# Superresolution light microscopy of the *Drosophila* histone locus body reveals a core–shell organization associated with expression of replication–dependent histone genes

James P. Kemp Jr.[a], Xiao-Cui Yang[a], Zbigniew Dominski[a,b], William F. Marzluff[a,b,c,d], and Robert J. Duronio[a,c,d,e,*]

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

**ABSTRACT** **The histone locus body (HLB) is an evolutionarily conserved nuclear body that regulates the transcription and processing of replication-dependent (RD) histone mRNAs, which are the only eukaryotic mRNAs lacking a poly-A tail. Many nuclear bodies contain distinct domains, but how internal organization is related to nuclear body function is not fully understood. Here, we demonstrate using structured illumination microscopy that *Drosophila* HLBs have a "core–shell" organization in which the internal core contains transcriptionally active RD histone genes. The N-terminus of Mxc, which contains a domain required for Mxc oligomerization, HLB assembly, and RD histone gene expression, is enriched in the HLB core. In contrast, the C-terminus of Mxc is enriched in the HLB outer shell as is FLASH, a component of the active U7 snRNP that cotranscriptionally cleaves RD histone pre-mRNA. Consistent with these results, we show biochemically that FLASH binds directly to the Mxc C-terminal region. In the rapid S-M nuclear cycles of syncytial blastoderm *Drosophila* embryos, the HLB disassembles at mitosis and reassembles the core–shell arrangement as histone gene transcription is activated immediately after mitosis. Thus, the core–shell organization is coupled to zygotic histone gene transcription, revealing a link between HLB internal organization and RD histone gene expression.**

**Monitoring Editor**
Diane Lidke
University of New Mexico

Received: Oct 21, 2020
Revised: Mar 1, 2021
Accepted: Mar 10, 2021

## INTRODUCTION

The nucleus is a highly dynamic and yet well-organized cellular compartment containing many structures that have been studied using a variety of microscopic approaches. Among these are nuclear bodies (NBs), which are composed of nucleic acids and proteins that partition via liquid–liquid phase separation (LLPS) into relatively large, nonmembranous structures visible by light microscopy (reviewed in Mao *et al.*, 2011; Zhu and Brangwynne, 2015; Stanek and Fox, 2017; Sawyer *et al.*, 2018). Their formation is associated with a wide range of cell biological processes, including ribosome biosynthesis (Hernandez-Verdun *et al.*, 2010), proteolysis (Lafarga *et al.*, 2002), transcription and processing of specific RNAs (Nesic *et al.*, 2004; Strzelecka *et al.*, 2010; Duronio and Marzluff, 2017; Hur *et al.*, 2020), storage of different classes of RNAs and proteins (Galganski *et al.*, 2017; Fox *et al.*, 2018), gene pairing and silencing (Grimaud *et al.*, 2006; Lanzuolo *et al.*, 2007), stress responses (Sahin *et al.*, 2014; Scherer and Stamminger, 2016), and control of cell cycle progression (Tsai and Pederson, 2014). Despite this wide range of known processes, detailed understanding of the functional significance of NB formation, organization, and biophysical properties is lacking for most bodies. Many NBs contain internal substructure, with different factors localizing to distinct domains within the body (reviewed in [Sawyer *et al.*, 2019; Lafontaine *et al.*, 2021]). For example, mammalian nucleoli have a tripartite phase structure in

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E20-10-0645) on March 31, 2021.

*Address correspondence to: Robert J. Duronio (duronio@med.unc.edu).

Abbreviations used: HCC, histone cleavage complex; HLB, histone locus body; LLPS, liquid-liquid phase separation; NB, nuclear bodies; RD, replication dependent; SLBP, stem loop biding protein; SIM, structure illumination microscopy.

which the three domains function independently in the transcription and processing of rRNAs and assembly of ribosome subunits (Feric *et al.*, 2016; Lafontaine *et al.*, 2021). Other bodies such as paraspeckles and PML bodies have a "core–shell" organization in which some factors occupy a central core domain that is surrounded by a shell domain containing a distinct set of factors (Boisvert *et al.*, 2000; Hands *et al.*, 2014; West *et al.*, 2016; Fox *et al.*, 2018). Although the core–shell arrangement is thought to be functionally significant, whether it is a cause or a consequence of the biochemical processes associated with particular NBs is often unclear. Here we describe a core–shell organization of the *Drosophila* histone locus body (HLB) that is associated with the transcription and processing of replication-dependent (RD) histone mRNAs.

HLBs are nuclear bodies that form exclusively at clusters of RD histone genes in animal cells (Duronio and Marzluff, 2017). In mouse and human cells these clusters reside at two different genomic locations (Marzluff *et al.*, 2002), whereas *Drosophila melanogaster* contains a single large cluster on chromosome 2 comprising ~100 copies of a tandemly arrayed 5kb repeat unit containing each of the 5 RD histone genes (H1, H2a, H2b, H3, and H4; Lifton *et al.*, 1978; McKay *et al.*, 2015; Bongartz and Schloissnig, 2019). HLBs contain factors necessary for transcription and processing of RD histone mRNAs, which, unlike all other mRNAs, terminate in a conserved stem loop structure rather than a poly-A tail (Marzluff and Koreski, 2017). These factors include several that function only in histone mRNA biogenesis (Duronio and Marzluff, 2017). Generation of this specialized mRNA 3′ end requires factors that interact with cis elements within nascent RD histone pre-mRNA, including stem loop binding protein (SLBP), which binds the RD histone mRNA 3′ UTR stem loop, and U7 snRNP, which interacts with a sequence downstream of the stem loop. The Lsm11 subunit of U7 snRNP directly binds the N-terminal region of FLASH. This interaction forms a surface that recruits the histone cleavage complex (HCC), which contains the scaffolding protein Symplekin bound to CPSF73/CPSF100 nuclease and catalyzes an endonucleolytic cleavage four nucleotides downstream of the stem loop to generate the 3′ end of mature RD histone mRNA (Dominski *et al.*, 2005; Sabath *et al.*, 2013; Yang *et al.*, 2013; Skrajna *et al.*, 2017, 2018). The recently solved cryo-EM structure of these processing factors bound in their active configuration to RD histone pre-mRNA provided important mechanistic insight into RD histone mRNA 3′ end formation (Sun *et al.*, 2020). However, how these factors are organized within relatively large (~1 µm) HLBs in vivo, particularly relative to factors involved in RD histone gene transcription, is unknown. In addition, although some *Drosophila* HLB factors (e.g., U7 snRNP and FLASH) are constitutively localized to HLBs (Liu *et al.*, 2006; White *et al.*, 2007, 2011; Yang *et al.*, 2009), others such as the HCC are only recruited to the HLB when RD histone genes are expressed during the S phase (Tatomer *et al.*, 2014). Whether this dynamic recruitment depends on a particular biochemical organization or biophysical property of HLBs is an important outstanding question.

NBs and other biomolecular condensates often display properties consistent with liquid–liquid phase separation (LLPS) (Banani *et al.*, 2017; Shin and Brangwynne, 2017). We recently demonstrated that in *Drosophila* embryos the HLB undergoes LLPS, and that this behavior can be modeled based on phase separation of a single component (Hur *et al.*, 2020). The best candidate for this single component is Mxc, the *Drosophila* ortholog of human NPAT required for HLB assembly (White *et al.*, 2011). Mxc and NPAT only accumulate at RD histone genes and represent the most definitive marker of HLBs in animal cells (Ma *et al.*, 2000; Zhao *et al.*, 2000; White *et al.*, 2011; Kaya-Okur *et al.*, 2019). Mxc (1837 aa) and NPAT (1427 aa) are large proteins composed mostly of intrinsically disordered regions. Each protein is required for RD histone mRNA expression, yet neither has a predicted sequence-specific DNA binding domain, and they likely function as a scaffold for recruitment of transcription and processing factors needed for histone mRNA synthesis (Wei *et al.*, 2003; Ye *et al.*, 2003; Miele *et al.*, 2005; White *et al.*, 2011). Mxc contains a self-interacting region located at the N-terminus that includes a LisH domain responsible for Mxc oligomerization and HLB formation (Terzo *et al.*, 2015). FLASH is another large, intrinsically disordered protein in HLBs that is essential for histone pre-mRNA processing (Yang *et al.*, 2009). The C-terminus of NPAT binds directly to FLASH (Yang *et al.*, 2014), and in *Drosophila* the C-terminal region of Mxc is required for recruiting FLASH to the HLB and for efficient histone pre-mRNA processing (Terzo *et al.*, 2015; Tatomer *et al.*, 2016).

Despite the knowledge of HLBs gained from standard confocal microscopy, this approach does not provide sufficient resolution to examine how Mxc/NPAT and other factors are organized within the HLB. We hypothesized that these large NBs (0.5–1 µm in diameter for fly and mammalian HLBs) have substructure that is important for biological function. Specifically, we wished to determine whether histone mRNA biogenesis occurred within the HLB, or whether the HLB was adjacent to the histone genes. We addressed these issues using structured illumination microscopy and found that in early *Drosophila* embryos the HLB is configured in a core–shell arrangement with a different region of Mxc present in each domain. Nascent histone mRNA and RNA pol II are found in the central core domain and the essential pre-mRNA processing factor FLASH resides in the outer shell domain. These results reveal coincidence of HLB internal organization and histone gene expression and suggest that some processing factors might move between domains within the HLB.

