histones are necessary for replication-dependent histone pre-mRNA processing. Mol Cell 28, 692–699.
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, 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.
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.
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 J, Hlavacek WS (2011). Scaffold-mediated nucleation of protein signaling complexes: elementary principles. Math Biosci 232, 164–173.
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.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137.
Zhao J, Dynlacht B, Imai T, Hori T, Harlow E (1998). Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S- phase entry. Genes Dev 12, 456–461.
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.