## RESULTS

### Superresolution microscopy reveals substructure within histone locus bodies

To gain insight into the organization of HLBs, we first compared several different high-resolution light microscopy techniques to determine which could provide the maximal resolution with conventional preparation approaches and available detection reagents in order to best capture possible internal substructure. Because some microscopy techniques are sensitive to out-of-focus light and not compatible with thick samples, we focused our analyses on the *Drosophila* syncytial blastoderm embryo, where the nuclei are located at the surface. We imaged carefully staged cycle 14 embryos because their nuclei have activated zygotic histone gene expression and have a longer S-phase than earlier cycles, and thus contain abundant "active" HLBs (Hur *et al.*, 2020). The nuclei of these embryos are amenable to several different types of microscopy with high signal-to-noise ratio and can be situated relatively close to a coverslip, within 10–20 µm. We found using standard laser scanning confocal microscopy (reported by the manufacturer to give a lateral resolution of 200 nm) and an antibody against the C-terminal 169 amino acids of Mxc (MXC-C) (White *et al.*, 2011) that HLBs appeared primarily as roughly spherical blobs with areas of higher and lower intensity but without well-defined regions (Figure 1A). We next used Zeiss Airyscan microscopy (reported by the manufacturer to give a lateral resolution of 140 nm) and found that MXC-C signal accumulated in defined regions, in which we began to resolve a central region that was depleted of C-terminal Mxc staining (Figure 1B). Finally, we used structured illumination (SIM; manufacturer theoretical resolution of

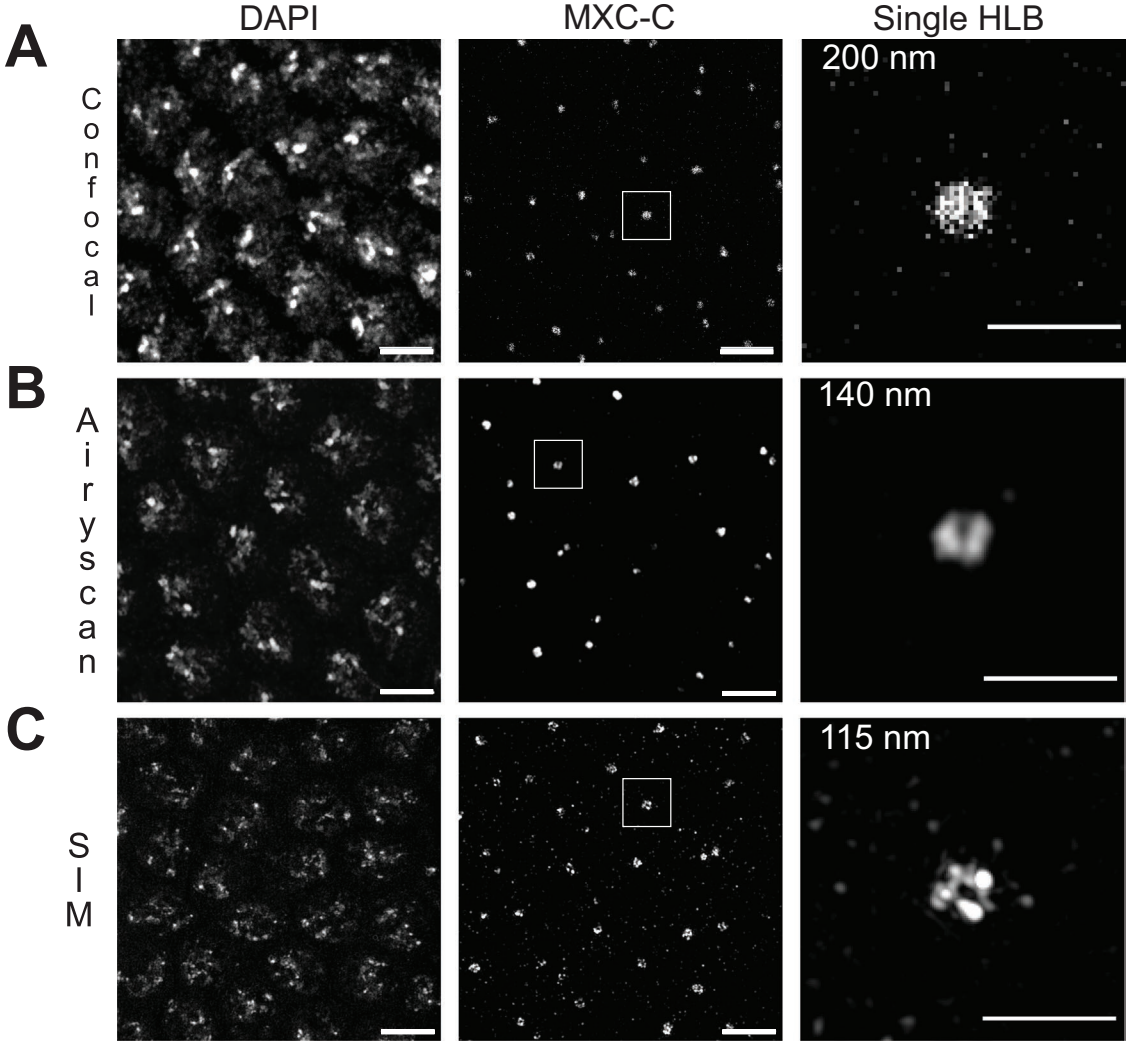


FIGURE 1: Increasing resolution of light microscopy reveals substructure within *Drosophila* HLBs. Maximum-intensity images of syncytial nuclear cycle 14 *Drosophila* embryos stained with antibodies that recognize the C-terminal 169 amino acids of the HLB factor Mxc (MXC-C) and DAPI. A single HLB is magnified in the right-hand column with the theoretical resolution of the microscope indicated in the upper left of each panel. (A) Standard laser scanning confocal image. (B) Airyscan processed confocal image. (C) SIM reconstructed image. Scale bars are 5 μm in the left and middle panels and 2 μm in the right panel.

115 nm) and found that MXC-C staining appeared as a shell with a "hole" in the middle (Figure 1C). Because SIM gave the best resolution, we used this technique to further elucidate the localization of factors within the HLB.

### The Mxc C-terminus and FLASH have a similar pattern of localization within the histone locus body

Mxc is a large protein (1837 aa) and thus has the potential for different domains to be resolved from one another by light microscopy. We therefore asked whether detection of the Mxc N-terminus would result in a structure like that obtained with the MXC-C antibodies. To do this we used flies expressing GFP-MXC from the endogenous *mxc* locus (Hur *et al.*, 2020). Note that in these flies all Mxc molecules are N-terminally tagged by GFP, and that their development and fertility is indistinguishable from those of wild type. Using SIM to image embryos from GFP-MXC flies revealed GFP signal in the center of the HLB, with MXC-C staining remaining unchanged and forming a "shell" around the GFP signal (Figure 2, A and B). We also used an antibody raised against the $NH_2$-terminal 178 amino acids of FLASH (FLASH-N; Yang *et al.*, 2009) and found an HLB staining pattern similar to that of MXC-C, namely a shell with a hole in the center that was filled with GFP signal (Figure 2, A–C). FLASH-N and MXC-C staining largely overlapped (Figure 2C). Line scans across many HLBs demonstrate that FLASH-N shows staining profiles more similar to MXC-C than to GFP-MXC (Figure 2D and Supplemental Figure 2). This is consistent with previous observations that the C-terminal portion of Mxc is required for FLASH recruitment to the HLB (Terzo *et al.*, 2015) and that the C-terminus of NPAT binds directly to FLASH (Yang *et al.*, 2014). Using an antibody that was raised against the C-terminal 178 amino acids of FLASH, we observed a shell structure similar to that obtained using the FLASH-N antibody (Figure 2E). Thus, unlike Mxc, most of the entire FLASH protein localizes to the HLB shell domain.

To quantify further the relative locations of GFP-MXC, MXC-C, and FLASH, we first used Imaris software to segment either the FLASH-N signal or the MXC-C signal into surfaces across 48 individual HLBs from different embryos. We then used the "Spots" function to localize peak signal intensities in 3D space for GFP-MXC,

MXC-C, and FLASH-N within these same individual HLBs (Figure 2F). After this segmenting, we measured the shortest distances from each of the peak intensities to the rendered surface for FLASH (Figure 2G) or the rendered surface for MXC-C (Figure 2H). From this analysis we found that the GFP-MXC peak intensities were primarily localized positive distances away from (>0.0 µm) the space enclosed by the FLASH-N or MXC-C rendered surfaces (Figure 2, G and H). Conversely, both the MXC-C and FLASH-N peak intensities were primarily located within (≤0.0 µm from) the space enclosed by either the MXC-C or FLASH-N surfaces (Figure 2, G and H). Note also that as would be expected from our analysis, all of the peak intensities for FLASH-N were located within (i.e., ≤0.0 microns from) the rendered surface of FLASH (Figure 2G), and all of the peak intensities MXC-C were located within (i.e., ≤0.0 µm from) the rendered surface of MXC-C (Figure 2H). These data are consistent with the Mxc N-terminus (GFP-MXC signal) occupying a distinct location within the center of the HLB relative to both FLASH and the Mxc C-terminus (MXC-C), which largely overlap.

To test whether the lack of MXC-C and FLASH staining within the center of the HLB was a result of the antibodies simply not penetrating the HLB interior, we used an antibody against GFP. We found that anti-GFP and GFP-MXC had similar staining patterns that were largely overlapping (Figure 2, I and J). We then segmented the anti-GFP signal into a "surface" and measured the distances from this surface to peak intensity spots from anti-GFP and GFP-MXC signals for 109 HLBs. Consistent with the line scans, these measurements revealed that the majority of GFP-MXC peak intensities were localized within the space enclosed by the anti-GFP surface (Figure 2K), indicative of penetration of the anti-GFP antibody into the interior of the HLB. These results establish that IgG molecules can access the HLB core in fixed embryos. Because GFP-MXC is fully functional, we conclude that the conjugated GFP provides an accurate representation of the location of the native Mxc N-terminus within the HLB.

The GFP-MXC and MXC-C data are consistent with a model in which the N- and C-terminal regions of Mxc, a large protein containing 1837 amino acids, predominantly reside in different regions of the HLB that can be resolved by light microscopy. To further test this model, we used CRISPR to engineer different versions of endogenous Mxc tagged at its C-terminal end. The first tag we tested was a fusion to the red fluorescent protein mScarlet. MXC-mScarlet signal often appeared as a doughnut or hollow shell within the HLB similar to MXC-C (Figure 3A). This signal also overlapped extensively with anti-RFP staining (Figure 3, A–C), indicating concordance between two different methods of detection and suggesting that the C-terminus of Mxc is depleted from the center of the HLB (Figure 3A). However, the SIM reconstructions we obtained from MXC-mScarlet were consistently more diffuse and less precise than the results of antibody labeling (Figure 3A), and sometimes the MXC-mScarlet signal lacked an obvious "hole" even when this was apparent with anti-RFP antibodies (Figure 3B). We quantified the overlap between the anti-RFP and mScarlet signals by segmenting the anti-RFP signal into a surface and then assigning peak intensity "spots" from both the anti-RFP and mScarlet signals and measuring the shortest distance from these spots to the anti-RFP surface rendering. All the anti-RFP "spots" were within (<0.0 µm) the space enclosed by this surface, as expected, and 97% of the mScarlet "spots" were also within the anti-RFP surface space (Figure 3C). These data suggest that even though the mScarlet results in more diffuse reconstructions, it is still highly enriched in the domain of the anti-RFP signal. We next tagged the C-terminus of endogenous Mxc with APEX2 enzyme (Lam *et al.*, 2015) and stained embryos simultaneously with anti-APEX2 and anti-FLASH antibodies. (Note that we are not using the enzymatic activity of APEX2 to biotinylate proteins in this experiment.) SIM imaging of these HLBs revealed anti-APEX2 signal in the same configuration as MXC-C and largely overlapping with FLASH-N (Figure 3D), consistent with the core–shell model. Finally, to test if we could resolve the two ends of Mxc using a similar method, we analyzed HLBs in embryos derived from GFP-MXC/MXC-mScarlet mothers. In theory, the HLBs in these embryos will contain a 1:1 mixture of each fluorescent protein. As described above, we found using FLASH-N antibodies that the FLASH signal formed a shell around the Mxc-GFP signal (Figure 3, E–G). We also detected nonuniformity in both the GFP-MXC and the MXC-mScarlet signals, including "holes" in the MXC-mScarlet signal. However, unlike with the MXC-C antibody, most of the MXC-mScarlet signal overlapped with the GFP-MXC signal and was within the FLASH-N shell (Figure 3, E–G). Thus, two different methods of detecting the Mxc C terminus (Ab staining and fusion to a fluorescent protein) give somewhat different results (see *Discussion*).

As a further test of the core–shell model and to avoid the caveat that our fixation conditions might not accurately represent HLB organization in vivo, we imaged HLBs in live embryos expressing only the MXC-mScarlet protein (Figure 4). Our goal was to determine whether we could detect "holes" in the mScarlet signal, which would indicate that the C terminus of Mxc was more enriched toward the outside of the HLB. For live embryo imaging, we are unable to use SIM and thus obtain the level of resolution that technique affords. We therefore imaged unfixed live cycle 13 (Figure 4A) or cycle 14 (Figure 4B) embryos using a Leica SP8 LIGHTNING microscope, which provides higher resolution than standard confocal microscopy. Using this approach, we observed nonuniformity of the MXC-mScarlet signal within HLBs, including "holes" toward the center (Figure 4, C–E). The mScarlet signal bleached quite rapidly during imaging (Figure 4C), and as the fluorescence waned these holes became more apparent, suggesting smaller amounts of MXC-mScarlet protein within the centermost regions of the HLB (Figure 4D). Taken together with the results from fixed embryos, our data indicate that the HLB of early *Drosophila* embryos assumes a core–shell configuration in which the Mxc N-terminus predominantly localizes to the central core domain, while FLASH predominantly localizes to the outer shell domain. The Mxc C-terminus is more enriched in the HLB shell than the core, with the MXC-C antibodies revealing this difference more clearly than MXC-mScarlet.

### The C-terminal region of Mxc binds directly to the C-terminal region of FLASH

We showed previously that the C-terminal region of FLASH is essential to localize FLASH to the HLB (Burch *et al.*, 2011). We also previously demonstrated using in vitro pull-down assays that the C-terminal 16 amino acids of the human NPAT protein bind directly to the SANT domain at the C-terminus of human FLASH (Yang *et al.*, 2014). Thus, one explanation for the enrichment of the Mxc C-terminus in the shell domain of the *Drosophila* HLB is that Mxc and FLASH interact directly. The MXC-C antibody immunoprecipitated FLASH from a nuclear extract derived from cultured Kc cells (Figure 5A, lane 4). To determine whether coimmunoprecipitation results from direct binding between FLASH and Mxc, we used a series of $^{35}$S-labeled in vitro translated proteins containing different regions of the C-terminus of Mxc and found that the last 296 amino acids of Mxc (aa 1542–1837) were the smallest fragment that efficiently bound a bacterially expressed GST fusion protein containing the last 178 amino acids of FLASH (FLASH178C; Figure 5B). To further map the determinants on Mxc that bind FLASH178C, we used a series of C- and N-terminal truncations of the Mxc-296 fragment in the GST pull-down assay

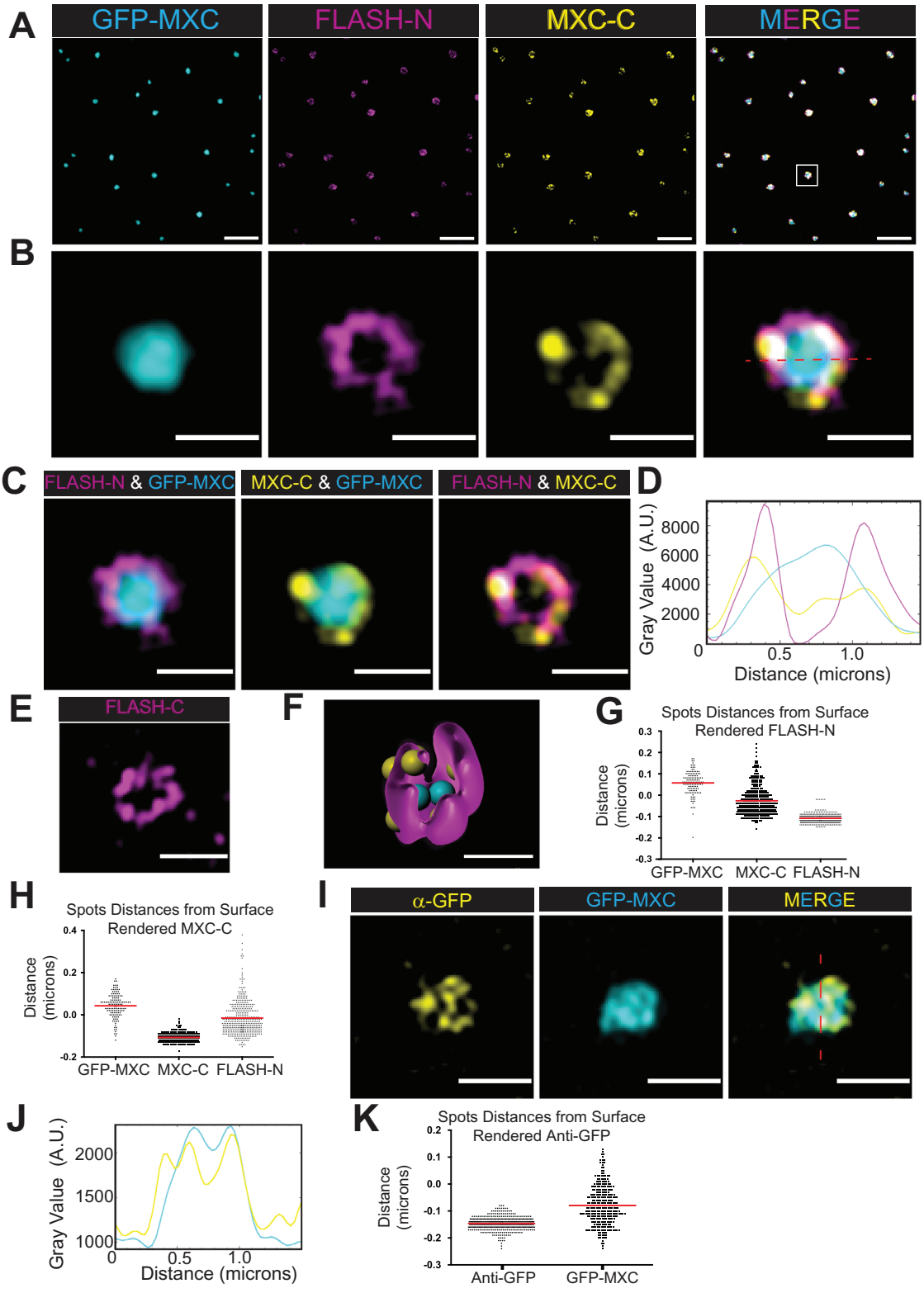


FIGURE 2: Mxc and FLASH visualization define a core–shell organization of the *Drosophila* HLB. (A) SIM image of the middle *Z*-section of a cycle 14 GFP-MXC *Drosophila* embryo with HLBs stained for FLASH-N and MXC-C. Scale bars are 5 μm. (B) High-magnification view of a selected HLB from the box in panel A merge. Scale bar is 1 μm. (C) HLB from B showing different pairs of fluorescent signals. (D) A representative line scan (dashed line in panel B merge) indicating relative level of GFP-MXC (cyan), FLASH-N (magenta), and MXC-C (yellow; *y*-axis) measured across a single HLB (*x*-axis). (E) Representative middle *Z*-section of an HLB stained for FLASH-C. Scale bar is 1 μm. (F) Image of a single HLB that was segmented using Imaris software with FLASH-N signal rendered as an enclosed surface (magenta) and the peak intensities of MXC-C (yellow), GFP-MXC (cyan), and FLASH-N (magenta) rendered as spheres. Scale bar = 1.0 μm. (G) Plot of the shortest distances from individual spots of peak intensities (each dot) to the rendered FLASH-N surface

(Figure 5, C and D). We mapped the FLASH interaction domain to between amino acids 1542 and 1745 of the full-length 1837 aa Mxc protein (Figure 5, D and E). Fragments between amino acids 1609 and 1837 or 1542 and 1686 bound FLASH178C with lower affinity. Interestingly, the *mxc*G46 nonsense mutation has a stop codon in the center of this domain that truncates the Mxc protein after 1643 aa (Figure 5, B and E) and results in a failure of FLASH to accumulate in HLBs (White *et al.*, 2011; Terzo *et al.*, 2015; Tatomer *et al.*, 2016). Removal of the last 133 amino acids of FLASH also results in failure of FLASH to accumulate in HLBs (Burch *et al.*, 2011; Tatomer *et al.*, 2016). These data demonstrate that FLASH binds the same region of Mxc that is enriched in the HLB shell domain, consistent with our SIM microscopy data indicating that FLASH localizes to the HLB shell.

### Transcription occurs in the core domain of the histone locus body

Because the N-terminus of NPAT and Mxc are necessary for expression of RD histone genes (Wei *et al.*, 2003; Terzo *et al.*, 2015), we determined whether the nascent histone RNA and active RNA Pol II also localized to the centers of HLBs. We have observed nascent histone transcripts previously (Lanzotti *et al.*, 2002, 2004; Tatomer *et al.*, 2016). To precisely localize these nascent transcripts relative to HLBs, we performed RNA FISH with a probe derived from the ~400nt histone H3 coding region (Hur *et al.*, 2020) in combination with FLASH-N staining of cycle 14 wild-type embryos, reasoning that this approach would give us the best opportunity to see structure within a HLB that is actively transcribing histone mRNA. SIM imaging of these embryos revealed enrichment of H3 FISH signal in the HLB core surrounded by FLASH staining (Figure 6, A and B). Because nascent histone H3 mRNA was enriched in the HLB core, we next tested whether active RNA Pol II also was enriched there by staining GFP-MXC cycle 14 embryos (Figure 6C) with antibodies that recognize phospho-Ser5 on the pol II C-terminal tail (P-Pol II). Others have shown previously with conventional confocal microscopy of cultured S2 cells and nurse cells in the ovary that Pol II is enriched at *Drosophila* histone genes and can be detected at HLBs (Isogai *et al.*, 2007; Nizami *et al.*, 2010; Guglielmi *et al.*, 2013). We found using SIM imaging that the P-Pol II signal is enriched in the center of the HLB like the GFP-MXC signal (Figure 6D). In the same embryos, FLASH-N largely surrounded the P-Poll II signal, although a portion of the P-Pol II staining also overlapped FLASH-N in the outer shell as did nascent H3 mRNA (Figure 6, D and E). We conclude from these data that transcriptionally active RD histone genes are localized primarily within the core of the *Drosophila* HLB.

### Histone locus bodies attain a core–shell organization coincident with transcription

When examining populations of younger embryos, we noticed variance in HLB staining patterns with different antibodies as these embryos progressed through S phase and mitosis of the syncytial nuclear cycles. The SIM staining patterns were consistent with live imaging of GFP-MXC in these embryos, in which we observed HLB disassembly at each mitosis and reassembly during the subsequent interphase (Terzo *et al.*, 2015). In addition, early during the S phase of each cycle, the HLB rapidly grows and attains a particular steady state size that is dependent on the number of histone genes, which seed HLB assembly (Hur *et al.*, 2020). Using SIM imaging of nuclear cycle 10, which is the earliest cycle where we can consistently observe focal staining for FLASH, we observed that the P-Pol II was mostly diffuse and lacked an obvious accumulation within FLASH-N foci that was distinct from the overall staining within the nucleus (Figure 7A). We first observe P-Pol II puncta that clearly overlap with FLASH-N foci in nuclear cycle 11 (Figure 7A). This result is consistent with zygotic RD histone gene activation in cycle 11 when fully mature HLBs first appear (Edgar and Schubiger, 1986; White *et al.*, 2007, 2011; Terzo *et al.*, 2015; Hur *et al.*, 2020). During cycles 12 and 13, we observed P-Pol II puncta that were larger and more pronounced than the diffuse nuclear P-Pol II puncta (Figure 7A). These large P-Pol II puncta overlap with FLASH-N puncta and represent HLBs (Figure 7A, NC13). At higher magnification, we detected the characteristic core–shell HLB configuration in cycles 11–13 with FLASH-N in the shell and P-Pol II within the core, but also overlapping with FLASH (Figure 7B). In cycle 10,, the shell configuration of FLASH-N was absent, and the small focus of FLASH-N lacked high levels of P-Pol II (Figure 7B). Thus, the FLASH-N foci observed during cycle 10 likely represent immature "proto-HLBs" that are not associated with RD histone gene expression (Salzler *et al.*, 2013). Taken together, these data suggest that the core–shell HLB organization correlates with transcriptionally active RD histone genes.

To examine this possibility further, we imaged at different times during cycle 14, which has a longer S phase and total cycle length relative to earlier cycles (Figure 7C). Early during cycle 14 when S phase occurs, we observed three classes of staining (Figure 7D): 1) small puncta of FLASH-N lacking the shell configuration and without high levels of P-Pol II; 2) larger FLASH-N puncta with a clear but less well-defined shell configuration overlapping P-Pol II; and 3) large FLASH-N puncta in a clearly defined shell configuration with high levels of P-Pol II in the core surrounded by FLASH-N. The organization of these three types of HLBs and the localization of P-Pol II within them mirrors the staining of nuclear cycles 10–13 as zygotic RD histone transcription is initiated (compare Figure 7B, D). Thus, because transcription is aborted during mitosis (Shermoen and O'Farrell, 1991) and reinitiated upon entry into the subsequent interphase, we interpret the different HLB configurations within cycle 14 as reflecting rapid assembly of the HLB after mitosis, together with the activation of RD histone gene expression when S phase begins immediately after the completion of mitosis 13. These data also support our conclusion that the core–shell configuration represents an "active" *Drosophila* HLB in which RD histone mRNA synthesis is occurring within the core domain.

## DISCUSSION

Here we used superresolution light microscopy to examine HLB structure in *Drosophila* embryos. We conclude from our data that when RD histone genes are expressed, the HLB is arranged into a core–shell configuration like that observed in several other nuclear bodies (Woulfe *et al.*, 2007; Lang *et al.*, 2010; Hands *et al.*, 2014; West *et al.*, 2016; Wheeler *et al.*, 2016; Nakano *et al.*, 2017; Sawyer *et al.*, 2019; Lafontaine *et al.*, 2021). In addition, the core–shell configuration is most apparent when histone genes are actively expressed, and thus represents "active" HLBs.

for GFP-MXC, MXC-C, and FLASH-N across 48 HLBs. (H) Plot of distances from the same spots in the same HLBs as in F but using the surface rendering for MXC-C. (I) Representative anti-GFP staining (left, yellow) of a single HLB from a cycle 14 GFP-MXC embryo, compared with visualizing the GFP (center, cyan) and the merged image (right). (J) Graph of line scan from dotted line in panel H merge. Note that the GFP-MXC (cyan) and anti-GFP (yellow) signals largely overlap. (K) Plot of distances of GFP-MXC and anti-GFP maximum–signal intensity spots to the anti-GFP rendered surface.

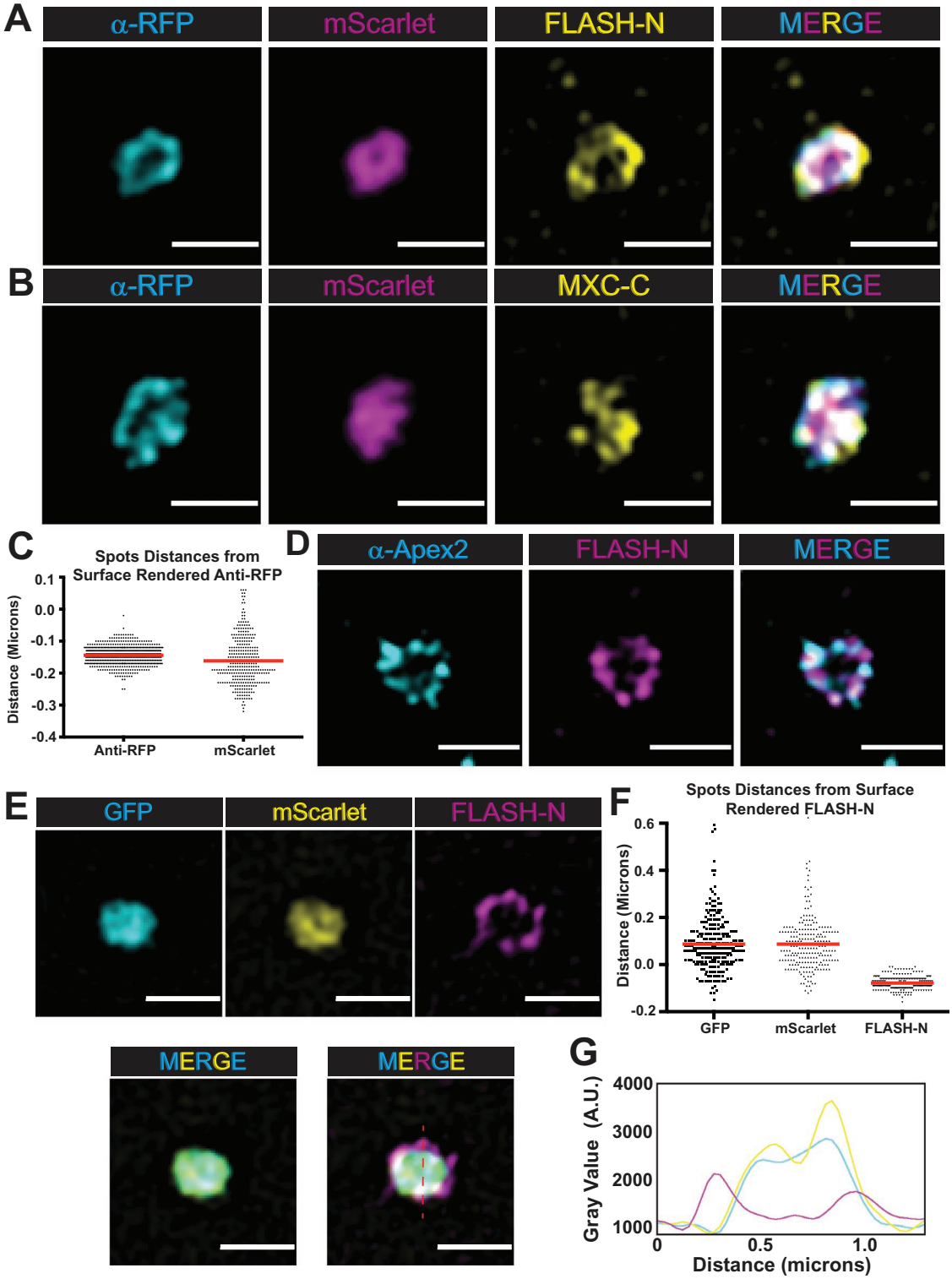


FIGURE 3: Protein tags at the C-terminus of MXC localize in the outer shell domain of the *Drosophila* HLB. (A) Anti-RFP, mScarlet fluorescence, and anti-FLASH-N staining of a cycle 14 HLB. (B) Anti-RFP, mScarlet fluorescence, and MXC-C staining of a cycle 14 HLB. (C) Plot of distances from anti-RFP and mScarlet spots to anti-RFP surface rendered HLBs. (D) Representative anti-Apex2 and FLASH-N staining of a selected HLB from MXC-Apex2 embryos. All images are the middle *Z* section of the HLB. (E) Representative anti-FLASH-N, GFP-MXC, and MXC-mScarlet fluorescence of a cycle 14 HLB. GFP-MXC and MXC-mScarlet and all three are also shown as merged images. (F) Plot of distances from anti-FLASH-N, GFP-MXC, and MXC-mScarlet spots to anti-FLASH-N surface-rendered HLBs. (G) A representative line scan (dashed line in panel E merge) indicating relative levels of GFP-MXC (cyan), FLASH-N (magenta), and MXC-C (yellow; *y*-axis) measured across a single HLB (*x*-axis). Scale bars are 1 μm.

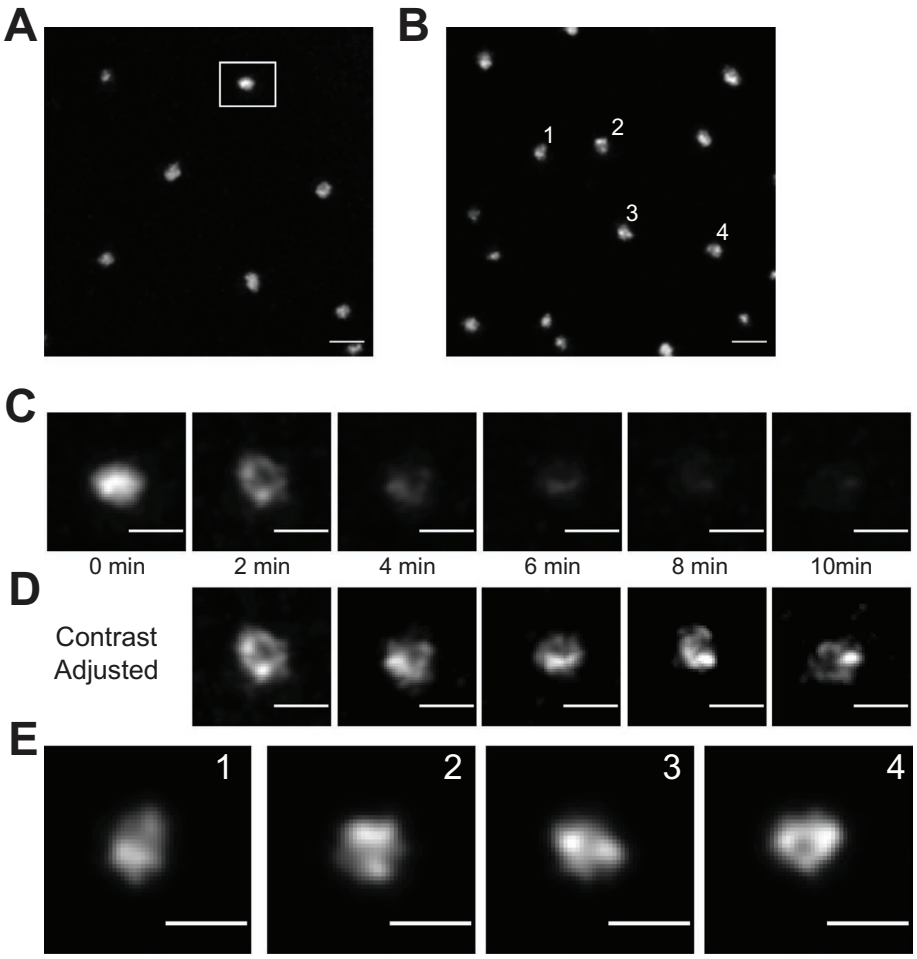


**FIGURE 4:** Live imaging demonstrates nonuniform distribution of MXC-mScarlet within HLBs. (A) Maximum-intensity projection of the first time point of a midnuclear cycle 13 MXC-mScarlet embryo. Scale bar 5 µm. (B) Maximum-intensity projection of the first time point of a late nuclear cycle 14 MXC-mScarlet embryo. Scale bar 5 µm. (C) Maximum-intensity projection of the labeled time points of a single HLB boxed in panel A demonstrating the rapid photo bleaching of Mxc-mScarlet. No thresholding was performed. (D) Contrast adjusted time points from C showing that "holes" become more apparent after some bleaching occurs. (E) The middle *z* plane of the four HLBs indicated in panel B, demonstrating the nonuniform distribution of MXC-mScarlet within HLBs including the decrease in signal at their centers.

The core–shell HLB configuration is based on immunofluorescence of fixed embryos using antibodies against the HLB factors Mxc and FLASH, as well as analysis of three different CRISPR-engineered alleles encoding Mxc proteins tagged at either the $NH_2$ or COOH terminus. Visualizing HLBs in fixed diploid cells of the early fly embryo using antibodies against Mxc and FLASH and conventional confocal microscopy results in uniformly stained spherical HLBs in which Mxc and FLASH staining are coincident (White *et al.*, 2007, 2011). This type of staining is also apparent when live embryos expressing GFP-Mxc are imaged (Terzo *et al.*, 2015; Hur *et al.*, 2020). In marked contrast, these two proteins can be resolved from one another using SIM, with a shell of FLASH surrounding a core of Mxc. This organization was most apparent when GFP-Mxc was visualized in conjunction with either of two different antibodies recognizing N- and C-terminally located epitopes on the FLASH protein.

Our SIM data also suggest that the $NH_2$ and COOH regions of Mxc can occupy distinct regions within HLBs, although here the data are more complex. These two regions of Mxc are most clearly resolved by SIM when GFP fused to the $NH_2$ terminus of Mxc is used (i.e., either GFP-Mxc fluorescence or anti-GFP antibodies are used to recognize GFP-Mxc) together with an antibody raised against the last 169 residues of the 1837–amino acid Mxc protein. The signal from this Mxc antibody overlapped extensively with the signal obtained using FLASH antibodies or with antibodies recognizing APEX2 when it was fused to the COOH terminus of Mxc. However, fusing mScarlet to the COOH terminus of Mxc results in a signal that substantially overlaps with GFP-Mxc, although we can detect depletion of the Mxc-mScarlet signal from the center of the HLB in both fixed and live embryonic cells. Thus, antibodies and fluorescent fusion proteins give slightly different results when detecting the COOH terminal end of the Mxc protein. The length of the two IgG molecules used in our primary antibody bound by labeled secondary antibody detection scheme is ~17 nm (Weber *et al.*, 1978). This distance is smaller than the maximal distance of ~200 nm we measured from the GFP-Mxc signal peaks to the MXC-C signal peaks in line scans (Figure 2D and Supplemental Figure 2) or the mean distance of ~40 nm from GFP-Mxc spots to the surface-rendered MXC-C, which is an underestimate of the true distance between peak signals (Figure 2H). Therefore, it is unlikely that antibody detection artifactually "extends" the Mxc COOH signal farther to the outside of the HLB. We also cannot rule out that the Mxc fusion proteins subtly distort the normal organization of the HLB, even though all three of these CRISPR alleles are viable and fertile when homozygous. Consequently, some caution must be applied when various means of detecting proteins within nuclear bodies by superresolution light microscopy are used. Nevertheless, in aggregate, our results suggest that the N-terminus of Mxc resides within the HLB core, while the C-terminus is more closely associated with the shell, as defined by the location of FLASH. This interpretation of the SIM data is consistent with our biochemical observation that the C-terminal region of Mxc binds directly to FLASH. We note that this proposed arrangement of Mxc is similar to an observation made by examining NEAT1 lncRNA in paraspeckles, which also organize into a core–shell configuration. In this case, the central region of NEAT1 lncRNA resides in the paraspeckle core domain, whereas the 5′ and 3′ ends of NEAT1 are localized to the shell domain and may coordinate the processing of primary miRNAs (Souquere *et al.*, 2010; West *et al.*, 2016; Jiang *et al.*, 2017).

Several distinct observations allow us to conclude that the core–shell HLB configuration correlates with actively transcribed RD histone genes. First, the initial appearance of the core–shell configuration during embryonic development occurs precisely when zygotic RD histone gene expression begins. Second, both active RNA polymerase II and nascent histone H3 transcripts are enriched in the HLB core domain relative to the shell domain, demonstrating that histone

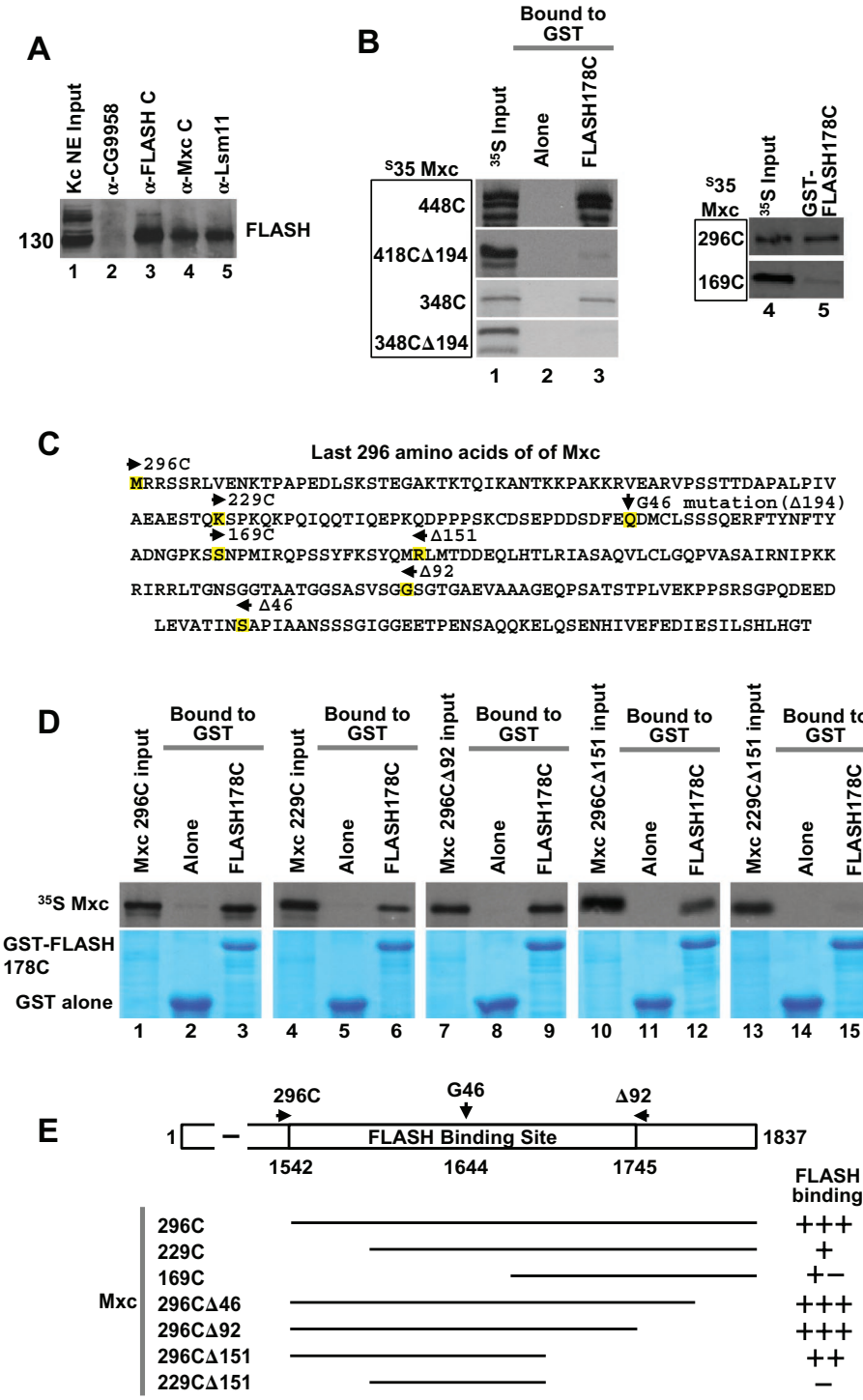


FIGURE 5: A C-terminal portion of FLASH interacts with a C-terminal portion of Mxc. (A) Immunoprecipitation of nuclear extract from cultured Kc cells s using the indicated antibodies. Anti-CG9958 (Snapin) serves as a negative control, and anti-Lsm11, a component of the U7 snRNP that directly binds FLASH, serves as a positive control. The immunoprecipitants were resolved by SDS-gel electrophoresis and analyzed by Western blotting with the antibody to the C-terminal region of FLASH. Lane 1 contains 5% of the extract used for each IP lane. (B) The indicated C-terminal fragments of Mxc were labeled with $^{35}$S methionine by in vitro translation and incubated with GST alone or GST fused to the 178 C-terminal amino acids of FLASH (FLASH178C). The bound proteins were detected by autoradiography. The proteins analyzed were as follows: lane 1: input: 448C (the C-terminal 448 aa of Mxc), 418XΔ194, 348C, 384CΔ194; lane 2: bound to GST; lane 3: bound to FLASH178C; lane 4: input; 296C, 169C; lane 5: bound to FLASH178C. The input (lanes 1,4) was 20% of the amount analyzed in lanes 2, 3, and 5. (C) The sequence of the last 296 amino acids of Mxc (starting just after the AT-hook) is shown. The starts (forward arrows) or ends (reverse arrows) of the Mxc fragments analyzed for binding FLASH178C are indicated. Also shown is the position of the *mxc*G46 nonsense mutation (Q to stop codon). This protein supports formation of HLBs that do not contain FLASH (Terzo *et al.*, 2015; Tatomer *et al.*, 2016). (D) The indicated labeled Mxc fragments were analyzed as in panel B. Input: lanes 1, 4, 7, 10, 13; bound to GST: lanes 2, 5, 8, 11, 14; bound to FLASH178C: lanes 3, 6, 9, 12, 15. Below is shown the stained gel of the binding assay. The input was 20% of the amount used for binding. (E) Schematic of the region of Mxc required for binding the C-terminus of FLASH. We mapped the greatest affinity FLASH binding domain to between amino acids 1542 and 1745 of Mxc. Below the top schematic is a diagrammatic list of the subfragments of the Mxc296C protein that were tested for binding to FLASH178C. To the right of this list is a summary of the relative binding of these fragments to FLASH178C with +++ > + > + - > -.

mRNA biosynthesis occurs within the core. Third, during cycle 14, we observe that the HLB transitions from a small body lacking serine 5–phosphorylated RNA Pol II immediately after mitosis 13 to a larger body later in S phase 14 with the core–shell configuration and high levels of nascent transcripts in the core. This observation is consistent with cessation of transcription and disassembly of the HLB at each syncytial mitosis (White *et al.*, 2007; White *et al.*, 2011; Terzo *et al.*, 2015; Hur *et al.*, 2020), a general feature of nuclear bodies (Rai *et al.*, 2018), followed by reactivation of histone gene transcription during the subsequent interphase, which lacks a G1 phase. The small bodies we observed shortly after the completion of mitosis 13 are reminiscent of the incomplete "proto-HLBs" containing small amounts of Mxc and FLASH that we previously observed in cycle 10 before histone gene transcription or when histone transcription is blocked (Salzler *et al.*, 2013; Hur *et al.*, 2020). These observations are consistent with a "seed and grow" model of nuclear body assembly (Dundr, 2011), as also described for the nucleolus in early fly embryos (Falahati *et al.*, 2016; Falahati and Wieschaus, 2017), with the histone genes providing the seed for the HLB (Hur *et al.*, 2020). In addition, a dependency between transcription and HLB assembly has been observed in fish embryos and is thus evolutionarily conserved (Heyn *et al.*, 2017; Arias Escayola and Neugebauer, 2018).

We were surprised by our observation that the essential histone pre-mRNA

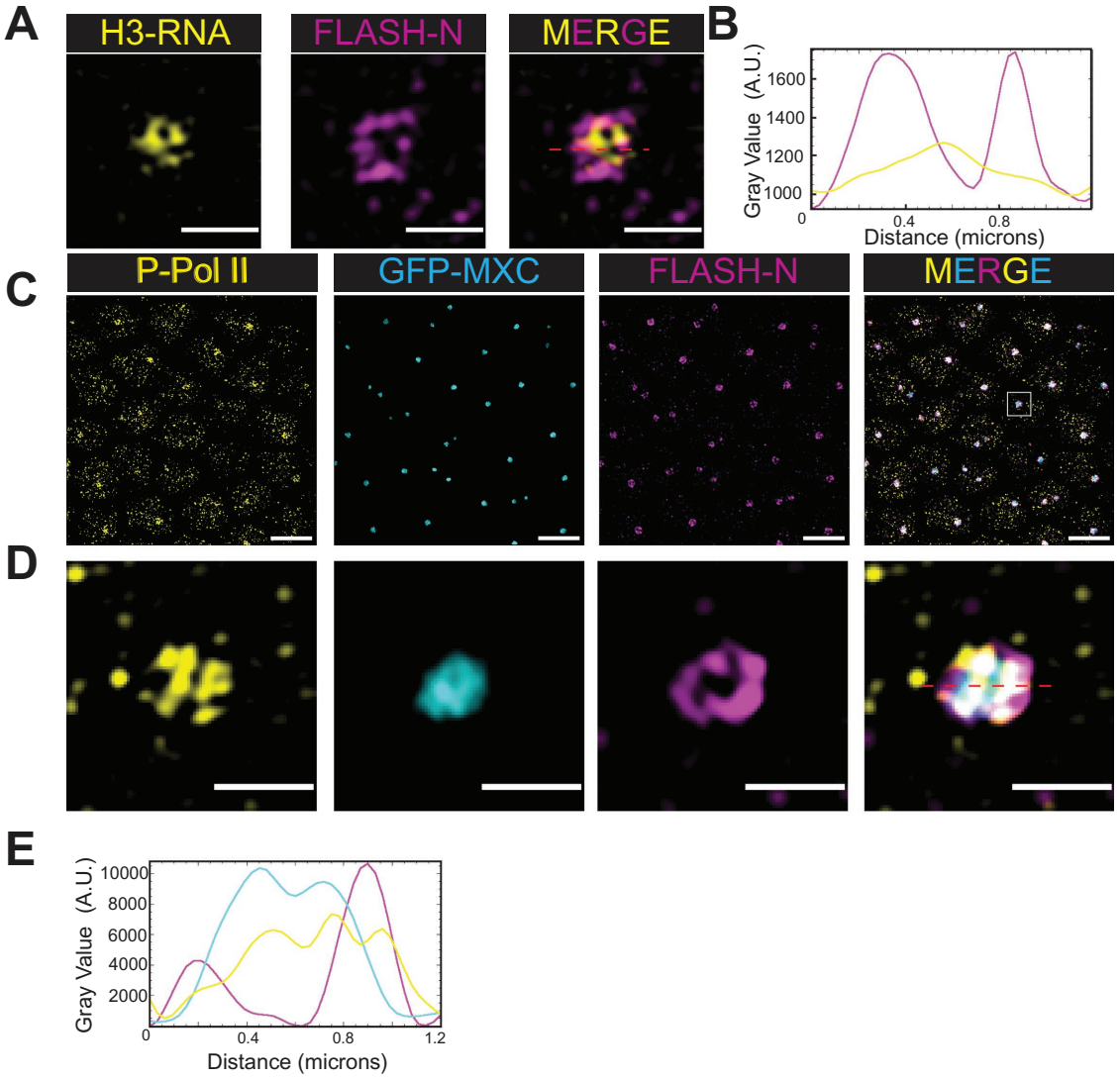


**FIGURE 6:** Histone mRNA and active Pol II are enriched in the *Drosophila* HLB core. (A) Histone H3 RNA FISH (left, yellow) and FLASH-N staining (middle, magenta) from a representative HLB of a cycle 14 embryo. The majority of the H3 nascent transcripts are in the center of the HLB. Scale bar 1 μm. (B) Line-scan graph of the fluorescent intensity of H3 RNA (yellow) and FLASH-N (magenta) across the red dotted line in the merged image in panel A. (C) Maximum-intensity projection of a cycle 14 GFP-MXC embryo (cyan) stained for P-Ser5-Pol II (P-Pol II, yellow) and FLASH-N (magenta). Note the large P-Pol II foci colocalizing with GFP-MXC and FLASH-N in nuclei with the single HLB in each cell because the homologous chromosomes are paired. Scale bars are 5 μm. (D) The middle *Z* plane of staining of a representative HLB, highlighted with a box in C. Scale bar 1 μm. (E) Line scan of the fluorescent intensity of P-Pol II (yellow), GFP-MXC (cyan), and FLASH-N (magenta) across the line drawn in the merged image in panel D.

processing factor FLASH is primarily localized to the HLB shell domain, unlike the nascent transcripts, which are enriched in the core domain. RD histone genes are small (less than 500 nts long), with the polymerase pausing after transcription of the 3′ processing signals (Adamson and Price, 2003) and transcription terminating close to the processing site (Tatomer *et al.*, 2016), resulting in tight coupling of transcription, pre-mRNA processing, and transcription termination, placing these activities in close physical proximity. Perhaps RD histone pre-mRNA processing occurs at the interface of the core and shell within the *Drosophila* HLB, analogous to the transcription and initial processing of rRNAs at the border between the FC core and the DFC shell domains of the nucleolus (Yao *et al.*, 2019; Lafontaine *et al.*, 2021). It is also possible that when RD histone mRNA is being synthesized during the S phase, only a small amount of U7 snRNP bound to FLASH is actively translocated from the shell to the core to bind and process the nascent pre-mRNA. The active U7 snRNP containing RD histone cleavage complex may be present in the HLB only during the S phase (Tatomer *et al.*, 2014). The bulk of FLASH located in the shell may play an important structural role in the HLB, as depletion of FLASH can attenuate HLB assembly (White *et al.*, 2011; Tatomer *et al.*, 2016). We note that a core–shell arrangement can arise independent of active gene transcription, as PML bodies have a well-developed core–shell organization but do not contain RNA in the core domain (Boisvert *et al.*, 2000; Hands *et al.*, 2014).

Our study of the HLB focused only on the situation where histone genes are constitutively active in the S phase, and before the introduction of gap phases to the cell cycle. In subsequent cell cycles,

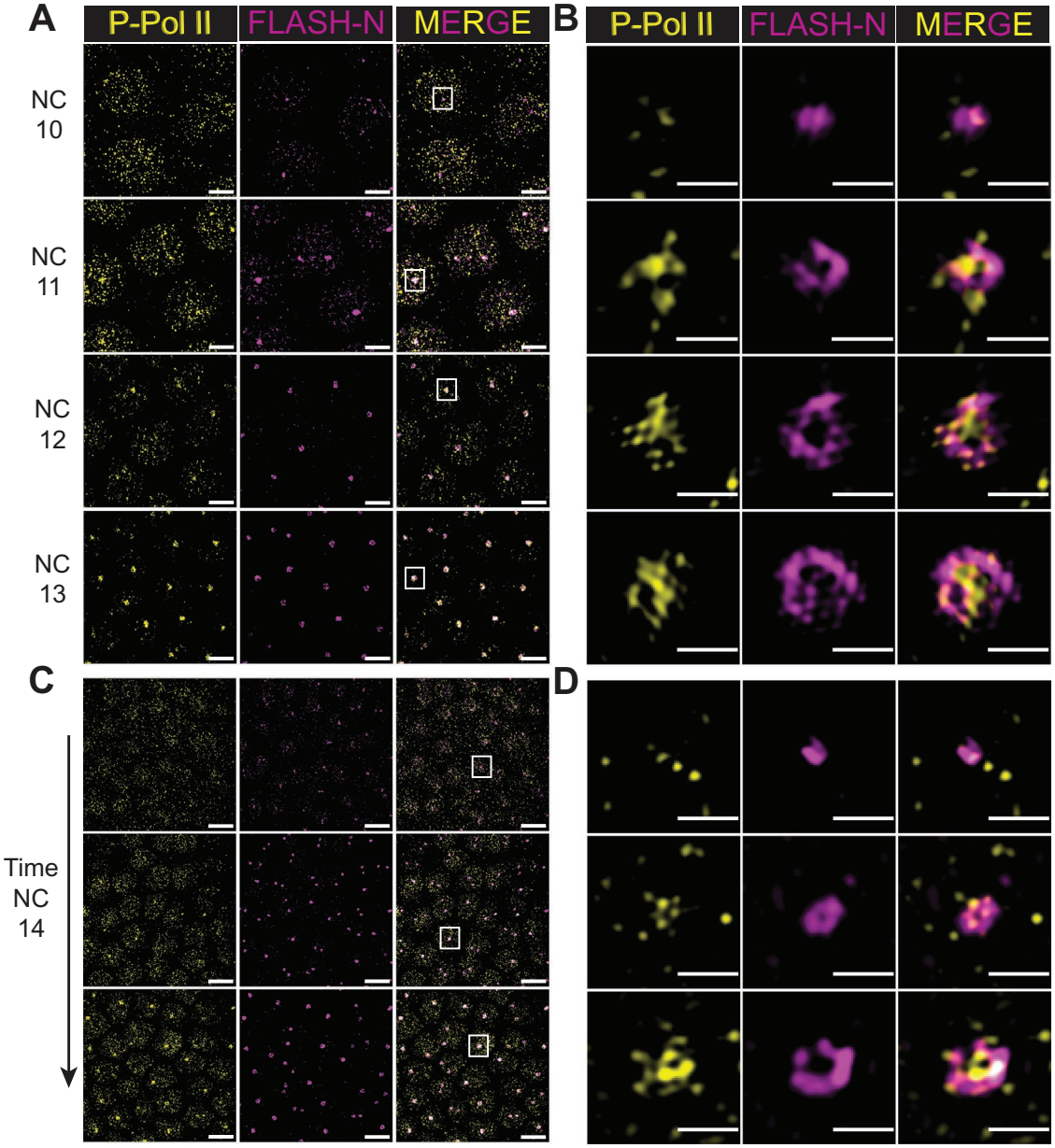


FIGURE 7: The *Drosophila* HLB core shell structure is dynamic and forms coincident with RD histone gene transcription. (A) Maximum-intensity projections of P-Pol II (yellow) and FLASH-N (magenta) staining of syncytial embryos from nuclear cycles 10–13. Scale bars 5 μm. (B) Higher-magnification images of the middle *Z*-plane from the foci highlighted with boxes in panel A. RNA Pol-2 is associated with the FLASH-N foci starting in cycle 11. Scale bars are 1 μm. (C) Maximum-intensity projections of P-Pol II and FLASH-N staining in embryos as they progress into nuclear cycle 14. Scale bars are 5 μm. (D) High-magnification images of the middle *Z*-plane from the HLBs highlighted with boxes in the merged image of panel C. Scale bars are 1 μm.

HLBs are present during the G1 phase, but there is no histone mRNA expression (Liu *et al.*, 2006; White *et al.*, 2007). Histone gene expression is activated as cells enter the S phase, and a critical step is the phosphorylation of Mxc by Cyclin E/Cdk2 (Lanzotti *et al.*, 2004; White *et al.*, 2007, 2011). Phosphorylation of Mxc by Cyclin E/Cdk2 may result in alteration of the structure of the HLB as part of the activation of histone gene expression, which likely involves both activation of transcription and activation of histone pre-mRNA processing by recruitment of the pre-mRNA cleavage module to U7 snRNP. Thus, we anticipate that the approaches described here will allow us to correlate changes in HLB structure during the cell cycle with histone gene expression.

## MATERIALS AND METHODS

Request a protocol through *Bio-protocol*.

### Antibodies

The primary antibodies are as follows:

Secondary antibodies used were Goat anti-Rabbit CF405M (Biotium), Goat anti-Rabbit Alexa 405 Plus, Goat anti-Guinea Pig Alexa Fluor 647, Goat anti-Rabbit Alexa Fluor 488, Goat anti-Guinea Pig Alex Fluor 555, Goat anti-Rat Alexa Fluor 488 (Invitrogen), Goat anti-Rat Cy3, Goat anti-Mouse Cy3, and Goat anti-Rabbit Cy5. Samples were mounted in Vectashield (Vector Labs) and sealed with CoverGrip (Biotium).

| Primary Antibody | Concentration | Reference/Manufacturer |
|---|---|---|
| Rabbit anti-MXC: C-terminal | 1:2000 | (White *et al.*, 2011) |
| Guinea pig anti-MXC: C-terminal | 1:8000 | (White *et al.*, 2011) |
| Guinea pig anti-*Drosophila* FLASH: N-terminal | 1:10000 | (White *et al.*, 2011) |
| Rabbit anti-*Drosophila* FLASH: C-terminal | 1:5000 | This Manuscript |
| Rat anti-phospho-serine5 PolII | 1:200 | Sigma-Aldrich Cat# MABE954 |
| Rabbit anti-GFP | 1:1000 | Abcam Cat# ab6556 |
| Rabbit anti-RFP | 1:500 | Rockland Cat # 600-401-379 |
| Rabbit anti-APX2 | 1:1000 | Thermo Fisher Cat# PA5-98320 |

### Buffer compositions

Embryo wash buffer: 120 mM NaCl, 0.07% TX-100.

Phosphate-buffered saline (PBS): 140 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.4

(PBST): 140 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.4, 0.1% Triton X-100.

### Other reagents

Embryos were fixed in 4% formaldehyde in PBST that was diluted from a 16% stock (Polysciences Cat#18184). RNA FISH probes were designed against the H3 coding sequence as described previously (Hur *et al.*, 2020). Rabbits were immunized with a GST fusion protein containing the last 178 amino acids of *Drosophila* FLASH (Pocono Farms), and positive serum was confirmed by staining of embryonic HLBs that were identified using a CRISPR engineered allele of the *mxc* locus expressing GFP-Mxc (Hur *et al.*, 2020).

### *Drosophila* strains

Endogenously tagged MXC-mScarlet, and MXC-Apex2 strains were CRISPR engineered in a *yw* background as described previously for GFP-MXC (Hur *et al.*, 2020). Briefly, *yw; nos:cas9/CyO* flies were coinjected with a plasmid, pCFD3, encoding a gRNA that recognizes the 3′ prime end of the *mxc* gene near the stop codon and a repair template in pUC19 that includes the protein tag plus a linker sequence and a mutated PAM site. Both Apex2 and mScarlet were placed after the last *mxc* codon with a linker between the end of the *mxc* ORF and the inserted protein tag. The target sequence for the gRNA was GTAAACGATAACAACTTCAA-TGG and in the repair template the PAM site was mutated to TAA. The linker sequence used was CACCGGTATACCAGCTTGTACAAAAAGGCGGGAAGC-GCAGCGGCCCCATTTACT. Sequences for the repair templates can be found in Supplemental Figure 3.

### Embryo collections, fixation, and staining

Embryos 1–4 h old were collected on apple juice agar plates at 25°C. The embryos were placed in 1.5-ml tubes, dechorionated in 1.0 ml 50% bleach, washed in 1.0 ml embryo wash buffer, and then fixed in 1.0 ml 4% formaldehyde in PBS with 50% heptane for 20 min at room temperature on a nutator. The aqueous phase was removed and replaced with 0.5 ml of MeOH. The tubes were vigorously shaken for 30 s to remove the vitellin membranes from the embryos. The MeOH and heptane were removed, fresh 0.5 ml MeOH was added, and the tube was gently inverted three times. Then 0.5 ml PBST was added and embryos were nutated for 10 min at RT. PBST (0.5 ml) was added again to the embryos (total 1.5 ml) and nutated for 10 min at RT. After the embryos settled, the buffer was removed and 1.0 ml PBST was then added and nutated for 10 min. The PBST was removed and replaced with primary antibodies diluted into PBST and nutated overnight at 4°C. The embryos were then washed in PBST three times for 10 min at RT before addition of 1.0 ml secondary solution in PBST and nutation for 45 min at room temperature. Embryos were then washed three times in PBST, mounted in Vectashield, and sealed with CoverGrip.

### Imaging and image analysis

For Figure 1, standard confocal and Airyscan images were acquired on a Zeiss 880. SIM images were acquired on a Nikon NSIM equipped with an Andor iXon 897 camera. Live imaging of MXC-mScarlet embryos was performed on a Leica SP8 LIGHTNING system where the pinhole was set to 0.9 Airy units in order to increase the resolution, yet not increase collection times to the point of severe bleaching or risk large movements of the HLBs during collection. This allowed a complete Z-stack of HLBs in the field of view in less than 30 s. Then another time point was taken 30 s after the start of the previous. The zoom was also increased to 10× in order to decrease collection time. In all cases, both fixed and live, embryos were only imaged on the side closest to the coverslip and the nuclear cycle stage was determined by nuclear density. All stainings were repeated from at least three different embryo collections from the same genotype flies. For the P-PolII and RNA FISH staining, only embryos where we could visually see puncta in wide field were used for SIM imaging. FIJI was used to draw line scans across HLBs and measure the intensity of the respective fluorophore across the middle *z* section. Using FIJI, each channel was then overlaid onto a single graph and pseudocolored to match the image it was taken from. The thresholding for all images was adjusted in FIJI to the background fluorescence. A comparison of an image without the threshold to its counterpart with the threshold is included in Supplemental Figure 1. To quantify maximum-intensity spots within a surface in Imaris, the HLB signal from one channel was segmented as an enclosed surface while the other channels were all segmented as spots. The distances from the spot center to the nearest location on the rendered surface was measured. A negative measurement indicated that the spot was inside the segmented surface volume, and a positive measurement indicated that the spot was outside the surface volume. This process was repeated for each spot in each channel, including spots calculated from the channel used to render the surface volume. These measurements were imported into GraphPad and bar graphs showing individual data points were made with the mean highlighted using a red line.

### Protein binding assays

Preparation of nuclear extracts from *Drosophila* culture cells and immunoprecipitation protocols have been described (Sabath *et al.*, 2013). Binding assays were performed by labeling the desired fragment of Mxc protein by in vitro translation in rabbit reticulocyte

lysates, followed by addition of a GST-fusion protein, GST-FLASH178C, containing the C-terminal end of FLASH (Yang *et al.*, 2014). The bound proteins were recovered by binding to glutathione beads, analyzed by SDS-gel electrophoresis, and detected by autoradiography as previously described (Burch *et al.*, 2011).


## ACKNOWLEDGMENTS

We thank Tony Perdue of the UNC-CH Department of Biology microscopy core for his help with SIM and LSM 880 imaging, Ashlesha Chaubal for helpful comments on the manuscript, Joe Gall for anti-Lsm11 antibodies, and Mia Hoover for help screening MXC insertion alleles. This work was supported by NIH grant R01GM058921 to R.J.D. and W.F.M.


## REFERENCES


Adamson TE, Price DH (2003). Cotranscriptional processing of *Drosophila* histone mRNAs. Mol Cell Biol 23, 4046–4055.

Arias Escayola D, Neugebauer KM (2018). Dynamics and function of nuclear bodies during embryogenesis. Biochemistry 57, 2462–2469.

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

Boisvert FM, Hendzel MJ, Bazett-Jones DP (2000). Promyelocytic leukemia (PML) nuclear bodies are protein structures that do not accumulate RNA. J Cell Biol 148, 283–292.

Bongartz P, Schloissnig S (2019). Deep repeat resolution—the assembly of the *Drosophila* histone complex. Nucleic Acids Res 47, e18.

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.

Dominski Z, Yang XC, Marzluff WF (2005). The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. Cell 123, 37–48.

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

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.

Edgar BA, Schubiger G (1986). Parameters controlling transcriptional activation during early *Drosophila* development. Cell 44, 871–877.

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

Falahati H, Wieschaus E (2017). Independent active and thermodynamic processes govern the nucleolus assembly in vivo. Proc Natl Acad Sci USA 114, 1335–1340.

Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP (2016). Coexisting liquid phases underlie nucleolar subcompartments. Cell 165, 1686–1697.

Fox AH, Nakagawa S, Hirose T, Bond CS (2018). Paraspeckles: where long noncoding RNA meets phase separation. Trends Biochem Sci 43, 124–135.

Galganski L, Urbanek MO, Krzyzosiak WJ (2017). Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res 45, 10350–10368.

Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. Cell 124, 957–971.

Guglielmi B, La Rochelle N, Tjian R (2013). Gene-specific transcriptional mechanisms at the histone gene cluster revealed by single-cell imaging. Mol Cell 51, 480–492.

Hands KJ, Cuchet-Lourenco D, Everett RD, Hay RT (2014). PML isoforms in response to arsenic: high-resolution analysis of PML body structure and degradation. J Cell Sci 127, 365–375.

Hernandez-Verdun D, Roussel P, Thiry M, Sirri V, Lafontaine DL (2010). The nucleolus: structure/function relationship in RNA metabolism. Wiley Interdiscip Rev RNA 1, 415–431.

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 54, 379–394.

Isogai Y, Keles S, Prestel M, Hochheimer A, Tjian R (2007). Transcription of histone gene cluster by differential core-promoter factors. Genes Dev 21, 2936–2949.

Jiang L, Shao C, Wu QJ, Chen G, Zhou J, Yang B, Li H, Gou LT, Zhang Y, Wang Y, *et al.* (2017). NEAT1 scaffolds RNA-binding proteins and the microprocessor to globally enhance pri-miRNA processing. Nat Struct Mol Biol 24, 816–824.

Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10, 1930.

Lafarga M, Berciano MT, Pena E, Mayo I, Castano JG, Bohmann D, Rodrigues JP, Tavanez JP, Carmo-Fonseca M (2002). Clastosome: a subtype of nuclear body enriched in 19S and 20S proteasomes, ubiquitin, and protein substrates of proteasome. Mol Biol Cell 13, 2771–2782.

Lafontaine DLJ, Riback JA, Bascetin R, Brangwynne CP (2021). The nucleolus as a multiphase liquid condensate. Nat Rev Mol Cell Biol 22, 165–182.

Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, Mootha VK, Ting AY (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat Methods 12, 51–54.

Lang M, Jegou T, Chung I, Richter K, Munch S, Udvarhelyi A, Cremer C, Hemmerich P, Engelhardt J, Hell SW, Rippe K (2010). Three-dimensional organization of promyelocytic leukemia nuclear bodies. J Cell Sci 123, 392–400.

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.

Lanzotti DJ, Kupsco JM, Marzluff WF, Duronio RJ (2004). String(cdc25) and cyclin E are required for patterned histone expression at different stages of *Drosophila* embryonic development. Dev Biol 274, 82–93.

Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. Nat Cell Biol 9, 1167–1174.

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, Koreski KP (2017). Birth and death of histone mRNAs. Trends Genet 33, 745–759.

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.

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.

Nakano Y, Takahashi-Fujigasaki J, Sengoku R, Kanemaru K, Arai T, Kanda T, Murayama S (2017). PML nuclear bodies are altered in adult-onset neuronal intranuclear hyaline inclusion disease. J Neuropathol Exp Neurol 76, 585–594.

Nesic D, Tanackovic G, Kramer A (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. J Cell Sci 117, 4423–4433.

Nizami ZF, Deryusheva S, Gall JG (2010). Cajal bodies and histone locus bodies in *Drosophila* and *Xenopus*. Cold Spring Harb Symp Quant Biol 75, 313–320.

Rai AK, Chen JX, Selbach M, Pelkmans L (2018). Kinase-controlled phase transition of membraneless organelles in mitosis. Nature 559, 211–216.

Sabath I, Skrajna A, Yang XC, Dadlez M, Marzluff WF, Dominski Z (2013). 3′-End processing of histone pre-mRNAs in *Drosophila*: U7 snRNP is associated with FLASH and polyadenylation factors. RNA 19, 1726–1744.

Sahin U, Ferhi O, Jeanne M, Benhenda S, Berthier C, Jollivet F, Niwa-Kawakita M, Faklaris O, Setterblad N, de The H, Lallemand-Breitenbach V (2014). Oxidative stress-induced assembly of PML nuclear bodies controls sumoylation of partner proteins. J Cell Biol 204, 931–945.

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.
Sawyer IA, Bartek J, Dundr M (2018). Phase separated microenvironments inside the cell nucleus are linked to disease and regulate epigenetic state, transcription and RNA processing. Semin Cell Dev Biol.
Sawyer IA, Sturgill D, Dundr M (2019). Membraneless nuclear organelles and the search for phases within phases. Wiley Interdiscip Rev RNA 10, e1514.
Scherer M, Stamminger T (2016). Emerging role of PML nuclear bodies in innate immune signaling. J Virol 90, 5850–5854.
Shermoen AW, O'Farrell PH (1991). Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. Cell 67, 303–310.
Shin Y, Brangwynne CP (2017). Liquid phase condensation in cell physiology and disease. Science 357, 1253.
Skrajna A, Yang XC, Bucholc K, Zhang J, Hall TMT, Dadlez M, Marzluff WF, Dominski Z (2017). U7 snRNP is recruited to histone pre-mRNA in a FLASH-dependent manner by two separate regions of the stem-loop binding protein. RNA 23, 938–951.
Skrajna A, Yang XC, Dadlez M, Marzluff WF, Dominski Z (2018). Protein composition of catalytically active U7-dependent processing complexes assembled on histone pre-mRNA containing biotin and a photo-cleavable linker. Nucleic Acids Res 46, 4752–4770.
Souquere S, Beauclair G, Harper F, Fox A, Pierron G (2010). Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. Mol Biol Cell 21, 4020–4027.
Stanek D, Fox AH (2017). Nuclear bodies: news insights into structure and function. Curr Opin Cell Biol 46, 94–101.
Strzelecka M, Trowitzsch S, Weber G, Luhrmann R, Oates AC, Neugebauer KM (2010). Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. Nat Struct Mol Biol 17, 403–409.
Sun Y, Zhang Y, Aik WS, Yang XC, Marzluff WF, Walz T, Dominski Z, Tong L (2020). Structure of an active human histone pre-mRNA 3'-end processing machinery. Science 367, 700–703.
Tatomer DC, Rizzardi LF, Curry KP, Witkowski AM, Marzluff WF, Duronio RJ (2014). *Drosophila* Symplekin localizes dynamically to the histone locus body and tricellular junctions. Nucleus 5, 613–625.
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.
Tsai RY, Pederson T (2014). Connecting the nucleolus to the cell cycle and human disease. FASEB J 28, 3290–3296.
Weber K, Rathke PC, Osborn M (1978). Cytoplasmic microtubular images in glutaraldehyde-fixed tissue culture cells by electron microscopy and by immunofluorescence microscopy. Proc Natl Acad Sci USA 75, 1820–1824.
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.
West JA, Mito M, Kurosaka S, Takumi T, Tanegashima C, Chujo T, Yanaka K, Kingston RE, Hirose T, Bond C, *et al*. (2016). Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. J Cell Biol 214, 817–830.
Wheeler JR, Matheny T, Jain S, Abrisch R, Parker R (2016). Distinct stages in stress granule assembly and disassembly. Elife 5, e18413.
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.
Woulfe JM, Prichett-Pejic W, Rippstein P, Munoz DG (2007). Promyelocytic leukaemia-immunoreactive neuronal intranuclear rodlets in the human brain. Neuropathol Appl Neurobiol 33, 56–66.
Yang XC, Burch BD, Yan Y, Marzluff WF, Dominski Z (2009). FLASH, a pro-apoptotic 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, Debski J, Kaus-Drobek M, Dadlez M, Marzluff WF, Dominski Z (2013). A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. Mol Cell Biol 33, 28–37.
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.
Yao RW, Xu G, Wang Y, Shan L, Luan PF, Wang Y, Wu M, Yang LZ, Xing YH, Yang L, Chen LL (2019). Nascent Pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. Mol Cell 76, 767–783.e711.
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.